

Сентимент и стадное поведение частных инвесторов: кластерный анализ российского фондового рынка

Максим Сергеевич Файзулин

E-mail: faizulin.maxi@yandex.ru, ORCID: 0000-0003-3273-9005

Национальный исследовательский университет

«Высшая школа экономики», г. Москва, Российская Федерация

Аннотация

В данной работе сентимент (настроение) частных инвесторов и расхождение мнений пользователей инвестиционных онлайн-платформ анализируются как факторы возникновения стадного поведения на фондовом рынке на разных кластерах, сформированных на основе биржевой информации по акциям российских эмитентов. Актуальность исследования заключается в определении степени расхождения мнений среди участников рынка и их влияния на поведение различных групп инвесторов на российском фондовом рынке в зависимости от принятия инвесторами рисков. На основе данных российского фондового рынка за период с 2019 по 2023 г. было проанализировано 66 обсуждаемых российских акций. Для проведения сентимент-анализа применен алгоритм сбора и автоматической классификации текстовых массивов за счет методов машинного обучения, а также построены в качестве прокси настроения частных инвесторов метрики логарифмического сентимента и общерыночного расхождения мнений по обсуждаемым акциям. Тестирование гипотез исследования базируется на использовании методов кластеризации и квантильного регрессионного анализа. В итоге определена значимая роль расхождения мнений пользователей онлайн-платформ в формировании стадного поведения частных инвесторов по акциям с меньшей средней доходностью и волатильностью. Также на основе полученных результатов обнаружено, что кластеризация активов позволяет определить отрицательную связь расхождения мнений и стадного поведения по акциям с более высоким уровнем волатильности и доходности. Тестирование гипотезы о влиянии общего уровня настроений инвесторов позволило определить, что именно расхождение мнений среди участников рынка является важным фактором в объяснении стадного поведения по активам обсуждаемых компаний.

Ключевые слова: стадное поведение, дивергенция мнений, сентимент инвесторов, российский фондовый рынок

JEL: G41, G11

Для цитирования: Файзулин М. С. Сентимент и стадное поведение частных инвесторов: кластерный анализ российского фондового рынка // Финансовый журнал. 2024. Т. 16. № 4. С. 95–113. <https://doi.org/10.31107/2075-1990-2024-4-95-113>.

© Файзулин М. С., 2024

<https://doi.org/10.31107/2075-1990-2024-4-95-113>

Sentiment and Herd Behavior of Private Investors: A Cluster Analysis of the Russian Stock Market

Maxim S. Faizulin

HSE University, Moscow, Russian Federation

faizulin.maxi@yandex.ru, <https://orcid.org/0000-0003-3273-9005>

Abstract

In this paper, the sentiment of private investors and the divergence of opinions of users of online investment platforms are analyzed as factors in the emergence of herd behavior on the stock market in different clusters. This clusters are formed on the basis of stock exchange information on shares of Russian issuers. The relevance of the study lies in determining the significance of consensus periods and their impact on the behavior of different groups of investors on the Russian stock market depending on different levels of risk taking by investors. Based on the Russian stock market data for the period from 2019 to 2023, 66 discussed Russian stocks were analyzed. To conduct sentiment analysis, an algorithm was applied to collect and automatically classify textual data using machine learning methods. As proxies for private investor sentiment, metrics of logarithmic sentiment and market-wide divergence of opinions on the stocks under discussion were constructed. The testing of the research hypotheses was based on the implementation of clustering and quantile regression analysis methods. As a result, the significant role of divergence of opinions of Internet users in the formation of herd behavior of private investors on shares with lower average return and risk level was determined. It was also found that asset clustering helps to determine the opposite behavior of investors on stocks with higher volatility level. Consensus of opinions in the market is a signal to deviate from the general market trend. Finally, the test of the hypothesis about the significant influence of investor sentiment did not provide sufficient evidence that this kind of sentiment is important for determining herd behavior in the market.

Keywords: *herd behavior, divergence of opinions, investor sentiment, Russian stock market*

JEL: *G41, G11*

For citation: *Faizulin M.S. (2024). Sentiment and Herd Behavior of Private Investors: A Cluster Analysis of the Russian Stock Market. Financial Journal, 16 (4), 95–113 (In Russ.).*

<https://doi.org/10.31107/2075-1990-2024-4-95-113>.

© Faizulin M.S., 2024

ВВЕДЕНИЕ

Инвестиционное поведение участников на финансовых рынках связано со многими факторами, которые влияют на принятие решений на рынке. Такие факторы играют важную роль в формировании ликвидности, уровня риска и доходности как отдельных активов, так и рынка в целом. Отдельные группы инвесторов подвержены следованию настроениям на рынке, полагаясь на мнение и опыт других рыночных игроков. Это формирует стадное поведение частных инвесторов, которое выражается в их подражании друг другу при формировании торговой активности по различным акциям.

Одним из важных факторов в формировании стадного поведения является настроение частных инвесторов. В ряде исследований авторы доказывают, что настроение частных инвесторов способно объяснять не только динамику доходности активов, но и уровень стадного поведения [Vieira, Pereira, 2015; Economou et al., 2018; Choi, Yoon, 2020; Li et al., 2023].

Особенность настоящей работы заключается в изучении настроения частных инвесторов как фактора, который способствует возникновению кластерного стадного поведения

участников рынка. Измерение настроения разделяется на две категории: учет общего уровня рыночных настроений и учет расхождения мнений среди пользователей инвестиционных онлайн-платформ Tinkoff Pulse, MFD и SmartLab. Размещенные на этих платформах пользовательские сообщения по российским акциям были извлечены и классифицированы на основе Stacking-ансамбля по трем тональностям. Это позволило построить метрики сентимента как прокси настроения частных инвесторов. Для более детального изучения феномена стадности была произведена кластеризация акций с использованием нескольких имплементаций алгоритма K-means и протестированы гипотезы исследования по методике квантильного анализа.

ОБЗОР ЛИТЕРАТУРЫ

Активное изучение стадного поведения инвесторов началось с развитием информационных технологий в 2000-х гг. и продолжилось во время мирового финансового кризиса в 2008 г. В первых исследованиях отмечается, что «стадность» может быть вызвана как рациональным, так и иррациональным поведением инвесторов [Bikhchandani, Sharma, 2000]. Рациональность поведения объясняется стремлением инвесторов поддерживать свою репутацию, доказывая свою позицию на рынке как верную. Это приводит к стадному поведению других участников рынка, которые верят, что следуют за опытным гуру, и ожидают получить выигрыш в будущем [Prendergast and Stole, 1996; Trueman, 1994]. В то же время иррациональное поведение инвесторов объясняется тем, что отдельные группы участников рынка не анализируют фундаментальную информацию об активе и просто следуют за динамикой рынка, ожидая, что вся информация об активе уже заложена в этой динамике [De Long et al., 1990; Froot et al., 1992]. Такое поведение также объясняется памятью инвесторов, которые на основе своего опыта знают, что некоторые акции той или иной компании могут продолжить рост, что и побуждает инвесторов по-прежнему покупать такие активы. Этот феномен был объяснен психологическими факторами — принятием желаемого за действительное, то есть веры в будущий рост рынка без учета каких-либо сопутствующих рисков [Bikhchandani, Sharma, 2000].

В дальнейшем были проведены исследования по возникновению явления стадности в экстремальные периоды, такие как финансовые кризисы, резкий рост или падение уровня волатильности фондовых рынков, изменение ликвидности [Espinosa-Méndez, Arias, 2020; Duygun et al., 2021; Ferreruela, Mallor, 2021; Rubesam, Gerson de Souza, 2022]. В данных исследованиях стадное поведение измеряется как разность между динамикой ожидаемой доходности портфеля и фактической доходностью рынка. Если разрыв в динамике минимизируется, то утверждается, что уровень стадного поведения становится выше.

Так, в работе [Rubesam, Gerson de Souza, 2022] изучается стадное поведение инвесторов во время пандемии COVID-19 на десяти различных фондовых рынках, включая США, Австрию, Китай, Францию и другие страны. Особенностью указанного исследования является объяснение стадности участников рынка с учетом не только биржевой информации, но и макроэкономических показателей. Авторы обнаружили, что лишь по рынкам Италии, Швеции и США присутствуют признаки стадности. Стадность инвесторов во время пандемии отрицательно связана с более жесткими правительственными мерами, ограничивающими мобильность граждан, и положительно связана с мерами экономической поддержки.

Анализ стадного поведения рыночных участников на фондовых рынках Испании и Португалии в моменты кризисных ситуаций проводится также в исследовании [Ferreruela, Mallor, 2021]. Авторы доказывают, что стадное поведение проявляется с большей интенсивностью в периоды до финансового кризиса, исчезает во время кризиса, а потом вновь появляется, хотя и с меньшей интенсивностью. При анализе полного периода пандемии

COVID-19 такое поведение не обнаруживается. Отмечается, что выделение дней с высокой рыночной волатильностью позволяет фиксировать стадное поведение во время кризиса COVID-19.

Однако такие исследования не включают изучение настроений частных инвесторов, которые могут иметь значимое влияние на объяснение формирования стадного поведения. Авторы используют в качестве прокси настроений инвесторов биржевые данные, что может стать проблемой упущения возможных факторов, на которые участники рынка обращают свое внимание. Это ставит под вопрос решение проблемы эндогенности в моделях, которые были построены авторами предыдущих исследований.

Один из примеров взаимосвязи настроения и стадного поведения инвесторов представлен в работе [Filip, Pochea, 2023], где была выявлена и описана зеркальная симметрия между стадностью и индексами настроений на фондовых рынках США и Европы. Авторы обнаружили, что в периоды потрясений индексы настроений отрицательны, а стадность на рынке растет. То есть следование инвесторов за рыночной динамикой становится более выраженным в периоды пессимизма. Это объясняется тем, что в такие периоды инвесторы избегают риска падения рыночной доходности активов и готовы принимать мнение большинства на рынке. В спокойные периоды степень стадности снижается, а восприятие инвесторов становится оптимистичным. Это объясняется уверенностью инвесторов в продолжении рыночного тренда и снижении рисков роста волатильности цен на различные финансовые активы. Однако авторы определяют сентимент инвесторов как условные ожидания, рассчитанные на основе избыточной доходности. Такой подход не принимает во внимание реальных мнений участников рынка, которые характеризуются не только их активностью на рынке, но и склонностью учитывать мнение инвестиционного общества в онлайн-среде.

Исследование взаимосвязи между стадным поведением и настроением инвесторов проводится также на китайском фондовом рынке. Так, в работе [Li et al., 2023] авторы использовали сентимент-метрику в качестве прокси измерения настроения частных инвесторов. Для ее построения были собраны текстовые сообщения из онлайн-платформы Sina Weibo и автоматически размечены при помощи модели рекуррентной нейронной сети LSTM. На основе результатов анализа авторы отмечают, что учет количества пользователей на онлайн-платформе является важным критерием в определении стадного поведения инвесторов. Например, когда число подписчиков больше 10 000, но не превышает 1 000 000, частота появления новых сообщений пользователей в социальных сетях способствует увеличению уровня стадного поведения инвесторов на китайском фондовом рынке.

Другим примером изучения взаимосвязи между сентиментом инвесторов и их стадным поведением является работа по анализу данной проблемы на фондовом рынке Южной Кореи [Yoon, Oh, 2022]. Авторы разработали собственную метрику измерения сентимента рыночных участников на основе постов, публикуемых пользователями на онлайн-платформе Naver Financial. Результаты исследования показали, что более положительный сентимент является важным фактором в формировании стадного поведения у инвесторов. Это объясняется тем, что частные инвесторы чувствительны больше к потерям, чем к выигрышам и склонны доверять мнению большинства в моменты роста рынка.

Однако недостатком таких работ является малое количество инструментов для измерения сентимента частных инвесторов. Так как рыночная среда подвержена воздействию со стороны аналитиков с различным уровнем инвестиционного опыта, то могут возникать эффекты конфликта мнений, когда среди пользователей онлайн-платформ наблюдается расхождение позиций по отдельным активам. Весьма вероятно, что на рынке могут присутствовать различные группы инвесторов, которые отличаются разной степенью принятия риска в ситуациях, когда консенсус среди участников рынка находится на низком или высоком уровнях.

МЕТОДОЛОГИЯ И ДАННЫЕ

В данном исследовании тестируется несколько гипотез по изучению настроения инвесторов и их поведения на фондовом рынке.

H1. Общий уровень настроения частных инвесторов является важным фактором в формировании стадного поведения на российском фондовом рынке.

H2. Инвесторы склонны следовать за мнением большинства во время высокой степени расхождения мнений среди участников рынка.

H3. Кластеризация акций позволяет определить группы активов, по которым частные инвесторы готовы принимать на себя дополнительный риск в моменты высокой общерыночной дивергенции мнений.

Для измерения уровня настроений используются метрики общерыночной дивергенции (расхождения) мнений и сентимента частных инвесторов. Помимо этого, учитываются не только все акции, по которым идут обсуждения на инвестиционных онлайн-платформах (Tinkoff Pulse, MFD и SmartLab), но и кластеры таких активов, с помощью которых возможно определить разность силы влияния общерыночного сентимента и дивергенции мнений частных инвесторов. На данный момент не наблюдается подобных исследований в изучении вопроса дивергенции мнений и стадности по отдельным группам активов. Настоящее исследование заполняет выявленные исследовательские пробелы с целью добавления научного знания в область эмпирических работ по анализу стадного поведения частных инвесторов.

В прикладных работах по изучению стадного поведения участников рынка используются различные методы измерения стадности инвесторов, например на основе модели LSV [Lakonishock et al., 1992], изменения стоимости портфеля (PCM), стандартного отклонения (CSSD) и абсолютной разницы между доходностью портфеля и рынком (CSAD). Однако в последнее время в исследованиях все чаще встречается использование методики на основе метрики CSAD [Li et al., 2023; Filip, Pochea, 2023], так как другие методики ограничиваются предпосылками о том, что инвесторы уже имеют доступ к высокочастотным данным и к более широкой информации о рыночной ситуации в целом. Методика измерения стадности поведения в классическом варианте представляет собой среднее значение разности между модулем доходности портфеля и модулем доходности рынка в целом. Но в текущем исследовании применяется метод расчета без учета модуля для фиксации отрицательной динамики доходности активов и рынка. Это позволяет учитывать две стороны (отрицательную и положительную) отклонения поведения частных инвесторов от общерыночной динамики. Методика расчета доходности активов и рынка представлена в формуле 1, методика расчета метрики CSAD – в формуле 2.

$$R_t = \ln \left(\frac{\text{Close Price}_t}{\text{Close Price}_{t-1}} \right), \quad (1)$$

$$\text{CSAD}_t = \frac{1}{n} \sum_{i=1}^n (R_{i,t} - R_{m,t}), \quad (2)$$

где Close Price_t – цена актива / значение фондового индекса на момент закрытия биржи; R_t – дневная доходность в момент времени t ; n – общее количество акций в выборке; $R_{i,t}$ – доходность i -го актива в момент времени t ; $R_{m,t}$ – рыночная доходность в момент времени t . В текущем исследовании в качестве рыночной доходности выступает изменение индекса Мосбиржи (IMOEX).

Всего в выборку исследования вошло 66 акций российских эмитентов (см. Приложение 1), по которым шли обсуждения на трех инвестиционных онлайн-платформах: Tinkoff Pulse, MFD, SmartLab. Отобранные платформы являются наиболее востребованными среди русскоязычной аудитории, интересующейся инвестиционной темой. Tinkoff Pulse – самая молодая и популярная платформа, которая открылась осенью 2019 г.

Поэтому анализируемые данные по сентименту, активам и рынку охватывают период с 07.10.2019 по 28.11.2023.

Всего с трех онлайн-платформ было собрано более 4 млн текстовых сообщений на основе разработанных парсеров. Для решения задачи классификации текстовых данных были сформированы три группы тональностей: положительная (1) — если сообщение носит характер одобрения фондового актива; нейтральная (0) — если сообщение не относится к активу или имеет равновесное значение оптимизма и пессимизма одновременно; негативная (−1) — если сообщение характеризует явное неодобрение анализируемого актива. С целью векторизации и токенизации текстовых данных был использован алгоритм TF-IDF (term frequency, inverse document frequency). После предобработки текстовых сообщений для их классификации было отобрано 1 942 831 сообщение (Tinkoff Pulse: 765 032; MFD: 823 693; SmartLab: 354 106). Для классификации текстовых данных обучался Stacking-ансамбль, состоящий из моделей машинного обучения двух уровней. В данном ансамбле предсказания моделей первого уровня передаются на вход для обучения мета-классификатора (модель второго уровня). В качестве моделей первого уровня используются алгоритмы машинного обучения: Random Forest, XGBoost, AdaBoost, Linear SVC, Logistic Regression и Naïve Bayes. Лучшим вариантом мета-классификатора была выбрана модель логистической регрессии. Полученный результат классификации на тестовой выборке показал точность на уровне 61,68%, что превысило показатель точности предобученных моделей RuBERT и FinBERT.

На основе полученных результатов классификации текстовых данных для измерения общего уровня настроений частных инвесторов используется логарифм отношения положительных сообщений к отрицательным (см. формулу 3). В ряде исследований [Sul et al., 2017; Wu et al., 2014] такой метод признается наиболее устойчивым вариантом измерения сентимента частных инвесторов. Также в текущем исследовании рассчитывается среднерыночный сентимент S_mkt_t по компаниям в выборке (формула 4).

$$S_log_{i,t} = \ln \frac{1+N_{i,t}^{pos}}{1+N_{i,t}^{neg}}, \quad (3)$$

$$S_mkt_t = \frac{1}{n} \sum_{i=1}^n \ln \frac{1+N_{i,t}^{pos}}{1+N_{i,t}^{neg}}, \quad (4)$$

где $S_log_{i,t}$ — сентимент i -й компании в момент времени t ; $N_{i,t}^{pos}$ — количество позитивных сообщений по компании i в момент времени t ; $N_{i,t}^{neg}$ — количество негативных сообщений по компании i в момент времени t ; n — общее количество акций в выборке.

С целью измерения уровня консенсуса в рыночной среде рассчитывается показатель дивергенции мнений, который основывается на разности мнений пользователей российских онлайн-платформ (формула 5). Также для определения общего уровня разногласий между различными группами участников обсуждений вычисляется среднерыночный уровень дивергенции мнений (формула 6).

$$DivO_{i,t} = \frac{1+N_{i,t}^{pos}}{N_{i,t}^{pos}+N_{i,t}^{neg}} - \frac{1+N_{i,t}^{neg}}{N_{i,t}^{pos}+N_{i,t}^{neg}}, \quad (5)$$

$$DivO_mkt_t = \frac{1}{n} \sum_{i=1}^n \frac{1+N_{i,t}^{pos}}{N_{i,t}^{pos}+N_{i,t}^{neg}} - \frac{1+N_{i,t}^{neg}}{N_{i,t}^{pos}+N_{i,t}^{neg}}. \quad (6)$$

Однако при учете общего уровня сентимента и разногласий может наблюдаться проблема пропущенных данных по причине выходных или праздничных дней, когда фондовая биржа не работает. В таком случае происходит смещение оценки влияния сентимента частных инвесторов, так как не учитывается их настроение за пропущенные периоды

времени. По этой причине производится модификация метрик $S_mkt_{i,t}$ и $DivO_mkt_{i,t}$ за счет учета предыдущих настроений по отобранным активам. Данная модификация основывается на использовании пятидневного временного лага с фиксацией веса 20% для каждого прошлого торгового дня на скользящем окне длиной пять дней. С учетом того, что прошлые настроения могут оказывать более слабое влияние на доходность акций, производится коррекция весов путем умножения их на экспоненциальную константу, соответствующую наблюдению t , как показано в формулах 7 и 8:

$$ES_mkt_t = S_mkt_t + \sum_{i=1}^5 e(0.2)^{-i} S_mkt_{t-i}, \quad (7)$$

$$EDivO_mkt_t = DivO_mkt_t + \sum_{i=1}^5 e(0.2)^{-i} DivO_mkt_{t-i}. \quad (8)$$

Для кластеризации российских акций используется алгоритм кластеризации K-means с несколькими модификациями: с выбором наибольшего кластера (LC Bisecting K-means) и с выбором кластера с наибольшим значением критерия инерции (BI Bisecting K-means). Эти два дополнительных алгоритма позволяют итеративно выбирать центры кластеров на основе прошлой итерации путем разделения предыдущего кластера на два новых, пока не будет достигнуто целевое количество кластеров. Стандартный принцип алгоритма K-means заключается в создании таких центров заданных кластеров в самом начале алгоритма в случайном порядке [Arthur, Vassilvitskii, 2007].

С целью определения оптимального количества кластеров C_i производится разделение x_i наблюдений на неравные кластеры, каждый из которых описывается средним значением μ_i . Средние значения принято называть центроидами. Алгоритм K-means стремится выбрать такие центроиды, которые минимизируют критерий инерции (формула 9):

$$\sum_{i=0}^n \min_{\mu_i \in C} (\|x_i - \mu_i\|^2). \quad (9)$$

Основным правилом для определения оптимального числа кластеров является «метод локтя», при котором снижение суммы ошибок внутри кластеров становится менее выраженным.

После определения оптимального числа кластеров по отобранным акциям требуется проверить предположение о том, что стадное поведение не является причиной возникновения дивергенции мнений или изменения общерыночного сентимента частных инвесторов. Для этого проводится тест Грейнджера на причинность на основе VAR-модели с числом лагов, равным двум (формулы 10–12). Для каждого уравнения и каждой эндогенной переменной, которая не является зависимой переменной в модели VAR, производится тест Вальда на равенство нулю всех коэффициентов эндогенных переменных. Для каждого уравнения в модели VAR в качестве эндогенных и зависимых переменных выступают метрики настроений частных инвесторов и уровня стадности на рынке.

$$y_t = c_1 + \sum_{i=1}^2 \alpha_{1,i} y_{t-i} + \sum_{i=1}^2 \beta_{1,i} x_{t-i} + \sum_{i=1}^2 \gamma_{1,i} z_{t-i} + \epsilon_{x,t}, \quad (10)$$

$$x_t = c_2 + \sum_{i=1}^2 \alpha_{2,i} y_{t-i} + \sum_{i=1}^2 \beta_{1,i} x_{t-i} + \sum_{i=1}^2 \gamma_{1,i} z_{t-i} + \epsilon_{x,t}, \quad (11)$$

$$z_t = c_2 + \sum_{i=1}^2 \alpha_{2,i} y_{t-i} + \sum_{i=1}^2 \beta_{1,i} x_{t-i} + \sum_{i=1}^2 \gamma_{1,i} z_{t-i} + \epsilon_{x,t}. \quad (12)$$

Также для оценки влияния сентимента и дивергенции мнений на стадное поведение частных инвесторов в данной работе используется квантильный регрессионный анализ (формула 13), на основе которого производится сравнение со стандартной регрессионной моделью, оцененной на основе МНК (формула 14).

$$Q_{CSAD_t(\theta|\sum_{i=1}^8\beta_i\mu_{i,t})} = \alpha + \beta_1ES_mkt_t + \beta_2EDivO_mkt_t + \beta_3EDivO_mkt_t^2 + \beta_4D_t + \sum_{i=1}^3\mu_{i,t} + \beta_8CSAD_{t-1} + \varepsilon_{i,t}, \quad (13)$$

$$CSAD_t = \alpha + \beta_1ES_mkt_t + \beta_2EDivO_mkt_t + \beta_3EDivO_mkt_t^2 + \beta_4D_t + \sum_{i=1}^3\mu_{i,t} + \beta_8CSAD_{t-1} + \varepsilon_{i,t}, \quad (14)$$

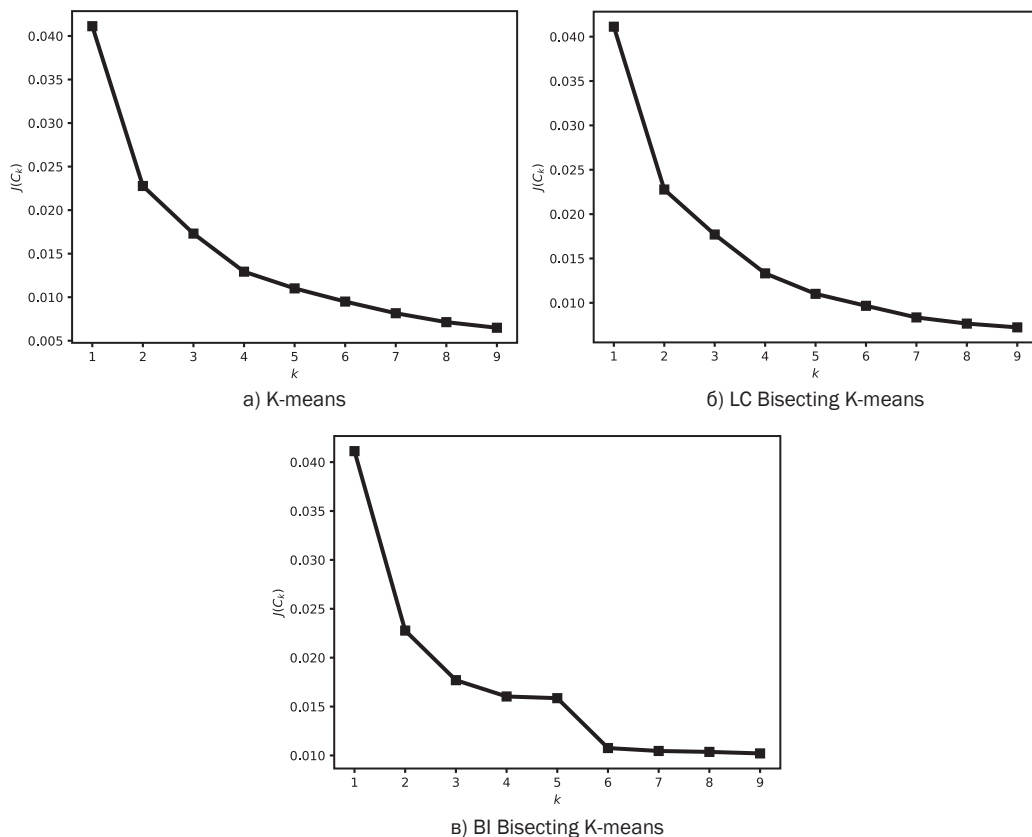
где θ – квантили на уровне 0,25; 0,50; 0,75; $\tau_{i,t}$ – вектор контрольных переменных; $\mu_{i,t}$ – вектор контрольных экзогенных переменных изменения значения индекса Мосбиржи (IMOEX), индекса волатильности RVI и уровня стресса АКРА на российском фондовом рынке; D – дамми-переменная динамики рынка (1 – если рыночная доходность положительна в момент времени t, 0 – в остальных случаях). Также производится контроль зависимой переменной за счет ее лага $CSAD_{t-1}$, чтобы убедиться, что полученные результаты не являются ложным пересчетом потенциальной автокорреляции в зависимой переменной.

РЕЗУЛЬТАТЫ

С целью поиска оптимального числа кластеров по анализируемым акциям производится применение трех версий алгоритма K-means на девяти вариантах в зависимости от количества кластеров (рис. 1).

Рисунок 1

Итеративный подход в определении оптимального числа кластеров



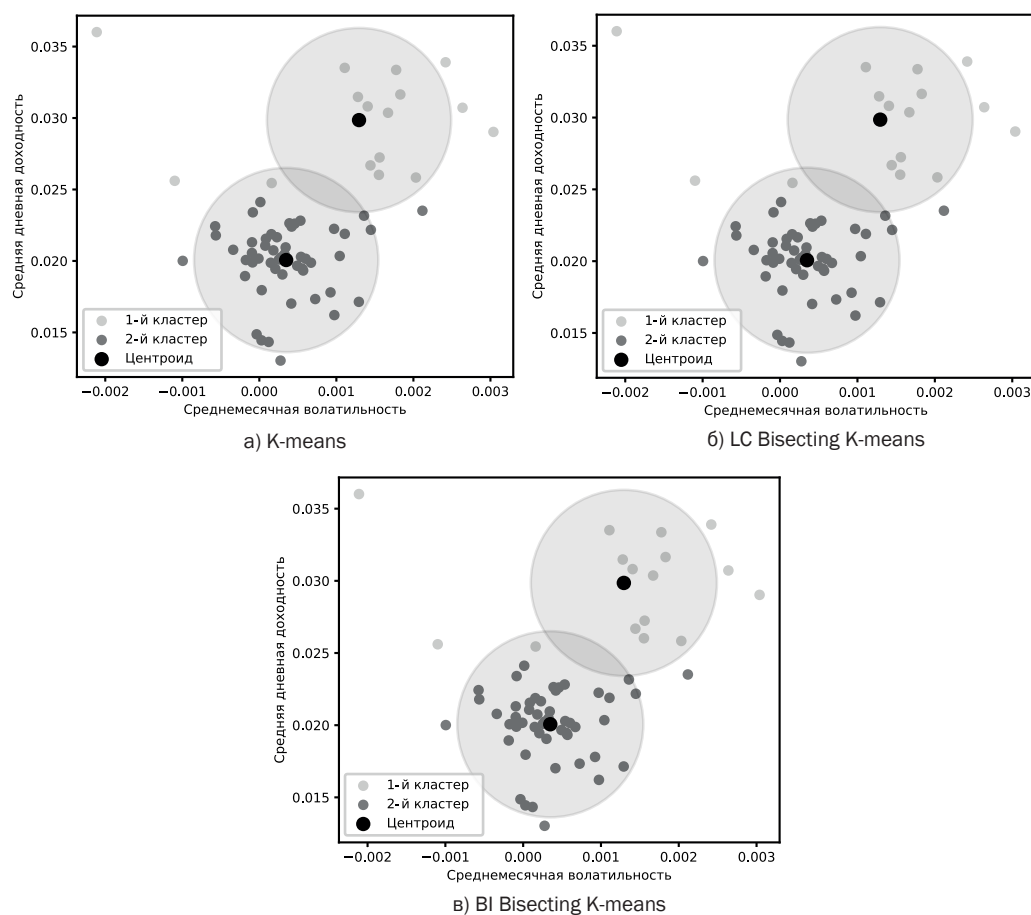
Источник: рисунок автора.

Исходя из представленных результатов девяти итераций кластеризации российских акций, можно сделать вывод, что наиболее оптимальным вариантом является разделение 66 активов на две группы. Такой подбор количества кластеров обосновывается «методом локтя» на основе резкого сокращения суммы ошибок внутри каждого кластера как явный сигнал к выбору числа кластеров. Выбор большего числа кластеров, несмотря на постепенное снижение ошибки в кластерах, может привести к смещениям в оценке параметров регрессионных моделей по причине слишком частого разделения выборочных данных и малого числа наблюдений в каждом из множеств кластеров.

По всем трем версиям алгоритмов были получены одинаковые размеры кластеров и относящихся к ним активов (рис. 2). Так, по первому кластеру наблюдаются более высокие среднегодовые доходности и уровни волатильности акций. В то же время во второй кластер попали активы с менее высокими значениями доходности акций и волатильности их цен. Однако отдельные активы находились на достаточно большом расстоянии от центроидов. В таком случае подобные выбросы оставались в выборке в том кластере, к которому был изначально отнесен актив.

Рисунок 2

Результаты кластеризации акций российских эмитентов



Источник: рисунок автора.

В данном исследовании анализируются два неравных по размерам кластера с 16 и 50 акциями компаний соответственно. Для тестирования наличия значимых различий между

отобранными группами акций проводится непараметрический тест Манна – Уитни на основе расчета U-критерия (табл. 1).

Таблица 1

Оценка различий между двумя кластерами российских акций

Кластер	Средняя дневная доходность	Среднемесячная волатильность
1-й кластер (16 акций)	0,0013	0,0299
2-й кластер (50 акций)	0,0003	0,0200
Критерий Манна – Уитни (U, p-value)	651,5000 (0,0002)	800,0000 (0,0000)

Источник: расчеты автора.

Благодаря разделению основной выборки активов на две группы вводятся две дополнительные метрики стадного поведения частных инвесторов: CSAD_0 (для кластера с 16 акциями) и CSAD_1 (для кластера с 50 акциями). Также добавляется обозначение метрики общерыночной стадности – CSAD_All (для всех 66 акций). Описательная статистика всех переменных представлена в табл. 2, а корреляционная матрица – в Приложении 2.

Таблица 2

Описательная статистика

Переменная	Наблюдения	Среднее	Стандартное отклонение	Минимум	Максимум
CSAD_All	1030	0,000	0,024	-0,326	0,502
CSAD_0	1030	0,001	0,028	-0,374	0,555
CSAD_1	1030	0,000	0,044	-0,487	0,801
ES_mkt	1030	2,729	0,798	0,615	5,544
EDivO_mkt	1030	1,289	0,346	0,306	2,373
EDivO_mkt ²	1030	1,782	0,905	0,093	5,629
IMOEX	1030	0,000	0,020	-0,405	0,183
Direction	1030	0,553	0,497	0,000	1,000
AKRA	1030	1,007	0,123	0,457	3,370
RVI	1030	1,002	0,067	0,718	1,469

Источник: расчеты автора.

Для исключения причинной связи между прошлыми значениями стадного поведения и текущим настроением частных инвесторов проводится тест Грейнджера на причинность, результаты которого представлены в табл. 3. Тестирование гипотезы о том, что эндогенные факторы являются значимыми в объяснении зависимой переменной, производится на основе теста Вальда.

Таблица 3

Результаты теста Грейнджера на причинность

Объясняемая переменная	Эндогенные переменные	chi2	p-value
CSAD_All	ES_mkt	13,743	0,001
CSAD_All	EDivO_mkt	12,823	0,002
CSAD_All	EDivO_mkt ²	15,162	0,001
CSAD_All	ES_mkt, EDivO_mkt, EDivO_mkt ²	38,647	0,000
ES_mkt	CSAD_All	0,994	0,608
ES_mkt	EDivO_mkt	9,261	0,010
ES_mkt	EDivO_mkt ²	3,908	0,142
ES_mkt	CSAD_All, EDivO_mkt, EDivO_mkt ²	10,067	0,122
EDivO_mkt	CSAD_All	1,361	0,506
EDivO_mkt	ES_mkt	9,017	0,011

Объясняемая переменная	Эндогенные переменные	chi2	p-value
EDivO_mkt	EDivO_mkt ²	4,557	0,102
EDivO_mkt	ES_mkt, CSAD_All, EDivO_mkt ²	10,363	0,110
EDivO_mkt ²	CSAD_All	0,43	0,807
EDivO_mkt ²	ES_mkt	13,196	0,001
EDivO_mkt ²	EDivO_mkt	13,304	0,001
EDivO_mkt ²	ES_mkt, CSAD_All, EDivO_mkt	14,088	0,029
CSAD_0	ES_mkt	12,857	0,002
CSAD_0	EDivO_mkt	11,812	0,003
CSAD_0	EDivO_mkt ²	14,036	0,001
CSAD_0	ES_mkt, EDivO_mkt, EDivO_mkt ²	28,315	0,000
ES_mkt	CSAD_0	0,524	0,769
ES_mkt	EDivO_mkt	8,944	0,011
ES_mkt	EDivO_mkt ²	3,713	0,156
ES_mkt	CSAD_0, EDivO_mkt, EDivO_mkt ²	9,590	0,143
EDivO_mkt	CSAD_0	0,809	0,667
EDivO_mkt	ES_mkt	8,691	0,013
EDivO_mkt	EDivO_mkt ²	4,315	0,116
EDivO_mkt	ES_mkt, CSAD_0, EDivO_mkt ²	9,803	0,133
EDivO_mkt ²	CSAD_0	0,119	0,942
EDivO_mkt ²	ES_mkt	12,955	0,002
EDivO_mkt ²	EDivO_mkt	13,030	0,001
EDivO_mkt ²	ES_mkt, CSAD_0, EDivO_mkt	13,770	0,032
CSAD_1	ES_mkt	11,717	0,003
CSAD_1	EDivO_mkt	10,391	0,006
CSAD_1	EDivO_mkt ²	16,811	0,000
CSAD_1	ES_mkt, EDivO_mkt, EDivO_mkt ²	85,016	0,000
ES_mkt	CSAD_1	2,228	0,328
ES_mkt	EDivO_mkt	9,100	0,011
ES_mkt	EDivO_mkt ²	4,142	0,126
ES_mkt	CSAD_1, EDivO_mkt, EDivO_mkt ²	11,32	0,079
EDivO_mkt	CSAD_1	2,489	0,288
EDivO_mkt	ES_mkt	9,019	0,011
EDivO_mkt	EDivO_mkt ²	4,778	0,092
EDivO_mkt	ES_mkt, CSAD_1, EDivO_mkt ²	11,508	0,074
EDivO_mkt ²	CSAD_1	1,478	0,478
EDivO_mkt ²	ES_mkt	13,506	0,001
EDivO_mkt ²	EDivO_mkt	13,454	0,001
EDivO_mkt ²	ES_mkt, CSAD_1, EDivO_mkt	15,161	0,019

Источник: расчеты автора.

На основе результатов проведенного теста следует подтвердить предположение о том, что настроение частных инвесторов является причиной общерыночного и кластерного стадного поведения. Однако такое поведение частных инвесторов не следует считать значимым в объяснении будущего настроения рыночных участников. При этом наблюдается сильная статистическая значимость при включении всех трех метрик сентимента с целью объяснения будущего стадного поведения инвестора. Это доказывает целесообразность одновременного включения таких факторов в модели для тестирования гипотез исследования.

Стадное поведение по всем обсуждаемым российским акциям в целом слабо зависит от сентимента частных инвесторов. Их общий уровень настроений не позволяет подтвердить гипотезу H1 без разделения активов на отдельные кластеры (табл. 4).

Результаты анализа стадного поведения по всей группе активов

Переменная	Объясняемая переменная: CSAD_All			
	q = 0,25	q = 0,50	q = 0,75	МНК
	Коэффициент (Ст. ошибка)	Коэффициент (Ст. ошибка)	Коэффициент (Ст. ошибка)	Коэффициент (Ст. ошибка)
const.	0,050 *** (0,014)	0,040 *** (0,012)	0,053 *** (0,012)	0,088 *** (0,016)
ES_mkt	-0,002 (0,014)	0,003 (0,009)	-0,003 (0,011)	-0,009 (0,010)
EDivO_mkt	0,004 (0,027)	-0,009 (0,014)	-0,014 (0,020)	0,012 (0,017)
EDivO_mkt ²	0,000 (0,004)	0,002 (0,004)	0,009 ** (0,004)	0,005 (0,004)
D	0,002 (0,002)	0,001 (0,002)	0,001 (0,003)	0,003 (0,003)
IMOEX	-0,688 *** (0,083)	-0,704 *** (0,108)	-0,792 *** (0,173)	-0,923 *** (0,164)
AKRA	-0,001 (0,004)	-0,001 (0,003)	-0,003 (0,003)	-0,010 * (0,005)
RVI	-0,055 *** (0,010)	-0,038 *** (0,011)	-0,037 *** (0,011)	-0,078 *** (0,015)
CSAD_All(t - 1)	-0,302 *** (0,074)	-0,274 *** (0,069)	-0,26 *** (0,074)	-0,357 *** (0,086)
Pseudo R ²	0,325	0,319	0,357	-
R ²	-	-	-	0,778
Наблюдения	1030	1030	1030	1030
Равенство коэффициентов по всем квантилям и МНК регрессии (p-value)				
H0: t(0,25) = t(0,50)	0,194	0,194	-	-
H0: t(0,25) = t(0,75)	0,001	-	0,001	-
H0: t(0,50) = t(0,75)	-	0,005	0,005	-
H0: t(OLS) = t(0,50)	-	0,000	-	0,000
Machado-Santos Silva тест на гетероскедастичность, chi2 (p-value)	345,174 (0,000)	461,901 (0,000)	562,297 (0,000)	-
Breusch-Pagan / Cook-Weisberg тест на гетероскедастичность, chi2 (p-value)	-	-	-	437,370 (0,000)

Примечание: * 10%-ный уровень значимости; ** 5%-ный уровень значимости; *** 1%-ный уровень значимости.
Источник: расчеты автора.

Возникновение более высоких значений дивергенции мнений (EDivO_mkt²) на рынке приводит к снижению стадного поведения на более высоком квантиле (q=0,75). Однако измерение дивергенции мнений в обычном виде (EDivO_mkt) не позволяет значимо объяснить уровень стадности. Определение расхождения инвесторов во мнениях не выявляет причину изменения их поведения. В то же время измерение общего уровня дивергенции мнений является положительным фактором, который объясняет абсолютную разность доходностей на более высоких квантилях. Возможно, это явление связано с наличием в выборке таких активов, по которым инвесторы готовы брать на себя риск и идти вразрез с мнением большинства, что может свидетельствовать о присутствии на рынке групп участников с разным принятием риска.

Устойчивую и сильную связь с показателем стадности демонстрирует также индекс волатильности RVI, рост которого ведет к возникновению стадного поведения инвесторов

на российском рынке. Это может означать, что на фоне увеличения рыночного риска инвесторы стремятся учитывать поведение других инвесторов, создавая при этом одинаковые паттерны движения доходностей обсуждаемых акций.

Среди акций первого кластера, то есть акций с более высокой средней доходностью и волатильностью (табл. 1), квадрат разности мнений является значимым фактором, который имеет положительную связь со стадностью частных инвесторов. Иначе говоря, большой уровень разногласий на рынке ведет к более высокому смещению доходности такой группы акций в сравнении с общерыночной динамикой (табл. 5).

Таблица 5

Результаты анализа стадного поведения по первому кластеру активов

Переменная	Объясняемая переменная: CSAD_0			
	q = 0,25	q = 0,50	q = 0,75	МНК
	Коэффициент (Ст. ошибка)	Коэффициент (Ст. ошибка)	Коэффициент (Ст. ошибка)	Коэффициент (Ст. ошибка)
const.	0,065 *** (0,012)	0,066 *** (0,011)	0,076 *** (0,012)	0,125 *** (0,016)
Log_Sent	-0,009 (0,026)	-0,006 (0,015)	-0,018 (0,015)	-0,019 (0,016)
Mkt_Div0	0,014 (0,041)	-0,008 (0,026)	0,011 (0,030)	0,021 (0,029)
Mkt_Div0 ²	0,003 (0,008)	0,010 * (0,005)	0,015 *** (0,005)	0,012 ** (0,005)
Direction	0,001 (0,002)	0,001 (0,002)	0,003 (0,003)	0,004 (0,004)
IMOEX	-0,716 *** (0,083)	-0,765 *** (0,102)	-0,906 *** (0,182)	-1,005 *** (0,234)
AKRA	-0,003 (0,004)	-0,002 (0,003)	-0,002 (0,005)	-0,018 * (0,011)
RVI	-0,067 *** (0,010)	-0,054 *** (0,010)	-0,061 *** (0,011)	-0,103 *** (0,015)
CSAD_0(t - 1)	-0,123 ** (0,057)	-0,094 (0,035)	-0,040 (0,041)	-0,224 * (0,123)
Pseudo R ²	0,205	0,202	0,224	-
R ²	-	-	-	0,587
Наблюдения	1030	1030	1030	1030
Равенство коэффициентов по всем квантилям и МНК регрессии (p-value)				
H0: t(0,25) = t(0,50)	0,027	0,027	-	-
H0: t(0,25) = t(0,75)	0,005	-	0,005	-
H0: t(0,50) = t(0,75)	-	0,100	0,100	-
H0: t(OLS) = t(0,50)	-	0,001	-	0,001
Machado-Santos Silva тест на гетероскедастичность, chi2 (p-value)	184,846 (0,000)	267,719 (0,000)	338,098 (0,000)	-
Breusch-Pagan / Cook-Weisberg тест на гетероскедастичность, chi2 (p-value)	-	-	-	27,480 (0,000)

Примечание: * 10%-ный уровень значимости; ** 5%-ный уровень значимости; *** 1%-ный уровень значимости. Источник: расчеты автора.

Это может означать, что частные инвесторы в случае снижения консенсуса на рынке готовы покупать активы с большим риском, ожидая, что другая часть инвесторов в аналогичные периоды времени будет больше склонна продавать такие акции. Подобное поведение описывает случай, когда на рынке присутствуют участники, которые склонны платить премию за риск в моменты роста дивергенции мнений. Также наблюдается отрицательное

влияние индекса RVI на стадность по акциям первого кластера. Стоит также отметить значения коэффициентов по данному индексу, которые варьируются от $-0,103$ до $-0,054$, что в несколько раз выше показателя дивергенции $E\text{DivO_mkt}^2$. Это также может означать, что на рынке присутствуют отдельные группы инвесторов, которые готовы рисковать по более волатильным активам в моменты не только снижения консенсуса мнений, но и роста волатильности по рынку в целом. Вместе с тем немалая часть инвесторов опасается нестабильности на рынке и заинтересована следовать за рыночной динамикой.

Однако по акциям второго кластера с более низкими доходностями и уровнем волатильности наблюдается противоположная ситуация, когда общий уровень дивергенции мнений отрицательно объясняет уровень стадного поведения частных инвесторов (табл. 6).

Таблица 6

Результаты анализа стадного поведения по второму кластеру активов

Переменная	Объясняемая переменная: CSAD_1			
	q = 0,25	q = 0,50	q = 0,75	МНК
	Коэффициент (Ст. ошибка)	Коэффициент (Ст. ошибка)	Коэффициент (Робастная ст. ошибка)	Коэффициент (Ст. ошибка)
const.	0,050 ** (0,021)	0,041 *** (0,015)	0,055 *** (0,015)	0,090 *** (0,020)
Log_Sent	0,026 (0,022)	0,014 (0,020)	0,008 (0,026)	0,031 (0,019)
Mkt_DivO	-0,016 (0,038)	0,005 (0,036)	-0,004 (0,046)	-0,035 (0,032)
Mkt_DivO ²	-0,016 ** (0,007)	-0,014 ** (0,006)	-0,006 (0,007)	-0,014 * (0,007)
Direction	0,003 (0,002)	0,003 (0,002)	0,002 (0,002)	0,002 (0,002)
IMOEX	-1,719 *** (0,164)	-1,742 *** (0,147)	-1,742 *** (0,024)	-1,723 *** (0,092)
AKRA	-0,002 (0,008)	0,002 (0,006)	0,007 (0,008)	-0,006 (0,010)
RVI	-0,083 *** (0,015)	-0,064 *** (0,012)	-0,062 *** (0,01)	-0,101 *** (0,017)
CSAD_1(t - 1)	-0,265 *** (0,048)	-0,267 *** (0,037)	-0,250 *** (0,029)	-0,329 *** (0,064)
Pseudo R ²	0,401	0,394	0,4158	-
R ²	-	-	-	0,813
Наблюдения	1030	1030	1030	1030
Равенство коэффициентов по всем квантилям и МНК регрессии (p-value)				
H0: t(0,25) = t(0,50)	0,341	0,341	-	-
H0: t(0,25) = t(0,75)	0,266	-	0,266	-
H0: t(0,50) = t(0,75)	-	0,938	0,938	-
H0: t(OLS) = t(0,50)	-	0,006	-	0,006
Machado-Santos Silva тест на гетероскедастичность, chi2 (p-value)	84,969 (0,000)	27,150 (0,000)	3,429 (0,180)	-
Breusch-Pagan / Cook-Weisberg тест на гетероскедастичность, chi2 (p-value)	-	-	-	15,260 (0,000)

Примечание: * 10%-ный уровень значимости; ** 5%-ный уровень значимости; *** 1%-ный уровень значимости. Для регрессии на 75-м квантиле по причине наличия проблемы гетероскедастичности остатков используются робастные стандартные ошибки.

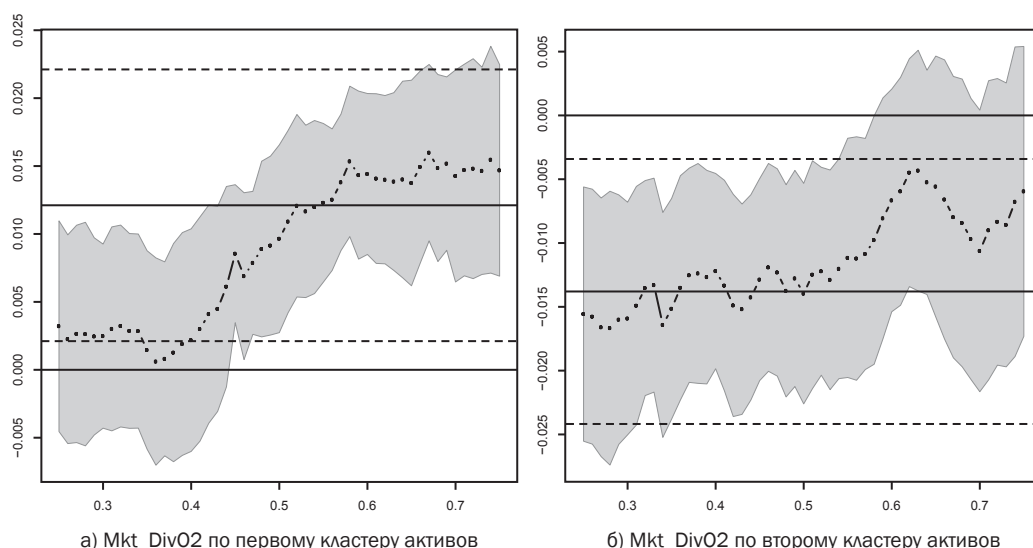
Источник: расчеты автора.

Следует отметить, что общий уровень сентимента, структурный сдвиг в моменты роста или падения рынка (Direction) и разность мнений с учетом направленности дивергенции (Mkt_DivO) не являются значимыми факторами в объяснении стадного поведения частных инвесторов. Это может объясняться тем, что для инвесторов более важно учитывать общее состояние мнений на рынке, которое выражается в консенсусе различных групп рыночных участников. Помимо этого, в среднем (на основе МНК-оценок, а также квантилей 0,25 и 0,50) рост дивергенции мнений по обсуждаемым акциям второго кластера является сигналом для частных инвесторов, которые начинают копировать поведение других участников рынка, что ведет к повышению уровня стадного поведения.

На рис. 3 представлены графики значений коэффициентов при показателе квадрата дивергенции мнений для оценки отличий влияния общего уровня дивергенции (Mkt_DivO²) по двум кластерам на стадное поведение частных инвесторов.

Рисунок 3

Динамика значений коэффициентов при показателе Mkt_DivO²



Примечание: серым цветом обозначен доверительный интервал на уровне 0,95.
 Источник: рисунок автора.

Еще одно важное обстоятельство — значимое отрицательное влияние прошлого уровня стадного поведения на текущее его состояние (см. табл. 4, 6). Следует предположить, что в моменты снижения стадного поведения в прошлом (показатель CSAD_{1(t-1)} растет) может происходить рост доходности по обсуждаемым активам, который на следующий торговый день привлечет других участников рынка. В дальнейшем эти участники рынка, вероятно, будут склонны копировать поведение инвесторов во втором кластере акций, что приведет к большему уровню стадного поведения. Такое поведение способно увеличить доходность рынка в целом и сократить разрыв между доходностью акций с более низким риском и общерыночной доходностью на текущий момент времени. Это также подтверждается тем, что индекс Мосбиржи (IMOEX) имеет отрицательную взаимосвязь со стадным поведением розничных участников по двум кластерам активов. Подобное поведение интуитивно объясняется склонностью инвесторов к оптимизму. То есть рост значения фондового индекса заставляет инвесторов верить в общий рост рынка и следовать за данной динамикой, что увеличивает стадность среди участников рынка, которые покупают акции вслед за положительной динамикой индекса IMOEX.

ЗАКЛЮЧЕНИЕ

Полученные результаты дают основание утверждать, что сентимент частных инвесторов является важным фактором их стадного поведения на российском фондовом рынке. Однако измерение общего уровня настроения участников рынка не позволяет доказать его объяснительную силу и определить степень влияния этого настроения на формирование рыночной стадности.

Гипотеза H1 отвергается как по двум кластерам российских акций, так и по выборке в целом. В то же время гипотеза H2 находит свое подтверждение только по акциям с более низким уровнем волатильности и доходности в среднем. Инвесторы готовы следовать за мнением других участников рынка в моменты роста уровня разногласий на исследуемых онлайн-порталах. Такое поведение объясняется страхом инвесторов идти против рыночной динамики, когда консенсус мнений в онлайн-среде начинает снижаться (чем выше разность мнений, тем больше склонность инвесторов доверять общерыночной динамике).

Кластеризация обсуждаемых акций российских компаний позволяет подтвердить гипотезу H3, доказывая наличие групп инвесторов, которые готовы принимать на себя риск по более волатильным активам в моменты роста дивергенции мнений. Некоторые частные инвесторы готовы использовать такие моменты в период низкого консенсуса на рынке, в то время как другие пытаются следовать рыночной динамике и продают более рискованные акции, а затем покупают акции с меньшим уровнем волатильности. В то же время часть инвесторов покупает более рискованные акции в периоды низкого консенсуса. Такое предположение о страхе инвесторов было также сделано в работе по исследованию стадного поведения на фондовом рынке США [Filip and Pochea, 2023].

Данные результаты анализа влияния сентимента частных инвесторов на стадное поведение розничных участников рынка частично согласуются с рядом исследований [Choi K. H., Yoon, 2020; Li et al., 2023]. Однако в текущей работе заполняются пробелы по изучению прокси разногласий между инвесторами за счет учета дивергенции мнений пользователей отечественных онлайн-платформ.

Полученные результаты позволяют предположить, что на российском фондовом рынке могут возникать манипулятивные и заведомо ложные информационные паттерны, мотивирующие инвесторов совершать выгодные для других участников рынка ошибочные торговые действия. Такие действия способны достаточно быстро сформировать стадное поведение в отношении отдельных видов активов. Исследование данной проблемы позволит эмпирически расширить область научного знания о работе финансовых рынков и возникновении локальных аномалий, которые становятся причиной неэффективности рыночной среды.

Одно из ключевых направлений продолжения исследований по этой теме — анализ влияния крупных геополитических событий, таких как начало специальной военной операции в 2022 г., на поведение розничных инвесторов. Разделение выборки на периоды до и после подобных событий может предоставить ценную информацию об изменениях в поведении инвесторов, особенно в контексте вытеснения иностранных инвесторов розничными в 2022 г. Такой анализ позволит глубже понять динамику рынка и адаптацию стратегий инвесторов к новым реалиям.

Дальнейшее исследование может также включать углубленный анализ выявленных кластеров активов и квантилей распределения показателя стадности. Изучение изменений этих параметров во времени и их реакции на различные рыночные условия может дать дополнительную информацию о поведении розничных инвесторов. Это направление исследований не только расширит понимание роли розничных инвесторов на российском фондовом рынке, но и заложит основу для разработки более эффективных стратегий управления рисками и регулирования рынка.

Список источников / References

1. Arthur D., Vassilvitskii S. (2007). k-means++: The advantages of careful seeding. In *Soda*, 7, 1027–1035. <https://doi.org/10.1007/s11222-024-10390-z>.
2. Bikhchandani S., Sharma S. (2000). Herd behavior in financial markets. *IMF Staff papers*, 47 (3), 279–310.
3. Choi K.H., Yoon S.M. (2020). Investor sentiment and herding behavior in the Korean stock market. *International Journal of Financial Studies*, 8 (2), 34. <https://doi.org/10.3390/ijfs8020034>.
4. De Long J.B., Shleifer A., Summers L.H., Waldmann R.J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, 45 (2), 379–395. <https://doi.org/10.1111/j.1540-6261.1990.tb03695.x>.
5. Duygun M., Tunaru R., Vioto D. (2021). Herding by corporates in the US and the Eurozone through different market conditions. *Journal of International Money and Finance*, 110, 102311. <https://doi.org/10.1016/j.jimonfin.2020.102311>.
6. Economou F., Gavriilidis K., Goyal A., Kallinterakis V. (2015). Herding dynamics in exchange groups: Evidence from Euronext. *Journal of International Financial Markets, Institutions and Money*, 34, 228–244. <https://doi.org/10.1016/j.intfin.2014.11.013>.
7. Espinosa-Méndez C., Arias J. (2021). COVID-19 effect on herding behaviour in European capital markets. *Finance research letters*, 38, 101787. <https://doi.org/10.1016/j.frl.2020.101787>.
8. Ferreruela S., Mallor T. (2021). Herding in the bad times: The 2008 and COVID-19 crises. *The North American Journal of Economics and Finance*, 58, 101531. <https://doi.org/10.1016/j.najef.2021.101531>.
9. Filip A.M., Pochea M.M. (2023). Intentional and spurious herding behavior: A sentiment driven analysis. *Journal of Behavioral and Experimental Finance*, 38, 100810. <https://doi.org/10.1016/j.jbef.2023.100810>.
10. 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*, 47 (4), 1461–1484. <https://doi.org/10.1111/j.1540-6261.1992.tb04665.x>.
11. 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](https://doi.org/10.1016/0304-405X(92)90023-Q).
12. Lakonishok J., Shleifer A., Vishny R.W., Hart O., Perry G.L. (1992). The structure and performance of the money management industry. *Brookings Papers on Economic Activity. Microeconomics*, 339–391.
13. Li T., Chen H., Liu W., Yu G., Yu Y. (2023). Understanding the role of social media sentiment in identifying irrational herding behavior in the stock market. *International Review of Economics & Finance*, 87, 163–179. <https://doi.org/10.1016/j.iref.2023.04.016>.
14. Prendergast C., Stole L. (1996). Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning. *Journal of political Economy*, 104 (6), 1105–1134. <https://doi.org/10.1086/262055>.
15. Rubesam A., Gerson de Souza R.J. (2022). Covid-19 and herding in global equity markets. *Journal of Behavioral and Experimental Finance*, 35, 100672. <https://doi.org/10.1016/j.jbef.2022.100672>.
16. Sul H.K., Dennis A.R., Yuan L. (2017). Trading on twitter: Using social media sentiment to predict stock returns. *Decision Sciences*, 48 (3), 454–488. <https://doi.org/10.1111/deci.12229>.
17. Trueman B. (1994). Analyst forecasts and herding behavior. *The review of financial studies*, 7 (1), 97–124. <https://doi.org/10.1093/rfs/7.1.97>.
18. Vieira E.S., Pereira M.S.V. (2015). Herding behaviour and sentiment: Evidence in a small European market: Comportamiento gregario y sentimiento: la evidencia sobre un pequeño mercado europeo. *Revista de Contabilidad-Spanish Accounting Review*, 18 (1), 78–86. <https://doi.org/10.1016/j.rcsar.2014.06.003>.
19. Wu D.D., Zheng L., Olson D.L. (2014). A decision support approach for online stock forum sentiment analysis. *The IEEE Transactions on Systems, Man, And Cybernetics: Systems*, 44 (8), 1077–1087. <https://doi.org/10.1109/TSMC.2013.2295353>.
20. Yoon J., Oh G. (2022). Investor herding behavior in social media sentiment. *Frontiers in Physics*, 1079. <https://doi.org/10.3389/fphy.2022.1023071>.

Перечень тикеров обсуждаемых российских акций

Тикер	Тикер	Тикер
AFKS	LSRG	PIKK
AFLT	MAGN	PLZL
AKRN	MGNT	POLY
ALRS	MOEX	RASP
AMEZ	MRKC	RKKE
APTK	MRKP	RNFT
AQUA	MRKU	ROSN
BSPB	MRKV	RTKM
CBOM	MRKY	RUAL
CHMF	MRKZ	SBER
CHMK	MSNG	SELG
DVEC	MSRS	SNGS
ELFV	MTLR	SVAV
FEES	MTSS	TATN
FESH	MVID	TGKA
GAZP	NKHP	TGKB
GMKN	NKNC	TRMK
GTRK	NLMK	TTLK
HYDR	NMTP	UWGN
IRAO	NVTK	VSMO
LKOH	OGKB	VTBR
LNZL	PHOR	YNDX

Источник: составлено автором на основе текстовых сообщений пользователей онлайн-платформ инвестиционной направленности.

Корреляционная матрица

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) CSAD_All	1.000									
(2) CSAD_0	0.944***	1.000								
(3) CSAD_1	0.899***	0.785***	1.000							
(4) ES_mkt	-0.024	0.016	-0.081***	1.000						
(5) EDivO_mkt	-0.026	0.014	-0.083***	0.998***	1.000					
(6) EDivO_mkt ²	-0.022	0.020	-0.083***	0.990***	0.983***	1.000				
(7) Direction	-0.277***	-0.238***	-0.351***	0.140***	0.141***	0.151***	1.000			
(8) IMOEX	-0.727***	-0.640***	-0.801***	0.165***	0.169***	0.163***	0.512***	1.000		
(9) AKRA	-0.031	-0.052*	0.015	-0.051*	-0.053*	-0.045	-0.051*	-0.109***	1.000	
(10) RVI	-0.192***	-0.187***	-0.109***	-0.079**	-0.086***	-0.075**	-0.189***	-0.204***	0.085***	1.000

Примечание: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Источник: составлено автором.

Информация об авторе

Максим Сергеевич Файзулин, аспирант факультета экономических наук, научный сотрудник Центра финансовых исследований и анализа данных НИУ ВШЭ, г. Москва

Information about the author

Maxim S. Faizulin, Post-graduate Student of the Faculty of Economic Sciences, Researcher at the Center for Financial Research and Data Analysis, HSE University, Moscow

Статья поступила в редакцию 06.03.2024
Одобрена после рецензирования 21.05.2024
Принята к публикации 08.08.2024

The article submitted March 3, 2024
Approved after reviewing May 21, 2024
Accepted for publication August 8, 2024