Мы живем в мире, где главным становится не факт, а интерпретация. Массы вовлечены не только в политические процессы, но и в экономические, финансовые. Ряд подвижек в жизни человечества (включая пандемию COVID-19) существенно изменил роль коммуникаций, каналы коммуникаций и роль частных (розничных) инвесторов на фондовом рынке. Наша книга посвящена этим подвижкам и попыткам построить разные метрики, улавливающие отношение участников инвестиционного рынка к происходящим событиям. Мы показываем, как на психологических искажениях и распознавании тональности информационных и коммуникационных сообщений профессионалы могут заработать.

ISBN: 978-5-16-018208-7
9 785160 182087

СЕНТИМЕНТ ИНВЕСТОРОВ И АНОМАЛИИ В ПОВЕДЕНИИ БИРЖЕВЫХ ХАРАКТЕРИСТИК ИНВЕСТИЦИОННЫХ АКТИВОВ

Искусственный интеллект

НАУЧНАЯ МЫСЛЬ

Финансы и экономика

Т.В. Теплова, Т.В. Соколова, М.С. Файзулин, А.В. Куркин

СЕНТИМЕНТ ИНВЕСТОРОВ И АНОМАЛИИ В ПОВЕДЕНИИ БИРЖЕВЫХ ХАРАКТЕРИСТИК ИНВЕСТИЦИОННЫХ АКТИВОВ

НАУЧНАЯ МЫСЛЬ
СЕРИЯ ОСНОВАНА В 2008 ГОДУ

Национальный исследовательский университет
«Высшая школа экономики»

Т.В. ТЕПЛОВА
Т.В. СОКОЛОВА
М.С. ФАЙЗУЛИН
А.В. КУРКИН

СЕНТИМЕНТ ИНВЕСТОРОВ И АНОМАЛИИ В ПОВЕДЕНИИ БИРЖЕВЫХ ХАРАКТЕРИСТИК ИНВЕСТИЦИОННЫХ АКТИВОВ

МОНОГРАФИЯ

Москва
ИНФРА-М
2022

УДК 336.7(075.4)
ББК 65.26
Т34

Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ)

Монография подготовлена при поддержке Российского научного фонда в рамках гранта 22-18-00276, https://rscf.ru/project/22-18-00276/

Рецензенты:

Абрамов А.Е., кандидат экономических наук, руководитель лаборатории анализа институтов и финансовых рынков Института прикладных экономических исследований Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации;

Ацканов И.А., кандидат экономических наук, старший управляющий активами УК «Атон-менеджмент»

Теплова Т.В.

Т34 Сентимент инвесторов и аномалии в поведении биржевых характеристик инвестиционных активов : монография / Т.В. Теплова, Т.В. Соколова, М.С. Файзулин, А.В. Куркин. — Москва : ИНФРА-М, 2022. — 199 с. — (Научная мысль).

ISBN 978-5-16-018208-7

Монография научных сотрудников Центра финансовых исследований и анализа данных (ЦФИАнД) факультета экономических наук (ФЭН) НИУ ВШЭ впервые обобщает описанные в академической литературе эмпирические результаты тестирования различных гипотез о поведении инвестиционных активов под влиянием метрик инвестиционного внимания и сентимента, а также демонстрирует авторские приемы анализа текстовых сообщений в различных каналах коммуникаций. Представлен большой статистический материал по участию розничных инвесторов на сегментах инвестиционного рынка, включая криптовалюты и NFT, а также рекомендации по построению инвестиционных стратегий с учетом сентимента на рынке в целом и по отдельным активам.

Ранее в академической литературе при изучении поведения разных классов инвесторов подчеркивалась особая роль непрофессиональных (розничных, частных) участников рынка и их склонность инвестировать «с толпой». В качестве каналов получения информации рассматривались деловые издания (СМИ), отчеты инвестиционных домов и брокеров, новостная лента информационных агентств. Отмечались различные искажения рационального поведения и были ярко описаны примеры поведенческих особенностей принятия решений, приводящие к аномалиям в поведении биржевых характеристик активов. Ковид-кризис, развитие социальных сетей и инвестиционных сообществ существенно поменяли каналы коммуникаций инвесторов. Все большее влияние на поведение активов, наряду с фундаментальными факторами и новостной лентой, стали оказывать мнения инвестиционных гуру (блогеров, активистов чатов в мессенджерах). Появились мемные акции и криптоактивы, аналитики стали разрабатывать различные метрики общерыночного сентимента и мониторить тональность обсуждений по отдельным акциям.

Будет интересна финансовым аналитикам, экономистам и социологам, специалистам IR-служб компаний, анализирующих каналы коммуникаций на инвестиционном рынке и пытающихся понять феномен мемных активов и «умных тикеров», возрастающую роль инвестиционных сообществ. Будет полезна аспирантам и студентам, изучающим финансы и компьютерные науки, для выбора тем будущих исследований и развития экономико-математического аппарата.

УДК 336.7(075.4)
ББК 65.26

© Теплова Т.В., Соколова Т.В., Файзулин М.С., Куркин А.В., 2022

ISBN 978-5-16-018208-7

ФЗ № 436-ФЗ	Издание не подлежит маркировке в соответствии с п. 1 ч. 2 ст. 1

Научное издание

Теплова Тамара Викторовна,
Соколова Татьяна Владимировна,
Файзулин Максим Сергеевич,
Куркин Алексей Викторович

СЕНТИМЕНТ ИНВЕСТОРОВ И АНОМАЛИИ В ПОВЕДЕНИИ БИРЖЕВЫХ ХАРАКТЕРИСТИК ИНВЕСТИЦИОННЫХ АКТИВОВ

МОНОГРАФИЯ

ООО «Научно-издательский центр ИНФРА-М»
127214, Москва, ул. Полярная, д. 31В, стр. 1
Тел.: (495) 280-15-96, 280-33-86. Факс: (495) 280-36-29
E-mail: books@infra-m.ru http://www.infra-m.ru

Подписано в печать 18.11.2022.
Формат 60×90/16. Бумага офсетная. Гарнитура Petersburg.
Печать цифровая. Усл. печ. л. 12,44.
Тираж 500 экз. Заказ № 00000

ТК 798635-1925519-181122

Отпечатано в типографии ООО «Научно-издательский центр ИНФРА-М»
127214, Москва, ул. Полярная, д. 31В, стр. 1
Тел.: (495) 280-15-96, 280-33-86. Факс: (495) 280-36-29

Оглавление

Авторский коллектив……..4
Введение……...5
Глава 1. Прокси-сентимент заметности отдельных активов и диагностируемые эффекты внимания……..16
1.1. Меняющаяся роль частных инвесторов на рынке……...16
1.2. Эффекты внимания, метрики заметности как прокси-сентимента и исследования по выявлению связи с биржевыми характеристиками активов……32
1.3. Статичные метрики сентимента: нейминг, алфавитизм и эмпирическое тестирование гипотезы «умного тикера» на примере акций Гонконга……41
1.4. Метрики наблюдений за успешными инвесторами……50
1.5. Демо-исследование сопоставления разных метрик заметности на доходность построенных портфелей……..52
Глава 2. Общерыночные метрики сентимента. Прокси-метрики сентимента на основе макроданных……63
2.1. Ранее проведенные исследования по прокси-метрикам сентимента на основе макроданных……63
Глава 3. Тональность публичных медиа-коммуникаций, обсуждений на интернет-платформах и форумах……75
3.1. Текстовый анализ в построении метрик сентимента……...75
3.2. Проблема недостатка сентимент-метрик по текстовому анализу и новые подходы: дивергенция мнения частных инвесторов и поведение фондового рынка..……90
3.3. Эмпирическое исследование дивергенции сентимента частных инвесторов на платформе «Тинькофф Пульс»……...97
3.4. Анализ влияния сентимента частных инвесторов на волатильность доходности акций российских компаний……...114
Глава 4. Криптоактивы и криптовалюты как объект исследований влияния сентимента……141
4.1. Обзор рынка криптовалют по состоянию на конец 2021 г. – середину 2022 г.……..141
4.2. Рынок NFT в 2018–2022 гг……...145
4.3. Метрики сентимента для криптовалют и цифровых активов……..148
4.4. Авторские индексы сентимента для рынка криптовалют и NFT……153
4.5. Эмпирический анализ факторов, влияющих на рынок NFT-активов……163
Заключение……...178
Список литературы……..185

Авторский коллектив

Теплова Тамара Викторовна, доктор экономических наук, профессор, директор Центра финансовых исследований и анализа данных Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ), ординарный профессор базовой кафедры инфраструктуры финансовых рынков НИУ ВШЭ.

Соколова Татьяна Владимировна, кандидат физико-математических наук, заместитель директора Центра финансовых исследований и анализа данных Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ), доцент базовой кафедры инфраструктуры финансовых рынков НИУ ВШЭ.

Файзулин Максим Сергеевич, стажер-исследователь Центра финансовых исследований и анализа данных Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ), аспирант факультета экономических наук НИУ ВШЭ.

Куркин Алексей Викторович, стажер-исследователь Центра финансовых исследований и анализа данных Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ), аспирант факультета экономических наук НИУ ВШЭ.

Введение[1]

Информационная эффективность фондовых рынков является базовой парадигмой современной финансовой экономики. Именно приходящая на рынок информация мотивирует игроков к активным действиям на рынке и меняет цены активов. Однако каналы получения информации и глубина информированности участников рынка, особенно со стороны не очень финансово грамотных трейдеров (часто они ассоциируются с частными инвесторами – физическими лицами, т.н. «физиками», розничными инвесторами, в академической литературе – «шумовыми трейдерами»), со временем меняются. Многочисленные исследования последних лет в области финансовой экономики принимают во внимание увеличение каналов получения информации, трансформацию традиционных источников получения данных, важную роль коммуникаций между частными инвесторами, минуя профессиональных аналитиков и брокеров, суждения и рекомендации которых могут быть предвзяты. Все большая роль отводится такому фактору, как известность компании и ее лидера (кейсы Илона Маска, Билла Гейтса, Кэти Вуд тому подтверждение, а также возможность привлекать капитал через SPAC). Еще один важный фактор, который перестали игнорировать исследователи – эмоциональная привязанность клиентов к продуктам и брендам компании, что существенно влияет на цены акций (Stambaugh et al., 2015; Li, 2015; Gao & Yang, 2018). Социальные настроения, социальная активность на базе межличностных коммуникаций, объединение действий инвесторов через различные сообщества стали не только привлекать внимание исследователей (Zhou et al., 2021, анализируют влияние Twitter на поведение акций в США, а Afzali & Martikainen, 2021, рассматривают финансовые последствия обсуждений в социальных сетях в европейских странах), но и регуляторов, например, центральных банков стран.

Классическая финансовая теория базируется на ряде предпосылок относительно поведения инвесторов (рациональность) и степени развития финансового рынка (информационная открытость и равный доступ к информации, конкурентность участников и т.п.). Относительно новое направление поведенческих финансов учитывает как высокую неопределенность на рынках, так и асимметрию в доступе к информации и подчеркивает, что поведение участников может отклоняться от рациональности, инвесторы могут искаженно воспринимать происходящие события и их освещение в разных каналах коммуникаций и, соответственно, принимать решения, не максимизирующие их выгоды (текущие и долгосрочные), а руководствоваться некими иными целями и мотивами (самореализоваться, отыграться и т.п.). Рис. 1 (как эмоции берут верх) ярко демонстрирует заблуждения

[1] Обзор ряда работ во введении выполнен студентами факультета экономических наук Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) Ерохиным Д.А. и Байрамкуловой Л.М. под руководством Т.В. Тепловой.

непрофессиональных инвесторов. Ранние (до 1990-х гг.) трактовки инвестиционного выбора базировались на предположении, что розничные инвесторы (так называемые шумовые трейдеры, noise traders) составляют незначительную часть биржевых торгов и не оказывают существенного влияния на равновесную доходность акций (equilibrium stock returns). Это действительно так, особенно на ряде классов активов, например, на глобальном облигационном рынке. Глобальный облигационный рынок - это рынок институциональных инвесторов с лотами торгов по 200 тыс. долларов США. Заметим, что в ряде стран (Россия, Южная Корея, Испания) развит облигационный рынок и для частных инвесторов, но доля физических лиц (частных, индивидуальных или розничных инвесторов) в объеме торгов не превышает 50% (по России для сегмента высокодоходных облигаций (HY) присутствие физических лиц доходит до 100%). Как распределены средства домохозяйств по странам, и какие активы лидируют по биржевым торгам, будет рассмотрено в п. 1.1. Здесь подчеркнем, что роль частных инвесторов существенно меняется с 2020-х гг. и среди детерминант поведения ряда активов значимую роль начинают играть метрики сентимента (а не только внимания, что отмечалось ранее).

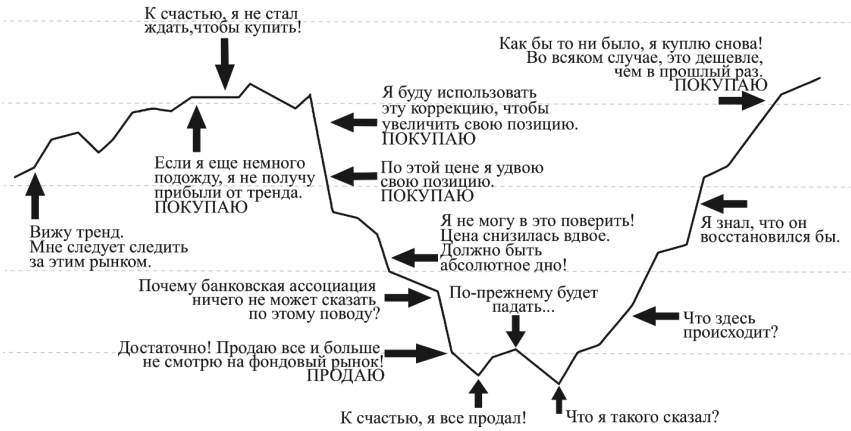

Рис. 1. Эмоциональное обоснование инвестиционных действий частных инвесторов (без анализа макрофакторов и фундаментальных показателей)

Одно из первых теоретических опровержений базовой предпосылки финансовой экономики о рациональности инвесторов в ценообразовании активов представлено в работе De Long et al. (1990), где авторы обосновывают, что скоррелированный сентимент шумовых трейдеров может менять равновесную цену акции на рынке. В работе Scharfstein and Stein (1990) ставится вопрос о важности учета стадного поведения инвесторов. В работе Schmeling (2006) подчеркивается важность различия поведения институциональных инвесторов (так называемых «умных

денег», smart money, которые руководствуются положениями классической финансовой экономики с акцентом на значимость фундаментальных факторов в оценке активов) и розничных инвесторов (шумовых трейдеров, для которых рациональность не является основой принятия решений). Важный вывод (Schmeling, 2006) – поведением частных инвесторов (мотивацией их принятия решений, факторами, которые выступают триггерами активных действий на рынке) пренебрегать нельзя. Еще один вывод – профессиональным (институциональным) инвесторам удается абстрагироваться от поведенческих искажений в биржевых торгах на фондовом рынке, что подтверждается и в более поздних работах (Itzkowitz et al., 2017).

В большом количестве работ Бабера и Одиина, включая работу (Barber & Odean, 2006), подчеркивается, что розничные инвесторы более вероятно будут покупать акции, которые привлекут их внимание какими-то новостями или событиями, или узнаваемыми и любимыми продуктами-услугами, чем на основе перебора фундаментальных факторов (stock-picking на основе макро-, отраслевых и фундаментальных для компании факторов). Следствием игнорирования правил финансовой экономики (анализ риска и доходности, диверсификация капитала) становятся для розничного инвестора такие ошибки как высокая концентрация портфеля в отдельных бумагах, нежелание фиксировать убытки и закрывать убыточные позиции, игнорирование не очень яркой (не обсуждаемой публично) негативной информации, усугубляя тем самым потери.

Эффект «толпы» присутствует на фондовом рынке. Barber et al. (2005) фиксируют сильную корреляцию между покупками и продажами отдельных инвесторов как в течение месяца, так и на более длинных горизонтах. Этот эффект наблюдается как по клиентам мелких брокеров (розничные брокерские компании), так и по клиентам-инвесторам крупных брокерских домов. Kumar and Lee (2006) на схожих данных крупных брокерских компаний отмечают, что приход инвесторов на рынок и уход с него также взаимосвязаны и мотивируются изменением доходности акций небольших компаний.

Barber et al. (2008) выстроили свое исследование на изучении американских акций, в объеме торгов которыми высока доля физических лиц, т.е. на специфической подвыборке акций. Вывод авторов – эти инвесторы-физические лица являются шумовыми трейдерами, т.е. их поведение далеко от рациональности. Акции, активно продаваемые частными лицами (индивидуальными инвесторами), опережают акции, активно покупаемые физическими лицами, на целых 13,5% в следующем году. Hvidkjaer (2008) приходит к аналогичным выводам для рынка США на горизонте инвестирования до трех лет. Т.е. «инвестировать с толпой» - плохое решение. Но вопрос, а что мотивирует частных инвесторов повторять действия других участников рынка, всегда ли это самостоятельное решение, или же на инвестора оказывается некое «мягкое давление», до 2010-х гг. оставался открытым.

Эти работы, а также парадокс мемных акций, мемных криптовалют (Dogecoin Илона Маска) и кейс Gamestop, положили начало активному изучению поведения частных инвесторов и их возможного значимого влияния за счет эмоционального восприятия приходящей информации на биржевые характеристики ценных бумаг и инвестиционных активов (доходность на разных временных горизонтах, волатильность, торговую активность и иные характеристики). Сначала работы были сконцентрированы на рынке акций США, а в последние годы все больше работ в качестве объекта исследований выбирают рынок Китая. Наша коллективная монография впервые поднимает вопросы роли частных инвесторов на рынке акций в РФ. Отдельный параграф посвящен криптовалютам и криптоактивам (NFT).

Перечислим типичные «провалы частных инвесторов» по работе Barber and Odean (2011): а) результаты инвестирования индивидуальных инвесторов проигрывают традиционным бенчмаркам по доходности (одна из причин – низкие транзакционные издержки широких индексов фондового рынка, но это сожаление о потерях заставляет принимать больше риска), б) розничные инвесторы закрывают прибыльные инвестиции, сохраняя при этом убыточные («эффект диспозиции» как очень распространенный тип поведенческих искажений (bais)), в) сильно страдают от ограниченного внимания (анализа) и опираются на прошлые результаты своих решений о покупках, г) готовы из раза в раз повторять свои прошлые действия, которые принесли успех (не учитывая, что рыночная ситуация могла поменяться) и быстро забывают провалы, без анализа ошибок, д) склонны держать недиверсифицированные портфели акций. Все эти ошибки пагубно влияют на финансовое благополучие инвесторов.

В работах Barber et al. (2005, 2008, 2009) по рынку США на трех биржевых площадках NYSE, ASE и Nasdaq, а также в работе Jackson (2004) по рынку Австралии с 1991 по 2002 гг., продемонстрировано через анализ покупок и продаж в торговых системах инвесторов с малыми объемами торгов, что трейдинг индивидуальных инвесторов является высококоррелированным, что порождает эффект систематического влияния этой группы инвесторов на ценообразование. Фактически сделан вывод, что наблюдается стадность инвестирования со стороны розничных (частных) инвесторов, и пренебрегать поведением индивидуальных инвесторов, которые торгуют слаженно, ошибочно в плане предсказания поведения цены.

Современные исследования (Kuo et al., 2015; Adachi et al., 2017; De Bortoli et al., 2019) подтверждают нерациональность поведения частных инвесторов. В работе (Yan et al., 2016) подтверждаются факты искажения рационального поведения частных инвесторов, и делается важный парадоксальный вывод, касающийся уже рынка в целом - информационная эффективность рынка снижается.

Открытыми остаются несколько вопросов: 1) какие метрики могут корректно отразить настроения инвесторов, в особенности – разных групп

инвесторов, 2) какие классы активов или сектора, сегменты наиболее подвержены влиянию шумовых (не рациональных) инвесторов, и есть ли страновые различия, 3) какие стили инвестирования (факторные стратегии) на рынке акций находятся под давлением настроений индивидуальных инвесторов (акции роста или стоимости, акции крупных или мелких эмитентов). Насколько моментум эффекты проявляются более ярко по акциям с большой долей торгов физических лиц.

Наше исследование позволило сформулировать важный вывод: стадная чрезмерная реакция, то есть коррелированные действия частных инвесторов (шумовых трейдеров), основанные на чрезмерно оптимистичных или пессимистичных ожиданиях, формируют устойчивые взаимосвязи между сентиментом на финансовых рынках и наблюдаемыми биржевыми характеристиками активов (доходностью, волатильностью, торговой активностью).

Мы различаем общерыночный сентимент (public sentiment) и повышенное внимание и сентимент по отдельным инвестиционным активам. Яркий пример изменения общерыночного сентимента – кризисные явления в экономике, нарастание геополитических рисков или пандемийные страхи (например, волны паники при пандемии Covid-2019).

Сентимент инвесторов (investor sentiment) традиционно ассоциируется с поведением розничных инвесторов (Lee et al., 1991), хотя общерыночный сентимент учитывается при принятии решений всеми участниками рынка. Можно предположить, что реакция розничных инвесторов может быть более эмоциональной, слабее опирающейся на долгосрочные зависимости.

Особый интерес к инвестиционному выбору розничных инвесторов породили три процесса XXI века: 1) технологические инновации, позволившие через удобные приложения в смартфонах осуществлять торговые сделки в любом месте (в метро, кафе, из дома; без потребности в стационарном компьютере и специальных биржевых приложениях (например, Quik для российского инвестора)), 2) развитие социальных мессенджеров и платформ, которые сформировали инвестиционные сообщества как по отдельным классам активов (акции, облигации, криптовалюта и криптоактивы, деривативы), так и по глубине анализа, с появлением «местных гуру», активистов, способных объединять частных инвесторов для совершения согласованных действий (pump and dump), 3) коронакризис и локдауны в большинстве стран, а также денежная помощь населению как компенсация ограничений в потреблении ряда товаров и услуг (велики были выплаты в США, ЕС). В п. 1.1 мы внимательно посмотрим на долю частных инвесторов в биржевых торгах на разных рынках капитала и в РФ и на те работы, которые демонстрируют отличие поведения институциональных инвесторов (smart money) от частных. Приведем примеры объяснений особенного поведения частных инвесторов.

Коронакризис действительно изменил многие процессы в экономике, политике и в поведении инвесторов. Исследование Smales (2021) на основе

анализа фондовых индексов указывает на изменившееся поведение инвесторов и поведение цен активов в зависимости от нарастания пандемийных страхов. С 2021 года в академической литературе появился целый пласт работ, изучающих влияние пандемийных страхов на инвестиционное поведение участников финансового рынка. Прокси внимания, фиксируемая по Google search volume (GSV) по слову «коронавирус», имеет статистически значимую взаимосвязь с доходностью рыночных индексов стран G7 (Smales, 2021) с учетом влияния количества заболеваний и макроэкономических индикаторов. Рост GSV на одно стандартное отклонение приводит к снижению средней доходности от 2,14% (Токийская фондовая биржа) до 10,03% (Миланская фондовая биржа). Более того, количество поисковых запросов в рассматриваемом периоде является существенным фактором для объяснения роста волатильности котировок. Эти результаты подтверждают гипотезу увеличения значимости новостного потока при принятии инвестиционных решений в ситуациях потрясений на рынке.

Проблемой в количественной оценке влияния сентимента на биржевые характеристики акций или иных активов является построение прокси-переменных модели для измерения ненаблюдаемых явно настроений. Существующие исследования демонстрируют применение различных прокси на базе обработки доступной для инвесторов и экспертов информации. Этим ранее разработанным метрикам, а также авторским построениям, и посвящена наша работа.

В своей коллективной монографии мы разделяем метрики сентимента, которые демонстрируют повышенное внимание инвесторов к тем или иным инвестиционным активам, и, соответственно, могут выступать значимыми объясняющими переменными в прогнозных или объяснительных моделях (прокси-метрики заметности актива, attention), и метрики «чистого сентимента», которые учитывают тональность восприятия тех или иных событий на рынке. Это восприятие доводится до частного инвестора по разным каналам коммуникаций: в газетных или журнальных статьях, сообщениях и комментариях периодически раскрываемой компаниями отчетности, в обращениях топ-менеджмента к инвесторам, аналитических отчетах брокерских компаний, в обсуждениях в социальных сетях. Главный момент здесь - в любую новость можно добавить оттенки, которые усилят или сгладят эффект восприятия той или иной новости. Метрики «чистого сентимента», наряду с прямыми рекомендациями на покупку или продажу, формируют отдельный пласт исследований, который активно стал развиваться в последние годы.

Далее кратко перечислим основные направления построения таких прокси, а в параграфах книги покажем демо-примеры их построения и выводы по ранее проведенным исследованиям относительно возможности инвестору заработать на них.

Традиционными с 1990-х гг. являются прокси-сентимента (речь идет об общерыночном сентименте, как общем настроении на рынке относительно

принятия риска), отражающие спрос на те или иные инвестиционные активы. Например, количество IPO и привлеченные на сделках IPO деньги, скидки в цене закрытых фондов (Lee et al., 1991; Swaminathan, 1996; Neal & Wheatley, 1998). Мы относим такие прокси-сентимента к макрофакторам и называем их общерыночным сентиментом (им посвящена глава 2 книги).

Мы различаем метрики внимания, когда исследователь не имеет возможности оценить (спрогнозировать) восприятие инвесторами наблюдаемых событий прямым образом (инвесторы не высказывают своего отношения) и метрики собственно сентимента («чистого сентимента»), когда исследователь имеет дело с выраженным тем или иным образом отношением участников рынка к происходящим изменениям. Количество поисковых запросов, как наиболее частую прокси-переменную, мы трактуем как проявление внимания инвесторов (соответственно – заметности компании или актива), повышенного интереса (п. 1.2). Примером таких работ в связке с поведением акций являются: Michaux (2019), Behrendt et al. (2020). В работе Behrendt et al. (2020) продемонстрировано на примере 447 акций, торгующихся на NASDAQ/NYSE, что количество поисковых запросов в Wikipedia может быть принято в исследовании в качестве прокси эффекта внимания. Более того, положительная динамика этих поисковых запросов на данном сайте может использоваться инвесторами для построения прибыльной торговой стратегии. Alanyali et al. (2013) делают вывод о наличии корреляции между объемом торгов акциями компании и частотой упоминания этой компании в Financial Times.

К метрикам внимания как особым прокси-сентимента мы относим такие метрики как: число новостей о той или иной компании, число публикаций и аналитических отчетов («аналитическое покрытие» со стороны профессиональных участников), число «твиттов» и поисковых запросов.

Первой работой, в которой поисковые запросы как метрика внимания (attention) используются для тестирования портфельных построений, стала работа Joseph et al. (2011). Далее в п. 1.5 приведен пример проведения такого исследования (по работе Тепловой Т., Миковой Е. и Мажова М.).

Заметим, что поисковые запросы Google по специально подобранным словам-триггерам позволяют предсказать наступление пандемии (как было с заболеваемостью коронавирусом), а также изменение потребительских расходов. De Vries et al. (2012) подчеркивают, что можно предсказать популярность бренда с помощью информации об онлайн-общении на его фан-страницах. Asur and Hubermann (2010) использовали количество комментариев в блоге для прогнозирования кассовых сборов.

В качестве подвида прокси-сентимента в категории повышенного внимания мы выделяем наблюдаемые инвесторами биржевые характеристики (растущий объем торгов или растущая волатильность).

Как отдельное исследовательское направление мы выделяем проблему построения метрик общерыночного сентимента (начиная с самых простых, как в работе (Schmeling, 2009), где в качестве прокси-сентимента по стране

взят показатель потребительской инфляции (страновой индекс CPI), так и интегральных, как в работах (Baker & Wurgler, 2006 и далее) и тестирование их значимости в решении двух задач: 1) объяснение поведения тех или иных финансовых активов и наблюдаемые различия в их биржевых характеристиках, 2) прогнозные построения по ним. Метрики общерыночного сентимента могут быть как наблюдаемыми на финансовом рынке, так и опросными (в ряде исследований (Charoenrook, 2003; Lemmon & Portniaguina, 2006; Schmeling, 2006, 2008) опросы потребительского поведения рассматриваются как прокси сентимента, которые могут объяснить поведение ценных бумаг).

Отметим, что настроения инвесторов важны не только для прогнозирования отдельных акций или индексов, но и для предсказания кризисных явлений. Fu et al. (2019) разработали составной индекс на основе рыночных индикаторов и метрик настроения инвесторов и продемонстрировали через применение логит-модели возможность применения этого индекса для прогнозирования кризисов фондового рынка.

Baker and Wurgler (2006) построили показатель общерыночных настроений инвесторов на основе нескольких наблюдаемых на рынке переменных (дисконты в оценке по закрытым фондам, количество состоявшихся IPO, торговый оборот по ценным бумагам и другие рыночные показатели). Авторы оригинально применили метод главных компонент.

Еще одно направление построения метрик общерыночного сентимента – применение микро-торговых данных (биржевые заявки разных участников торгов), как в работе Kumar and Lee (2006), которые используют данные брокеров, или Barber et al. (2008), которые используют данные TAQ/ISSM.

Еще одно распространенное направление построения общерыночного сентимента – опросы инвесторов: Brown and Cliff (2004, 2005), Menkhoff and Rebitzky (2008). В ряде случаев анкеты заполняют сами участники рынка (Ding et al., 2014; Singal and Tayal, 2020). Примером таких прокси-метрик может выступать The UBS/GALLUP Index of Investor Optimism. Индекс строится на основе опросов (случайно выбранных) инвесторов с состоянием более $10 000. В течение первых двух недель каждого месяца UBS/Gallup проводит 1 000 опросов инвесторов.

Важный вывод по исследованию общерыночного сентимента – положительные и негативные настроения по разному распределены по инвестиционному сообществу. Одна из работ, подтверждающая этот вывод, - TsaiI-Chun (2017) по Тайваньской фондовой бирже. Автор с помощью модели векторной авторегрессии (VAR) проверяет гипотезу о том, что распространение настроений инвесторов является асимметричным в оптимистичных и пессимистических рыночных условиях. В благоприятных рыночных условиях, когда институциональные инвесторы настроены оптимистично, диффузионным эффектом настроений

инвесторов можно пренебречь. А вот пессимистические настроения распространяются широко.

Тренд последних лет – текстовая обработка новостной ленты, информационных и аналитических каналов, сообщений топ-менеджмента компаний, сообщений в мессенджерах, обмен мнениями членов инвестиционных сообществ. Этапами этой текстовой обработки является не только ее сбор (парсинг) с большим числом текстовых символов, но и классификация обработанных текстовых материалов теми или иными способами по тональности (чаще всего по трем классам: отрицательные, нейтральные и положительные) и подсчет различных показателей на основе этих данных. Подробное применение таких метрик будет далее показано в нашей работе (глава 3). В этом направлении учитываются мнения различных стейкхолдеров: мнение (внимание) аналитиков (аналитическое покрытие), мнение и прогнозы топ-менеджеров и мажоритарных собственников компаний, бенефициаров, самих инвесторов (институциональных и частных).

Интересна работа Zhu et al. (2019), в которой построен индекс, рассчитываемый на основании ключевых слов в статьях крупнейших американских газет.

Освещение деятельности компании в средствах массовой информации (СМИ) является важным элементом привлечения капитала. В статье Liu et al. (2011) показана значимость влияния «СМИ покрытия» IPO на доходность акций и их ликвидность на длительных временных горизонтах. Авторы приходят к выводу, что имидж в СМИ является важным фактором успеха публичных компаний. Отметим, что аналогичный вывод можно сделать и относительно имиджа компании и ее профессионального помощника в размещении на публичном рынке долга (организатора размещения) на сегменте ВДО (high yield) российского облигационного рынка. Чем более открыто руководство компании к частным инвесторам, чем сильнее контакт в инвестиционном сообществе телеграм чата, тем быстрее и дешевле компания может провести размещение облигаций.

Мы отмечаем, что в последние годы традиционные средства информирования (СМИ) и коммуникаций уступают социальным сетям. В работе Liang et al. (2020) на дневных данных акций Шанхайской фондовой биржи (SSEC) представлено сопоставление прогностических способностей трех индексов настроений, основанных на позитивно и отрицательно «окрашенных» постах в социальных сетях, новостях из газет и на Интернет-сайтах, Для моделей прогнозирования авторы используют гетерогенную авторегрессию (HAR), а в качестве объясняющих переменных выступают позитивные и негативные оценки разных метрик сентимента. Авторы показывают, что ежедневные значения индексов настроений в социальных сетях и интернет-СМИ оказывают значительное влияние на волатильность фондового рынка, в то время как индекс настроений, построенный на основе традиционных газет, не выявил влияния. Два индекса настроений (социальных сетей и новостей в

Интернете) могут значительно повысить точность прогнозов на краткосрочных временных интервалах. Точность прогнозов средне- и долгосрочной волатильности может повысить только индекс настроений, построенный на основе новостей Интернет-СМИ.

Отметим интересную работу по тональности сообщений топ-менеджеров компаний (Xu et al., 2022), где строится авторская сентимент-метрика на основе корпоративных новостей. Авторы утверждают, что на основе тона новостных сообщений менеджеров компании можно предсказать доходность инвестирования в акции рассматриваемой фирмы в следующем месяце. Но есть один нюанс: на основе эмпирических тестов авторы приходят к выводу, что качество прогнозов выше в периоды оптимистичных общерыночных настроений на рынке, т.е. важен общерыночный сентимент.

В работе Song et al. (2017) авторы демонстрирует выгоды построения торговых стратегий на отслеживании настроений инвесторов. Авторский прием – выявление «шоков сентимента». Наложение шоков на индикаторы трендов, по мнению авторов, позволяет предсказать динамику цен акций. Авторы предлагают стратегию открытия длинных позиций для 25% «лучших компаний» (с сильным положительным сентиментом) и коротких позиций для 25% худших компаний (с сильным отрицательным сентиментом). В более ранних работах эти же авторы строили торговые стратегии с длинными и короткими позициями по бумагам, отсортированным по эффекту внимания. В результате такие портфели значительно превышали доходность индекса S&P 500 (Song et al., 2015).

Отдельное направление исследований – сентимент по сообщениям в Твиттер (Nishimura & Sun, 2021). Сеть коммуникаций Твиттер (https://twitter.com) появилась в 2006 году и быстро завоевала популярность у пользователей благодаря формату коротких сообщений длиной в 140 символов (т.н. твитов), что гарантирует получение новостей в предельно концентрированном виде. В век больших скоростей эта лаконичность приобретает особую ценность. Вторая особенность авторских твитов – яркие, эмоционально окрашенные, запоминающиеся фразы (иногда с нецензурной лексикой для усиления эмоционального фона), что позволило этой сети стать прародителем многочисленных «мемов», «перепост» которых осуществлялся в другие популярные социальные сети и мессенджеры. Особую популярность Твиттер приобрел в США, став значимым каналом коммуникаций политических лидеров с электоратом (Wang et al., 2016). Дональд Трамп, как известный пользователь Твиттер, влиял своими сообщениями не только на симпатии электората, но и на инвесторов (Nishimura & Sun, 2021), которые строили даже портфельные стратегии по «любимчикам» Трампа. Роль Твиттера в политических кампаниях активно исследована в работах Lee et al. (2020), Frame and Brachotte (2015). Анализ твитов Д. Трампа и Х. Клинтон в ходе президентских выборов США 2016 г. позволил в работе (Lee & Lim, 2016) показать особенности гендерных различий PR-стратегий. Многие авторы

высказывают озабоченность потенциальной возможностью манипулирования через социальные сети, что демонстрируется в работе (Reisach, 2021).

Интересная попытка рассмотреть влияние сентимента на активы товарного рынка представлена в работе (Shen et al., 2017). Авторы выделяют такие эмоции, как оптимизм, страх и радость, и оценивают их влияние на доходность активов товарного рынка (commodity market). Авторы показали краткосрочное влияние эмоционального фона на доходность сырьевого товара за следующие пять дней. Текстовые данные были собраны на платформе-агрегаторе TRMI, в котором объединяются множество новостных заметок и постов из различных источников. Далее текст обрабатывался системой искусственного интеллекта MarketPsych, которая с помощью лексического анализа позволяет определять настроение на рынке. Имея данные длиной в 14 лет, авторы построили эконометрические модели TGARCH и VAR и обнаружили, что такие эмоции, как оптимизм, страх и радость, оказывают значительное влияние на доходность отдельных сырьевых товаров, но не оказывают влияние на доходность индекса товарного рынка.

«Сила слов» на примере твитов Трампа в контексте санкционного давления и поведения курса рубля рассмотрена в работе Афанасьева и соавторов (2019).

Глава 1
ПРОКСИ-СЕНТИМЕНТ ЗАМЕТНОСТИ ОТДЕЛЬНЫХ АКТИВОВ И ДИАГНОСТИРУЕМЫЕ ЭФФЕКТЫ ВНИМАНИЯ

1.1. МЕНЯЮЩАЯСЯ РОЛЬ ЧАСТНЫХ ИНВЕСТОРОВ НА РЫНКЕ[1]

На глобальном рынке можно выделить две принципиально разные модели участия частных инвесторов на инвестиционном (и, в частности, на фондовом) рынке. В первой модели доминируют профессиональные инвесторы: инвестиционные фонды (хедж фонды, private equity), как в США, или пенсионные и страховые компании, как в Германии, и доля розничных инвесторов исторически не превышала 30%. Заметим, что денежные средства физических лиц на брокерских счетах в США страхуются на $250 тысяч (в РФ не страхуются средства на брокерских счетах, есть страховка Агентства по страхованию вкладов (АСВ) только по банковским депозитам на сумму 1 400 тысяч рублей). В развивающихся странах доля проникновения частных инвесторов на инвестиционный рынок может быть еще меньше. Например, в Бразилии перед коронакризисом доля розничных инвесторов на фондовом рынке составляла менее 0,5%. Во второй модели доля частных инвесторов достигает 80% (материковый Китай). Для Китая интерес частных инвесторов к фондовому рынку и недвижимости, как инвестиционному активу, объясняется монополизацией банковского сектора (превалируют государственные банки) с крайне низкими процентными ставками. Представление о страновых различиях в сберегательном выборе частных инвесторов (домохозяйств) дает Рис. 1.1.

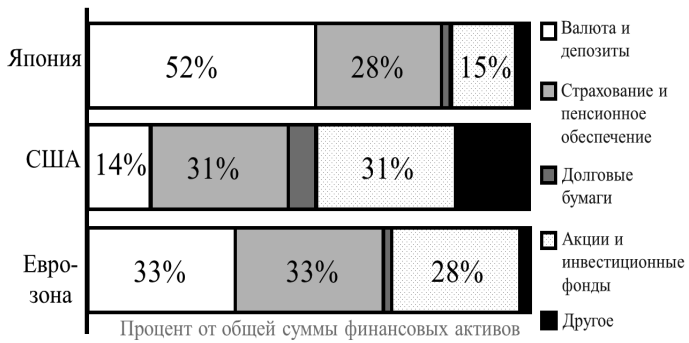

Рис. 1.1. Разнообразие финансовых активов, которыми владеют домохозяйства по странам мира (2019 г.)

[1] Обзор ряда работ в данном параграфе выполнен студентом факультета экономических наук НИУ ВШЭ Ерохиным Д.А. под руководством Т.В. Тепловой.

В развивающихся странах доля проникновения частных инвесторов на инвестиционный рынок может быть еще меньше. Например, в Бразилии перед коронакризисом доля розничных инвесторов на фондовом рынке составляла менее 0,5%. Доля вложений в акции в общем распределении сбережений населения на начало 2022 г.: США – 42%, Европа – 30%, ЮАР – 23%, Индия, Казахстан, Бразилия - менее 3%.

Вторая модель предполагает большое вовлечение частных инвесторов в отдельные сегменты рынка ценных бумаг. Яркие примеры – Китай, Россия с марта 2022 года на фоне ухода с рынка иностранных инвесторов, которые исторически занимали большую долю как во владении ценными бумагами, так и в оборотах торгов.

Доля частных инвесторов в структуре торговой активности на фондовых рынках различных стран показана в Табл. 1.1. Также в Табл. 1.1 указаны источники данных.

Таблица 1.1

Доля индивидуальных инвесторов в структуре торговой активности на фондовых рынках

Фондовый рынок	*Доля розничных инвесторов*	*Год*	*Источник*
NASDAQ (США)	34% акций во владении (включая ETF)	2020	Mackintosh (2020)
Shanghai Stock Exchange (Китай)	20,7% акций во владении	2020	SSE Market and Infrastructure Report (2020)
Tokyo Stock Exchange (Япония)	37% оборота акций	январь 2021	Monthly TSE Equities Report (2020)
Hong Kong Stock Exchange (Гонконг)	20,3% оборота акций	2019	HKEX Cash Market Transaction Survey (2019)

Источник: сбор Ерохина Д.А. под руководством Тепловой Т.В. в рамках подготовки выпускной квалификационной работы (ВКР) на факультете экономических наук (ФЭН) Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) в 2021 г.

Специфика Шэньчжэньской фондовой биржи Китая по сравнению с развитыми фондовыми рынками, - широкое участие индивидуальных инвесторов (Alhaj-Yaseen, Yau, 2018). Авторы выявили на китайском рынке акций типа А явно выраженное стадное поведение, которое фиксировалось как до либерализации рынка, так и после (к концу 2007 года почти 97% компаний, котирующие на рынке акции класса А, завершили реформу (Li et al., 2011)).

Институциональные инвесторы, как правило, опираются при принятии решений на стандартные процедуры, описанные в классических учебниках

по инвестициям, макропрогнозы и фундаментальные показатели компаний, соблюдают тайминг по техническому анализу. В большинстве академических работ предполагается, что частные инвесторы ориентируются в большей степени на новостную повестку и на мнения других инвесторов, брокеров и аналитиков, у них достаточно короткий (спекулятивный) горизонт инвестирования.

Подвижки в этих двух моделях произошли в пандемию Covid-19 и локдауны. Падение рынков весной 2020 года и последующее резкое восстановление позволили пришедшим на рынок миллионам новых инвесторов получить позитивный опыт. В помощь частным инвесторам пришли социальные сети и удобные платформы для совершения сделок (приложения на смартфоне). В результате, в странах Европейского союза доля инвестиций в портфеле сбережений населения составила не менее 40%, в США - более 70%.

За последние 10 лет изменилось и инвестиционное поведение россиян, по крайней мере, состоятельной когорты. К концу 2021 года доля инвестиций в портфеле сбережений россиян за последние два года увеличилась с 13% до 20%. Доля инвесторов (с очищением от двойного счета) достигла 25% экономически активного населения страны к концу первого квартала 2022 года, тогда как в 2020 году доля уникальных клиентов на брокерском обслуживании составила 12% экономически активного населения РФ. Социальная сеть «Пульс», где инвесторы делятся своими инвестиционными стратегиями, результатами их управления, общаются друг с другом, насчитывает по оценкам Оливера Хьюза на 2021 г. 1,5 млн активных пользователей (клиентов Тинькофф-брокер). В IPO «Совкомфлота», по данным Московской биржи, частные инвесторы подали 41,2 тыс. заявок. Порядка 35 тысяч инвесторов-держателей облигаций эмитента потеряли деньги из-за дефолта публичной компании «Обувь России» в 2021-2022 гг.

Портрет начинающего российского инвестора получен в результате исследования ЦБ РФ «Портрет клиента российского брокера»[1].

Объем денежных средств частных инвесторов на брокерских счетах РФ к середине 2021 г. достиг 7,2 трлн рублей, к сентябрю 2022 г. снизился до 5,4 трлн рублей. Для сравнения, на банковских депозитах находится порядка 30 трлн рублей средств физических лиц. По данным Банка России, с апреля по сентябрь 2021 г. частные инвесторы обеспечили более 85% нетто-покупок акций на рынке.

На конец августа 2022 г. число физических лиц, имеющих брокерские счета на Московской бирже, достигло отметки 21,2 млн человек (20,8 млн человек в июле 2022 г.). Всего частными инвесторами открыто 34,2 млн брокерских счетов по состоянию на 30 июня 2022 г. (физические лица продолжают открывать порядка 400-600 тысяч счетов в месяц). Количество открытых индивидуальных инвестиционных счетов (ИИС, которые предо-

[1] https://www.cbr.ru/analytics/rcb/cl_broker/

ставляют налоговые льготы по НДФЛ) составило 5,6 млн (+49 тысяч только за август 2022 г.). На фондовом рынке сделки совершили в августе более 2,1 млн человек (1,9 млн человек в июле 2022 г.). Заметим, что по оценкам Anderida Financial Group, большая часть брокерских счетов российского рынка — пустая (около 65%). Порядка 20% счетов не превышают 15 тыс. руб. Количество активных инвесторов в РФ на середину 2022 г. не превышает 9,4 млн человек (оценка по 2021 г. – около 5 млн. человек). На облигационном рынке число «слонов», входящих в отдельные выпуски ВДО с суммой порядка 3-5 млн руб., не превышает 300 человек. Аналитики вводят отсечку по разделению частных инвесторов от институциональных в первичных размещениях на уровне 10 млн руб.

По данным Минфина, представленным на правительственном совещании 30 августа 2022 г.[1], замороженные международные резервы в результате специальной военной операции (СВО) составили $300 млрд; замороженные евробонды Минфина и корпораций РФ - 4,6 трлн рублей; замороженные розничные продукты и депозиты - 563 млрд рублей. После февраля 2022 г. падение капитализации фондового рынка составило 40%; снижение банковского капитала - 10%, под санкциями оказались 80% активов банковской системы. В иностранных акциях российские инвесторы потеряли в 2022 г. порядка 1,3 трлн руб. (замороженные активы). На конец марта 2022 г. общая стоимость ценных бумаг иностранных компаний на счетах частных инвесторов составила около 3 трлн руб. До СВО российские частные инвесторы вложили в иностранные акции и облигации около 36% всех своих инвестиций. Как отмечает ЦБ РФ, на увеличение объема вложений физических лиц в акции иностранных эмитентов в 2021 г. повлияли, в том числе, запуск торгов иностранными ценными бумагами на Московской бирже и бурное их развитие на СПБ Бирже.

Количество клиентов, передавших управление своими активами в доверительное управление (ДУ) составляло в 2021 г. 0,8 млн лиц, и 0,9 млн человек наблюдалось и по I, и по II кварталам 2022 г. Стоимость портфелей в ДУ в 2021 г. составляла 1,8 трлн руб., а в 2022 г. снизилась до 1,6 трлн руб. к концу I квартала 2022 г. и до 1,5 трлн на конец II квартала (из-за валютной переоценки и укрепления рубля).

Активы физлиц на брокерском обслуживании за II квартал 2022 г. сократились с 6,7 до 5,4 трлн руб. По состоянию на конец марта 2022 г. доля всех акций («замороженных» иностранных и торгуемых российских) в портфеле частных лиц составляла 46%, а доля облигаций - 42%.

После введения санкций против России и оттока нерезидентов структура инвесторов на российском фондовом рынке сильно изменилась. До февраля 2022 г. во владении активами и в объемах торгов высока была доля иностранных инвесторов. Контрсанкции также нанесли существенный урон иностранным инвесторам. На конец января 2022 г. американские взаимные

[1] https://www.bloomberg.com/news/articles/2022-09-14/russia-quietly-adds-up-direct-losses-from-financial-sanctions

фонды и ETF владели российскими акциями и облигациями на сумму более 71 млрд долл. (данные Morningstar Direct). До СВО более 50% оборота на торгах акциями приходилось на долю институциональных инвесторов, по многим бумагам – нерезидентов. В конце февраля 2022 г. ЦБ остановил торги на Мосбирже и СПб бирже из-за обвала рынка после начала СВО России в Украине. Торги возобновились только в конце марта после перерыва на месяц. Иностранные инвесторы не получили допуска к торгам.

В общем объеме торгов акциями (объемы по месяцам см https://www.moex.com/ru/ir/interactive-analysis.aspx#) доля частных инвесторов составила на конец сентября феноменальные 79% (76,2% на конец августа 2022 г., на 30 июня 2022 г. - 74%, на февраль 2022 г. – 43%, в середине 2021 г. - 40%). Доля частных инвесторов в объеме торгов облигациями – 26%, на спот-рынке валюты – 26,4%, на срочном рынке – 70,5%. Только по облигациям торговые обороты вышли на уровень января 2022 г., в отличие от акций и срочного рынка (Табл. 1.2).

Интересна структура спроса инвесторов на выпуски облигаций в юанях (которые начали приходить на российский рынок со стороны российских эмитентов осенью 2022 г. в рамках дедолларизации финансового рынка): 6% - розничные инвесторы (ритейл), 72% - банки, 2% - инвестиционные компании, 20% - управляющие компании (источник: Отчет Акра от 19 октября 2022 г. по долговому рынку (корпоративные облигации)).

Таблица 1.2

Среднедневные объемы торгов по популярным инвестиционным инструментам и активные физлица

Период	*Среднедневной объем торгов, млрд*			*Уникальные физлица*	*Активные физлица*	*Уникальные физлица ДУ*
	облигации	*срочный*	*акции / ДР / паи*			
2021	73,0	619,6	117,4	16 779 069	2 620 855	327 432
Январь	37,6 (17,8%)	794,7 (42,9%)	200,6 (41,2%)	17 405 817	2 841 195	345 411
Февраль	50,6	815,1	231,9	18 128 529	3 133 569	361 205
Март	21,1 (36,1%)	116,3 (68,9%)	74,2 (61,7%)	18 881 687	1 990 756	360 205
Апрель	20,5	198,2	48,4	19 323 213	2 362 460	374 023
Май	22,1 (27,0%)	309,8 (68,6%)	33,2 (72,6%)	19 805 248	1 872 734	377 979
Июнь	29,6 (25,2%)	252,1 (72,0%)	43,3 (74,0%)	20 434 402	1 971 238	390 501
Июль	34,5	304,4	36,6	20 820 354	1 930 771	394 494

По российскому рынку представление о доходности популярных стратегий инвестирования (акции, облигации и др.) дает Табл. 1.3.

Таблица 1.3

Статистика по физическим лицам на инвестиционном рынке по 30 стандартным стратегиям: показатели доходности и риск-доходности на 30.06.2022 г.

Показатель	*Количество стратегий, ед.*	*Количество клиентов - физлиц, тыс. ед.*	*Объем портфелей физлиц, млрд руб.*	*Доля от объема портфелей топ-30, %*	*Доходность (3 мес), %*	*Доходность (12 мес.), %*	*Коэффициент Шарпа*
Весь рынок	601	866,6	300,3	-	-	-	
Топ-30 стандартных стратегий	30	392,6	210,3	100	-19,0	-21,4	-0,8
В том числе по объектам вложений							
облигации резидентов	6	239,7	63,2	30	9,2	-0,8	-0,1
российские акции	2	31,4	17,6	8	-14,1	-39,3	-1,5
облигации иностранных эмитентов	6	18,7	29,4	14	-39,1	-40,5	-1,0
иностранные акции	3	12,2	9,0	4	-43,9	-37,8	-0,8
структурные продукты	13	90,5	91,1	43	-22,6	-18,6	-0,6
В том числе по типам инвесторов							
для квалифицированных инвесторов	11	10,1	58,5	28	-36,9	-36,0	-1,0
для неквалифицированных инвесторов	19	382,5	151,9	72	-8,8	-13,6	-0,7
из них стратегии ИИС	8	298,6	73,7	35	5,4	-6,8	-0,6
В том числе по способу инвестирования							
прямые вложения	12	17,8	54,6	26	-32,7	-34,9	-1,0

Показатель	*Количество стратегий, ед.*	*Количество клиентов - физлиц, тыс. ед.*	*Объем портфелей физлиц, млрд руб.*	*Доля от объема портфелей топ-30, %*	*Доходность (3 мес), %*	*Доходность (12 мес.), %*	*Коэффициент Шарпа*
вложения через ПИФ, структурные облигации или ETF	18	374,8	155,8	74	-12,5	-15,0	-0,7

Источник: Обзор ключевых показателей профессиональных участников рынка ценных бумаг» от Банка России, сентябрь 2022 г.

Ежемесячно Мосбиржа выделяет самые популярные акции в портфелях частных инвесторов.

В августе 2022 г. среди самых популярных бумаг были:

- акции Газпрома (33,9%),
- обыкновенные и привилегированные акции Сбербанка (19,8% и 5,6%, соответственно),
- акции Норникеля (9,5%),
- акции ЛУКОЙЛа (8,9%),
- акции Яндекса (5,3%),
- привилегированные акции Сургутнефтегаза (5,1%),
- акции Роснефти (4,6%),
- акции МТС (3,7%),
- акции НОВАТЭК (3,7%).

Этот набор практически не меняется по месяцам и формирует так называемый «народный портфель».

Из Обзора ЦБ РФ следует, что «на фоне существенного падения курса доллара к рублю и снижения рынка акций» по большинству инвестиционных стратегий (Табл. 1.3) наблюдалась отрицательная доходность.

Единственным классом активов, который показал рост по итогам II квартала 2022 г., стали облигации резидентов.

В структуре объемов торгов (Рис. 1.2) прослеживается перенос интереса и частных, и институциональных инвесторов к облигационному рынку в 2022 г.

Рис. 1.2. Структура биржевых торгов на фондовом рынке РФ с 2019 по 2022 гг.: квартальная статистика

Уровни доходности портфелей розничных инвесторов за прошлые годы, сформированных по фондовым индексам акций и облигаций дает Табл. 1.4.

Таблица 1.4

Номинальная доходность и риск инвестирования на фондовом рынке РФ (портфели акций и облигаций) с 2003 по 2022 гг.

Доля акций	*0%*	*10%*	*20%*	*30%*	*40%*	*50%*	*60%*	*70%*	*80%*	*90%*	*100%*	*Депозиты*	*Инфляция*
Доля облигаций	*100%*	*90%*	*80%*	*70%*	*60%*	*50%*	*40%*	*30%*	*20%*	*10%*	*0%*		
Доходность в процентах													
2003	18.1	21.9	25.8	29.6	33.5	37.3	41.2	45.0	48.9	52.7	56.6	13.3	12.0
2004	6.5	6.6	6.7	6.7	6.8	6.9	7.0	7.0	7.1	7.2	7.3	11.6	11.7
2005	7.1	14.9	22.7	30.4	38.2	46.0	53.8	61.5	69.3	77.1	84.9	10.2	10.8
2006	9.0	15.1	21.1	27.2	33.3	39.4	45.5	51.6	57.6	63.7	69.8	8.5	9.1
2007	6.5	7.1	7.8	8.4	9.0	9.6	10.2	10.8	11.4	12.0	12.7	8.7	11.8
2008	-11.7	-17.2	-22.7	-28.2	-33.7	-39.2	-44.7	-50.1	-55.6	-61.1	-66.6	7.8	13.2
2009	23.7	33.9	44.2	54.4	64.6	74.9	85.1	95.4	106	116	126	11.4	8.9
2010	13.9	15.0	16.1	17.3	18.4	19.6	20.7	21.9	23.0	24.1	25.3	11.0	8.7

Доля акций	*0%*	*10%*	*20%*	*30%*	*40%*	*50%*	*60%*	*70%*	*80%*	*90%*	*100%*	*Де-по-зи-ты*	*Ин-фля-ция*
Доля облигаций	*100%*	*90%*	*80%*	*70%*	*60%*	*50%*	*40%*	*30%*	*20%*	*10%*	*0%*		
2011	6.2	4.1	2.1	0.0	-2.1	-4.1	-6.2	-8.3	-10.3	-12.4	-14.5	6.8	6.2
2012	8.6	8.6	8.6	8.6	8.7	8.7	8.7	8.7	8.7	8.7	8.8	7.7	6.5
2013	8.8	8.6	8.3	8.1	7.8	7.6	7.3	7.1	6.8	6.6	6.3	8.4	6.5
2014	-1.4	-1.5	-1.5	-1.6	-1.6	-1.6	-1.7	-1.7	-1.8	-1.8	-1.8	7.6	11.3
2015	18.3	19.7	21.1	22.5	23.9	25.3	26.7	28.1	29.5	30.9	32.3	15.0	12.9
2016	11.2	13.4	15.5	17.7	19.8	22.0	24.2	26.3	28.5	30.6	32.8	9.8	5.4
2017	12.3	11.0	9.8	8.5	7.3	6.0	4.8	3.6	2.3	1.1	-0.2	8.0	2.5
2018	4.2	5.7	7.2	8.7	10.2	11.7	13.1	14.6	16.1	17.6	19.1	7.0	4.2
2019	14.5	16.9	19.3	21.7	24.1	26.5	28.9	31.3	33.7	36.1	38.5	6.5	3.1
2020	8.3	8.9	9.6	10.2	10.9	11.5	12.2	12.9	13.5	14.2	14.8	5.4	4.9
2021	-0.1	2.1	4.3	6.5	8.7	10.9	13.1	15.2	17.4	19.6	21.8	4.2	8.4
6 мес. 2022	2.1	-2.2	-6.5	-10.9	-15.2	-19.6	-23.9	-28.3	-32.6	-36.9	-41.3	7.1	11.8
Среднегодовая доходность	8.2	9.4	10.4	11.3	12.1	12.8	13.3	13.7	14.0	14.1	14.0	9.0	8.7
Совокупная доходность	369	475	590	710	830	945	1049	1134	1192	1215	1195	435	407
Показатели, учитывающие риск													
Стандартное отклонение	7.6	10.2	13.3	16.6	20.0	23.5	27.0	30.5	34.0	37.6	41.1	2.6	
Коэффициент Шарпа	-0.10	0.04	0.11	0.14	0.16	0.16	0.16	0.16	0.15	0.14	0.12		

Источник: ИК Доход

После допуска к торгам нерезидентов из дружественных стран (сентябрь 2022 г.) давление на динамику цен акций было крайне слабым. Причина в том, что этим нерезидентам на осень 2022 г. запрещено продавать акции многих стратегических компаний. Нерезиденты недружественных стран не

имеют возможности совершать действия с российскими акциями и облигациями. Оптимистичные ожидания по фондовому рынку РФ связаны с возможным приходом денег «депозитчиков» на фоне снижения ключевой ставки ЦБ и общим снижением ставок на облигационном рынке (порядка 240 млрд рублей до конца 2022 г.). Оценки на сентябрь 2022 г. показывают, что средняя максимальная ставка по вкладам топ-10 российских банков, привлекающих наибольший объем депозитов, снизилась до 6,83%. На этом фоне дивидендная доходность индекса Мосбиржи (около 12,9% на следующие 12 месяцев выплат компаниями российского рынка) может представлять интерес для физических лиц. Негативным фактором для физических лиц является закрытие финансовой отчетности крупных компаний-голубых фишек.

Присутствие частных инвесторов на облигационном рынке РФ в 2020–2021 гг. также претерпело изменения. Хотя на фоне больших объемов ОФЗ и иностранных инвесторов роль частных инвесторов в целом по рынку не была ярко выражена, но в отдельных выпусках корпоративных облигаций знаменательным трендом стал приход розничных инвесторов-физических лиц. В 2022 г. этот тренд продолжился. Уже по 2020 г. организаторы размещений отмечали, что в сравнительно небольших выпусках размером 1–5 млрд рублей с кредитными рейтингами BBB+, A и выше частные инвесторы забирали зачастую аллокацию у институциональных инвесторов, покупая в рамках первичных размещений около половины объема выпуска («Киви Финанс» 001Р-01 - 42% выпуска, «Максима Телеком» БО-П01 - 46%, АО «Аэрофьюэллз» 001Р-01 - 73%) и докупая после «в стакане». Еще больше физических лиц (обычно их диагностирование в первичных размещениях идет по сумме в пределах 10 млн рублей на одного участника размещений) наблюдается в сегменте высокодоходных облигаций (ВДО, без рейтинга или с рейтингом ниже BB+). По большинству выпусков ВДО частные инвесторы формируют практически 100% объем держания и торгов.

СВО 2022 г. изменила инвестиционное поведение российских инвесторов-частных лиц и на валютном рынке, когда доллар и евро все чаще стали называть «токсичными», инвесторы вынуждены были обратить внимание на юань, несмотря на то, что эту безналичную валюту практически невозможно конвертировать в «бумажную». Если в апреле 2022 г. доля китайской валюты на биржевом рынке составляла 6%, то в августе - уже 26%.

Заметим, что с начала 2022 г. по начало сентября китайский юань обесценился по отношению к рублю на 26%. По отношению к доллару в летние месяцы юань также обесценился, достигнув минимума за два года.

Россияне во II квартале 2022 г. на фоне резкого укрепления рубля (на бирже евро за второй квартал подешевел к рублю на 42%) инвестировали в иностранную валюту 439 млрд рублей, из них в юани вложено 16 млрд рублей. В покупках валюты доминировал евро, значительная часть приобретенной валюты была переведена на зарубежные счета.

Для сравнения, во II квартале 2022 г. в золото инвестировано всего 3 млрд руб.

Заметим, что в конце февраля-марте (начало СВО) россияне продали валюты на 570 млрд руб. (после начала СВО курс на валютном рынке доходил до 100 рублей за доллар).

Процессы в РФ, запущенные СВО, изменили инвестиционный ландшафт в ближнем зарубежье. Аналитики отмечают, что для Казахстана начало 2022 г. обеспечило самый значительный рост за всю историю фондового рынка страны. Доля граждан, которая торгует на фондовом рынке Казахстана - менее 5,1% от экономически активного населения Казахстана. Однако 2022 г. стал рекордным по открытию брокерских счетов (количество счетов, которые открыли розничные инвесторы в Казахстане, на конец I полугодия 2022 г. составило 367 тысяч, источник: издание Kapital.kz со ссылкой на данные Агентства Республики Казахстан по регулированию и развитию финансового рынка (АРРФР)). СПБ биржа и крупные брокеры, в том числе Freedom Holding, обсуждают идею создания торговой площадки в Казахстане на базе Международного финансового центра «Астана» (МФЦА).

Для сравнения, в Бразилии на начало 2022 г. число розничных инвесторов на фондовом рынке достигло 3,5 млн человек. В Индии только 3–4% населения страны вкладывают свои финансовые сбережения в акции, на начало 2022 г. открыто более 46 млн уникальных счетов частных инвесторов[1].

Яркий пример резких движений цены акций Газпрома летом 2022 г. - скриншот «фейковой новости», которая якобы вышла на сайте РБК, о том, что совет директоров «Газпрома» отложил выплату дивидендов (как позже оказалось, СД проголосовал за дивиденды, а общее собрание отклонило это решение). Когда появилась новость о том, что будут рекордные дивиденды по решению СД (т.е. новость исходно «фейковая»), акции «Газпрома» подорожали до 295 рублей. А после отказа на общем собрании акционеров цена обвалилась.

Последние годы дают много примеров как нерациональных действий розничных инвесторов, так и ярких «войн против умных денег» в ситуации слаженных действий физлиц. Самая яркая история на американском рынке акций – GameStop Inc, когда целенаправленные действия розничных инвесторов-пользователей Reddit (на сайте Reddit и других был призыв к инвестиционному сообществу наказать «шортистов») против коротких позиций институциональных инвесторов стали основной причиной существенного увеличения волатильности и появления short squeeze (когда шортисты вынуждены были принудительно закрывать позиции, теряя деньги) на торгах акциями с тикером NASDAQ:GME. Предыстория этого кейса такова, что по акциям до этого малоизвестной компании GameStop Inc фиксировалась

[1] https://www.bricsmagazine.com/ru/articles/strana-na-vyrost

большая короткая позиция: многие инвесторы, особенно институциональные, ставили на то, что акции компании будут снижаться. «Умные деньги» не видели фундаментальных перспектив у этой компании. Но тут в игру вступили розничные инвесторы, объединенные идеей «наказать шортистов» и «смести их с рынка». В период с 21.01.2020 г. по 04.02.2020 г. цена закрытия акции поднялась на бирже с $43.03 до $347.51, после чего вернулась к значению $90. Институциональные инвесторы понесли огромные убытки и обращались за защитой к ФРС США.

В работе Umar et al. (2021) анализируется взаимосвязь между доходностью мемных акций компании GameStop и индикаторами настроения инвесторов, такими как количество публикаций в социальной сети Твиттер и соотношение «пут-колл» (отношение количества опционов на продажу и покупку). Для анализа применяется метод вейвлет-когерентности, который, в отличие от традиционного моделирования временных рядов, фиксирует взаимодействие между двумя временными рядами как во временной, так и в частотной областях. Выборка исследования Umar et al. (2021) базируется на ежедневных данных за период с 1 по 30 января 2020 г., охватывающий противостояние между розничными инвесторами-пользователями Reddit, которые покупали акции компании, базируясь на иррациональных предпосылках, и институциональными инвесторами, которые осуществляли продажи акций из-за слабых фундаментальных показателей. Коллективные действия розничных инвесторов привели к 21-кратному росту акций GameStop в течение одного месяца, что нанесло огромный ущерб институциональным инвесторам.

В работе Umar et al. (2021) делается вывод, что настроения розничных инвесторов в социальных сетях могут привносить неэффективность на рынок. Такая метрика, как соотношение «пут–колл» значимо положительно влияло на доходность акций GameStop до кульминации истории с завышением цен и являлось одним из факторов стремительного роста цены. Авторы подчеркивают важность возможности ограничения регулирующими органами коротких продаж без покрытия, поскольку они могут привести рынок к существенной неэффективности.

Одна из причин ярких проявлений нерациональности, с одной стороны, и массированных атак на институциональных инвесторов, с другой, - приток розничных инвесторов на фондовые рынки. Согласно РБК, «количество брокерских счетов утроилось в 2020 г. по сравнению с 2019 г., писал Bloomberg со ссылкой на данные портала по сравнению услуг от различных брокеров BrokerChooser. Ритейловые инвесторы после начала коронакризиса обеспечили две трети оборота торгов в Южной Корее. В Саудовской Аравии практически каждый третий имеет брокерский счет. В США в 2020 г. доля розничных инвесторов на фондовом рынке достигла 20%»[1]. Руководитель отдела по инвестиционному консультированию «Тинькофф

[1] https://www.rbc.ru/newspaper/2021/03/05/6040bd3c9a794720759d07df

Инвестиции» Кирилл Комаров в статье РБК отмечает риски прихода физических лиц в том, что инвесторы могут кооперироваться и совершать похожие сделки в инструментах с недостаточной ликвидностью. «Это может привести к росту волатильности не только в отдельных акциях, но и на рынке в целом. Особенно опасен этот феномен в случае серьезных кризисов и потрясений - все мы помним март 2020 г., когда инвесторы массово начали в панике продавать активы». К. Комаров отмечает, что синхронные действия частных инвесторов уже заметны на котировках не самых крупных российских компаний: «по одной из версий, именно они могли стать причиной взлета акций «Белуги» (которые затем, впрочем, обвалились) и других эмитентов в феврале 2020 г.». За несколько дней февраля 2021 г. на Московской бирже прошли совокупные обороты, равные общему free-float акций «Белуги».

Кейс акций «Белуги» (дистрибьютер алкоголя, продает водку, виски, коньяк, бренди, ром и прочие напитки) стал ярким, но не единственным «примером сильного роста малоликвидных бумаг. В апреле 2009 г. низколиквидные акции девелопера РТМ взлетали в десять раз за две недели, потом также быстро падали. Необычно сильные движения в ценах в период ковид-паники фиксировались в акциях компаний «Абрау-Дюрсо», «Красный Октябрь», «Русгрэйн», «Левенгук», «Квадра» (источник: РБК).[1]

«Акции «Белуга» в январе-феврале 2021 г. сначала взлетели почти на 200% за несколько дней, а затем рухнули на 64%. 17 февраля 2021 г. объем торгов по акциям «Белуги» превысил 1 млрд руб. По объемам торгов, они обогнали такие «голубые фишки», как АФК «Система», ММК, «Северстали», префы «Сургутнефтегаза» и бумаги самой Мосбиржи. При этом в 2020 г. средний дневной торговый оборот акциями «Белуги» составил чуть более 4 млн руб. 19 февраля 2021 г. акции «Белуги» пробили отметку в 6400 руб, хотя еще 15 февраля они стоили всего 2200 руб. Однако в тот же день котировки начали рушиться: за два неполных торговых дня они почти вернулись на уровень начала недели - до 2469 руб. за бумагу» (источник: РБК[2]).

При этом никаких ярких корпоративных новостей, связанных с компанией, не было. Финансовые результаты компания обычно публикует в марте. 25 января 2021 г. компания опубликовала операционные результаты по 2020 г. с указанием роста объемов отгрузки (+10,7%) и продажам партнерских импортных брендов (+42,1%). В феврале 2021 г. в СМИ появилась информация о возможной продаже «Белугой» птицефабрики «Михайловская» мощностью 30 тыс. тонн группе агропредприятий (ГАП) «Ресурс». Но эта новостная лента вряд ли могла утроить цену акции. Возможное объяснение аномального роста и падения цены «Белуги» - разгон через Телеграм каналы. В декабре 2020 г. компании были посвящены большие посты

[1] https://quote.rbc.ru/news/article/6030e4459a7947587a748e1d

[2] https://quote.rbc.ru/news/article/6030e4459a7947587a748e1d

в Телеграм каналах инвестиционной тематики, и подчеркивался «рост винного сегмента, недооцененность акций «Белуги» по отношению к конкурентам, развитие ее розничной сети «Винлаб» и предполагаемый интерес к покупке компании со стороны ее конкурентов».

С началом пандемии Ковид-2019 биржи мира испытали существенный приток новых клиентов. Многие инвесторы в торговле акциями нашли альтернативу ставкам на спортивные события, количество которых снизилось в период пандемии. Соответственно, такие инвесторы стали искать активы, которые имеют некоторые «лотерейные» характеристики. Появился отдельный «класс активов» - мемные акции.

Специфика инвестиционного поведения розничных инвесторов наглядно продемонстрирована в работе Barber and Odean (2009), предложивших в качестве прокси экстраординарные события, наблюдаемые во время торгов той или иной акцией (резкие изменения объема торгов, аномальная однодневная доходность, повышенный поток информационного фона). Используя данные о торгах американских брокерских компаний двух типов, авторы указывают на существование диспропорций в buy-sell структуре торгов. Брокеры с широкой базой розничных клиентов являются чистыми покупателями акций в ситуации экстраординарных событий и повышенного внимания. Объем покупок на 17,67% больше объема продаж для акций из дециля с наибольшим отклонением объема торгов от среднего, на 11,13% больше для акций из дециля с самой высокой доходностью за предыдущий день. Для акций с более широким новостным покрытием количество заявок на покупку больше почти на 17%. Эта повышенная активность розничных клиентов не претерпевает коррекции даже при добавлении фильтра по размеру компаний (рыночной капитализации). Barber and Odean (2009) показывают, что значения вышеуказанных диспропорций существенно ниже для профессиональных инвесторов. С точки зрения авторов, это может быть объяснено лучшим пониманием происходящих на рынке событий, а также зачастую наличием сформулированных стратегий торговли, которым придерживаются профессионалы и которые ограничивают возможности эмоциональных покупок или продаж.

Несколько теоретических конструкций претендуют на объяснение поведения частных инвесторов, которые отходят от рационального выбора. Одна конструкция (Caplin et al., 2011) подчеркивает важность учета поведенческого искажения статуса-кво (status-quo bias). Авторы экспериментальным путем показывают, что задача поиска индивидом оптимального решения может быть приостановлена при появлении удовлетворяющего его предпочтениям первого решения (при достижении некого «резервного» уровня полезности). Решение такого инвестора (которого устроил промежуточный результат, некий статус-кво) не будет совпадать с оптимальным. Инвесторов следует рассматривать как ограниченно рациональными из-за сужения анализируемого информационного множества и принятия промежуточного уровня удовлетворения.

Также на выбор инвестора может влиять эффект эмоциональной привязки (anchoring bias), «якорения». Для частного инвестора та или иная акция может иметь дополнительную субъективную ценность (эмоциональную привязку, как, например, с ярыми последователями Илона Маска и покупкой акций Тесла), из-за чего готовность купить или продать определенный актив может быть искажена. В соответствии с данным эффектом инвестор может искаженно интерпретировать изменения внешних факторов, влияющих на акцию, и принимать нерациональные решения. Например, инвестор может откладывать продажу убыточной позиции из-за того, что акция была в его портфеле с момента его формирования и на тот момент выглядела привлекательной. Времени и способностей провести аналитику всех акций в портфеле у инвестора может не быть (особенно, когда в соответствии с рекомендациями профессионалов инвестор формирует портфель из 100-200 бумаг как широко диверсифицированный). А внешняя среда (новости, аналитические заметки, новые риски для компании) может указывать на то, что в дальнейшем данный актив принесет убытки. Но инвестор ищет новые активы в свой портфель, считая, что «старые активы» уже прошли анализ и диверсификация спасет его от больших убытков.

Согласно работе Абрамова А.Е. и др. (2021), «в США количество фондируемых счетов клиентов у трех крупнейших брокеров (Charles Schwab, Fidelity Investments и Robinhood) выросло с 51 млн в 2019 г. до 84 млн в сентябре 2021 г., т.е. в 1,6 раза» (Рис. 1.3).

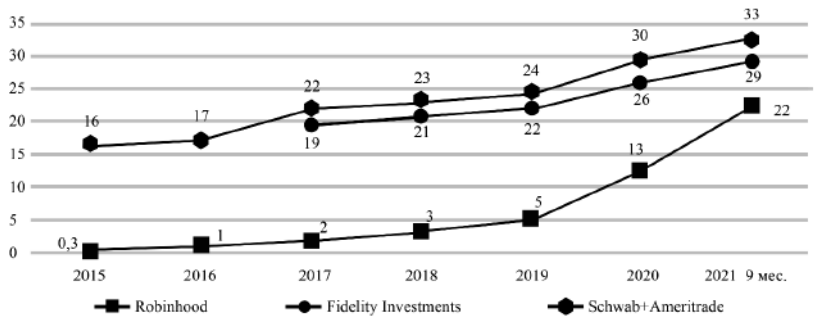

Рис. 1.3. Количество фондируемых счетов частных лиц у крупнейших онлайн брокеров в США с 2015 г. по сентябрь 2021 г., млн

Источник: расчеты А.Е. Абрамова и соавторов по данным отчетности финансовых компаний (Абрамов и др., 2021).

Согласно работе Абрамова А.Е. и др. (2021), «на российском фондовом рынке приток частных инвесторов идет даже более активно, чем в США (Рис. 1.4). С 2019 г. по октябрь 2021 г. общее число счетов клиентов брокеров на Московской бирже выросло с 3,9 млн до 15,3 млн, или в 3,9

раза, а число счетов активных клиентов брокеров на Санкт-Петербургской бирже – с 87 тыс. до 825 тыс., или в 9,5 раза. За тот же период существенно увеличилось число пайщиков открытых и биржевых ПИФов – с 0,5 млн до 3 млн, или в 6 раз». Банки «Тинькофф», «Сбер» и ВТБ обеспечили 82,5% прироста счетов клиентов.

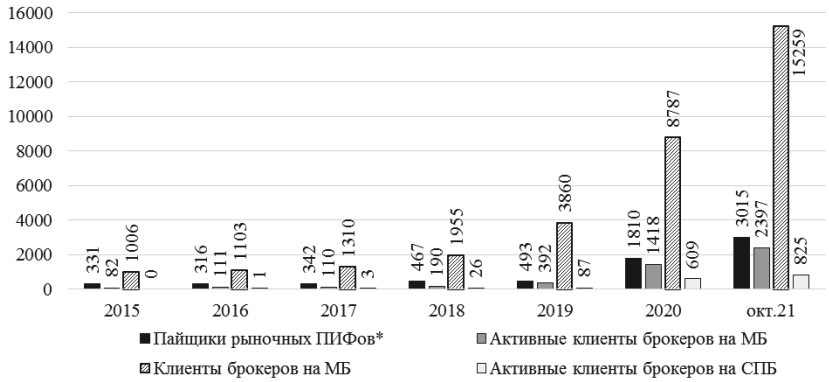

Рис. 1.4. Количество зарегистрированных и активных счетов клиентов брокеров на Московской бирже (МБ), количество активных счетов клиентов брокеров на Санкт-Петербургской бирже (СПБ) и численность пайщиков в открытых и биржевых паевых инвестиционных фондах (ПИФах) с 2017 г. по октябрь 2021 г., тыс.

* в 2021 г. данные по состоянию на июнь месяц.

Источник: расчеты Абрамова А.Е. и соавторов по данным Московской и Санкт-Петербургской бирж и информационно-аналитического ресурса "Эксперт-РА" (Абрамов и др., 2021).

«Согласно данным Банка России, по итогам II квартала 2021 г. доля физических лиц со счетами без активов у брокеров составляла 62%», т.е. фондируемыми являлись только 5,1 из 13,5 млн счетов, что искажает картину участия физических лиц на фондовом рынке (Абрамов и др., 2021).

Если по итогам 2016 г. общая сумма инвестиций граждан на брокерских счетах составила около 520 млрд рублей, то по итогам 2020 г., то есть спустя 4 года, она выросла в 10 раз — до 5,3 трлн рублей. Практически на этот уровень опустилась величина активов на начало октября 2022 г.

«С 2017 г. в России происходят значительные изменения структуры финансовых активов домашних хозяйств» (Табл. 1.5): снижение доли депозитов в банков и доли сбережений в виде пенсионных и страховых резервов, а также рост прямых вложений в акции и облигации, что объясняется стремлением получить более высокую доходность (Абрамов и др., 2021).

Таблица 1.5

Структура рыночных финансовых активов домашних хозяйств в России с 2017 г. по июнь 2021 г., %

Показатель	*2017*	*2018*	*2019*	*2020*	*01.07.2021*
Наличные средства	24,4	25,1	23,1	26,1	26,0
Депозиты	57,1	55,5	55,2	50,9	49,6
Акции	4,4	4,3	5,4	6,5	7,1
Облигации	2,4	3,1	3,8	4,4	4,6
Паи инвестиционных фондов	2,3	2,7	2,9	3,4	3,8
Пенсионные и страховые резервы	9,5	9,3	9,5	8,7	8,8
Финансовые активы всего	100	100	100	100	100

Источник: расчеты Абрамова А.Е. и соавторов (Абрамов и др., 2021)

1.2. ЭФФЕКТЫ ВНИМАНИЯ И МЕТРИКИ ЗАМЕТНОСТИ КАК ПРОКСИ-СЕНТИМЕНТА И ИССЛЕДОВАНИЯ ПО ВЫЯВЛЕНИЮ СВЯЗИ С БИРЖЕВЫМИ ХАРАКТЕРИСТИКАМИ АКТИВОВ[1]

В условиях современного потока информации в результате появления интернета и широкого доступа инвесторов к различным платформам и сообществам инвесторов непрофессиональные участники физически не могут отслеживать и обрабатывать всю информацию, поступающую на рынок. Зачастую им сложно не только отобрать информацию, просмотрев большое количество источников, но и аналитически обработать ее (понять о какой компании идет речь, оценить потенциальное влияние на будущее поведение цены).

Поэтому согласно теории внимания (заметности, видимости, «attention theory), предложенной в ряде работ Barber and Odean (2006, 2007, 2008), инвесторы из тысячи акций, торгующихся на бирже, концентрируются на тех, которые тем или иным образом привлекают их наибольшее внимание и далее «якорятся в них». Аналитики отмечают немалое число казусов, когда инвесторы по ошибке принимают положительные новости относительно другой компании за акции той, которая совпадает по названию или быстрее находится в торговом терминале. Яркий пример 2022 г. – взлет цены акций машиностроительной публичной компании с крайне неликвидными акциями «Звезда» MCX:ZVEZ[2]. Президент РФ Путин В.В. 18 августа 2022 г. сообщил, что государство поддержит судоверфь «Звезда» на Дальнем Востоке (непубличная компания, будет обеспечить танкерами новый СПГ-проект «Новатэка»), а инвесторы по ошибке начали скупать акции одноименной

[1] Обзор ряда работ в данном параграфе выполнен студентом факультета экономических наук НИУ ВШЭ Мажовым М. под руководством Т.В. Тепловой.

[2] https://ru.investing.com/news/stock-market-news/article-2176589, https://ru.investing.com/equities/zvezda

машиностроительной компании, цена подскочила с 3,1 до 4,12 руб. за акцию на конец рабочего дня. Акции производителя двигателей «Звезда» на Мосбирже в моменте росли на 31%.

Аналогичная история с покупкой акций публичной компании со схожим названием вместо непубличного стартапа имела место и в США в 2013 г. Инвесторы по ошибке стали скупать акции малоликвидной компании после выхода на рынок информации, что Google поглощает за $3,2 млрд стартап Nest (разработка умных термостатов и других компонентов систем умного дома). «В итоге инвесторы, еще вчера не слышавшие это название, в надежде на рост стоимости акций Nest купили их так много, что ценные бумаги выросли на 1900%. Проблема в том, что купленный Google стартап Nest Labs был частной компанией и не выпускал акции для торговли на бирже. Все это время люди покупали акции другой компании – тикер NEST относился к фирме Nestor Inc. Эта компания разрабатывала системы контроля трафика для правительственных заказчиков. В конечном итоге акции NEST были выведены с биржи»[1]. «Более того, к моменту старта всей этой истории фирма уже обанкротилась и распродала все свои активы, но ее акции еще можно было купить. И как оказалось, подобные всплески стоимости этих акций случались и раньше. Еще до поглощения Google, акции NEST взлетали в стоимости на 10000% – после новостей о релизе новой системы обнаружения дыма от Nest Labs»[2].

Кейс обвала цены на 4% акции Electronic Arts 19 июля 2019 г. из-за потери лицензии популярного футбольного клуба в FIFA 20 можно найти по ссылке: https://habr.com/ru/company/iticapital/blog/467865/. Компания потеряла лицензию на использование названия и изображения формы итальянского «Ювентуса» (многократного чемпиона Италии, в составе которого играет звезда мирового футбола Криштиан Роналду). Инвесторы решили, что с потерей бренда «Ювентуса» из игры-симулятора исчезнет и Роналду. «Однако, на самом деле права на использование реальных составов команд итальянской лиги у EA осталось, так что изменилось лишь название команды и ее форма. Поэтому впоследствии цена акций скорректировалась»[3].

Еще один казусный случай - взлет и падение акций Nintendo в 2016 г. на волне популярности Pokemon GO. «С ростом популярности игры начали дорожать и акции корпорации. Стоимость ценных бумаг Nintendo меньше чем за неделю выросла на 50%, а общий объем торгов ими превысил $10 млрд. Интересный момент заключался в том, что разработкой игры занимались The Pokemon Company and Niantic – именно эти названия появляются на стартовых экранах приложения. Nintendo также участвовала в проекте, но ее доля не превышала 32%. Руководство японской компании было вынуждено публично заявить о том, что именно разработкой игры Nintendo не

---

[1] https://habr.com/ru/company/iticapital/blog/467865/

[2] https://habr.com/ru/company/iticapital/blog/467865/

[3] https://habr.com/ru/company/iticapital/blog/467865/

занималась. Вскоре после этого «узнавшие правду» инвесторы начали тотальную распродажу акций, которая привела к падению их стоимости на 30%. Таким образом, капитализация Nintendo все равно осталась в плюсе по итогам этой истории».[1]

Заметим, что менеджмент компании зачастую отлично понимает, как «раскрутить свои акции» и проводит ребрендинг и ренейминг, используя популярные у инвесторов слова-приманки. В 1990-х гг. «в моде» был ренейминг с использованием «Internet», «.net» и «.com» (Рис. 1.5).

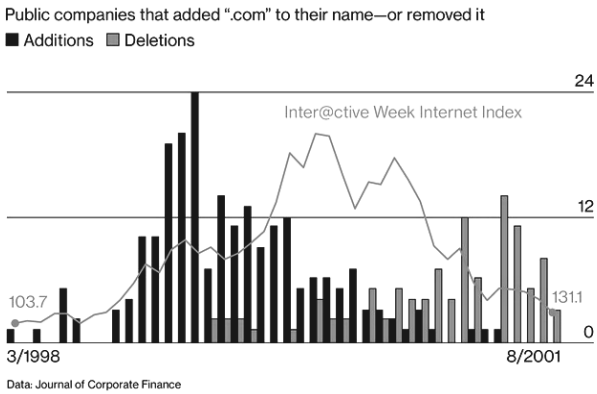

Рис. 1.5. Ренейминг публичных компаний США с акцентом на Интернет привязку в 1999-2001 гг.

Источник: https://www.bloomberg.com/news/articles/2018-06-18/the-stock-pop-from-a-crypto-rebrand-doesn-t-last, 18 июня 2018

В 2017-2018 гг. популярным приемом ребрендинга стало обыгрывание слова «блокчейн»[2], когда даже компании, далекие от IT технологий, подчеркивали свою близость к майнингу криптовалют и технологии блокчейна. «Когда в декабре 2017 г. табачная компания Rich Cigars Inc. стала Intercontinental Technology, Inc. и сообщила, что планирует заниматься майнингом, котировки ее акций возросли на 2233%. Такой же вираж совершили еще несколько компаний из области медицины, строительства, пищевой и легкой промышленности»[3], Рис. 1.6.

[1] https://habr.com/ru/company/iticapital/blog/467865/

[2] https://habr.com/ru/company/iticapital/blog/414955/

[3] https://www.bloomberg.com/news/articles/2018-06-18/the-stock-pop-from-a-crypto-rebrand-doesn-t-last

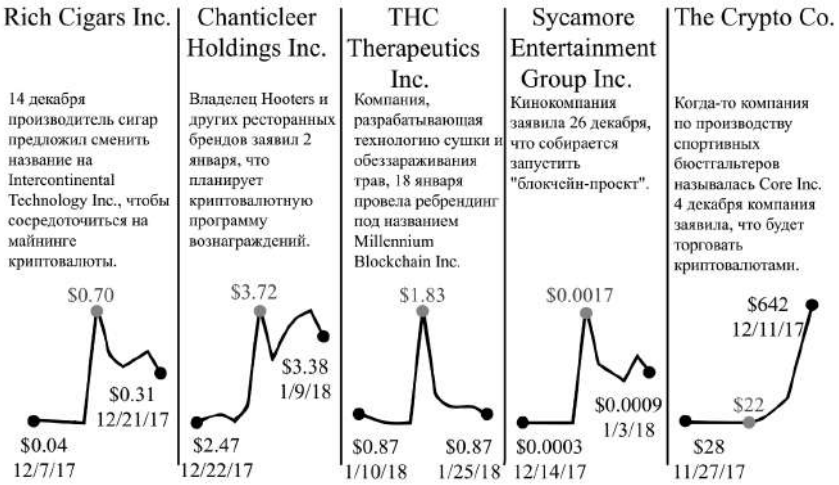

Рис. 1.6. Кейсы взрывного роста цен акций на ребрендинге по слову «блокчейн»
Источник: Блумберг, 18 июня 2018 г.

«Эффект внимания» - действительно феномен, который на экспериментах «коктейльной вечеринки» был продемонстрирован Cherry (1953). Суть выявленного эффекта заключается в способности узнавать определенную полезную информацию среди голосового шума и сосредотачиваться на одном важном для слушателя диалоге в оживленной обстановке. Феномен заключается в том, что слушатель способен услышать свое имя из игнорируемых ранее обсуждений среди различных групп людей и даже повторить вслух то, что было доступно только для одного его уха. Предполагается, что первоначально информация поступает в мозг через органы чувств и сохраняется в сенсорной памяти. В мозге есть механизм фильтра, который обеспечивает селекцию и предохраняет канал от перегрузки. Фильтр пропускает только желаемую информацию (или ослабляет нежелательную), которая может выбираться на основе ее физических характеристик, таких как местоположение и громкость.

Позднее идеи наличия фильтров в обработке информации и доказательства ограниченной рациональности частных инвесторов развили Канеман и Тверски. Объем доступной информации намного больше, чем могут воспринять индивиды, и поэтому внимание является дефицитным ресурсом (Kahneman, 1973). Заметим, что в классической теории финансов и ценообразования финансовых активов делается допущение о том, что инвесторы обладают бесконечными когнитивными ресурсами. В действительности, у инвестора нет возможности следить за всей поступающей информацией,

что может привести к риску игнорирования существенной информации, что, в свою очередь, может стать причиной временных рыночных неэффективностей. Ограниченная рациональность инвесторов растет по мере увеличения доступа и скорости потока публичной информации (информационные базы данных Блумберг, Томсон Рейтер, журналы, газеты, сайты инвестиционных компаний, приложения брокеров, социальные сети и т.д.).

Klemola (2020) на основе поисковых запросов как прокси-метрики внимания показал на данных по акциям США, как неожиданные изменения в интересе инвесторов к компании влияют на последующую рыночную премию. Для этого были собраны данные по поисковым запросам из Google Trends, на основе которых рассчитывались еженедельные значения метрики внимания, включаемой с лагом в регрессионную модель. Автор установил, что акции менее устойчивых с финансовой точки зрения компаний более чувствительны к изменениям запросов инвесторов. Неожиданный рост позитивных (негативных) настроений инвесторов предсказывает положительную (отрицательную) рыночную премию акций американских компаний.

Рынок акций не единственный класс инвестиционных активов, на котором проявляются эффекты внимания. Goddard et al. (2015) подтвердили положительную статистически значимую связь между вниманием инвесторов и волатильностью обменных курсов среди семи валютных пар. Показатель внимания инвесторов на валютном рынке строится с использованием индекса поисковых запросов. В работе показано, что на основе метрики внимания инвесторы могут строить прогнозы будущей волатильности доходности валюты и формировать прокси-фактор риска на валютном рынке.

Ограниченное внимание оказывает существенное влияние на ликвидность на рынке ценных бумаг. Corwin and Coughenour (2008) подтвердили, что чем выше метрика внимания у ценной бумаги, тем выше уровень ликвидности. Bank et al. (2011) доказали, что увеличение поисковых запросов в Интернете приводит к росту торговой активности.

Повышенное внимание инвесторов к активу влечет рост волатильности его доходности. Aouadi et al. (2013) исследуют влияние внимания инвесторов на торговую активность и волатильность на рынке акций Франции. Была обнаружена высокая корреляция между вниманием инвесторов и объемом торгов, причем внимание оказывает существенное влияние на неликвидность и волатильность фондового рынка после учета влияния кризиса.

Интересный результат выявлен по взаимосвязи эффекта внимания (готовности инвесторов отслеживать информацию) с календарными эффектами на фондовом рынке. Chakrabarty et al. (2018) выявили, что на решения инвесторов существенно влияют день недели и дни, когда несколько компаний одновременно публикуют плохую отчётность, что авторы рассматривают как прокси показатель внимания (обострения или, наоборот, «замыливания зрения»). В работе Dellavigna and Pollet (2009) сопоставляется разница между эффектами объявлений о прибыли (earnings announcements) в пятницу и в другие дни недели. Интересный вывод - инвесторы в пятницу

менее внимательны, что сопровождается уменьшенной немедленной реакцией на 15% и увеличенной отложенной реакцией (дрифтом) на 70% на публикацию негативных финансовых показателей. Таким образом, предвыходной день порождает фактор ограниченного внимания (потери бдительности) и проявляется в недостаточной реакции инвесторов, что объясняет наблюдаемый дрифт цены после объявления прибыли и убытков компании.

Исследовательская проблема заключается в том, что нет однозначного вывода относительно взаимосвязи повышенного (и упущенного) внимания и доходности ценных бумаг на разных временных горизонтах. Более того, существуют три конкурирующие теории, подкрепляемые эмпирическими результатами.

«Теория осведомленности инвестора» (the investor recognition hypothesis) Мертона (Merton, 1987), утверждает, что акции компаний, привлекающие меньше внимания, должны демонстрировать более высокую ожидаемую доходность. Отсутствие внимания (аналитического покрытия, обсуждений) фактически должно рассматриваться как принятие инвестором большего идиосинкратического (уникального) риска. Мертон предложил равновесную модель ценообразования активов в условиях неполноты информации, способную объяснить часть вариации ожидаемых доходностей, не покрываемой рыночным риском. Стоимость компании растет по мере увеличения степени ее признания инвесторами при сохранении остальных финансовых показателей неизменными. Для вывода модели делается ключевое допущение о том, что инвесторы при построении своих оптимальных портфелей ориентируются только на те ценные бумаги, информацией о которых они владеют. Если инвесторы не информированы о каких-то акциях, хотя они торгуются на бирже, они не будут включать их в свои портфели. Но если небольшая группа инвесторов знает об определенной ценной бумаге, а остальные не знают о ней, то для обеспечения равновесной цены на эту акцию, информированные инвесторы должны занять большие недиверсифицированные позиции в этом активе. Возрастающий при этом специфический (уникальный, идиосинкратический) риск должен быть вознагражден рынком в терминах доходности. Этот эффект аналогичен компенсации риска низкой ликвидности. За принятие на себя повышенного риска инвесторы требуют премии, связанной с позициями в этой ценной бумаге. Степень «признания инвесторами» конкретной акции или облигации может быть определена по количеству инвесторов, владеющих информацией об этом активе. Таким образом, чем более заметен актив, тем на меньшую доходность (при прочих равных) может рассчитывать инвестор в равновесии.

Альтернативная «теория внимания» (attention theory) была предложена позднее Barber and Odean (2009). Инвесторы предпочитают покупать акции, которые в большей степени привлекают их внимание, прокси которого могут выступать анормальные объемы торгов, экстремальные уровни доходностей или большой новостной фон. Для инвесторов приобретает важность не только само содержание рыночной информации, но и способы ее представления. Например, Hirshleifer et al. (2004) показывают, что одна и та же

информация, но представленная менее заметным образом (сноски вместо текста), может повлиять на восприятие инвесторов и на цены активов, соответственно. Huberman and Regev (2001) обнаружили, что просто упоминание небольшой биотехнологической компании в газете New York Times оказало выраженный положительный эффект на ее акции.

Аналитики отмечают и другие эффекты, когда заметно влияние на биржевые характеристики акций. Акции тех компаний, которые показывали свою рекламу во время трансляций SuperBowl с 1969 по 2001 гг., демонстрировали значимую положительную доходность (Fehle et al., 2005). Аналогично, акции тех компаний, которые упоминались в японской ТВ-программе «ProjectX», демонстрировали тенденцию к росту в период после трансляций (Takeda & Yamazaki, 2006). Meta (2015) доказывает, что если средства массовой информации обращают внимание на конкретную акцию, то это вероятно привлечет внимание инвесторов и повлечет за собой изменение цены. Barber and Odean (2000, 2008) эмпирически подтверждают наличие эффекта внимания на фондовом рынке США, обнаружив положительную связь между доходностью актива и уровнем внимания к этой бумаге.

Суть теории внимания сводится к тому, что происходит нарушение равновесия между спросом и предложением по рассматриваемой ценной бумаге. Индивидуальные инвесторы склонны покупать те акции, которые «более яркие», привлекают их внимание, так как у них нет достаточно времени и ресурсов для тщательного исследования всего множества торгующихся на бирже акций. Но возникает асимметрия, потому что такая проблема анализа не наблюдается в ситуации с продажей акций, потому как инвесторы уже владеют этими акциями и имеют достаточно информации по ним. Gervais et al. (2001) обнаружили, что акции с высоким объемом торгов в течение дня или недели имеют тенденцию к изменению цены в положительную сторону. Более высокая популярность акции, вызванная большим объемом торгов, увеличивает количество потенциальных покупателей, но не количество потенциальных продавцов, что приводит к росту доходности.

Barber and Odean (2008) предполагают, что, если внимание и полезность инвесторов являются ортогональными или, как минимум, отрицательно коррелированными величинами, то покупка акций с повышенным вниманием может привести к падению их ожидаемой доходности (напомним, что в теории Мертона акции с повышенным вниманием претендуют на пониженную доходность).

Третье объяснение в поведении доходности - «теория выдающихся событий» (salience theory, заметных событий) Bordalo et al. (2012). Индивиды, принимающие решения, преувеличивают вероятность заметных событий, если они знают о потенциальной возможности их наступления. Под «эффектом заметных событий» (salience) авторы понимают непропорциональное взвешивание разной информации, когда высокая вероятность приписывается наступлению результата (получения дохода или убытков) именно по заметным, обсуждаемым событиям. Остальной поток информации не игно-

рируется, но получает меньший вес в прогнозе общего итога. Таким образом, влияние «заметных обсуждаемых событий» меняет спрос на рискованные активы (сдвигает равновесие), поскольку экстремальные выплаты приобретают непропорциональный вес в процессе рыночного ценообразования. Barber and Odean (2008) предполагают, что покупка акций, инициированная повышенным вниманием со стороны большинства инвесторов (некое выдающееся событие), может временно привести к росту цены (моментум эффект), но в дальнейшем последует реверсия. Seasholes and Wu (2007) обнаружили статистически значимую отрицательную связь между выдающимися событиями на рынке и возвратом к среднему (mean reversion in prices). Феномен переоценки акций при IPO (эффект апсайда первого дня торгов) как раз укладывается в рамки «теории выдающихся событий». Более того, доходность сделок частных инвесторов по ценным бумагам с повышенным уровнем внимания снижается в последующие дни после событий, привлекающих всеобщее внимание и обсуждение (Yuan, 2015).

Таким образом, эффект внимания положительно связан с доходностью акций в краткосрочном периоде и отрицательно в долгосрочном периоде (Blankespoor et al., 2018).

Повышенное внимание инвесторов к активу влечет рост волатильности его доходности. В работе (Aouadi et al., 2013) исследуется влияние внимания инвесторов на торговую активность и волатильность на рынке акций Франции. Была обнаружена высокая корреляция между вниманием инвесторов и объемом торгов, причем внимание оказывает существенное влияние на неликвидность и волатильность фондового рынка после учета влияния мирового финансового кризиса.

Открытый вопрос – адекватная метрика внимания. В существующих работах можно встретить достаточно большое разнообразие исследуемых прокси-метрик внимания. В начальных работах по данной тематике исследовались, в том числе, расходы на рекламу (Grullon et al., 2004; Chemmanur & Yan, 2009; Lou, 2014), новости и их заголовки (Barber & Odean, 2009; Yuan, 2015), объем торгов (Gervais et al., 2001; Barber & Odean, 2009; Hou et al., 2009), охват медиа (media coverage) (Engelberg & Parsons, 2011; Drake et al., 2012; Yuan, 2015), число выходящих аналитических отчетов (Drake et al., 2012). Однако, все перечисленные прокси имеют свои недостатки. Например, расходы на рекламу, новости и охват медиа могут быть измерены только в момент выхода новостей, поэтому не могут гарантировать того, что значительное число инвесторов будет следить за этими источниками информации. Увеличение объема торгов может быть результатом манипуляций на рынке, осуществляемых группой инвесторов, что также делает этот показатель не лучшим показателем внимания инвестиционного сообщества (Tantaopas et al., 2016). Далее в п. 1.3 и 1.5 продемонстрированы эмпирические тесты эффектов внимания на разных рынках капитала.

1.2.1. Метрики внимания на поисковых запросах[1]

Попытка перейти от общерыночного сентимента к сентименту по отдельным акциям или инвестиционным активам на динамической основе породила большое исследовательское направление анализа поисковых запросов в качестве прокси-метрики заметности. Da et al. (2011) оценивают внимание инвесторов с помощью частоты поисковых запросов в Google (традиционно сокращенная метрика GSV – Google Search Volume) и доказывают, что наблюдается положительная и статистически значимая взаимосвязь между прокси фактором внимания инвесторов и доходностью акций. Открытым остается вопрос выбора площадки для подсчета поисковых запросов (например, по РФ – Яндекс или Google), а также названия компании, потому что не всегда поиск может быть реализован по тикеру или ISIN коду биржи. Отдельные поднаправления – исследования не только доходности как зависимой переменной, но и другие биржевые характеристики (волатильность, ликвидность акций), а также прогнозы динамики цены на разных временных горизонтах (некая попытка оценить «память инвесторов»).

Bank et al. (2011), Takeda and Wakao (2014) и Adachi et al. (2017) использовали для получения данных по запросам имя компании вместо тикера, как и в работе Da et al. (2011). Часть авторов исследовала данные по отдельным компаниям (например, Joseph et al., 2011 и Takeda & Wakao, 2014), тогда как другие авторы изучали эффект внимания на уровне композитных индексов (Tantaopas et al., 2016).

Исследователи исходят из того, что метрика GSV отражает в первую очередь степень внимание со стороны розничных (частных, ритейловых) инвесторов, т.к. институциональные инвесторы в большинстве случаев прибегают к другим, более сложным алгоритмам поиска информации (например, с использованием баз данных Bloomberg, Refinitiv, Capital IQ и других).

В работе Aouadi et al. (2013) на данных 209 европейских фирм показано, что GSV имеет положительное влияние на ликвидность акций, входящих во французский индекс CAC 40. Также найдена статистически значимая положительная взаимосвязь между показателем GSV и стандартным отклонением доходности акций (прокси волатильности). Найденная взаимосвязь сохраняет положительный знак в течение всего рассматриваемого периода 2004-2010 гг., т.е. она не обусловлена существенными колебаниями цен активов во время мирового финансового кризиса 2008-2009 гг.

Behrendt et al. (2020) идут еще дальше и пытаются строить торговые стратегии на базе информации о количестве поисковых запросов в Wikipedia. Сигналом к покупке является рост количества запросов, сигналом к продаже – спад интереса инвесторов, выраженный в снижении запросов.

---

[1] Обзор ряда работ в данном подпараграфе выполнен студентом факультета экономических наук НИУ ВШЭ Ерохиным Д.А. под руководством Т.В. Тепловой.

Авторы показывают, что портфели, составленные из акций с повышенным вниманием в Wikipedia, показывают более высокую среднюю доходность на отрезке с 2010 по 2017 гг. по сравнению со средней доходностью случайного портфеля. Результаты являются устойчивыми к изменению частоты разбалансировки (дневная / недельная) и включению в рассмотрение кризисного периода. Парадоксально, но работа доказывает, что рыночным инвесторам для принятия решения может быть достаточно краткого описания деятельности компании (но это считая, что на следующем шаге инвесторы не будут погружены в анализ финансовой отчетности на сайте самой компании).

1.3. СТАТИЧНЫЕ МЕТРИКИ СЕНТИМЕНТА: НЕЙМИНГ, АЛФАВИТИЗМ И ЭМПИРИЧЕСКОЕ ТЕСТИРОВАНИЕ ГИПОТЕЗЫ «УМНОГО ТИКЕРА» НА ПРИМЕРЕ АКЦИЙ ГОНКОНГА

1.3.1. Статичные метрики сентимента: нейминг, алфавитизм[1]

Отдельным направлением исследований эффекта заметности является анализ сентимента, формируемого не внешним информационным фоном, а осознанными действиями эмитента. Это относительно статичные метрики сентимента. Эффект внимания может формироваться названием компании (запоминающимся, порождающим «правильные эмоции»), местом первой буквы названия или тикера в списке, из которого инвестор выбирает акции (алфавитизм), или собственно оригинальными биржевыми тикерами компаний-эмитентов. Заметим, что это статичные метрики сентимента, так как достаточно редко компании проводят ренейминг. Рекомендации в рамках этого направления исследований полезны управляющим портфелями и директорам компаний. Это, по сути, темы инвестиционного маркетинга. Исследования в качестве зависимых переменных рассматривают относительную рыночную оценку компании (Xing et al., 2018), доходность инвестирования на разных временных горизонтах (Head et al., 2009, Baer et al., 2020), показатели ликвидности (часто используется коэффициент Амихуда, как в работе Green et al., 2013).

Примеры «умного тикера» и динамика цены портфеля, составленного из акций этих эмитентов, по сравнению с рыночным индексом показаны на Рис. 1.7.

Benos and Jochec (2013) доказывают, что по акциям биржи NASDAQ, которые имели слова «American» или «USA» в названии, наблюдается «патриотическое искажение» выбора в двух кейсах «всплеска патриотизма». Портфель таких компаний генерировал статистически значимую избыточную положительную доходность после Корейской войны и террористических актов в США в 2001 г. Месячная оценка модельного коэффициента

[1] Обзор ряда работ в данном подпараграфе выполнен студентом факультета экономических наук НИУ ВШЭ Ерохиным Д.А. под руководством Т.В. Тепловой.

альфа (оцененной избыточной доходности) диагностируется авторами в 0,5% для четырехлетнего периода после обозначенных двух событий.

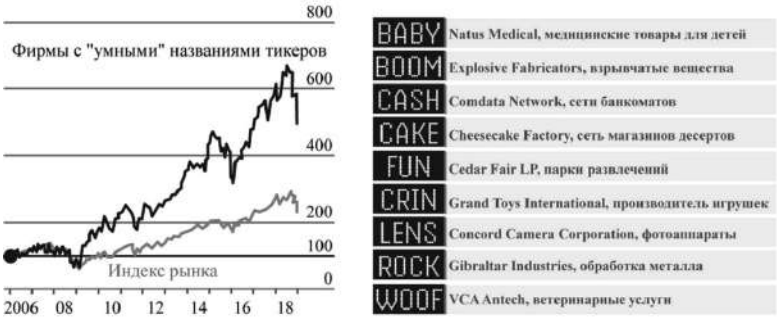

Рис. 1.7. Анализ поведения портфеля, составленного из «умных тикеров»
Источник: The Economist. URL: https://www.economist.com/graphic-detail/2019/10/07/woof-cake-boom-stocks-with-catchy-tickers-beat-the-market

В работе Ma et al. (2020) показано как «страдают китайские компании», в названии которых имеются слова «technology» и «electron», на отрезке торгового конфликта между Китаем и США в 2018 г. Портфель акций таких компаний показал статистически значимый модельный убыток (потери относительно нормальных уровней доходности) в 1%.

Концепция «умного тикера» была предложена в работе в работе Head et al. (2009). К «умным тикерам» на американском рынке автор относит аббревиатуры, которые формируют легко произносимое слово английского языка (например, «CAR» по компании AvisBudgetGroupInc.) и, тем самым, создают определенную ассоциацию с бизнес-моделью эмитента (как «BEER» у BigRockBrewery). Построение словаря «умных тикеров» авторы осуществляют на базе опроса профессиональных участников рынка. Авторский портфель, составленный из 22 акций, которые были отмечены респондентами как наиболее запоминающиеся, показал избыточную дневную доходность в размере 0,05% на отрезке анализа 1984-2005 гг. В работе Baer et al. (2020) построенный портфель из сформированного в работе Head et al. (2009) словаря продолжает обыгрывать индекс и в последующие годы (2006-2018 гг.), но при этом значение дневной альфы снижается в два раза (до 0,026%). Однако, портфель, составленный по результатам нового опроса, имеет альфу 0,019% в четырехфакторной модели Кэрхарда (Carhart, 1997), которая оказалась незначимой на уровне 5%.

Последующие работы в области «умных тикеров» строятся на более формализованных приемах формирования словарей и отбора акций. Green et al. (2013) используют лингвистический алгоритм, который учитывает условную вероятность нахождения комбинации букв тикера в английском

языке (метрика «Englishness», которая впервые протестирована в работе Travers et al., 1978) и ряд других характеристик названий.

Green et al. (2013) строят авторскую метрику легко-произносимости (Fluency), которая учитывает: 1) длину тикера, 2) значение Englishness и 3) наличие такого слова в словарях. Результаты показывают, что метрика Fluency положительно влияет на ликвидность акций. При прочих равных увеличение метрики Fluency (как прокси-сентимента заметности) на одну условную единицу ведет к росту оборачиваемости акции (Turnover) на 5,02%, а также к снижению коэффициента неликвидности Амихуда на 4,61%. Также при прочих равных увеличение Fluency на одну единицу приводит к росту капитализации медианной компании из выборки на $3.75 млн.

Схожие результаты получены и для кейса применения иного метода определения нейминговых привязок (кода), когда опросы строятся на непрофессиональных участниках рынка (Xing et al., 2016). Изменение метрики запоминаемости кода приводит к увеличению значения Q-Тобина фирмы на 3,4%.

Заметим, что в работе Anderson and Larkin (2019) доказывается по выборке акций NASDAQ, что эффект нейминга проявляется в разной степени на зависимой переменной объема торгов по розничным и по институциональным инвесторам.

Теплова и Микова (2019) в авторском исследовании с Е. Глинской демонстрируют, что эффекты нейминга выявляются и на развивающихся рынках капитала (Индия, Бразилия, Китай). На индийском фондовом рынке акции, в названии которых присутствует фрагмент «Indi», способны генерировать 20% дополнительной оборачиваемости (соответствующая переменная значима на уровне 10%). На бразильском рынке акции, тикеры которых составляют полное слово, имеют на 0,3% меньше дней с нулевой доходностью (коэффициент значим на уровне 5%), а также на 79% большую оборачиваемость торгов (коэффициент значим на уровне 15%).

Влияние нейминга на ликвидность акций также обнаружено и на китайском фондовом рынке. Jin et al. (2021) показали, что акции с более коротким тикером (названием) торгуются с премией по отношению к другим. Портфель, составленный из акций с тремя символами в названии с равными весами, генерирует дополнительную доходность в размере 0,25% в месячном выражении на горизонте с 2004 по 2017 гг., не объясняемую традиционными факторами. Заметим, что величина премии снижается при включении в модель дополнительных прокси видимости акции (в работе используются доля национальных фондов в структуре держателей, аналитическое покрытие и «бычье» настроение рынка). Отметим, что на азиатских биржевых площадках акции имеют числовой код и сокращенное название, записанное иероглифами (аналог тикеров на западных биржах). В общей сложности на китайской бирже 11,2% акций имеют названия из трех символов.

В работах Jacobs and Hillert (2015), Itzkowitz et al. (2015) демонстрируются эффекты алфавитизма на рынке США, ЕС и Японии, как результат стандартизации списка акций (в подавляющем большинстве случаев используется алфавитный порядок). Идея заключается в том, что при составлении портфеля в поле зрения инвестора в первую очередь попадают акции, название или тикеры которых начинаются на первые буквы алфавита. В исследованиях авторы проверяют гипотезу, что акции из начала алфавитного списка (первые 1%, 5% или 10% выборки) будут характеризоваться большей ликвидностью (зависимые переменные – объемы торгов в относительном выражении и коэффициент неликвидности Амихуда). По рынку США эффект алфавитизма проявляется не только в краткосрочных решениях, но и в долгосрочных, например, при распределении капитала между фондами по пенсионным планам 401(k) (Doellman et al., 2019). Авторы выявили, что при распределении активов инвесторы отдают предпочтения фондам, которые находятся в начале списка. Расчеты авторов показывают, что перемещение фонда в списке на одну позицию вверх способно привести к увеличению распределяемых в его пользу денег на 2,23%.

Hu and Li (2021) доказывают «условный алфавитизм» на китайских биржевых площадках. Проблема в том, что китайские иероглифы не имеют принципов четкого упорядочения, аналогичного западным алфавитам. Авторы проверяют гипотезу, что иероглифы, записанные меньшим количеством штрихов, обеспечивают большую привлекательность среди инвесторов ввиду заметности, простоты обработки информации, что выражается в повышенной ликвидности. Важный вывод авторов, который подтверждает результат Itzkowitz et al. (2015) по американскому рынку - данный эффект демонстрирует большую статистическую значимость для 2000-х гг. (в работе Hu & Li, 2021, это период с декабря 1990 по декабрь 2015 гг.). Отличие от классического алфавитизма заключается в том, что для китайских акций значимым является лексикографическая сложность всего тикера, а не отдельных символов (например, первого иероглифа по аналогии с первой буквой алфавита).

Еще один важный вывод работы Itzkowitz et al. (2017), выявленный по динамике акций США, в структуре владения которых есть различия между институциональными и розничными инвесторами: искажения в инвестиционном поведении из-за эффекта нейминга проявляются по-разному в торговле частных и институциональных инвесторов. Частные инвесторы в большей степени подвержены искажениям, т.к. в процессе выбора акций они в большей степени полагаются на эмоциональные привязки и запоминаемость акции (эффекты якорения). Дамми-переменные, отражающие эвристики нейминга и алфавитизма, являются значимыми только для объяснения ликвидности тех акций, в структуре владения которых преобладают розничные инвесторы.

1.3.2. Эмпирическое тестирование нейминга на бирже Гонконга

Исследование проведено Д.А. Ерохиным в ходе подготовки ВКР НИУ ВШЭ в 2021 г. (под руководством Т.В. Тепловой).

Особенность обработки данных для проведения исследования по компаниям биржи Гонконга заключается в том, что торгующиеся в Гонконге ценные бумаги имеют в информационных источниках, базах данных, торговых терминалах четырехзначные числовые коды. Во множестве используемых кодов зарезервированы блоки под отдельные инструменты (например, по состоянию на февраль 2020 г. под выпуск новых обыкновенных акций зарезервированы не использованные ранее коды в промежутках от 0001 до 2799, 3300-3999, 6030-6199, 6400-6749, 6800-7199, 8000-8999, 9400-9499, 9600-9999).

С ноября 1999 г. правилами Гонконгской фондовой биржи предусмотрен следующий порядок выбора кодов. Множество доступных кодов разделено на две группы: Normal Ballot и Special Ballot. Во второй группе собраны коды, которые могут иметь положительные эмоциональные привязки. Стандартный путь бесплатного получения кода исходит из случайного выбора. Эмитент имеет право выбрать любой код из пяти предложенных, которые сгенерированы случайным образом. Нестандартный путь на бирже Гонконга строится на альтернативном варианте платного выбора кода из особой группы (эмитент обязан внести определенную сумму в благотворительную организацию, величина взноса зависит от выпускаемого инструмента и секции рынка). В листинге на основной площадке Биржи взнос составляет 3 млн гонконгских долларов (эквивалент $391 тыс. по состоянию на 06.04.2021 г.).

Фондовый рынок Гонконга имеет две основные торговые площадки: основная площадка (Main Board) и Growth Enterprise Market (далее – GEM). GEM отличается упрощенным порядком листинга, в рамках которого предъявляется меньше требований к эмитенту, и позиционируется как площадка для сравнительно молодых компаний. Коды акций, торгующихся на GEM, зафиксированы Биржей в промежутке между 8000 и 8999 включительно.

Интересный момент, что число 8 считается счастливым из-за созвучия с глаголом «процветать». Напротив, число 4 считается не счастливым по причине созвучия со словом «смерть». Анализируя китайских инвесторов, можно предположить, что эти культурные особенности могут создавать специфические эмоциональные привязки, которые могут порождать нерациональность поведения инвесторов. Этот исследовательский вопрос и рассматривается в данном параграфе.

Учитывая то, что две первые цифры кода определены в соответствии с планом распределения кодов Биржи, случайным или сознательным выбором может считаться только порядок определения двух последних цифр тикера эмитента. Соответственно, в данной работе рассматривается распределение двух последних цифр числовых кодов, и проводятся тесты на соответствие их распределения равномерному.

В работе анализируются недельные данные о ценах и рассчитанные на их основе доходности для 200 акций, торгующихся на HKEX в период с 2014 по 2020 гг., характеризующихся относительно высокой ликвидностью. При этом рассматриваются исключительно те инструменты, данные по которым доступны на протяжении всего периода. Общее количество запоминающихся тикеров в рассматриваемой выборке оценено нами в 44, общая доля таких акций в итоговой выборке равна 22%. Из этих запоминающихся тикеров построен авторский портфель. Учитывая, что максимальная доля одной отрасли составляет 22,7%, можно признать портфель относительно диверсифицированным. Сравнение доходности авторского портфеля с бенчмарком (индексом Hang Seng) представлено на Рис. 1.8, а в Табл. 1.6 показаны оценки коэффициента Шарпа по двум портфелям.

В составе рыночного индекса только 14 акций из нами отобранных 44 (26%) классифицированы как запоминающиеся. При исключении данных акций из рассмотрения доходность условно рыночного портфеля на рассматриваемом периоде оценивается в 10,99%, что свидетельствует о существенном влиянии запоминающихся акций на результат. Кумулятивная доходность авторского портфеля превысила доходность бенчмарка на 12,75%. Подобный феномен теоретически может описываться аномальной доходностью отдельных акций, однако данное предположение отвергается при рассмотрении поведения акций, входящих в портфель (31 акция из 44 обыграла индекс; 70%). Коэффициент Шарпа авторского портфеля выше, чем аналогичный показатель для рыночного индекса. Такой поверхностный анализ позволяет предположить, что авторский портфель запоминающихся тикеров обыгрывает рыночный индекс в период с 2014 по 2020 гг. при использовании стратегии buy & hold. Более аккуратные тесты показывают, статистически значимая связь с доходностью с коррекцией на риск не подтверждается. Схожее исследование (Baer et al., 2020) для американского рынка приходит к такому же выводу.

Еще одна сделанная проверка касается повышенной ликвидности «запоминающихся тикеров». Ниже приведены модели для торгового оборота и коэффициента неликвидности Амихуда как объясняемых переменным по выборке гонконгских акций.

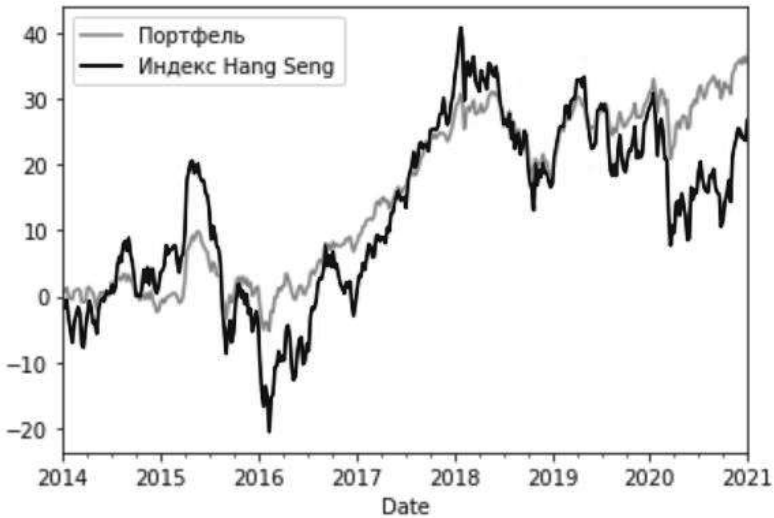

Рис. 1.8. Сопоставление накопленной доходности рыночного индекса HangSeng с доходностью авторского портфеля (акции с запоминающимися кодами), вертикальная ось – доходность в процентах

Таблица 1.6

Сопоставление авторского портфеля и индекса биржи Гонконга по накопленной доходности и коэффициенту Шарпа

Портфель	*Накопленная доходность (январь 2014 – декабрь 2020 гг.)*	*Коэффициент Шарпа*
Портфель из акций с запоминающимися тикерами	36,43	0,576
Индекс Hang Seng	26,83	0,251

Из списка торгуемых инструментов, доступных на официальном портале Гонконгской Биржи было отобрано 2205 прошедших листинг обыкновенных акций. Далее было построено распределение цифр, которые находятся на третьей и четвертой позициях кодов данных инструментов. Предполагается, что при отсутствии существенных предпочтений эмитентов при выборе тикера данные эмпирические распределения не должны существенным образом отличаться от равномерного. Графическое представление построенных распределений представлено на Рис. 1.9.

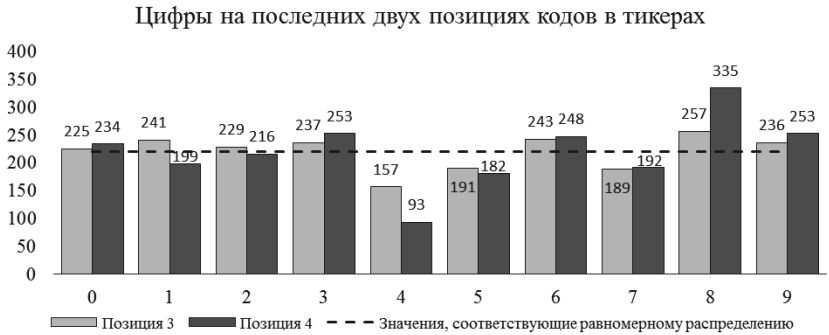

Рис. 1.9. Распределение цифр, находящихся на третьей и четвертой позициях числовых кодов-тикеров гонконгских акций

Обратим внимание на существенное отклонение количества цифр 4 и 8 от значений, которые соответствуют теоретическому равномерному распределению. Это может указывать на стремление эмитентов выбирать тикеры со «счастливым» числом и избегать «несчастливого» числа 4. Цифры 3 и 9, количество которых также значительно выше предсказываемого равномерным распределением, имеют преимущественно положительную эмоциональную окраску в китайской культуре.

Тесты Бройша-Пагана (выполнены Д. Ерохиным) показали, что панельная спецификация модели зависимости ликвидности акций от различных характеристик компании-эмитента, включая запоминающиеся тикеры, описывает существующие взаимосвязи лучше, чем обычная линейная модель, не принимающая к рассмотрению эффекты группы и времени. Была проведена проверка различных спецификаций модели на мультиколлинеарность с помощью показателей VIF. Все VIF не превышают 10, в силу чего сделан вывод о сохранении набора переменных (Табл. 1.7).

Таблица 1.7

Модельные оценки для объяснения различий двух метрик ликвидности по выборке гонконгских акций (спецификация модели с индикатором запоминаемости тикера - Memorable)

Переменные	*Модель 1 (Turnover)*	*Модель 2 (Amihud)*
Memorable	0.156* (0.0848)	-0.118** (0.0599)
Beta	0.183*** (0.0779)	-0.331*** (0.058)
Cap	1.075*** (0.0437)	-1.554*** (0.0360)
P/B	0.304***	-0.154***

Переменные	*Модель 1 (Turnover)*	*Модель 2 (Amihud)*
	(0.0415)	(0.0343)
Price	-0.0288*** (0.00352)	0.0225*** (0.00331)
StDev	0.572*** (0.0239)	-0.396*** (0.0241)
Pos_return	0.109*** (0.00827)	-0.0782*** (0.00799)
Leverage	-0.000808*** (0.000208)	0.001037*** (0.000175)
Crisis	-0.392*** (0.0126)	0.447*** (0.0131)
GEM	0.366*** (0.156)	-0.011 (0.107)
Constant	3.725*** (0.319)	-9.194*** (0.259)
R^2	0.333	0.389
Наблюдения	106 797	106 797
Кол-во акций	1 789	1 789

В скобках указаны кластеризованные стандартные ошибки.
Уровни значимости: *** p<0.01, ** p<0.05, * p<0.1.

Коэффициент при переменной Memorable (Табл. 1.7) является статистически значимым. При прочих равных запоминаемость тикера обеспечивает рост оборачиваемости на 15,6% и снижение коэффициента неликвидности Амихуда на 11,8%. Контрольные регрессоры (за исключением индикатора Index Member в обеих моделях) имеют достаточную объясняющую силу при анализе ликвидности. При прочих равных ликвидность выше для акций с большей капитализацией (переменная Cap), с большим значением мультипликатора «кратное балансовой оценке» P/B и волатильностью месячной доходности (StDev). Еще одна группа моделей построена для дамми-переменной, оценивающей позицию акции в верхней / нижней частях списка (Табл. 1.8).

Таблица 1.8

Оценка моделей для двух метрик ликвидности с дамми переменной места цифры тикера в списке

Переменные	*Модель 3 (Turnover)*	*Модель 4 (Amihud)*
Top 1%	-0.495 (0.326)	0.549 (0.234)
Top 5%	0.334** (0.151)	-0.142 (0.169)
Top 10%	-0.190 (0.14)	-0.035 (0.121)
Bottom 5%	0.358	-0.049

Переменные	*Модель 3 (Turnover)*	*Модель 4 (Amihud)*
	(0.427)	(0.454)
Bottom 1%	0.792 (0.511)	0.108 (0.504)

В скобках указаны кластеризованные стандартные ошибки.
Уровни значимости: *** p<0.01, ** p<0.05, * p<0.1.

Набор контрольных переменных моделей 3 и 4 такой же, как и по моделям 1-2 (Табл. 1.8). Скачок коэффициента при переменной Top 5% в модели 3 может быть обусловлен нахождением в верхних 2-5% списка акций компаний со сравнительно невысоким количеством торгующихся акций, что, в свою очередь, увеличивает показатель оборачиваемости. Заметим, что коэффициенты при дамми-переменных, отражающих позицию акций в списке, преимущественно не являются статистически значимыми, а также наблюдается существенный разброс значений коэффициентов. Таким образом, мы не выявили устойчивую связь между позицией гонконгской акции в отсортированном по возрастанию числового кода списке и ее показателями ликвидности акций.

1.4. МЕТРИКИ НАБЛЮДЕНИЙ ЗА УСПЕШНЫМИ ИНВЕСТОРАМИ[1]

Отдельный пласт работ, связанный с эффектом внимания – копирование состава портфеля, который инвесторы могут наблюдать по крупным инвестиционным фондам (хедж-фондам) или гуру-инвесторам. В этот пласт работ попадают и различные стратегии автоследования. Вместо собственного анализа и выбора инвестор готов копировать торговые действия неких лиц, которые привлекли его внимание рекламой своих успехов (зачастую ложных) или же доверием к нему других участников инвестиционного сообщества.

Относительно простой и бесплатный способ выявить интересные для инвестирования акции – изучить портфели топовых управляющих активами. По рынку США традиционно опираются на форму 13F. Данную форму подают институциональные инвесторы на ежеквартальной основе. Основным критерием обязанности составлять такую форму является капитал под управлением фонда (сумма активов должна превышать $100 млн.). Данная форма позволяет отслеживать покупки крупных фондов, однако она содержит в себе ряд существенных ограничений, которые накладывает SEC США: 1) в форме показаны только «длинные позиции», хотя хедж фонды часто используют «короткие продажи», 2) временной лаг подачи (отчёт подаётся в течение 45 дней после окончания квартала), 3) слабый контроль за достоверностью (Anderson & Brockman, 2018).

[1] Обзор ряда работ в данном параграфе выполнен студентом факультета экономических наук НИУ ВШЭ Соломоником М.А. под руководством Т.В. Тепловой.

Пример метрики внимания, построенной по форме 13F, где фиксируется отчетность фондов (точнее «движения попсовых фондов» в акциях) во II квартале 2022 г.

Berkshire Hathaway (Buffett)
новые покупки: ———
полная продажа: #VZ #RPRX
увеличение доли: #ALLY #OXY #ATVI #MKL #MCK
снижение доли: #USB #GM #KR #STOR
Bridgewater (Ray Dailo)
новые покупки: #RIVN #AMZN #SE
полная продажа: #BABA #JD
увеличение доли: #META GOOGL #CVS #MA
снижение доли: #PEP #LIN #PG
Elliott Management (Elliot)
новые покупки: #AJRD #PINS #PYPL #ERD #CAH
полная продажа: #TWTR #EVRG #APA
увеличение доли: #VAL #NE
снижение доли: ——-
Third Point (Dan Loeb)
новые покупки: #CL #AR #DIS #TMUS
полная продажа: #SPGI #INTU #AVTR #IQV
увеличение доли: #OVV #CVE #SMFR
снижение доли: #S #PCG #DHR #CSX
Appaloosa (David Tepper)
новые покупки: #CRM #BABA #NFLX #CEG #APTV #HCA #CZR
полная продажа: #XLE #XOP #GT #FCX #WYNN
увеличение доли: #KSS #MPLX #META
снижение доли: #MU #OXY

Как элемент стратегии автоследования за топовыми хедж-фондами может быть включение в портфель с большими весами тех акций, которые были куплены несколькими хедж-фондами. Дополняя портфель короткими позициями, аналитик может включить в портфель и короткие позиции по продаваемым активам (соответственно, если какие-то акции продают два и более фонда, то вес этих коротких позиций в портфеле увеличивается).

Часто отслеживаемые фонды: AQR Capital Management, BlackRock, BlueCrest Capital Management, Bridgewater Associates, D.E. Shaw & Co., Generation Investment Management, Lone Pine Capital LLC, Millennium Management, Point72 Asset Management, Renaissance Technologies, Sands Capital Management LLC, Two Sigma Investments, Wellington Management Company.

Исследовательский вопрос ставится следующим образом: можно ли, копируя набор акций в портфеле топового фонда или неким образом комбинируя стратегии фондов (например, составляя портфель из тех активов, которые присутствуют и увеличивают позиции в максимальном количестве фондов), получить доходность выше бенчмарка с учетом принимаемого риска.

В контексте этого направления интересна работа Brown and Schwarz (2020), где анализируются выгоды, которые могут получить инвесторы из анализа формы 13F. Эта форма подаётся множеством участников финансо-

вого рынка среди институциональных инвесторов, однако авторы приходят к выводу, что наиболее интересным с точки зрения копирования набора активов являются стратегии, которые используют информацию по хедж-фондам. Более того, авторы утверждают, что раскрываемая хедж-фондами информация способна оказывать значимое влияние на рынок.

Отчасти прибыльность таких стратегий копирования приводит Aragon et al. (2013) к выводу, что управляющим хедж-фондов выгодно отдалять дату раскрытия информации по форме 13F, так как некоторая часть дополнительной прибыли может быть получена в течение конфиденциального периода (Aragon et al., 2013). В основном запросы о продлении временного промежутка раскрытия информации связаны с менее ликвидными инструментами.

Инвесторам бывает полезно отслеживать действия и коммуникации не только топовых управляющих фондов, но и компаний-эмитентов. В высказываниях CEO инвестор может найти полезные сигналы, что показано в работе Федоровой и соавторов (2017).

1.5. ДЕМО-ИССЛЕДОВАНИЕ СОПОСТАВЛЕНИЯ РАЗНЫХ МЕТРИК ЗАМЕТНОСТИ НА ДОХОДНОСТЬ ПОСТРОЕННЫХ ПОРТФЕЛЕЙ

В работе Теплова и Микова (2019) приведено исследование студента программы «Финансовые рынки и финансовые институты» (ФРФИ) НИУ ВШЭ М. Мажова под руководством Т.В. Тепловой и Е.С. Миковой, в котором при помощи метода главных компонент строится авторский индекс внимания по компаниям российского рынка (в выборку были включены акции компаний, торгующиеся на Мосбирже с рыночной капитализацией на момент сбора данных, превышающей 100 млн. руб.) и сравнивается с наиболее популярным сейчас прокси показателем внимания - частотой поисковых запросов (SVI Google). В первоначальную выборку вошло 193 компании, капитализация которых составила 95% от капитализации всех компаний на бирже. Исследование охватывает период с января 2015 г. по март 2019 г.

Показатели для оценки эффекта внимания из аналитической платформы Блумберг:

- TURNOVER – оборот торгов акциями;
- TWITTER_PUBLICATION_COUNT – общее количество «твитов», связанных с компанией;
- NEWS_PUBLICATION_COUNT – общее количество новостей, связанных с компанией;
- TOT_ANALYST_REC – общее количество отчетов аналитиков по компании;
- EQY_REC_CONS – консенсус-прогноз аналитиков (по шкале от 0 до 5).

Помимо пяти представленных выше переменных исследование учитывало данные по поисковым запросам в Google для каждой из компаний

(Search Volume Index или SVI) по: 1) тикеру, 2) английскому и 3) русскому названию компании. Исследование строилось по относительной популярности запросов по 100-балльной шкале. Чтобы не выгружать поисковые запросы вручную, был использован пакет gtrendsR для RStudio, который автоматизирует загрузку данных по поисковым запросам и радикально упрощает получение необходимых показателей. С помощью этого пакета были загружены недельные данные по частоте поисковых запросов трех типов для каждой компании (по биржевым тикерам компаний и по названиям на английских и русских языках). После очистки собранных данных осталось 85 компаний, которые имеют полные наборы показателей для исследования. Описательная статистика по ним показана в Табл. 1.9.

Таблица 1.9

Описательная статистика основных показателей исследования эффекта внимания по запросам Google по 85 российским компаниям

Переменные	*Капитализация, млн. руб*	*Оборот, млн. руб*	*Количество новостей*	*Оценка аналитиков*	*Рекомендации аналитиков*	*B/M*	*SVI*	*Доходность, %*
Минимум	885	0,004	5	1	0	-505,2	0	-34,6
Максимум	5 990 162	199 050	5 592	5	115	839	281	359,8
Медиана	66 625	47	46	3,25	10	0,9	92	0
Среднее	367 350	1 922,9	137	3,33	21,09	4,3	92,96	0,3
Стандартное отклонение	768 301	6 704	276,2	1,22	25,48	0,02	45,40	5,6
Количество наблюдений	17 765							

Так как разбросы анализируемых прокси-внимания часто распределены не в соответствии с нормальным распределением, часто наблюдаются скошенности, то в ряде работ по исследуемой теме используются не только «чистые» показатели, но также логарифмы и оценки отклонений (Tan et al., 2019, Chai et al., 2019; Vozlyublennaia, 2014). Для сопоставления разных прокси-метрик внимания (запросов в Google) будут рассчитаны натуральные логарифмы всех факторов, кроме консенсус-оценки аналитиков, а затем будут посчитаны анормальные величины (оценки отклонений) по формуле (1):

$$AF_{it} = F_{it} - Median[F_{i-1},...,F_{i-8}] \quad (1)$$

где F_{it} - показатель *F* для компании *i* в неделю *t*, а $Median[F_{i-1},...,F_{i-8}]$ - медианное значение *F* для компании *i* за предыдущие 8 недель.

Анализ полной корреляционной таблицы построенных факторов позволил выделить факторы с наиболее высокой степенью корреляции, представленные в Табл. 1.10.

Таблица 1.10

Корреляционная матрица для выбранных факторов внимания

	Turnover	*Tweets*	*News*	*Reccomen-dations*	*Log (Turnover)*
Turnover	1	0,40	0,63	0,48	0,46
Tweets	0,40	1	0,72	0,33	0,29
News	0,63	0,72	1	0,51	0,54
Reccomendations	0,48	0,33	0,51	1	0,71
log(Turnover)	0,46	0,29	0,54	0,71	1

Исследовательская задача ставилась на сопоставлении двух прокси-метрик внимания: построенный авторский индекс методом главных компонент и поисковые запросы SVI. Построение показателя методом главных компонент на основе нескольких индикаторов внимания позволяет построить индекс, который охватывает сразу несколько источников внимания к акциям компании. Если число «твиттов» и новостей показывает интерес преимущественно со стороны индивидуальных инвесторов, то число рекомендаций аналитиков больше отражает интерес институциональных инвесторов. Оборот торгов по каждой акции является более явным индикатором интереса, вытекающим из трех остальных.

Показатель поисковых запросов SVI представляет исследовательский интерес по той причине, что в отличие от «твиттов», новостей или рекомендаций аналитиков, которые являются «пассивными» прокси, SVI – «активный». Другими словами, большое количество новостей, хотя и показывает повышение информационного освещения компании, но может демонстрировать не непосредственный интерес инвесторов к этой компании, а только лишь косвенный. В то время как увеличение поисковых запросов, связанных с компанией, отражает однозначное возрастание интереса к компании, так как оно вызвано ростом активного поиска информации со стороны заинтересованных индивидов. Тем не менее, потенциально существует риск того, что увеличение количества поисковых запросов может быть связано не только с интересом к акциям данных компаний. Отчасти по этой причине было принято решение собрать данные и по количеству поисковых запросов по тикерам компаний, потому что такие запросы демонстрируют интерес именно к акциям искомых компаний.

Среди построенных факторов встречается как показатель *Turnover*, так и *log(Turnover)*. На основе данных по корреляции нельзя сделать однозначных выводов о том, какой из этих факторов окажется наиболее подходящим для построения авторского составного индекса. Поэтому было принято решение построить две разновидности индекса (Табл. 1.11).

Таблица 1.11

Два варианта компонентов для авторского составного индекса внимания

Вариант 1	*Вариант 2*
Tweets	Tweets
News	News
Recommendations	Recommendations
Turnover	log(Turnover)

Для построения интегрального индекса использован метод главных компонент (PCA) – ортогональная линейная трансформация, которая переносит данные в новую систему координат таким образом, что наибольшая доля дисперсии всех исходных факторов объясняется первой координатой в новой системе (первой главной компонентой), вторая наибольшая доля – второй компонентой и так далее (Jolliffe, 2002).

Метод главных компонент работает следующим образом. Первоначально все факторы стандартизуются и центрируются. Для исходной матрицы *X*, содержащей в себе *p* измерений, подбирается такой набор весов $w_{(k)} = (w_1,...,w_p)_{(k)}$, который трансформирует каждый вектор $x_{(i)}$ матрицы *X* в новый вектор $t_{(i)} = (t_1,...,t_l)_{(i)}$, где $t_{k(i)} = x_{(i)} \times w_{(k)}$ для *i = 1,…,n*; *k = 1,…,l*.

Для того чтобы первая компонента содержала наибольшую возможную долю дисперсии всех исходных факторов, первый вектор весов $w_{(i)}$ должен удовлетворять условию (2):

$$w_{(1)} = \arg\max\left\{\frac{w^T X^T X w}{w^T w}\right\} \qquad (2)$$

После нахождения весов для первой компоненты, ее значения получаются с помощью формулы (3):

$$t_{1(i)} = x_{(i)} \times w_{(1)} \qquad (3)$$

Интерес для нас представляет только первая компонента, поскольку она максимизирует выборочную дисперсию проекции. В случае с первым вариантом (Табл. 1.11) эта доля составила 63,88%, а в случае со вторым – 64,06%. Следовательно, второй вариант индекса хоть и не существенно, но показывает себя лучше по данному критерию, и для дальнейшего исследования нами выбран именно он. Наибольший вес пришелся на фактор *News*, но при этом его перевес достаточно небольшой, и в целом все факторы имеют относительно равные веса. Этот факт позволяет нам предположить, что полученный индекс максимально широко охватывает аспекты проявления внимания, представленные входящими в индекс показателями. Кроме визуального представления весов математическая формула индекса внима-

ния или *AI* (сумма весов незначительно превышает 1 по причине округления) представлена в уравнении (4):

$$AI = 0{,}252 \times \mathrm{Re}\,c + 0{,}251 \times \log(Turnover) + 0{,}273 \times News + 0{,}225 \times Tweets \quad (4)$$

На следующем шаге после получения прокси эффекта внимания строится набор портфелей на основе двух прокси-метрик внимания по методологии Фамы и Френча (Fama & French, 1993, 1996, 2015) на основе сортировки акций по полученным прокси выделения акций в соответствующие перцентили. Анализ построенных портфелей позволяет проанализировать, присутствует ли на российском рынке эффект внимания, который может быть отслежен на основе монотонного изменения доходности от нижнего k% перцентиля к верхнему, или наоборот. Такая методология исследования использовалась в работе Joseph et al. (2011), являющейся по сути одной из первых работ, в которых поисковые запросы рассматриваются в качестве прокси-метрики эффекта внимания. Построение портфелей реализуется следующим образом:

- в начале недели *t* все компании сортируются по убыванию показателя внимания;
- на основе этой сортировки формируется *n* портфелей, размер которых зависит от выбранных перцентилей;
- по итогам недели *t* считается доходность каждого портфеля за данную неделю, взвешенная по капитализации входящих акций (5):

$$r_{pt} = \sum_{i=1}^{n} (w_{it} \times r_{it}) \quad (5)$$

где r_{pt} – доходность портфеля в неделю *t*, w_{it} – вес компании *i* в портфеле в неделю *t*, r_{it} – доходность компании *i* в неделю *t*;

- через неделю происходит пересмотр портфеля, и шаги (1)–(3) повторяются.

Таким образом, по результатам выполнения шагов для всего временного горизонта получены значения средних и накопленных доходностей для каждого портфеля. Накопленная доходность вычисляется по формуле (6):

$$R = \Pi_{i=1}^{t} (1 + r_i) \quad (6)$$

где r_i – доходность в неделю i.

Кроме того, для каждого портфеля оценивается t-статистика для определения значимости его средней доходности (7):

$$t = \frac{\bar{x}}{s / \sqrt{n}} \quad (7)$$

где $\bar{x}$ – среднее значение доходности, s – выборочное стандартное отклонение, n – размер выборки.

Для обеспечения достаточного уровня диверсификации и надежных результатов тестирования в работе строились 10 портфелей, следуя логике Joseph et al. (2011). Такое количество обусловлено двумя факторами – меньшее количество портфелей кажется недостаточно подробным разбиением, тогда как большее количество портфелей приведет к малому количеству компаний в каждом из портфелей, что может дать нерепрезентативные результаты. В Табл. 1.12 представлены накопленные доходности за весь период исследования, средние недельные доходности и значения t-статистики для каждого из портфелей.

Таблица 1.12

Результаты сортировки портфелей по прокси-метрике внимания к российским компаниям

Портфель	*Авторская метрика AI*			*Метрика поисковых запросов SVI*		
	накопленная доходность, %	*средняя недельная доходность, %*	*t-статистика*	*накопленная доходность, %*	*средняя недельная доходность, %*	*t-статистика*
Высокое внимание	93,9	0,35	1,92*	128,9	0,44	2,11
2	77,6	0,30	1,94*	77,6	0,30	1,82*
3	53,6	0,23	1,47	75,8	0,30	1,69*
4	64,0	0,26	1,66*	54,5	0,24	1,46
5	254,2	0,65	3,16***	138,4	0,45	2,56**
6	302,0	0,72	3,17***	97,8	0,36	1,93*
7	138,3	0,45	2,55**	39,4	0,19	1,08
8	85,3	0,32	2,16**	69,0	0,29	1,52
9	-21,3	-0,10	-0,69	78,3	0,32	1,59
Низкое внимание	-22,8	-0,11	-0,9	291,7	0,70	3,51***
Высокое – низкое	140,3	0,40	2,33**	-47,7	-0,25	-1,08

* - статистическая значимость на 10% уровне, ** - на 5% уровне, *** - на 1% уровне. Расчеты М. Мажова.

Результаты использования показателя AI. Показатели накопленной и средней доходности монотонно убывают при уменьшении индекса внимания в случае первых трех портфелей (средняя доходность уменьшается с 0.35% до 0.23%), четвертый же портфель показывает доходность, незначительно превышающую таковую у третьего портфеля (средняя – 0.26%). Портфели с пятого по восьмой показывают неожиданно высокую доходность, при этом для пятого и шестого портфеля эту доходность можно

назвать анормальной, превышающей доходность первого портфеля более, чем в три раза. Последние два портфеля показывают отрицательную доходность, что согласуется с логикой зависимости доходности от величины внимания, наблюдаемой по первым трем портфелям. Кроме того, была посчитана доходность арбитражного или самофинансируемого портфеля, включающего в себя занятие инвестором одновременно двух позиций: длинной позиции по акциям с самым высоким показателем внимания и короткой позиции по акциям с самым низким эффектом внимания. Стоит заметить, что для восьми из 11 портфелей доходность оказалась статистически значимой.

Результаты использования показателя поисковых запросов Google (SVI переменная). Тогда как показатель AI обеспечил монотонное убывание доходности для верхних трех портфелей, то SVI увеличил количество таковых портфелей до четырех: 0,44% у первого портфеля и 0,24% – у четвертого. Кроме того, явно анормальная доходность наблюдается в данном случае не только для портфелей из середины: наибольшим значением обладает десятый портфель со средней доходностью 0.7%. Для портфелей с пятого по девятый не наблюдается зависимости доходности от величины SVI. В отличие от показателя AI, комбинированный портфель из длинных и коротких позиций показал отрицательную доходность, вызванную аномально высокой доходностью портфеля с низким показателем внимания. Значимую доходность показали шесть из 11 портфелей.

Таким образом, результаты портфельного анализа показали, что доходность крайних портфелей существенно отличается и согласуется с логикой, заложенной в эффект внимания, однако серединные портфели ведут себя неоднозначно. Для элиминирования данного эффекта была предпринята попытка варьировать количество портфелей, выбор перцентилей и способы учета весов компаний в рамках портфелей. Однако эффект немонотонного изменения доходности по портфелям, полученным в ходе сортировки акций по эффекту внимания, сохранился.

После визуального, статистического и графического анализа полученных результатов, был проведен asset pricing тест на основе построения time-series регрессий и GRS-теста на полученных портфелях с использованием бета рыночного риска по аналогии с работой Fama and French (1996). Напомним, равновесная модель ценообразования CAPM имеет вид (8):

$$R_{it} - R_{ft} = \alpha_i + \beta_i (R_{mt} - R_{ft}) + \varepsilon_{it} \qquad (8)$$

где R_i – доходность *i*-го актива, R_f – безрисковая доходность, R_m – рыночная доходность.

Помимо однофакторной CAPM модели, доходность построенных нами портфелей, упорядоченных по эффекту внимания, протестирована при помощи трёхфакторной модели Фамы-Френча (9):

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \varepsilon_{it} \quad (9)$$

где R_i – доходность *i*-го актива, R_f – безрисковая доходность, R_m – рыночная доходность, *SMB* – премия за размер компании, *HML* – премия за потенциал роста компании (учет эффекта стоимости).

GRS-тест, впервые описанный в работе Gibbons et al. (1989), тестирует нулевую гипотезу вида (10), т.е. проверяет значения альфа на равенство нулю:

$$H_0 : \alpha_{ip} = 0 \ \ \forall i = 1,...,N \quad (10)$$

Тест проверяет гипотезу об одновременном равенстве нулю всех α для портфелей. В нашем случае отвержение гипотезы будет означать, что наша сортировка портфелей была произведена не случайным образом, а имела под собой основание, и потенциально прокси фактор внимания может быть включен в модель ценообразования, релевантную для российского рынка. В основе GRS-теста лежит тест Вальда, но при этом GRS-тест требует поправку на малую выборку и предполагает F-распределение (11):

$$J = \left(\frac{T}{N}\right)\frac{T-N-L}{T-L-1}\left(1+\bar{r}_p'\hat{\Omega}^{-1}\bar{r}_p\right)^{-1}\hat{\delta}_0'\hat{\Sigma}^{-1}\hat{\delta}_0 \sim F_{N,T-N-L} \quad (11)$$

На основании значений статистики и результатов тестирования портфелей можно сделать определенные выводы о наличии эффекта внимания и о его связи с доходностью по выборке российских акций.

Для реализации модели были посчитаны прокси факторы размера и стоимости SMB и HML для российского рынка, соответственно. Методология построения факторов аналогична Fama and French (1993).

Анализ построенных портфелей дает неоднозначные результаты. С одной стороны, наблюдается разница в доходности между портфелями. Верхние по рангу портфели показывают убывающую положительную доходность, а нижние показывают отрицательную, что соответствует логике. Но при этом серединные портфели выбиваются из ожидаемой логики и показывают избыточную доходность.

Проведение GRS теста (11) показывает, что для двух случаев сортировки гипотеза о равенстве всех альфа (α) нулю отвергается как минимум на 5% уровне значимости (Табл. 1.13). Такие результаты говорят о том, что существующие модели ценообразования не могут полностью объяснить различия в доходности портфелей, построенные на основе эффекта внимания. То есть, использование как показателя AI, так и SVI оправдано на российском рынке, но при этом природа эффекта еще не ясна, для чего следует провести дополнительное исследование.

Таблица 1.13

Результаты GRS-теста по двум прокси-метрикам внимания

	AI	*SVI*
GRS	4,08	2,19
P-value	0,00	0,02

Существует риск, что на результаты могут оказывать влияние перекрестные эффекты: иными словами, в серединные портфели могут попадать, например, акции малой капитализации, что увеличивает их доходность за счет повышенного риска. Для выявления таких эффектов была произведена ортогонализация, после чего осуществлена двойная сортировка портфелей с целью получения чистого эффекта внимания, для чего необходимо было выбрать второй фактор помимо прокси фактора риска внимания. Интересный вывод: для российского фондового рынка наибольшее значение имеет фактор размера, и именно он будет использован в качестве второго параметра для сортировки портфелей. Фактором размера в данном случае будет являться рыночная капитализация компании.

В Табл. 1.14 представлены результаты тестирования по доходности и статистики, полученные для портфелей, сформированных с помощью двойной сортировки. По вертикали портфели отсортированы по величине внимания (по AI или SVI), а по горизонтали – по размеру. Такое количество выделенных портфелей было выбрано в соответствии с работой Joseph et al. (2011).

Таблица 1.14

Результаты двойной сортировки портфелей российского рынка

Авторская метрика AI				*Метрика поисковых запросов SVI*			
Накопленная доходность, %							
	Крупные	Средние	Малые		Крупные	Средние	Малые
Высокое	89,0	150,4	83,5	Высокое	108,6	271,4	30,6
Среднее	69,3	278,5	294,1	Среднее	78,6	105,2	107,9
Низкое	-12,6	44,5	-9,9	Низкое	57,9	126,2	67,8
Средняя доходность (недельная), %							
	Крупные	Средние	Малые		Крупные	Средние	Малые
Высокое	0,33	0,50	0,3	Высокое	0,38	0,69	0,16
Среднее	0,27	0,67	0,7	Среднее	0,30	0,36	0,39
Низкое	-0,05	0,19	-0,03	Низкое	0,25	0,41	0,27
t-статистика							
	Крупные	Средние	Малые		Крупные	Средние	Малые
Высокое	2,08**	2,13**	2,96***	Высокое	2,22**	2,58**	0,91
Среднее	1,90*	3,69***	3,57***	Среднее	1,94*	2,67***	1,89*
Низкое	-0,48	1,79*	-0,23	Низкое	1,43	2,84***	1,83*

Примечание. Здесь и далее: * - значимость на 10% уровне, ** - на 5% уровне, *** - на 1% уровне.

Согласно полученным результатам (Табл. 1.14), можно отметить, что для больших компаний наблюдается монотонное снижение как средней, так и накопленной доходности при уменьшении двух построенных индексов (факторов модели) внимания. Если обратить внимание на значения t-статистики, то оценки средней доходности значимы на 10% уровне для больших компаний с высокими и средними значениями индекса внимания. Высокое значение p-value для больших компаний с низким уровнем AI (0,633) объясняется тем, что в течение 175 недель (из 209 всего) в этот портфель не попадало ни одной компании, то есть в эти недели портфель показывал нулевую доходность. Этот факт поддается логике и не должен удивлять: так как портфели строились по методологии Fama and French (1993), то сортировка происходила по двум показателям независимо, и поэтому в квантилях с низким эффектом внимания превалировали средние и малые компании, про которые, очевидно, выходит меньше новостей, пишется меньше «твиттов» и т.д.

Аналогичная ситуация наблюдается для портфеля малых компаний с высоким показателем AI: этот портфель оказывался пустым в течение 173 недель, что объясняется той же логикой, что была представлена выше. Но для данного портфеля значение средней доходности оказывается значимым даже на 1% уровне, и это объясняется тем, что в те недели, когда портфель не был пустым, он показывал среднюю недельную доходность 1,87%. Такая высокая доходность может объясняться тем, что малые компании, и без того имеющие более высокую доходность, чем крупные, резко получают большое количество внимания, и, тем самым, при выходе новостей или откликов в прессе привлекают к себе внимание даже тех инвесторов, которые раньше про эти компании могли не знать. В то же время, при использовании показателя SVI отсутствовали ситуации, когда какой-либо портфель оказывался пустым. Это обусловлено различиями в методологии расчета показателей AI и SVI. На основе этой информации можно сделать вывод, что показатель SVI более релевантен для использования в случаях нетипичных портфелей.

На следующем этапе проведена аналогичная процедура тестирования, которая была использована для простой сортировки и посчитана GRS-статистика в рамках трехфакторной модели Фамы-Френча для полученных портфелей. Названия столбцов (AI и SVI) в Табл. 1.15 обозначают тот фактор, по которому производилась сортировка.

Таблица 1.15

Результаты GRS-теста при двойной сортировке акций в портфелях

	AI	*SVI*
GRS	4,48	2,33
P-value	0,00	0,02

Исходя из результатов (Табл. 1.15), гипотеза о равенстве всех коэффициентов α нулю отвергается, как и для варианта с простой сортировкой.

У 12 из 18 портфелей коэффициент α оказался положительным, при этом отрицательные значения коэффициента, за исключением одного портфеля, наблюдаются только у портфелей, относящимся к группе с низким показателем внимания (Табл. 1.16). Этот результат также поддерживает гипотезу о наличии эффекта внимания на российском рынке, т.е. отбор акций в портфель по прокси-внимания позволяет получить повышенную доходность.

На основе описанных результатов можно сделать вывод, что на российском рынке можно признать существование эффекта внимания, но при этом обнаружить его удалось только для крупных компаний – минимальная капитализация компаний, попадавших в портфель «крупных», составила 70 млрд рублей.

Таблица 1.16

Значения константы (альфа) для портфелей при двойной сортировке

Авторская метрика AI				*Метрика поисковых запросов SVI*			
Коэффициент α							
	Крупные	Средние	Малые		Крупные	Средние	Малые
Высокое	0,132	0,129	0,156	Высокое	0,179	0,334	-0,105
Среднее	0,061	0,281	0,348	Среднее	0,129	0,071	0,048
Низкое	-0,246	-0,060	-0,297	Низкое	-0,015	0,048	-0,008
P-value							
	Крупные	Средние	Малые		Крупные	Средние	Малые
Высокое	0.000	0.501	0.168	Высокое	0.004	0.179	0.484
Среднее	0.660	0.042	0.035	Среднее	0.091	0.480	0.795
Низкое	0.026	0.536	0.008	Низкое	0,913	0.650	0.950

Глава 2
ОБЩЕРЫНОЧНЫЕ МЕТРИКИ СЕНТИМЕНТА. ПРОКСИ-МЕТРИКИ СЕНТИМЕНТА НА ОСНОВЕ МАКРОДАННЫХ

2.1. РАНЕЕ ПРОВЕДЕННЫЕ ИССЛЕДОВАНИЯ ПО ПРОКСИ-МЕТРИКАМ СЕНТИМЕНТА НА ОСНОВЕ МАКРОДАННЫХ[1]

К общерыночным метрикам сентимента относятся прокси-метрики, построенные на показателях макроуровня, потребительском поведении и по опросам.

Еще лорд Майнорд Кейнс указывал, что цена активов на рынке зависит не только от фундаментальных характеристик актива, но и во многом определяется общерыночными настроениями (оптимизмом или пессимизмом). Поведение инвесторов во многом определяется не рациональными расчётами и оценками, а «духом жизнерадостности». Как результат общей жизнерадостности возникающее «стадное чувство» приводит к иррациональным действиям, вызывающим экономически необоснованное изменение цен (например, пузыри).

Мы различаем общерыночные метрики настроений и метрики по отдельным ценным бумагам. Примером общерыночных метрик являются различные индексы, которые уже зарекомендовали себя для предсказания макроэкономической динамики и косвенно - поведения биржевых индексов акций. Например, индекс деловой активности или индекс потребительский настроений Мичиганского университета, индекс Американской ассоциации индивидуальных инвесторов (AAII), индекс настроений Investors Intelligence и индекс настроений инвесторов UBS/GALLUP (Ding et al., 2014; Sun et al., 2016; Dimpfl & Kleiman, 2019), индекс страха VIX (Smales, 2017; Long & Manyi Zhao, 2020).

Популярными прокси-метриками общерыночного сентимента являются (Qiu & Welch, 2004):

- CEFD (The Closed-End Fund Discount) – принимается за прокси сентимента инвесторов на финансовом рынке.
- The Consumer Confidence – является прокси сентимента инвесторов, а именно служит для идентификации потребителей как индивидуальных розничных инвесторов, которые являются соответствующими шумовыми трейдерами на рынке.
- The Michigan Consumer Confidence Index – индекс потребительского доверия, ежемесячно публикуемый Мичиганским университетом.

Недостатком таких индексов как прокси сентимент-метрик является ограниченность данных, которые используются в методике их формирова-

[1] Обзор ряда работ в данном параграфе выполнен студентами факультета экономических наук НИУ ВШЭ Плаксиной М.С. и Байрамкуловой Л.М. под руководством Т.В. Тепловой.

ния. Приведенные выше индексы рассчитываются в основном при помощи статистических данных (когда повышенная торговая активность на фондовом рынке или повышенный потребительский спрос могут рассматриваться как признак оптимизма домохозяйств, который теоретически должен вылиться и в спрос на фондовом рынке) или результатов опроса инвесторов.

Популярная метрика настроений – индекс страха. Smales (2017) по американскому рынку акций приходит к выводу о том, что VIX (американский индекс волатильности) является надежным показателем настроений. Аналогичны выводы и по китайскому рынку в работе Long and ManyiZhao (2020), где iVX делится на три временных горизонта: краткосрочные и среднесрочные колебания и долгосрочный тренд. Для создания индекса настроения для китайского рынка Long and ManyiZhao (2020) используют прокси-переменные на трех уровнях: макро-, мезо- и микроуровень. В качестве прокси переменных для макроуровня авторы выбрали: индекс цен производителей (PPI), индекс деловой активности (PMI), индекс удовлетворенности потребителей (SI), индекс потребительских цен (CPI), общий объем розничных продаж потребительских товаров (RSCD), индекс делового климата (CBCI), который является индикатором бума и спадов экономики Китая. Для мезоуровня авторы применяют четыре прокси переменные: дисконт по закрытым фондам (FDDR), коэффициент оборачиваемости, «кратное чистой прибыли» (Price-Earning ration) и количество новых инвесторов (NIA). Свой выбор авторы обосновывают более ранними аналогичным исследованием по китайскому рынку Xu and Zhou (2018).

Оценка сентимента инвесторов на основе биржевых данных может являться косвенной и неточной, так как процесс интерпретации настроения игроков рынка имеет вероятность к смещению в суждениях исследователей.

Проблема опросов и анкет – корректность задания вопросов и сложность учета межстрановых особенностей (потребительское поведение и корзина для измерения инфляции не совпадают по странам). Сбор мнений через анкетирование является трудоемким и сложным вариантом построения индекса настроения инвесторов, что также является недостатком данного способа получения данных для исследования сентимента. Поэтому в настоящее время популярность получили новые методы, основанные на искусственном интеллекте, которые призваны собирать и обрабатывать данные по сентименту за более короткие промежутки времени. Особое внимание многие исследователи уделяют Интернет-платформам и социальным сетям, которыми пользуется множество частных инвесторов, выражая свое мнение по различным активам финансового рынка.

Основной проблемой в изучении сентимента розничных инвесторов является ограничение индексов и метрик, впрямую характеризующих состояние настроения участников финансового рынка. В большинстве работ до начала 2000-х гг. сентимент рассматривается через прокси-метрики.

Потенциально прокси-метрикой общерыночных настроений могли бы стать индексы экономических настроений (Economic Sentiment Indicator,

ESI), но по разным странам применяются свои методики и объяснительная и прогнозная сила этих прокси не очень велика для фондового рынка. Заметим, что ESI для рынка США рассчитывается по методике Dow Jones, которая учитывает публикации в 15 ведущих деловых изданиях (когда количество негативных заметок превышает позитивные, авторы делают заключение об ухудшении экономического настроения). Для Германии ZEW Economic Sentiment индикатор рассчитывается ежемесячно Европейским центром экономических исследований (нем. Zentrum für Europäische Wirtschaftsforschung) на основе опросов 285 экономистов и аналитиков. Для РФ индикатор рассчитывается НИУ ВШЭ (Институт статистических исследований и экономики знаний) и является композитным (строится на трех факторах от четырех сфер экономики, а также на показателе потребительской уверенности).

Спрос на защитные активы (традиционно рассматриваемые такие активы-убежища (safe haven), как золото и другие драгоценные металлы) также может выступать прокси показателем неприятия инвесторами риска на глобальном рынке, т.е. метрикой негативного сентимента. В исследованиях 2020-х в качестве защитных активов также рассматриваются (Umar et al., 2023), кроме золота, серебро, криптовалюта биткоин, британский фунт (Pound Sterling, GBP), японская йена (JPY), китайский юань (CHY). Как показали исследования, спрос на защитные активы имеет предсказательную силу относительно волатильности различных инвестиционных активов, например, акций (Dai & Chang, 2021), нефти (Demirer et al., 2021) и валютные курсы (Wu et al., 2022).

Только с середины 2000х гг. появляются работы, где метрики сентимента выстраиваются на базе обработки текстовых сообщений в социальных сетях, т.е. на основе коммуникаций инвесторов между собой.

Изучение взаимосвязи прокси-метрик сентимента (по доверию потребителей, торговой активности и т.п.) дало интересные результаты. Baker and Wurgler (2006) показали, что авторские сентимент-метрики по-разному проявляются в поведении доходности акций, принадлежащих к разным инвестиционным стилям (акции роста, акции стоимости, акции различающиеся по размеру). Акции роста сложнее задействовать в арбитраже и труднее рыночно оценить, чем акции крупных компаний с долгой и стабильной историей зарабатывания прибыли и выплат дивидендов. Соответственно, сентимент в большей степени должен проявиться в акциях роста. Схожие гипотезы о влиянии размера компании на проявление сентимента в доходности акций обсуждается и в работах Brown and Cliff (2005) и Lemmon and Portnaguina (2006).

Работа Baker and Wurgler (2006) - самая цитируемая статья по сентименту в Journal of Finance. В работе эмпирически подтверждается гипотеза о высокой будущей доходности акций с низким сентиментом из групп малой капитализации и высокой волатильности. Авторы формируют составной индекс сентимента, который основан на общих вариациях шести базовых показателей настроений методом главных компонент: наблюдаемого дис-

конта по закрытым фондам, оборачиваемости акций NYSE, количества и средней доходности IPO в первый день торгов (доходность первого дня - это так называемый «эффект недооценки»), доли капитала в новых выпусках и дивидендной премии.

Работа (Baker & Wurgler, 2006) построена на ежегодных оценках авторского индекса настроений с1962 по 2001 гг. Дисконт для закрытых фондов (CEFD) представляет собой среднюю разницу между стоимостью чистых активов (СЧА) акций закрытых фондов и их рыночными ценами. Предыдущие исследования показывают, что CEFD обратно пропорциональна настроениям.

Baker and Wurgler (2006) получили зависимость (12) для рыночного сентимента (интегральной оценки авторского индекса):

$$\begin{aligned} Sentiment_t = &-0{,}241 CEFD_t + 0{,}242 TURN_{t-1} + 0{,}253 NIPO_t + 0{,}257 RIPO_{t-1} + \\ &+ 0{,}112 S_t - 0{,}283 P_{t-1}^{D-ND} \end{aligned} \quad (12)$$

Еще один пример построения интегрального общерыночного сентимента – работа Gao and Xie (2020) по китайскому рынку. Авторы анализируют возможность изменения настроений инвесторов прогнозировать краткосрочную избыточную доходность и аномальные объемы торгов на фьючерсном рынке. Применяется метод главных компонент. Авторы используют пять прокси:

- Сумма открытых позиций фьючерсов на фондовый индекс CSI 300 (OI).
- Дисбаланс купли-продажи (buy–sell imbalance, BSI). Сделка классифицируется как инициированная покупателем, если цена сделки равна или выше текущей лучшей цены спроса. Если цена сделки равна или ниже текущей лучшей цены предложения, сделка классифицируется как инициированная продавцом.
- Индекс Psychological Line (PSY) на фондовый индекс CSI 300. PSY считается качественным индикатором для прогнозирования краткосрочных разворотов рынка.
- Объем спотовых торгов фондовым индексом CSI 300 (TV).
- Регулировка скорости оборачиваемости фондового индекса CSI 300 (adjusting turnover rate, ATR). Пятый прокси – это индекс настроения фьючерсов на фондовые индексы, основанный на корректировке скорости спот-оборота. Если доходность выше нуля, скорректированная скорость оборота положительна, что указывает на «бычий» тренд на фондовом рынке. Если доходность ниже нуля, скорректированная скорость оборота отрицательная, что указывает на «медвежьи» настроения фондового рынка.

Двунаправленная связь между настроениями инвесторов, объемом торгов и избыточной доходностью исследуется в работе Gao and Xie (2020) с использованием модели векторной авторегрессии (VAR). В результате было обнаружено, что на дневном горизонте индекс настроений надежно

предсказывает избыточную доходность и аномальные объемы торгов. Более того, чувствительность доходности и объемов торгов к настроениям имеет более сильную положительную связь в периоды высокой волатильности и периоды низких маржинальных требований. Также настроения инвесторов служат причиной избыточной доходности и аномальных объемов торгов на дневном горизонте.

Для европейского рынка построение метрики общерыночного сентимента реализовано в работе Reis and Pinho (2020). Авторы строят метрику на месячных данных с января 1972 г. по июль 2019 г., используя следующие прокси: индекс потребительских настроений (CCI), индикатор экономических настроений, индекс волатильности (VSTOXX), цены золотых слитков и спред между доходностью десятилетних и трехлетних казначейских облигаций Германии.

Baker and Wurgler (2006), исследовав сентимент на основе общерыночных прокси-метрик, пришли к выводу, что важны не только стили инвестирования (крупные и малые компании, акции роста и стоимости), но и межстрановые различия. Культурные различия между народами (индивидуализм или коллективизм) могут играть значимую роль в относительной силе проявления поведенческих предубеждений. Значит, и проявление влияния «стадного» поведения может по-разному проявляться в динамике ценных бумаг на разных рынках. Это мотивирует исследования по отдельным странам, включая рынок России. Страновые исследования показывают несогласованные результаты.

Jackson (2004) не находит значимых доказательств влияния шумовых трейдеров на доходность акций на рынке Австралии, в то время как Schmeling (2007) подчеркивает значимость влияния настроений индивидуальных инвесторов на совокупную рыночную доходность (доходность индекса) в Германии.

Сложный исследовательский вопрос межстрановых исследований – какими метриками диагностировать различия по приверженности к коллективному поведению («стадности»). Использование метрики Hofstede (2001) достаточно спорно, хотя находит свое применение в большом числе работ, например, в работе Chui et al. (2008).

Можно предположить, что индивидуализм может порождать моментум эффекты в ценообразовании акций. Отсутствие индивидуализма (самостоятельности в принятии решений), т.е. коллективизм, может привести к предубеждениям, которые, скорее всего, порождают различные виды неэффективности рынка наряду с моментум эффектами. Соответственно, можно предположить, что в коллективистских странах (странах с коллективистским опытом, например, в бывших странах СССР) люди объединяются в устойчивые группы и могут инвестировать слаженно. Такая ситуация может привести к превалированию доминирующего мнения и, таким образом, к «стадной чрезмерной реакции» (Chui et al. 2008). Следовательно, можно ожидать, что фондовые рынки в коллективистских странах в большей степени подвержены влиянию настроений инвесторов, тогда как фон-

довые рынки в индивидуалистических странах, где люди склонны придавать большее значение собственной информации и мнению, более независимы от действий толпы, должны меньше зависеть от проявлений поведенческих предубеждений и меньше демонстрировать аномалии в ценообразовании на фондовом рынке.

Важный вывод ранее проведенных исследований по влиянию сегодняшнего сентимента на будущую доходность – отрицательная зависимость (снижающаяся доходность под влиянием положительного сентимента, измеряемого различными прокси-показателями на основе потребительских настроений) объясняется не только «стадностью» поведения инвесторов (коррелированным поведением между отдельными людьми), но и ошибками при необходимости принятия быстрых решений, по сути, сверх реакцией, например, на плохие новости. Реакция может быть настолько большой, что падение цены акций в день появления новости вызовет возможность для части инвесторов (например, институциональных) закупиться и заработать на следующий день или в следующую неделю. Таким образом, согласно эмпирическим результатам тестирования поведенческих аномалий в ценообразовании необходимо учитывать оба аспекта - стадность инвесторов и их чрезмерную реакцию на события или мнения других участников, что создает для определенной группы инвесторов возможность заработка.

В работе Jiang et al. (2018) исследование сентимента участников фондового рынка проведено на различных временных горизонтах, которые были получены за счет декомпозиции временных рядов. Авторы рассматривали прокси-метрики настроений инвесторов согласно работе Baker and Wurgler (2006). Самый короткий период составлял 3,6 месяца, а самый длинный - 47,74 месяца. Авторы определили, что сентимент инвесторов значимо объясняет доходность акций компаний-эмитентов только на долгосрочном горизонте инвестирования.

He et al. (2019) также используют готовую метрику настроений инвесторов – индекс Американской ассоциации индивидуальных инвесторов (AAII), рассчитываемый на основе опроса участников фондового рынка. Данный индекс измеряет процент людей, настроенных оптимистично, пессимистично или нейтрально по отношению к фондовому рынку в течение следующих шести месяцев. Было определено, что данная сентимент-метрика способна положительно объяснять доходность акций только в текущий момент наблюдения. Прошлое состояние сентимента оказалось незначимым в объяснении будущей доходности активов фондового рынка.

Похожие результаты в зависимости от учета памяти рынка с лагом в один день фиксируются в работе He et al. (2019). Авторы доказывают положительную связь между сентиментом и доходностью индекса S&P 500 в текущий момент наблюдения, однако влияние прошлого сентимента на будущую доходность (недельные наблюдения) отрицательно (положительный сентимент не позволяет инвестору заработать в будущем). Для объяснения доходности фондового рынка путем учета сентимента частных инвесторов авторы используют еженедельные доходности индекса S&P 500 и

данные по индексу AAII на основе сайта Американской ассоциации частных инвесторов с 24 июля 1987 по 10 декабря 2015 гг. (1481 неделя). Для поиска объяснений доходности индекса авторы используют TVA-GARCH-M модель. Также для определения причинной связи между сентиментом и индексом применяется тест Гренджера на основе регрессионного моделирования с МНК оценкой, который показал наличие влияния сентимента на доходность индекса S&P 500 и риск-премию на фондовом рынке.

Как попытка построить специфический отраслевой индекс сентимента на уровне отдельной компании интересна работа Bernile and Lyandres (2011). Авторы для измерения сентимента инвесторов используют ставки букмекерских компаний перед игрой определенной публично-торгуемой футбольной команды.

В работе Corredor et al. (2015) сопоставляются глобальные и локальные метрики сентимента в объяснении месячной ценовой реакции акций на текущий момент наблюдения. Заметим, что речь идет о прокси-сентименте, который строится по опросам потребителей. Авторы пришли к интересному выводу, что глобальные метрики сентимента значимы в объяснении динамики доходности по фондовым индексам, а вот использование локальных метрик сентимента (сентимент по каждой стране в выборке: Польша, Венгрия, Чехия, Великобритания, Германия, Франция) не поддерживает гипотезу исследования о наличии значимого эффекта в объяснении изменения доходности акций. Все биржевые данные работы Corredor et al. (2015) были получены с платформы Datastream в помесячном разрезе с 2001 по 2011 гг. Информация по локальным и глобальным метрикам сентимента (CCI, WEU, WEA) были получены с вебсайта Европейской комиссии. Локальные сентимент-метрики используются авторами на основе индекса CCI для ряда развивающихся европейских государств (Польша, Венгрия, Чехия) и экономически развитых стран (Великобритания, Германия, Франция). Авторы применяют предпосылку о том, что сентимент инвесторов тесно связан и может выстраиваться на оценках об экономическом состоянии стран (индекс промышленного производства, потребление товаров длительного пользования и уровень безработицы) в выборке исследования. Главная особенность этой работы и работы Schmeling (2008) - сентимент строится на индексе доверия потребителей (consumer confidence).

Глобальными метриками сентимента авторами признаются обобщенные показатели сентимента по доверию потребителей США и Европы (WEU and WEA), которые рассчитываются на основе авторской методики исследования (13), (14):

$$\begin{aligned} \textit{Worldwide EU (WEU)} &= 0.524 * \textit{US Consumer Confidence} + \\ &+ 0.524 * \textit{EU Consumer Confidence} \end{aligned} \tag{13}$$

$$\begin{aligned} \textit{Worldwide EA (WEA)} &= 0.533 * \textit{US Consumer Confidence} + \\ &+ 0.533 * \textit{EA Consumer Confidence} \end{aligned} \tag{14}$$

Первым шагом в качестве объясняемой переменной авторами были выбраны месячные доходности по каждой из трех развивающихся стран Европы для того, чтобы на основе МНК построить регрессию по выявлению наиболее сильной локальной сентимент-переменной. Далее авторы сопоставляют влияние сентимента в трех развивающихся странах с влиянием сентимента в наиболее экономически развитых странах Европы: Великобритании, Германии и Франции. Следующий этап в представленном исследовании заключается в тестировании портфелей с различными временными промежутками. Авторы сравнивают их доходность с ежемесячной доходностью рыночных индексов европейских стран, если бы активы сохраняли свои удельные веса 3, 6, 12 или 24 месяца. Данное тестирование выполняется также при помощи регрессионного моделирования на основе оценок МНК. В работе исследование строится по 14 европейским страновым индексам, а за базу построения метрики сентимента берется «Directorate Generale for Economic and Financial Affairs» (DG ECFIN).[1]

В работе Schmeling (2008) анализируется взаимосвязь между настроениями инвесторов по индексу потребительского поведения общерыночного и странового построения и доходностью индекса акций в 18 промышленно развитых странах. Автор приходит к важному выводу, что сентимент является важным предиктором ожидаемой доходности в среднем по странам. Зависимость отрицательная: высокий сентимент (по опросу потребительского поведения, т.е. оптимизм потребителей) порождает снижение доходности. Автор рассматривает несколько временных горизонтов будущей доходности и приходит к выводу, что сентимент имеет резко отрицательное влияние на доходность акций на всех прогнозных горизонтах (от 1 до 24 месяцев), хотя статистическая значимость ограничена уровнем 6% на самом длинном горизонте. Рассматривая инвестиционные стили, автор приходит к выводу, что сентимент эффекты для акций стоимости более выражены, чем для индекса в целом или акций роста, что поддерживает выводы работы Lemmon and Portnaguina (2006, p. 1526).

Автор строит исследование с применением панельной регрессии с фиксированными эффектами по 18 страновым индексам (Schmeling, 2008). В работе показано, что влияние сентимента на доходность индексов сильнее в странах с менее развитыми рыночными институтами и в странах, культурно более склонных к чрезмерной реакции инвесторов.

Опросы доверия потребителей строятся по стандартизированным на международном уровне наборе вопросов, который близок к вопросам, задаваемым в рамках опроса Мичиганского университета. В большинстве исследований в развитых странах используются эти стандартизированные вопросы для обеспечения международной сопоставимости. В частности, эти вопросы касаются прошлого и будущего финансового положения домохо-

[1] http://ec.europa.eu/economy_finance/indicators_en.htm

зяйства, прошлого и будущего по экономической ситуации в целом, учитываются крупные покупки товаров длительного пользования.

В исследовании Schmeling (2009) также по прокси-сентимента на основе опросов потребительского поведения при помощи регрессионного моделирования на панельных данных с использованием фиксированных эффектов выявлена связь между средней месячной доходностью страновых индексов (в выборке автора 18 рынков) и прошлыми настроениями потребителей на различных временных горизонтах (от 1 до 24 месяцев); средняя доходность была объясняемой переменной. В качестве прокси сентимента был выбран индекс CPI по каждой стране в выборке. Все данные составляют период с января 1985 по декабрь 2005 гг. в месячном разрезе.

В исследовании Schmeling (2009) обнаружено, что более высокие значения сентимента отрицательно объясняют будущую доходность (с лагами 1, 6, 12 и 24 месяцев) и наоборот. Будущую отрицательную доходность сентимент инвесторов значимо объясняет на временном горизонте 1 и 24 месяцев. В качестве метрики сентимента авторы использовали уверенность «покупателей» среди 18 различных стран как прокси для обозначения сентимента инвесторов на фондовых рынках.

В работе Huang et al. (2015) используют агрегированные месячные биржевые данные индекса S&P 500 с июля 1965 по декабрь 2010 гг. (546 наблюдений). Взаимосвязь с сентиментом авторы строят по скорректированному показателю доходности индекса, вычитая безрисковую ставку по каждому месяцу. Для формирования индекса сентимента авторы используют шесть прокси-переменных в помесячном разрезе: Close-end fund discount rate (CEFD), Share turnover (TURN), Number of IPO (NIPO), First-day returns of IPO's (RIPO), Dividend premium (PDND), Equity share in new issues (EQTI). Все данные о прокси-метриках сентимента инвесторов получены с сайта Jeffrey Wurgler.

Далее для определения индекса сентимента авторы используют предпосылку о том, что существует ненаблюдаемая компонента сентимента инвесторов, которая может объяснять ожидаемую доходность актива. Для определения данного сентимента авторы строят поэтапно две последовательные регрессии на основе МНК (полный цикл выведения сентимента представлен в работе Huang et al. (2015)). На основе полученных результатов, авторы представляют собственный индекс сентимента. Этот индекс получается на шаге оценивания зависимости (15):

$$S^{PLS} = -0{,}22CEFD + 0{,}16TURN - 0{,}04NIPO + 0{,}63RIPO + 0{,}07PDND + 0{,}53EQTI \quad (15)$$

Далее исследователи строят регрессии (однофакторную для объяснения доходности индекса S&P 500 с регрессором S^{PLS} и многофакторную с использованием всех шести прокси сентимент-метрик). Для объяснения доходности фондового индекса с использованием авторской композитной сентимент-метрики S^{PLS} авторы применяют три подхода к оценке данного

фактора: регрессия на основе МНК, модель с использованием смещенной оценки (Stambaugh) и Look-ahead bias-free модель (авторская разработка – двухшаговая регрессионная оценка сентимент-метрики S^{PLS}).

Прокси стадного поведения по корреляции доходности ценных бумаг строится в работе Choi and Yoon (2020). Авторы изучают взаимосвязь между стадным поведением инвесторов и их сентиментом на корейском фондовом рынке, анализируя KOSPI и KOSDAQ. Авторы используют Cross-Sectional Absolute Deviation (CSAD) и квантильную регрессию. Данные представляют собой выборку за период с января 2003 по декабрь 2018 гг. Все биржевые данные (доходности акций эмитентов) и сентимент инвесторов представлены в дневном разрезе на основе базы данных Informix[1]. Авторы пришли к другому способу определения прокси сентимента инвесторов: в качестве данной метрики используется индекс волатильности акций, входящих в индекс KOSPI200. Именно CSAD модель позволяет оценить стадное поведение инвесторов на фондовом рынке. Задействована предпосылка о том, что в случае нарастания стадного поведения доходность отдельных акций будет коррелировать с доходностью всего рынка. Следовательно, стадное поведение способно вызывать небольшое отклонение доходности в сравнении с доходностью отдельных акций, поэтому CSAD модель определяется как среднее значение суммы всех рассчитанных разниц доходностей между рынком и отдельными акциям.

Следует обратить внимание на основную модель в исследовании, которая основывается на модели ценообразования активов CAPM, но учитывает только доходность всего рынка или среднюю доходность активов в выборке исследования. Модель имеет вид (16):

$$CSAD_t = \alpha + \beta_1 \left| R_{m,t} \right| + \beta_2 R_{m,t}^2 + \varepsilon_t \qquad (16)$$

где $CSAD_t$ - мера перекрестной (cross-sectional) дисперсии доходности, $R_{m,t}$ - средняя рыночная доходность кросс-секции акций всех фирм в момент времени t, $R_{m,t}^2$ - нелинейный член, отражающий стадное поведение инвесторов. Если β_2 отрицателен и статистически значим, то это говорит о том, что на фондовом рынке существует стадное поведение. Однако, если β_2 положителен и значим, то на фондовом рынке существует обратное стадное поведение, имеющее тенденцию к большим покупкам на фондовом рынке. Нелинейность в данном случае авторы объясняют игнорированием инвесторами информации, которая выражается на основе сентимента и движением цен самих активов. То есть, инвесторы могут быть склонны к ведению большего числа сделок, провоцируя тем самым больший рост цен на акции и увеличение переменной CSAD.

Авторы пришли к следующим важным выводам: выделение растущих (восходящих) и нисходящих рынков показывает, что стадное поведение

[1] (https://www.ibm.com/products/informix

проявляется в периоды нисходящего тренда на фондовых рынках Кореи KOSDAQ и KOSPI. Согласно результатам метода квантильной регрессии, стадное поведение наблюдается в низких и высоких квантилях на фондовых рынках KOSPI и KOSDAQ. Это означает, что стадное поведение проявляется в экстремальных рыночных условиях. Следовательно, CSAD зависит от признаков экстремальных рыночных условий. В экстремальные периоды функционирования рынка инвесторы игнорируют свою собственную информацию и убеждения, следуют за движением рынка и подражают другим инвесторам. Анализ взаимосвязи между настроением инвесторов и стадным поведением показал, что настроение инвесторов имеет положительный эффект на фондовом рынке KOSDAQ, но не является значимым на фондовом рынке KOSPI.

Еще один пример построения прокси-сентимента по волатильности на рынке показан в работе Aharon (2021). Автор проверяет взаимосвязь между неопределенностью инвесторов на основе рыночных настроений и наличием стадного поведения участников фондового рынка. В качестве сентимента инвесторов в работе используется индекс волатильности VIX как метрика, отображающая степень страха участников фондового рынка. Данная метрика строится по CBOE. Однако для изучения феномена «стадности» инвесторов на рынке были построены 10 портфелей в зависимости от их размера. Данные портфели включают акции, которые котируются на NYSE, AMEX, NASDAQ. Данные по построению портфелей были получены на основе информации онлайн-ресурса Kenneth French. Анализируются дневные временные ряды с января 1990 по апрель 2019 гг.

Для определения стадного поведения участников фондового рынка авторы определяют временные моменты, когда возникает разрыв портфельной доходности с доходностью актива, входящего в данный портфель, на основе формулы (17):

$$CSAD = \frac{1}{N}\sum_{i=1}^{N}\left|R_{i,t} - R_{m,t}\right| \qquad (17)$$

где $R_{i,t}$ - доходность актива i в момент времени t, $R_{m,t}$ - средняя доходность рыночного портфеля, а N - количество фирм в общей выборке.

Далее выстраиваются регрессии на основе МНК оценки параметров моделей с включением дельты индекса волатильности VIX с целью учета влияния уровня страха инвесторов на рынке, и выявляется, как это может влиять на возникновение эффекта «стадности» на фондовом рынке. Помимо этого, исследователи строят квантильную регрессию и GARCH (1, 1) модель для тестирования эффекта страха и нелинейного эффекта портфельной доходности.

Нелинейная зависимость между дисперсией доходности и доходностью рынка больше, когда дельта VIX растет. Это говорит о том, что стадность проявляется сильнее, когда рынок перегружен настроениями, то есть воз-

никают завышенные ожидания инвесторов. В частности, стадность более заметна, когда ситуация на рынке ухудшается, и эта стадность может быть вероятным инструментом в формировании собственных стратегий, когда инвесторы пытаются защитить свои портфели, следуя мнению толпы.

Стадность и влияние индекса волатильности VIX проявляются не только в самых высоких квантилях распределения, но и в большинстве квантилей перекрестной абсолютной дисперсии доходности. Также было выявлено, что влияние индекса VIX на стадное поведение более выражено в периоды повышенной неопределенности, такие как финансовый кризис 2008 г. и европейский долговой финансовый кризис 2011 г.

Ниже приведен пример работы, где в качестве метрики сентимента использован показатель доли новости заданной тематики в общем новостном потоке. Изучение влияния информационного поля медиа пространства на состояние фондового рынка реализовано в работе Bossman et al. (2022) для определения силы эффекта покрытия новостных материалов, связанных с пандемией Covid-19. Для проведения исследования авторы использовали дневные доходности основных фондовых индексов стран, входящих в BRICS, а также метрики определения степени новостного покрытия по темам Covid-19 как для локальных измерений (LMCI – local media coverage indices), так и глобальных (WMCI – both world media coverage indices) за период с января 2020 по март 2022 гг. Каждый из представленных информационных индексов представляет собой долю новостей по теме пандемии Covid-19 в структуре общего числа новостей за отдельный временный промежуток времени. Данные по фондовым индексам были выгружены на основе системы Bloomberg, а информация по информационным индексам была получена за счет базы RavenPack, где был использован обобщенный индекс новостного покрытия MCI (Coronavirus Media Coverage Index). Методология исследования заключается в применении методов квадратичной вейвлет-когерентности и вейвлет-разности фаз с целью идентификации различных масштабных локализаций на основе анализа временных рядов доходности фондовых индексов и информационных метрик новостного покрытия с последующей визуализацией полученных результатов на основе тепловых карт. На основе полученных результатов авторы свидетельствуют об обнаружении связи между глобальным освещением в новостях пандемии Covid-19 и появлением высокого риска на фондовом рынке на квартальных и месячных временных горизонтах. Говоря о локальных информационных метриках, авторами не было выявлено значимой связи между динамикой фондовых рынков развивающихся стран и покрытием новостных СМИ по отдельно взятой стране из состава BRICS.

Глава 3
ТОНАЛЬНОСТЬ ПУБЛИЧНЫХ МЕДИА-КОММУНИКАЦИЙ, ОБСУЖДЕНИЙ НА ИНТЕРНЕТ-ПЛАТФОРМАХ И ФОРУМАХ

3.1. ТЕКСТОВЫЙ АНАЛИЗ В ПОСТРОЕНИИ МЕТРИК СЕНТИМЕНТА[1]

Классификация текста на группы по эмоциям (три или пять групп) основывается на использовании двух технологий, о которых речь пойдет ниже. Часто применяются методы искусственного интеллекта. Формально задача выглядит следующим образом: текстовому сообщению или некому эмодзи, стикеру присваивается вероятность принадлежности к соответствующему классу настроений. Первая технология (embedder) отвечает за то, чтобы тексты сообщений можно было трансформировать в набор признаков, используемых в дальнейшем при классификации. Его роль можно представить следующим образом: пусть человек хочет выучить язык. Каждый раз, когда он сталкивается с незнакомым словом, он открывает словарь и ищет всю информацию о нем, которая обычно включает в себя следующие элементы: транскрипция, значение, способы применения и т.д. Эти элементы написаны на понятном для человека языке, чтобы он смог, прочитав словарь, разобраться в том, что же значит новое слово и как его можно использовать.

В процессе классификации сообщений embedder выполняет роль словаря, который переводит подаваемые на классификацию предложения в набор чисел. Этот набор чисел есть элемент некоторого абстрактного пространства значений многих слов, используемых в нашем языке. Таким образом, каждое предложение кодируется 512-размерным вектором, понятным для компьютера. И тут возникают закономерные вопросы: зачем переводить сообщения в числа? Каким образом embedder обладает информацией обо всех словах в языке? Что делать с полученной классификацией?

На первый вопрос можно ответить так: пытаясь построить модель, способную отображать тексты в метрики сентимента, мы вынуждены переходить на язык чисел, поскольку нейронные сети - это математические модели, работающие с численными представлениями входных данных. В различных задачах, будь то работа с изображениями, аудио или текстом, на первом этапе все данные переводятся в числа.

На второй вопрос можно ответить так: исследователи из компании Google, используя большие данные и нейросеть, построили такой embedder и выложили его в открытый доступ. Ответ на третий вопрос: имея набор признаков, характеризующих конкретное предложение, мы можем построить обобщение всех предложений на основе этих признаков. Подобно тому как человек, обладающий переводом и транскрипцией неизвестного слова,

[1] Обзор ряда работ в данном параграфе выполнен студентом факультета экономических наук НИУ ВШЭ Соломоником М.А. под руководством Т.В. Тепловой.

учится его использовать при общении, мы, обладая векторным представлением предложения, будем учиться его классифицировать по сентименту. Вторая используемая нами технология - нейронная сеть. Она принимает на вход вектор признаков и выдает на выходе тональность сообщения. Нейросеть можно представить в виде графа. Каждый блок - это определенный слой, через который проходит вектор подаваемых признаков, модифицируя веса этого слоя таким образом, чтобы улучшалась точность предсказания на обособленной от обучения выборке сообщений. После обработки всех предложений (их векторных представлений) мы получаем обученную модель, способную по вектору признаков определять класс сообщения с определенной точностью. Точность классификации зависит не только от архитектуры модели, но и от качества разметки, используемой при обучении.

Нейронную сеть можно также представить в виде системы преобразования данных, имитирующей способ обработки информации в человеческом мозге. Система состоит из входного слоя, скрытого слоя и выходного слоя. Слои состоят из множества обрабатывающих компонентов, называемых нейронами. Нейроны обрабатывают данные и передают их на следующий уровень. В частности, нейроны соединены соответствующими связями между слоями, и каждое подключенное звено имеет числовой вес. Нейронные сети могут автоматически корректировать веса для улучшения результата. Многие работы по обработке текстовых сообщений применяют модель нейронных сетей для решения разных задач, включая прогнозирование динамики акций, прогнозирование кризисов на фондовом рынке.

С академической и практической точек зрения возникает следующий вопрос: каким образом определить, уместно ли использовать выстроенную архитектуру? Этот вопрос решается следующим образом: если алгоритм способен переобучиться, то есть практически идеально запомнить входные данные, то сложности используемой модели достаточно для моделирования связи между сообщениями и их тональностью. Если же переобучения не происходит, то необходимо искать другую, более подходящую архитектуру. Таким образом, обязательным элементом реализации нейронных сетей является проверка переобучением.

Rouf et al. (2021) приводят обзор различных методов анализа рынков и показывают, как развитие технологий обработки данных способствовало изменению приемов анализа. В статье сравниваются как сложные методы машинного обучения, такие как глубинные нейронные сети, так и простые линейные модели. Статья позволяет познакомиться и изучить инструментарий методов машинного обучения для анализа фондового рынка. Авторы также отмечают, что, несмотря на все новации, которые появились в последние годы, предсказание динамики цен активов по-прежнему остается весьма сложной задачей, даже с учетом современных приемов искусственного интеллекта.

Karlemstrand and Leckstrem (2021) применяют глубинные нейронные сети с полносвязными и LSTM слоями для анализа взаимосвязи текстов постов в

Твиттере с ценами акций. Для обучения нейронной сети используются исторические цены на акции, технические индикаторы и различная информация, касающаяся определенного поста в Твиттере, например, тональность, число подписчиков и признак, является ли аккаунт верифицированным. Для обучения нейронной сети требуется достаточно большой массив обучающих данных, который авторы собирали с помощью API Твиттера. Авторы получили интересный вывод, что при добавлении в модель факторов, касающихся дополнительной информации о посте в Твиттере, среднеквадратичная ошибка модели падает на 3%. Если же добавить факторы технического анализа, то ошибка улучшается еще на 11%. Авторы отмечают, что ограничением их исследования является то, что подобный анализ можно проводить только для акций, популярных среди пользователей Твиттера. Еще одна отмечаемая сложность – часы работы биржи. Биржа, в отличие от социальной сети, работает не круглосуточно, что может внести дополнительный шум при моделировании, так как информация, приходящая в процессе торгов и полученная в ночные часы, может оказывать разное влияние на участников.

Методология ИИ эластичной сети (Elastic Net) продемонстрирована в работах (Zou & Hastie, 2005, Афанасьев и др., 2019; Afanasyev et al. 2021).

В работе Jaggi et al. (2020) реализовано сложное с инструментальной точки зрения исследование, так как авторы сделали попытку проанализировать тексты с использованием современных алгоритмов глубинного обучения. Для анализа финансового жаргона была предложена новая модель FinALBERT, которая представляет собой модель ALBERT, обученную на текстах финансового лексикона. В качестве размеченных данных выступали сообщения на платформе Stocktwits (социальная сеть для инвесторов и трейдеров), которые были промаркированы на основе ежедневного изменения цен акций по 25 американским компаниям. Помимо FinALBERT, было протестировано множество других алгоритмов: как классические алгоритмы машинного обучения (логистическая регрессия, градиентный бустинг), так и различные вариации нейронных сетей. В работе приведены полученные по каждой из моделей метрики качества. В результате наилучшее качество наблюдалось у таких моделей, как наивный байесовский классификатор и случайный лес, которые превзошли по эффективности передовые алгоритмы глубинного обучения.

В работе Jing et al. (2021) нейросетевой подход LSTM применяется для прогнозирования цены закрытия акции Шанхайской фондовой биржи на один день вперед. В качестве входных данных используются технические индикаторы, полученные на основе котировок акций, и фактор настроений инвесторов на крупном фондовом форуме. Сверточная нейронная сеть (Convolutional Neural Network, CNN) применяется для классификации настроений инвесторов, которые извлекаются из сообщений на крупнейшем китайском инвестиционном форуме. Гибридная модель (Рис. 3.1) на основе нейросетевого подхода применяется для выявления краткосрочной

памяти (Long short-term memory, LSTM) для анализа технических индикаторов фондового рынка и результатов анализа настроений с первого шага.

Jing et al. (2021) приводит обзор работ по прогнозированию фондового рынка с помощью инструментов машинного обучения (Табл. 3.1). Популярные методы анализа текстовых сообщений включают SVM, наивный Байес и словарь оценочной лексики (Табл. 3.2).

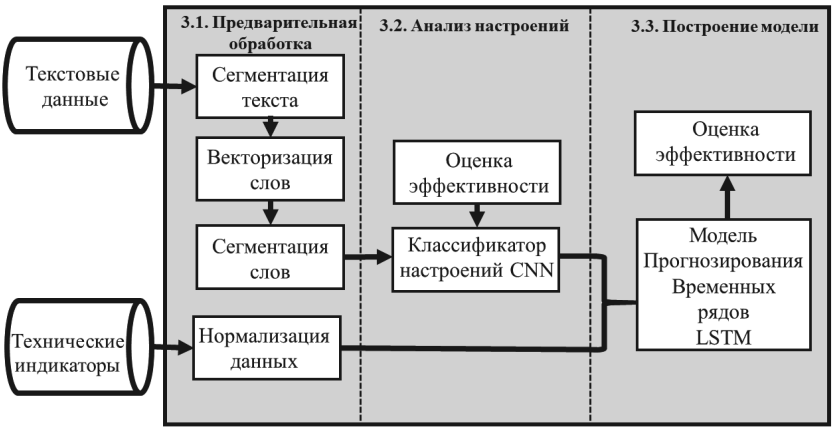

Рис. 3.1. Структура гибридной модели в работе Jing et al. (2021)

Таблица 3.1

Работы по прогнозированию фондового рынка с помощью инструментов машинного обучения

Литература	*Данные*	*Временной период*	*Частота*	*Метод*	*Прогнозируемый показатель*	*Показатели эффективности*
Hillebrand and Medeiros (2010)	NYSE	03.01.1995-31.12.2005	ежедневно	бэггинг, логлинейное, NN	волатильность	RMSE
Kara et al. (2011)	ISE National 100	02.01.1997-31.12.2007	ежедневно	SVM, ANN	направление движения (увеличение, уменьшение)	Accuracy
Kazem et al.	NASDAQ	12.09.2007-	еже-	SVR +	цена	MSE/MAPE

Литера-тура	*Данные*	*Временной период*	*Часто-та*	*Метод*	*Прогнози-руемый показа-тель*	*Показатели эффектив-ности*
(2013)		11.11.2011	дневно	FA		
Adebiyi et al. (2014)	NYSE	18.08.1988-25.02.2011	еже-дневно	ANN	цена	ошибка прогноза
Goçken et al. (2016)	BIST 100	08.06.2005-27.05.2013	еже-дневно	ANN, HS	индекс	MAE/ RMSE/ MARE / MSRE / RMSRE / MAPE / MSPE
Qiu et al. (2016)	Nikkei 225	ноябрь 1993 - июль 2013	ежемес ячно	ANN + GA	доход	MAE / NMSE / RMSE / MI
Zhang et al. (2016)	SZSE, NASDAQ	04.01.2010-22.05.2015	еже-дневно	SVM, Ada Boost, GA	тренд (по-купать, прода-вать, другое)	модифици-рованная метрика Accuracy
Chong et al. (2017)	KOSPI	04.01.2010-30.12.2014	5 мин	DNN	доход	MSE
Tashiro et al. (2019)	TSE	01.07.2013-30.06.2014	30 сек	улуч-шенная CNN	направ-ление движения	F1-мера, Precision
Chen and Ge (2019)	HKEX	01.01.2005-31.12.2017	еже-дневно	LSTM / Att LSTM	цена	Accuracy
Sirignano and Cont (2019)	NASDAQ	01.01.2014-31.03.2017	еже-дневно	LSTM	направ-ление движения	Accuracy
Jing et al. (2021)	SSE	три временных интервала	еже-дневно	LSTM + CNN	цена	MAPE

Источник: Jing et al. (2021)

Таблица 3.2

Работы по анализу настроений инвесторов в социальных сетях

Литература	*Тип текста*	*Источник*	*Временной период*	*Метод*	*Оцениваемый показатель*	*Показатели качества модели*
Li (2010)	корпоративные документы	Комиссия по ценным бумагам и биржам США	1994 - 2007	наивный Байес	класс: позитивный, нейтральный, отрицательный, неопределенный	доля успешного определения класса (success rate)
Huang et al. (2010)	новости, заголовки	Тайваньские электронные газеты	июнь 2005 - октябрь 2005	правило взвешенной ассоциации	уровень значимости	Precision / Recall
Groth and Muntermann (2011)	раскрытие информации	котируемые компании	не указан	SVM	класс: позитивный, негативный	Accuracy / Precision / Recall / F1-мера / AUC
Yu et al. (2013)	публикации	блоги, форумы, новости, микроблоги	01.07.2011 - 30.09.2011	наивный Байес	класс: позитивный, негативный	Accuracy, F-мера
Jin et al. (2013)	новости	Bloomberg	апрель 2010 - март 2013	латентное размещение Дирихле (LDA)	30 тем	Precision / Recall
Junque de Fortuny et al. (2014)	новости	Фламандские газеты	июнь 2005 - март 2012	лексикон	класс: позитивный, негативный	Accuracy / AUC
Li et al. (2014)	новости	FINET	январь 2003 - март 2008	словарь оценочной лексики	класс: позитивный, негативный, нейтральный	Accuracy
Wang et al. (2018)	публикации	Sina Weibo	2012 - 2015	словарь оценочной лексики	класс: позитивный, негативный, нейтральный	Accuracy / Precision / Recall / F1-мера / AUC

Источник: Jing et al. (2021)

3.1.1. Тональность публичных медиа-коммуникаций[1]

Tetlock (2007) количественно измеряет связь между СМИ и фондовым рынком. Используется ежедневный контент Wall Street Journal и анализируется тональность журнальных статей с помощью General Inquirer (GI). Автор показывает, что высокий пессимизм СМИ оказывает давление на рыночные цены с последующим возвращением к фундаментальным показателям. Аномально высокий или низкий пессимизм предсказывает высокий объем рыночных торгов.

Превалируют исследования заголовков и текстов деловых СМИ или новостных агрегаторов. Tetlock (2007) показал, что высокий пессимизм аналитических колонок в Wall Street Journal прогнозирует будущий разворот на падающем рынке акций США, а экстремальные значения пессимизма и оптимизма приводят к росту объема торгов. На наш взгляд, такой подход содержит критические допущения. Во-первых, авторы не раскрывают вопрос о том, являются ли все подписчики указанных изданий инвесторами на фондовом рынке США, и какова доля среди них частных инвесторов. Во-вторых, крайне затруднительно проверить реакцию подписчиков на опубликованные тексты, остаются открытыми вопросы о том, были ли в действительности прочитаны тексты и заголовки, и совершали ли подписчики сделки с акциями за период исследования в соответствии с тональностью статьи (Теплова и др., 2022).

Seong and Nam (2018) на данных корейского рынка показывают, какое влияние оказывают финансовые новости на крупнейшем новостном портале на динамику акций корейских компаний за 2014-2016 гг. Для этого с сайта Naver авторы выгрузили более миллиона записей, которые были промаркированы «1», если доходность акции компании в день выхода новости была положительна, и «0» в противном случае. Для векторизации текста использовался подход «мешка слов» (bag-of-words), после чего были сформированы признаки отнесения новости к тому или иному классу (методом TF-IDF), из которых были отобраны 10% с наивысшим значениями хи-квадрат. Авторы поставили исследовательскую задачу выявить, как информация по схожим компаниям помогает в прогнозировании доходности акции, для чего провели кластеризацию выборки. Поскольку при большом объеме данных наиболее популярна модель кластеризации k-средних, авторы использовали этот алгоритм для сегментации компаний и затем применяли метод обучения с использованием множественного ядра для прогнозирования доходности конкретной акции, объединяя информацию по компании и кластеру, к которому она принадлежит. В результате было показано, что полученная авторами модель более точно, по сравнению с альтернативными подходами, предсказывает движение котировок акций.

---

[1] Обзор ряда работ в данном подпараграфе выполнен студентами факультета экономических наук НИУ ВШЭ Ерохиным Д.А. и Соломоником М.А. под руководством Т.В. Тепловой.

В работе Yang et al. (2015) традиционные каналы коммуникаций - СМИ (крупные газеты, деловые журналы) - сопоставляются по влиянию на котировки акций в краткосрочном периоде с социальными сетями (блоги, форумы, Твиттер). Авторы применяли продвинутую технику анализа настроений, которая представляла собой наивный байесовский классификатор и выходила за рамки подсчета количества упоминаний. Авторы показали, что социальные медиа имеют более тесную связь с динамикой акций, чем традиционные СМИ. Еще одним выводом исследования стало обнаружение того факта, что различные площадки социальных сетей по-разному влияют на прогнозирование динамики акций, поэтому авторы отмечают важность дифференцирования источников текстовых данных.

3.1.2. Онлайн-платформы коммуникаций инвесторов и текстовый анализ как основа построения метрик сентимента

С 2010-х гг. появились работы, посвященные анализу сентимента инвесторов на базе социальных сетей (StockTwits, Twitter). Мы называем это направление анализа сентимента «прямым сентиментом» или «чистым сентиментом», так как участники рынка в переписке прямо выражают свое отношение к перспективам роста или падения акций и описывают свои действия. Базы коммуникаций позволяют отделить сообщения частных инвесторов от публикаций профессиональных участников (институциональных инвесторов). До 2020-х гг. подавляющее большинство работ фокусировалось на рынке США и на англоязычных обсуждениях (Oliveira et al., 2017; Al-Nasseri and Ali, 2018), в последние годы появились работы по китайским акциям (Dong & Gil-Bazo, 2020; Yang et al., 2020) и другим рынкам капитала.

Для социальной сети Твиттер сервис socialbakers.com реализует мониторинг и дает доступ к статистике в разрезе стран и заданных категорий (рубрикатор socialbakers.com) с ранжированием по числу подписчиков. Рубрикатор содержит восемь категорий, по каждой из них (развлекательные, новостные) можно выявить десять аккаунтов с наибольшим числом подписчиков. На основе этих данных может быть построено исследование влияния «гуру Твиттера» на поведение ценных бумаг. Например, у аккаунта ВЕСТИ более 3 млн подписчиков, у РИА Новости - 2,7 млн, у Дождя и Форбс Россия - по 2 млн, у Яндекса - 1,6 млн. В категории «Общество» есть аккаунты у Правительства РФ, Министерства иностранных дел, Государственной Думы и отдельных политиков, включая оппозиционных. Заметим, что аккаунты медийных и политических персон по числу подписчиков не уступают организациям.

Dong & Gil-Bazo (2020) на примере акций китайского рынка продемонстрировали связь настроения инвесторов и доходности акций. Авторами было проанализировано более 58 миллионов сообщений в национальной социальной сети Sina Weibo. Эта сеть интересна тем, что предоставляет возможность автоматически определять тональность текста, классифици-

руя сообщения на позитивные, нейтральные и негативные. После сбора данных по каждой акции авторы на ежедневной основе рассчитали авторский индекс сентимента, который, наряду с финансовыми характеристиками компании, выступал в качестве одного из регрессоров линейной модели прогнозирования доходности акций. В результате было обнаружено, что позитивные настроения инвесторов предсказывают повышение доходности акций в краткосрочном периоде, после чего происходит резкий спад котировок. Еще один вывод работы заключается в том, что значимая связь между сентиментом и доходностью акций в основном наблюдается при позитивном общерыночном новостном фоне.

Yang et al. (2020) рассчитывает авторский индекс настроения инвесторов, основанный на содержании сообщений на крупнейшем китайском финансовом форуме Eastmoney. Авторы диагностируют статистически значимое влияние индекса на объемы торгов (показатель ликвидности). Также показано, что более высокий индекс настроений существенно увеличивает риск снижения цены акций в ситуации общерыночного обвала цен.

Если текст в газетах относится к деловому письму, и сообщения можно разметить на позитивные, нейтральные и негативные классы с помощью готовых словарей и пакетов, то на площадках свободного общения в основном используется неформальный стиль, много сленговых выражений. Для неанглоязычных платформ это представляет отдельный этап текстовой обработки. Дополнительно усложняют задачу то, что сленг меняется со временем, может различаться по разным возрастным группам, поэтому часто авторы идут по пути применения самостоятельно обученной нейросети.

На 2022 г. примеры изучения сентимента на основе онлайн-платформ и обсуждений в мессенджерах представлены во многих работах. Однако, не все они подтверждают гипотезу о значимом влиянии обсуждений инвесторов тех или иных компаний на поведение цен акций (Kim et al., 2014). Скорее консенсус выводов связан с волатильностью: обсуждения в инвестиционном сообществе однозначно увеличивают торговую активность и волатильность акций (Lee et al., 2002; Huang et al., 2022). Интрига заключается в том, что в одних работах утверждается, что положительные мнения на онлайн платформах приводят к росту цены, и на следующий день инвестор может заработать положительную доходность, покупая акции с положительным сентиментом (McGurk et al. 2020; Wang et al. 2021), в других же выявляется отрицательная зависимость или доказывается, что положительные мнения не порождают рост цены, а важны только отрицательные высказывания (Huang et al., 2022). Поэтому исследовательский вопрос сохраняется. Открытыми остаются также вопросы значимости сентимента, построенного на текстовом анализе, относительно инвестиционных стилей, межстрановых сопоставлений.

Отчасти проблемой остается обработка текстов (методы сбора и классификации по тональности), а также выявление тех подгрупп активов, где обсуждения могут «двигать цены». Важно также учитывать таймфрейм (о каком временном горизонте построения инвестиционной стратегии идет

речь). В нашем обзоре ниже четко оговорены временные периоды, для которых делается вывод о целесообразности построения торговых стратегий.

В работе Das and Chen (2007) предложена авторская методология классификации настроений частных инвесторов из «биржевых сообщений» («stock message boards»). Методология включает в себя различные алгоритмы классификации, объединенные «схемой голосования». Достоинство, по мнению авторов, в том, что уровни точности аналогичны широко используемым классификаторам Байеса, но количество ложных срабатываний ниже. Временные ряды и перекрестная агрегация информации сообщения улучшают качество результирующего авторского индекса тональности, особенно при наличии сленга и двусмысленности в сообщениях. Эмпирическое тестирование подтверждает связь со стоимостью акций в техническом секторе, а также с объемами торгов и волатильностью.

Bollen et al. (2011) изучают текстовые сообщения в социальной сети Twitter с целью определения сентимента частных инвесторов по китайскому рынку на основе социальной сети Weibo. Авторы используют методику лексического анализа ICTCLAS 3.0 (Wanyun and Jie, 2013). Chen et al. (2014) используют информационный ресурс для частных инвесторов Seeking Alpha для определения связи между настроением пользователей платформы и ценовым поведением акций за период с 2005 по 2012 гг. Также используются и другие платформы, например, Sina и East Money (Li et al., 2014) для объяснения динамики доходности акций компаний-эмитентов. Для проведения анализа за счет текстовых данных, полученных из социальных онлайн-платформ, используется два инструмента разметки тональности сообщений пользователей Интернет-порталов: готовые или авторские словари настроений инвесторов. На основе таких словарей обучаются модели машинного обучения по идентификации классов сентимента.

Среди популярных готовых баз сентимента можно выделить General Inquirer, Loughran-McDonald (LM), SentiWordNet и HowNet. Loughran and McDonald (2011) разработали авторский словарь слов с негативной тональностью с целью оценки настроения частных инвесторов. Eickhoff and Muntermann (2016) используют готовые варианты сентимент-баз General Inquirer и Harvard IV-4 для изучения поведения розничных инвесторов на финансовых рынках. В исследованиях применение готовых баз упрощает задачу поиска данных, описывающих настроение пользователей онлайн-платформ, однако это приводит к другой проблеме – ограничение по количеству терминов и слов, которые относятся к финансовой сфере. Это потенциально может создать смещение оценок в моделях, снижая объективность выводов по проведенным работам в области изучения сентимента частных инвесторов.

Наблюдается малое количество работ, посвященных изучению сентимента на основе русскоязычных текстовых сообщений. Авторы работ по китайскому рынку акций решают подобную задачу путем создания собственных уникальных баз сентимента. Li et al. (2018), Borovkova and Tsiamas (2019) с применением машинного обучения определяют сентимент

инвесторов на основе социальных онлайн-платформ. По развивающимся рынкам многие исследования основываются на создании уникальных баз текстовых данных, их классификации и формировании авторских индексов. Такие работы отличаются использованием различных моделей на основе искусственного интеллекта по автоматической разметке текстовых данных, для каждого сообщения определяется класс тональности (Renault, 2017; Al-Nasseri, Ali, 2018; Wang et al., 2021; Bari and Agah, 2020; Ko and Chang, 2021).

Отметим библиотеку rulexicon, являющуюся словарем тональности русского языка. Комментарий по применению этого словаря дан в работе Федоровой и др. (2020). Это один из первых словарей, разделяющий экономические и финансовые тексты на русском языке по тональности.

Lee et al. (2002) исследуют на недельных данных взаимосвязь между сентиментом розничных участников рынка и уровнем волатильности фондового рынка на уровне фондовых индексов. Для определения сентимента инвесторов была использована база размеченных «писем» (сообщений) на основе информационного ресурса y Investors' Intelligence of New Rochelle. Каждое «письмо» оценивалось как «бычье» или «медвежье» в зависимости от рекомендаций аналитиков. Также был выделен третий класс сентимента, как «коррекция», когда аналитики информационного ресурса прогнозируют подъем рынка, но рекомендуют воздержаться от приобретения активов на нем. Период исследования - с 5 января 1973 г. по 6 октября 1995 г. В качестве основных бенчмарков состояния фондового рынка были выбраны три индекса: DJLA, S&P500, NASDAQ, которые были получены с CRSP.

В исследовании (Lee et al., 2002) используются недельные данные о доходности индекса и сентимента инвесторов. Для определения влияния сентимента участников фондового рынка на его поведение авторы используют GARCH-in-mean модель, которая содержит уровень настроения инвесторов на рынке и запаздывающие значения (лаг в 1 неделю) сентимента в модели. Авторы обнаружили, что изменения в настроениях отрицательно коррелируют с волатильностью рынка, то есть волатильность увеличивается (уменьшается), когда сентимент инвесторов становится более «медвежьим» («бычьим»). Значимость настроений для условной волатильности предполагает, что традиционные меры временной вариации риска упускают важный фактор. Однако было также обнаружено, что при более «бычьем» настроении инвесторов наблюдается рост избыточной доходности индексов фондового рынка США.

Kordonis et al. (2016) строят краткосрочные прогнозы цен акций с помощью агрегирования сообщений в сети Twitter (твиттов) за день, относящихся к определенной компании или финансовому индексу. Авторы выявляют тональность корпуса твиттов за день и оценивают корреляцию между выделенным сентиментом и волатильностью цен. Важно отметить, что имела место предобработка твиттов: они были очищены от шума, такого как теги пользователей, URL-ссылки и т.д.) и токенизированы, затем удалены стоп-слова (слова с нейтральным эмоциональным окрасом). Авторами была

сформирована обучающая выборка из твиттов, в которых присутствуют ключевые слова, по которым и определялась положительная, либо отрицательная тональность текста. Затем на этих данных были обучены модели машинного обучения Naive Bayes Bernoulli и Support Vector Machine. Использовался метод bag of words для выделения признаков из текстов, затем были оставлены только 10 000 самых значимых признаков. В результате авторам удалось выявить взаимосвязь между настроениями в Twitter и поведением цен на акции для топ-16 самых популярных технологических компаний по версии Yahoo Finance.

Напротив, Kim et al. (2014) делают вывод о том, что настроения инвесторов не влияют на будущую цену акций. Скорее цена акций влияет на будущее настроение инвесторов. Для анализа использовались тексты с площадки Yahoo Finance, всего около 32 миллионов сообщений о 91 компании. Сообщения поделены на пять групп настроений (“Strong Buy”, “Buy”, “Hold”, “Sell” and “Strong Sell”), по каждому сообщению есть данные об идентификаторе автора, времени публикации, содержании, местонахождении автора, поле, возрасте и коде фирмы. Для классификации сообщений используется Naïve Bayes алгоритм, для кодировки сообщений применяется bag of words, каждому сообщению присвоена бинарная переменная со значениями “купить” или “продать”. Для выявления влияния на цену акций используется кросс-секционный анализ и анализ временных рядов. В результате авторами не была найдена зависимость между настроениями инвесторов и поведением цен.

Пример портфельного отбора на базе сентимента представлен в работе Xing et al. (2018). Сентимент метрика используется для распределения активов в инвестиционном портфеле. Используется более 100 тысяч сообщений из социальной сети StockTwits более чем по 10 компаниям, а также ежедневные цены закрытия, объемы торгов и капитализация данных компаний. Для моделирования используются глубокие нейронные сети с LSTM слоями. Авторы рассмотрели три стратегии построения портфеля. Во-первых, the equal-weighted portfolio: акциям присваиваются равные веса (20%). Во-вторых, ARIMA portfolio: реинвестирование происходит каждый день в соответствии с прогнозируемыми ценами на основе модели ARIMA. В-третьих, Holt-Winters portfolio: происходит ежедневное реинвестирование на основе Holt-Winters additive smoothing method с изменяющимися во времени параметрами. В результате получилось добиться построения эффективного инвестиционного портфеля, который обыгрывал бенчмарки, построенные без использования сентимента.

Michaux (2019) исследует европейский рынок акций на предмет связи динамики котировок и постов в Твиттере. Автор также включает в исследование факторы капитализации рассматриваемых компаний и долю розничных инвесторов в торгах акциями. Получен интересный вывод, согласующийся с поведенческой экономикой, что именно негативные высказывания в Твиттере значимо влияют на поведение цены, а капитализация компании влияет на объем обсуждений ее на платформе.

Huang et al. (2020) исследуют взаимосвязь настроений игроков фондового рынка на китайской платформе Guba Eastmoney с доходностью и волатильностью акций. В статье предлагается новый алгоритм определения тональности текста, основанный на сглаженной по Лапласу нормализованной точечной кривой информации, сокращенно SO-LNPMI. В результате авторы получили, что статистически значима только зависимость между негативными настроениями на исследуемой платформе и доходностью фондового рынка, зависимость носит отрицательный характер. Положительные же настроения не имеют значимого влияния на доходность. Что касается волатильности, то авторы показывают, что положительные настроения отрицательно влияют на волатильность рынка, а отрицательные настроения, наоборот, положительно.

Yang et al. (2015) исходят из того, что в Твиттере сосредоточено крупное инвестиционное сообщество, которое способно оказывать значимое влияние на финансовые рынки. Оригинальность работы в том, что авторы разделяют сообщества по уровню их экспертности, заметности для других участников и влиянию на результаты торгов. Авторы выделили 50 авторитетных финансовых Твиттер-аккаунтов, из которых 25 - это наиболее успешные и известные трейдеры с наибольшим количеством подписчиков и публикаций, остальные 25 аккаунтов - это поставщики финансовых новостей, например, Bloomberg. По аудитории аккаунтов выстраивались сообщества. Участники сообщества фильтровались по ключевым словам, относящимся к финансовому рынку, что позволило в итоге сформировать группы из более чем 150 тысяч участников. Далее каждое сообщество было разделено на 3 группы: лидеры, компаньоны и случайные пользователи. Лидеры - это аккаунты, имеющие больше подписчиков, чем подписок, компаньоны или знакомые - это те, кто, в основном, подписан на тех, кто подписан на них. Случайные пользователи - это пользователи, имеющие больше подписок, чем подписчиков. В итоге исследователи показали, что влиятельные пользователи Твиттера имеют большое влияние на финансовое сообщество и на котировки цен акций. Самым значимым выводом стало то, что определенные люди, ведущие популярные аккаунты в Твиттере, могут влиять на поведение акций.

Сопоставление разных метрик сентимента, построенных на основе текстового анализа с применением искусственного интеллекта (ИИ) представлено в работе Wang et al. (2021), где изучалось влияние сентимента онлайн-платформы East Money на дневные биржевые характеристики акций, котирующихся на китайском фондовом рынке (125 акций компаний-эмитентов, котирующихся на Шанхайской фондовой бирже с 1 августа 2014 г. по 15 июня 2018 г.).

Wang et al. (2021) сопоставляют три сентимент-метрики (18)-(20):

$$S_{k,t} = \ln \frac{1 + M_{k,t}^{pos}}{1 + M_{k,t}^{nes}} \tag{18}$$

где $S_{k,t}$ – сентимент частных инвесторов на платформе East Money по каждой акции в выборке, $M_{k,t}^{pos}$ и $M_{k,t}^{neg}$ – количество позитивных / негативных сообщений на онлайн-платформе по каждой акции в выборке.

$$S_{k,t} = \ln \frac{1 + R_{k,t}^{pos}}{1 + R_{k,t}^{nes}} \tag{19}$$

где $R_{k,t}^{pos}$ и $R_{k,t}^{neg}$ – количество позитивных / негативных прочтений на онлайн-платформе по каждому посту, посвященной акции в выборке исследования.

$$B_t = \ln \frac{1 + Buy_t}{1 + Sell_t} \tag{20}$$

где B_t – индекс торгового дисбаланса, т.е. отношение количества заявок на покупку актива к количеству заявок на продажу (подразумевается, что отбираются наиболее большие сделки по акциям в выборке), t – период, равный одному торговому дню, Buy_t – объем сделок, направленных на покупку актива в момент времени t, $Sell_t$ – объем сделок, направленных на продажу актива в момент времени t.

Wang et al. (2021) строят модели машинного обучения на основе LSTM нейросети, задача которой автоматически классифицировать текстовые данные по тональности: негативные и позитивные. Тренировочная и валидационная базы данных разделены в соотношении 80% и 20% соответственно. Авторы также используют разработанный парсер, при помощи которого было получено 12 280 974 онлайн-комментариев, связанных с акциями, входящих в индекс CSI300. Биржевые данные включают дневные данные о сделках с акциями: объем торгов, цену закрытия, доходность акций, крупные сделки, капитализацию эмитента, бета-коэффициент и мультипликатор BV/P (обратное значение «кратное балансовой оценки») из базы данных CSMAR.

После удаления записей с недостающими значениями было отобрано 105 937 записей по 125 акциям, начиная с 1 августа 2014 г. по 15 июня 2018 г. Авторы строят регрессионные модели (с МНК оценками) на панельных данных для объяснения индекса торгового дисбаланса (соотношение заявок на покупку или продажу акций компаний-эмитентов), объема торгов по акциям в выборке и доходности по ним (три отдельные регрессии с указанными зависимыми переменными).

Результаты исследования (Wang et al., 2021) показали, что положительный сентимент частных инвесторов порождает положительную доходность акций на следующий день. Также положительный сентимент частных инвесторов оказывают значимое положительное влияние на объем торгов.

При этом лагированный объем торгов является хорошим предиктором самого объема торгов в краткосрочной перспективе на горизонте одного дня. Было выявлено, что сентимент непрофессиональных игроков китайского фондового рынка оказывает частичное влияние на дисбаланс ордеров на следующий день.

Приведенные примеры результатов исследований в области поиска связи между сентиментом частных инвесторов и ценовой реакцией акций компаний-эмитентов отражают неоднозначные выводы авторов, которые основываются на разнонаправленном направлении влияния сентимента на доходность активов фондового рынка. Используется широкий спектр готовых или авторских метрик сентимента.

Важный исследовательский вопрос прогнозных построений – временной горизонт, который на основе сентимента позволяет выстроить выигрышную инвестиционную стратегию. Частичный ответ на этот вопрос дает работа Behrendt and Schmidt (2018), где анализируются акции, входящие в индекс Dow Jones Industrial Average. Авторы приходят к выводу, что только совсем короткие инвестиционные интервалы (внутридневные таймфреймы в пять минут) обеспечивают выигрышность инвестиционных стратегий. Лаги переменных, отражающих количество записей и их эмоциональную окраску, имеют преимущественно положительное влияние на значение доходности, наблюдаемые в последующие пять минут. При рассмотрении более длинных периодов взаимосвязь ослабевает. Однако, Behrendt and Schmidt (2018) обнаружили, что метрика сентимента, основанная на анализе тональности публикаций в Twitter, оказывает статистически значимый эффект на значение внутридневной волатильности цены акций. С точки зрения профессиональных участников рынка данный результат является негативным, так как доказывает, что сентимент социальных сетей редко базируется на фундаментальных факторах и показателях компаний. Вывод работы авторов – реакция цены в большей степени базируется на эмоциях отдельных инвесторов. Отдельные участники сообщества могут существенно задавать тренд и менять настроение других инвесторов. По мнению авторов, более выигрышным может быть подход, при котором инвестиционная стратегия будет строиться на более длительном промежутке и сравниваться со стратегией, основанной на действиях институциональных инвесторов.

Еще один исследовательский вопрос – выделение отраслевых подгрупп, цены акций которых чувствительны к сентименту на онлайн платформах коммуникаций инвесторов. Отметим две работы, которые подчеркивают значимость отраслевой характеристики: Wittler (2018), Reboredo and Ugolini (2018).

В работе Wittler (2018) рассматривается влияние постов в Твиттере на поведение цен с акцентом на технологические компании (FAANG) в сопоставлении с компаниями из других секторов экономики. Автор отмечает, что в различные периоды может значительно изменяться активность поль-

зователей в Твиттере, что также влияет и на активность инвестиционного сообщества.

В работе Reboredo and Ugolini (2018) оценено влияние сентимента постов в Твиттере на компании, которые ведут свою деятельность в новых отраслях - в использовании возобновляемой энергии. Авторы приходят к выводу, что, по крайней мере, в данной области настроения в Твиттере не способны оказать значительного влияния на динамику цен, и использование текстового анализа не позволяет получить большую доходность, однако, возможно в других отраслях картина будет иная.

3.2. ПРОБЛЕМА НЕДОСТАТКА СЕНТИМЕНТ-МЕТРИК ПО ТЕКСТОВОМУ АНАЛИЗУ И НОВЫЕ ПОДХОДЫ: ДИВЕРГЕНЦИЯ МНЕНИЯ ЧАСТНЫХ ИНВЕСТОРОВ И ПОВЕДЕНИЕ ФОНДОВОГО РЫНКА

В Приложении А[1] представлены примеры исследований, цель которых состоит в изучении сентимента частных инвесторов и их влияние на биржевые характеристики. Авторами используются различные метрики сентимента и подходы к их применению для тестирования гипотез исследовательских работы.

Неоднозначность полученных результатов эмпирических исследований может быть объяснена рядом причин. Многие результаты исследований строятся на базе одной платформы, одного канала коммуникации, который, конечно, не покрывает всех инвесторов в рассматриваемые активы. Подобный подход способен оказывать смещение оценки влияния сентимента на биржевые характеристики, объясняя рыночные подвижки с недостаточным уровнем репрезентативности (Tetlock, 2007; Oliveira et al., 2013; Renault, 2017; Al-Nasseri & Ali, 2018).

Также в ряде исследований, где оценка сентимента производится через вычисления относительной доли положительных или негативных сообщений, упускается число сообщений по отдельным активам финансового рынка (Tetlock, 2007; Renault, 2017; Al-Nasseri & Ali, 2018). Другой особенностью исследований на текущем момент времени является превалирование обработки агрегированных текстовых данных по всем активам финансового рынка без их разделения по группам ликвидности акций или облигаций (Oliveira et al., 2013; Renault, 2017; Al-Nasseri & Ali, 2018). Лишь небольшая часть работ посвящена изучению связи между сентиментом и биржевыми характеристиками ценных бумаг путем учета уровня ликвидности самих активов (Oliveira et al., 2017).

Ранее исследования фактора настроения инвесторов финансового рынка выстраивались на основе доказательства однородности мнений игроков рынка, что позволяло выявить ценовую динамику и другие биржевые характеристики. В последние годы исследователи осознали важность учета

[1] Все приложения размещены по ссылке: https://fmlab.hse.ru/appendices

дивергенции мнений (Basak, 2005; Harris & Raviv, 1993; Harrison & Kreps, 1978; Miller, 1977; Hong & Stein, 2007).

Один из первых примеров значимого эффекта дивергенции мнений инвесторов представлен в работе Miller (1977), в которой утверждается, что мнение пессимистов не является значимым в объяснении цен акций, когда мнение оптимистов превалирует. Излишний оптимизм - часто предвестник возврата рыночной цены актива к своей справедливой стоимости. Интерес представляют модели, в которых учитывается дивергенция мнений инвесторов в объяснении ценовой реакции акций (Harrison & Kreps, 1978; Harris & Raviv, 1993; Basak, 2005).

Antweiler and Frank (2004) разработали метрику дивергенции мнений на основе текстовых данных, полученных из онлайн-платформ, с целью объяснения динамики объема торгов акциями компаний-эмитентов. Дивергенция может объяснить и возникновение аномальных объемов торгов по акциям компаний-эмитентов (Giannini et al., 2019). В работе (Sandner and Welpe, 2014) выявлена связь между дивергенцией сентимента частных инвесторов и доходностью акций на основе сообщений из социальной сети Twitter.

Общий недостаток всех вышеперечисленных работ по дивергенции – опора только на одну платформу обсуждений, т.е. поиск разногласий в оценках внутри определенной группы инвесторов, игнорируя возможность использования нескольких онлайн-платформ, т.е. нескольких сообществ инвесторов, различных по возрасту, интересам и т.п.

Hu and Tripathi (2016) выявили отрицательную зависимость доходности акций на текущий и будущий дни (t+1) в зависимости от сентимента дня t (как от положительных новостей, так и от отрицательных). Авторы выстраивают исследование на основании данных о дневной доходности акций, входящих в индекс ASX50. По этим акциям анализируются упоминания на австралийской онлайн-платформе для частных инвесторов HotCopper с 2014 г. по 2015 г. Для определения сентимента авторы выделяют из каждого сообщения на онлайн-платформе ключевые слова, относящиеся к активностям на фондовом рынке, к примеру, такие слова, как «buy», «sell», «hold». Далее на основе классифицированных сообщений выстраиваются два сентимент-индекса: «бычий» (21) и индекс «согласия» мнения частных инвесторов (22), который определяет степень разногласий между сообщениями по каждой акции.

$$Bullishness_{i,t} = \frac{M_{i,t}^{BUY} - M_{i,t}^{SELL}}{M_{i,t}} \cdot \ln(1 + M_{i,t}) \qquad (21)$$

где $M_{i,t}^{BUY}$ и $M_{i,t}^{SELL}$ - количество сообщений с призывом покупать и продавать активы на фондовом рынке соответственно, а $M_{i,t}$ – суммарное количество сообщений.

$$Agreement_{i,t} = 1 - \sqrt{\frac{M_{i,t}^{BUY} - M_{i,t}^{SELL}}{M_{i,t}}} \in [0,1] \quad (22)$$

Hu and Tripathi (2016) получили интересный результат: чем выше разногласие по индексу (22), тем выше доходность на текущий день, день *t+1* и *t+2*. Т.е. инвестирование «с толпой» не приносит прибыли инвестору.

Предположения о взаимосвязи настроения частных инвесторов и ценовой реакции акций компаний-эмитентов также находят подтверждение в исследовании Bouteska (2016), где выявлено наличие положительной связи между стандартным отклонением настроений розничных участников фондового рынка (стандартное отклонение рассчитывается на недельной основе по индексу AAII) и накопленной аномальной доходностью акций. Автор объясняет выявленную связь присутствием консервативной группой инвесторов на рынке. Настроение частных инвесторов было получено на основе данных American Association for Individual Investors (AAII), когда накопительная аномальная (избыточная) доходность рассчитывалась на основе использования событийного анализа с применением событийного окна размером три дня. Событием, ключевым для определения кумулятивной избыточной доходности, является день опубликования Earnings Restatement, то есть объявления по корректировке финансового отчета компаний-эмитентов. В качестве ключевой переменной в исследовании представлена дневная доходность акций эмитентов, на основе которой определяется избыточная доходность. Уникальность методики определения избыточной доходности по каждой акции заключается в способе расчета. Ожидаемая доходность определяется на основе взвешенной по стоимости рыночной доходности. Избыточная доходность акции рассчитывается как разность между фактической и ожидаемой доходности актива.

Вывод по отрицательной взаимосвязи сентимента частных инвесторов и доходности акций на основе месячных данных получен в работе Huang et al. (2015). Авторами предполагается наличие такого предубеждения, как предвзятость инвесторов. Причем, по мнению авторов, эта предвзятость в первую очередь связана с денежными потоками в будущем, чем ставками дисконтирования.

Ren et al. (2021) строят новостной индекс сентимента вне зависимости от биржевых характеристик акций. Цель работы состоит в изучении влияния различных типов информационных источников, таких как Sina Weibo и Sina Finance, на новостной сентимент по каждой компании, которая обсуждается в СМИ. Авторы строят серию сентимент-метрик (23)-(28) для определения настроения по каждому тикеру на изучаемых двух онлайн-платформах на основе дневных новостей и сообщений на платформах:

$$SentNews_{i,t} = \ln\left(\frac{1 + PN_{i,t}}{1 + NN_{i,t}}\right) \quad (23)$$

где $SentNews_{i,t}$ – прокси сентимента двух онлайн-платформ на основе новостных статей (PN – позитивных новостей, NN – негативных новостей);

$$SentWeibo_{i,t} = \ln\left(\frac{1+PW_{i,t}}{1+NW_{i,t}}\right) \tag{24}$$

где $SentWeibo_{i,t}$ – прокси сентимента социальной сети Weibo на основе количества микроблогов (PW – позитивные микроблоги, NW – негативные микроблоги).

Также Ren et al. (2021) строят отдельные сентимент метрики по определению количества позитивных или негативных сообщений по каждой из двух платформ:

$$SentWeiboP_{i,t} = \ln(1+PW_{i,t}) \tag{25}$$

где $SentWeiboP_{i,t}$ – прокси позитивного сентимента социальной сети Weibo на основе сентимента микроблогов (количество позитивных микроблогов);

$$SentWeiboN_{i,t} = \ln(1+NW_{i,t}) \tag{26}$$

где $SentWeiboN_{i,t}$ – прокси негативного сентимента социальной сети Weibo на основе сентимента микроблогов (количество негативных микроблогов);

$$SentNewsP_{i,t} = \ln(1+PN_{i,t}) \tag{27}$$

где $SentNewsP_{i,t}$ – прокси позитивного сентимента платформы Sina Finance на основе новостных статей по каждому тикеру;

$$SentNewsN_{i,t} = \ln(1+NN_{i,t}) \tag{28}$$

где $SentNewsN_{i,t}$ – прокси негативного сентимента платформы Sina Finance на основе новостных статей по каждому тикеру.

Для тестирования влияния различных сентимент-метрик на индекс сентимента в новостных статьях Ren et al. (2021) строят регрессионную модель с включением всех регрессоров (сентимент-метрик с лагом в 1 день), однако в качестве контрольных переменных используются метрики с лагом в 2 и 3 дня для включения эффекта «переходящего» сентимента, например, с пятницы на понедельник. Модель имеет вид (29):

$$
\begin{aligned}
SentNews_{i,t} &= \alpha_i + \alpha_t + \beta_1 SentNewsP_{i,t-1} \cdot SentWeiboP_{i,t-1} + \\
&+ \beta_2 SentNewsN_{i,t-1} \cdot SentWeiboP_{i,t-1} + \beta_3 SentNewsP_{i,t-1} \cdot SentWeiboN_{i,t-1} + \\
&+ \beta_4 SentNewsN_{i,t-1} \cdot SentWeiboN_{i,t-1} + \beta_5 SentNewsP_{i,t-1} + \beta_6 SentNewsN_{i,t-1} + \\
&+ controls + \varepsilon_{i,t}
\end{aligned} \tag{29}
$$

Всего в выборке Ren et al. (2021) присутствуют сообщения и новости по 2501 компании-эмитента акций за период с 2013 по 2014 гг. Для решения задачи определения сентимента по новостным статьям и микроблогам в социальной сети авторы использовали для сравнения готовую сентимент базу на основе платформы Sina, которая включает три класса тональности: негативный, нейтральный, негативный. Вместе с тем, авторы использовали модели BosonNLP и выбрали случайно 40 000 новостей для проведения автоматической разметки тональности тестовой выборки. Было обнаружено совпадение на 82,67% тональности текстовых данных с базой Sina. Таким образом, исследователи отобрали 60 млн. микроблогов с постами, которые содержали тикеры акций компаний-эмитентов.

Ren et al. (2021) приходят к выводу, что социальные медиа действительно влияют на появление настроений в СМИ для финансовых новостей. Согласованность настроений между реакцией социальных сетей и предыдущими новостными статьями усиливает устойчивость настроений в СМИ с течением времени. В отличие от этого, обнаружены свидетельства того, что социальные медиа снижают устойчивость настроений в СМИ с течением времени.

Shen et al. (2022) ставят целью изучить влияние сентимента новостных СМИ на доходность и волатильность китайского фондового рынка. Авторы анализируют доходность двух индексов: SZEI (композитный фондовый индекс Шеньчженя) и SSEI (композитный индекс Шанхайской фондовой биржи) на основе базы данных CSVMAR. Для определения сентимента авторы используют новости из базы данных GKG, полученной из системы GDELT. Были получены новостные статьи из таких разделов, как «Фондовый рынок», «IPO», «Экономический пузырь», используя Google BigQiery за каждый день для получения текстовых данных. Все биржевые и текстовые данные были получены с 1 июня 2016 г. по 31 декабря 2021 г. Для выявления тональности каждой новостной статьи на основе системы GDELT были токенизированы все статьи на слова, далее каждое слово было переведено на английский язык и автоматически оценено (на основе GDELT) в рамках сентимент-тональности. Непосредственно сама оценка тональности новостной статьи имеет вид (30):

$$
Tone_score_{t,j} = 100 \cdot \frac{\sum PositiveWords - \sum NegativeWords}{\sum TotalWords} \tag{30}
$$

Метрика *Tone_score* принимает значения от -10 до +10. Однако, для решения проблемы наличия одних и тех же уровней тональности новостных статей авторами было произведено усреднение совокупной тональности среди всех статей за отдельный день. Далее исследователи определяют текущую тональность как разницу средних значений тональности между сегодняшним днем и предыдущим (31), (32):

$$Average_tone_t = \frac{1}{M_t}\sum_{j=1}^{M_t} Tone_score_{t,j} \tag{31}$$

$$CurrentToneScore_{t,j} = Average_tone_t - Average_tone_{t=1} \tag{32}$$

Shen et al. (2022) также учитывают предпосылку о наличии расхождения мнений между различными новостными статьями и усредненным сентиментом среди всех новостных изданий в выборке за каждый день. Таким образом, авторы строят своего рода метрику дивергенции, а именно разброса сентимента среди новостных статей, используя следующую методику расчета (33):

$$Dispersion_t = \sqrt{\frac{1}{M_t}\sum_{j=1}^{M_t}(Tone_score_{t,j} - Average_tone_t)^2} \tag{33}$$

Разброс тональности статей *Dispersion* за определенный день меняется во времени и может передавать информацию о согласованности настроений всех новостных СМИ. Другими словами, дисперсия тональности может раскрыть важную информацию о рынке, когда усредненная тональность может быть замаскирована общей тональностью новостных статей за определенный момент времени. Когда дисперсия тональности новостей увеличивается, информационная неопределенность высока, и поэтому можно ожидать низких цен на акции. В противоположность этому, в случае нулевой дисперсии неопределенность отсутствует, и, следовательно, воздействие на доходность акций не ожидается.

Для изучения влияния данных метрик сентимента новостных СМИ на доходность фондового рынка Китая Shen et al. (2022) используют модель EGARCH (2,1). Результаты исследования показали, что оптимизм и меньшая дисперсия тональности новостных статей предполагают более высокую доходность рынка на следующий день (в модели были использованы лаги в 1 день). Что касается волатильности рынка, то пессимизм и большая дисперсия тональности новостных статей может говорить о более высокой волатильности рынка. Рыночная доходность и условная волатильность сильнее реагируют на негативные изменения настроений в новостных СМИ.

Изучение влияния тональности текстовых данных на биржевые характеристики является целью работы Федоровой и др. (2022), в которой авторы

пытаются найти взаимосвязь между сентиментом обсуждений на различных онлайн-платформах и ценовым поведением акций российских компаний-эмитентов и активов валютного рынка. Объект анализа - текстовые данные из социальной сети Twitter и новостные материалы онлайн-платформы РБК. Токенизация и разметка текстовых данных реализованы при помощи базы EcSentiThemeLex. Данная база является уникальным русскоязычным словарем, содержащим в себе финансовые тексты, классифицированные по пяти классам тональности. Для определения важности слова в работе авторы используют частотный метод TF-IDF, на основе которого определяется совокупный уровень тональности всего текста при помощи формулы (34):

$$S_i^j = \frac{\sum_{k=1}^{n} S_k^{word} TF - IDF(word)}{n} \tag{34}$$

где k – номер слова в тексте, S_k^{word} - тональность k-го слова в тексте, *TF-IDF(word)* – вес важности слова в тексте, n – число слов в тексте.

Далее Федорова и др. (2022) строят векторные авторегрессии с 50-дневным скользящим окном. В качестве параметров VAR-регрессий используются следующие переменные: r_t^{MOEX} - доходность индекса МосБиржи, r_t^{MIACR} - дневное изменение кредитной ставки MIACR, r_t^{curr} - дневное изменение валютного курса USD/RUB, r_t^{brent} - дневное изменение цены нефти марки Brent. Также в качестве метрик сентимента используются агрегированные переменные по новостным статьям издания РБК и микроблогам социальной сети Twitter.

Выводом работы Федоровой и др. (2022) стало наличие взаимного влияния тональности новостей и микроблогов в переходный период реакции временного ряда доходности российского рынка. Авторы работы выявили, что сила сентимента микроблогов становится все сильнее по своему влиянию на биржевые характеристики фондового и валютного рынка России в сравнении с влиянием сентимента отечественных новостных СМИ.

Внимание следует уделить и работе Huang et al. (2015), в которой делается упор на фундаментальные принципы эффективного рынка. Авторы оценивают возможную сентимент-составляющую как ненаблюдаемый фактор, который может объяснять ожидаемую доходность. Эконометрические расчеты подкреплены теорией о наличии премии за риск в формировании доходности.

3.3. ЭМПИРИЧЕСКОЕ ИССЛЕДОВАНИЕ ДИВЕРГЕНЦИИ СЕНТИМЕНТА ЧАСТНЫХ ИНВЕСТОРОВ НА ПЛАТФОРМЕ «ТИНЬКОФФ ПУЛЬС»

Анализ сентимента по текстам платформы «Тинькофф Пульс» проводился М. Файзулиным в двух направлениях: изучение влияния сентимента частных инвесторов на доходность акций российских компаний и на ликвидность (относительную долю торгов по акциям в выборке исследования). Представленные ниже результаты были получены при помощи регрессионного анализа несбалансированных панельных данных. Оценка коэффициентов моделей произведена на основе метода наименьших квадратов (далее – МНК) с фиксированными эффектами. Так как в процессе анализа была обнаружена гетероскедастичность и автокорреляции случайных отклонений, то был применен метод робастных стандартных ошибок (HAC).

Для определения степени смещения оценок регрессоров и их устойчивости в условиях изменения параметров моделей были построены уравнения с поочередным добавлением различных авторских метрик сентимента розничных инвесторов. Использовались такие контрольные переменные как: «Индекс страха» российского фондового рынка (RVI) как аналог индекса волатильности VIX, ставка по необеспеченным межбанковским кредитам (Ruonia), относительный объем торгов (для объяснения доходности акций).

Помимо этого, был проведен ADF-тест для проверки стационарности всех временных рядов, отобранных для настоящего исследования, а также тест Хаусмана для выбора требуемой методики оценки параметров регрессионных моделей. Результаты теста показали, что подход с использованием фиксированных эффектов является наиболее подходящим для МНК оценки уравнений на основе панельных данных по данной выборке.

В качестве объясняемых переменных были выбраны две биржевых характеристики: недельная доходность акций и объема торгов по ним. Всего в выборку вошли акции 131 компании-эмитента. В качестве контрольных переменных используются такие регрессоры, как недельная доходность индекса Мосбиржи и недельный прирост ставки межбанковского кредитования Ruonia.

В рамках анализа сентимента частных инвесторов по Тинькофф Пульс в качестве объясняющей переменной использовалась метрика из построенной линейки сентимент-метрик (Табл. 3.3). Сентимент определяется по текстовым данным - сообщениям пользователей платформы, которые отражают мнение непрофессиональных участников фондового рынка России.

Таблица 3.3

Рассматриваемые метрики сентимента

Метрика сентимента	*Формула расчета*	*Комментарии*
AG	$AG = 1 - \sqrt{1 - sent^2}$,	Индикатор *sent* показывает превышение позитивных сообщений над негативными. M_{buy} и M_{sell} показывают количество сообщений,

Метрика сентимента	*Формула расчета*	*Комментарии*
	$sent = \frac{M_{buy} - M_{sell}}{M_{buy} + M_{sell}}$	которые содержат слово «купить» или «продать» соответственно. Индикатор *AG* показывает степень согласованности мнений частных инвесторов
B, B_scaled	$B = M_{buy} - M_{sell}$ $Bscaled = \frac{B - B_{\min}}{B_{\max} - B_{\min}}$	B – разность между количеством позитивных и негативных сообщений. Bscaled – отмасштабированная метрика B, принимающая значения в диапазоне от 0 до 1
Blr	$Blr = \frac{M_{buy}}{M_{buy} + M_{sell}}$	Blr – доля позитивных сообщений в общем количестве позитивных и негативных сообщений
S	$S = \ln\left(\frac{1 + M_{buy}}{1 + M_{sell}}\right)$	S – относительное превышение доли позитивных сообщений над долей негативных
VA	$VA = Blr_t - Blr_{t-1}$	VA показывает разность долей позитивных сообщений за прошедший и текущий моменты наблюдений

Описательная статистика по выборке исследования

Основной особенностью исследования является использование недельных данных за период с 24.11.2019 г. по 15.05.2022 г., описательная статистика по которым приведена в Табл. 3.4.

Таблица 3.4

Описательная статистика переменных

Переменная	*Среднее*	*Медиана*	*Стандартное отклонение*	*Минимум*	*Максимум*
Return	-0,0003	0,0000	0,1250	4,6400	4,4100
IMOEX	-0,0019	0,0033	0,0435	-0,3170	0,1050
Sh_TV	0,0076	0,0000	0,0672	0,0000	0,9570
LnRVI	0,0085	-0,0077	0,1420	-0,3850	0,6390
Delt_RUONIA	0,0603	0,0250	0,9580	-2,9900	9,5300
AG	0,1500	0,0140	0,3190	0,0000	1,0000
B	-0,5340	1,0000	44,0000	0,0000	303,0000
B_scaled	0,5040	0,4880	0,2330	0,0000	1,0000
Blr	0,5390	0,5260	0,2060	0,0000	1,0000
S	0,1270	0,0690	0,5280	-2,5600	2,5600
VA	-0,0022	0,0000	0,2600	-1,0000	1,0000

Предварительный этап исследования

Предварительным этапом исследования является анализ закономерностей в поведении согласованности мнений частных инвесторов в зависимости от уровня волатильности фондового рынка. Так, российский аналог ин-

декса волатильности VIX, а именно «индекс страха» RVI, является бенчмарком для определения моментов повышенного беспокойства участников рынка. На Рис. 3.2 представлена динамика относительного недельного «индекса страха» RVI (показатель *LnRVI*) и среднерыночного состояния индекса согласованности мнений инвесторов *AG*, рассчитанного на основе среднего арифметического среди всех акций эмитентов выборки.

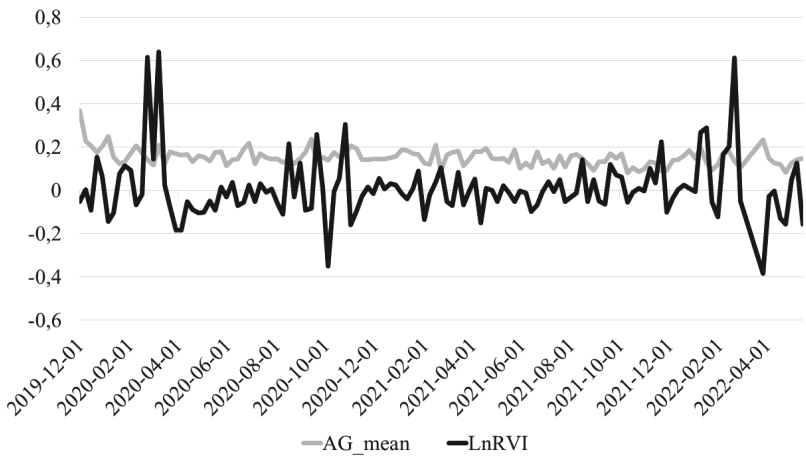

Рис. 3.2. Поведение «индекса страха» RVI и индекса согласованности мнений частных инвесторов (AG)

Наше опасение о зависимости согласованности мнений инвесторов от волатильности на рынке не подтверждается (Рис. 3.3). Тем не менее, мы проверили наличие в выборке таких эмитентов, по акциям которых может происходить снижение уровня дивергенции мнений частных инвесторов в моменты возникновения высокой динамики индекса RVI.

Т.е. мы предполагаем, что на рынке встречаются такие активы, по которым может возникать прочный консенсус мнений розничных инвесторов, когда происходит резкий рост волатильности по всему фондовому рынку.

Такое поведение может объясняться тем, что на рынке возможно присутствие непрофессиональных участников, которые придерживаются мнения большинства по отдельным эмитентам, когда по другим активам могут быть иные обсуждения, характеризующиеся высокой дивергенцией мнения розничных инвесторов.

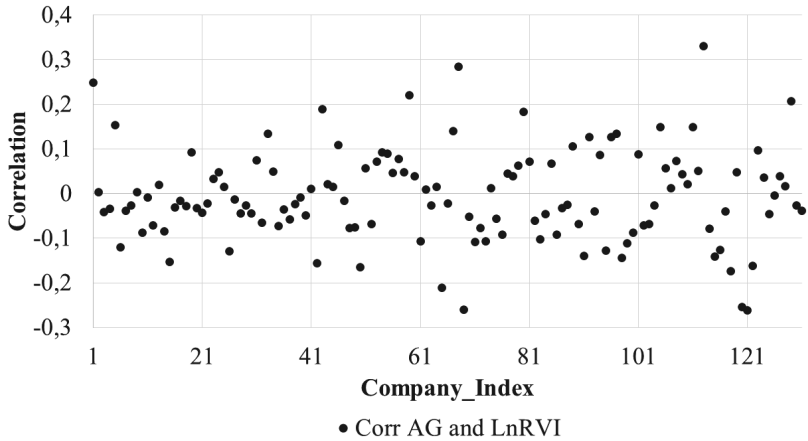

Рис. 3.3. Корреляционная зависимость между индексом согласованности мнений частных инвесторов (AG) и относительной динамикой «индекса страха» RVI (*Company_Index* совпадает с перечнем всех тикеров в Приложении Б, Табл. Б.1)

Для проверки данного предположения следует рассмотреть корреляцию между индексом согласованности мнений частных инвесторов (AG) и относительной динамикой «индекса страха» RVI (Рис. 3.3). Наличие отрицательной корреляции является сигналом к тому, что может возникать ситуация разнонаправленного поведения «индекса страха» и дивергенции мнений. Такими случаями могут являться моменты, когда при росте волатильности на российском фондовом рынке наблюдается низкое значение согласованности мнений частных инвесторов и наоборот. Из Рис. 3.3 видно, что примерно половина всех компаний из выборки исследования описывается отрицательной корреляцией между согласованностью мнений частных инвесторов и динамикой индекса RVI. Другими словами, когда возникает рост или падение индекса страха RVI, дивергенция мнений частных инвесторов начинается двигаться в противоположном направлении от «индекса страха».

Методология и результаты проверки гипотезы о влиянии различных метрик сентимента на биржевые характеристики акций на основе панельных данных

С целью рассмотрения вопроса о значимости сентимента частных инвесторов на платформе «Тинькофф Пульс» и согласованности мнений в онлайн-среде нами тестируется гипотеза, что метрики сентимента могут оказывать значимый эффект на доходность акций российских компаний-эмитентов и их ликвидность.

Следующая модель проверяется эмпирически (35):

$$Return_{i,t} = \alpha + \beta_1 IMOEX_t + \beta_2 Sh_TV_{i,t} + \beta_3 LnRVI_t + \beta_4 Delt_RUONIA_t + \\ + \sum \beta_{5,j} SENT_{j,i,t} + \varepsilon_{i,t} \quad (35)$$

где $Return_{i,t}$ – недельная доходность акций компаний-эмитентов, $IMOEX_{i,t}$ – логарифмическое изменение индекса Мосбиржи в момент времени *t*, $Sh_TV_{i,t}$ – доля объема торгов по *i*-ой акции в момент времени *t*, $LnRVI_{i,t}$ – логарифмическое изменение «Индекса страха» в момент времени *t*, $Delt_RUONIA_{i,t}$ – изменение необеспеченной ставки межбанковского кредитования (Ruonia) за период *t*; $SENT_{j,i,t}$ – серия поочередно использующихся сентимент-метрик (*AG*, *B*, *B_scaled*, *Blr*, *S*, *Sent-adj*, *VA*) в период *t*, *j* – порядковый номер сентимент-метрики в регрессии, $\varepsilon_{i,t}$ – случайная ошибка.

В Табл. 3.5 приведены результаты регрессионного анализа влияния сентимента платформы «Тинькофф Пульс» на доходность акций российских компаний-эмитентов.

Таблица 3.5

Регрессионный анализ влияния сентимента на платформе «Тинькофф Пульс» на доходность акций российских компаний-эмитентов

Зависимая переменная $Return_{i,t}$							
Переменная	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
const	0,0026	0,0026	-0,0224 ***	-0,0214 ***	-0,0285 ***	-0,0229 **	-0,0278 **
IMOEX	0,9708 ***	0,9596 ***	0,9461 ***	0,9466 ***	0,9471 ***	0,9486 ***	0,9444 ***
Sh_TV	0,1251 ***	0,1278 ***	0,1286 ***	0,1287 ***	0,1285 ***	0,1284 ***	0,0974 ***
LnRVI	0,0025	0,0023	0,0019	0,0019	0,0020	0,0021	0,0025
Delt_ RUONIA	-0,0017	-0,0017	-0,0016	-0,0016	-0,0016	-0,0016	-0,0016
AG	-0,0100 ***	-0,0100 ***	-0,0098 ***	-0,0093 ***	-0,0092 ***	-0,0095 ***	-0,0088 **
B		0,0001 ***	0,0001	0,0001	0,0001	0,0001	0,0001
B_scaled			0,0493 ***	0,0532 ***	0,0574 ***	0,0650 ***	0,0687 ***
Blr				-0,0055	0,0048	-0,0135	-0,0091
S					-0,0048	0,0078	0,0072
Sent_adj						-0,0065	-0,0069
VA							-0,0034

Зависимая переменная $Return_{i,t}$							
Пере-менная	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
AIC	-18958,98	-18967,91	-18995,52	-18994,48	-18993,19	-18992,14	-17068,35
BIC	-17928,10	-17929,45	-17949,47	-17940,85	-17931,99	-17923,36	-16002,13
R^2	0,1070	0,1077	0,1095	0,1095	0,1096	0,1096	0,1073
Obs	14 473	14 473	14 473	14 473	14 473	14 473	13 475

На основе приведенных результатов регрессионного моделирования (Табл. 3.5) обращаем внимание на модель (3), у которой наблюдаются наиболее низкие информационные критерии Акаике (AIC) и Шварца (BIC), несмотря на незначительную разницу коэффициента детерминации с моделью (5). Среди используемых метрик сентимента свою устойчивость показывают две метрики: показатель согласованности мнений (*AG*) и нормированный показатель превышения положительных сообщений над отрицательными (*B_scaled*).

Метрика согласованности мнений частных инвесторов на онлайн-платформе имеет закономерный знак. В случае, когда все инвесторы соглашаются, что нарастает «бычье» или «медвежье» настроение, доходность снижается. Наши расчеты показывают, что заработать вместе с толпой на недельном таймфрейме не удастся. Существует положительная связь между доходностью по акциям и увеличением положительного сентимента среди частных инвесторов (метрика *B_scaled*). Возникает противоречие относительно влияния описанных двух метрик: сентимент-метрика согласованности мнений (*AG*) оказывает эффект в два раза сильнее, чем переменная сентимента *B_scaled* с противоположным знаком коэффициента. Этот факт мотивирует нас углубить текущий анализ по отношению к каждой компании, выявив тех эмитентов, по которым может наблюдаться похожая ситуация.

Другим аспектом анализа сентимента частных инвесторов, проявляющих свою активность на онлайн-платформе «Тинькофф Пульс», является поиск детерминант доли объема торгов по *i*-ой акции в момент времени *t* по активам российского фондового рынка. Для этого используется регрессионная модель (36), в которую включена серия тех же сентимент-метрик:

$$Sh_TV_{i,t} = \alpha + \beta_1 IMOEX_t + \beta_2 Return_{i,t} + \beta_3 LnRVI_t + \beta_4 Delt_RUONIA_t + \\ + \sum \beta_{S,j} SENT_{j,i,t} + \varepsilon_{i,t} \quad (36)$$

Интересна модель (2) в Табл. 3.6: у нее минимальные значения информационных критериев AIC и BIC.

Таблица 3.6

Регрессионный анализ влияния сентимента на платформе «Тинькофф Пульс» на объем торгов акциями российских компаний-эмитентов

Зависимая переменная $Sh_TV_{i,t}$							
Переменная	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
const	0,0087 ***	0,0087 ***	0,0090 ***	0,0089 ***	0,0084 ***	0,0086 ***	0,0099 ***
IMOEX	-0,0022 **	-0,0015	-0,0014	-0,0014	-0,0014	-0,0013	0,0004
Sh_TV	0,0023 ***	0,0023 ***	0,0023 ***	0,0023 ***	0,0023 ***	0,0023 ***	0,0018 ***
LnRVI	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002
Delt_ RUONIA	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
AG	-0,0016 ***	-0,0016 ***	-0,0016 ***	-0,0016 ***	-0,0016 ***	-0,0016 ***	-0,0018 ***
B		0,0001	0,0001 *	0,0001 *	0,0001 *	0,0001 **	0,0001 **
B_scaled			-0,0005	-0,0006	-0,0003	0,0000	0,0001
Blr				0,0002	0,0009	0,0003	-0,0015
S					-0,0004	0,0000	0,0005
Sent_adj						-0,0002	-0,0004
VA							0,0010
AIC	-77088,21	-77088,82	-77086,97	-77085,01	-77083,23	-77081,30	-71246,58
BIC	-76057,33	-76050,35	-76040,92	-76031,38	-76022,03	-76012,52	-70180,36
R^2	0,9448	0,9448	0,9448	0,9448	0,9448	0,9448	0,9462
Obs	14 473	14 473	14 473	14 473	14 473	14 473	13 475

Среди всех сентимент-метрик наибольшую силу в объяснении различий ликвидности (объема торгов по акциям компании к совокупному объему торгов по всем активам в выборке) имеет метрика согласованности мнений (AG). Знак при переменной отрицательный. Согласованное понимание «качества актива» не мотивирует участников торгов к сделкам, и ликвидность снижается.

Полученные выводы позволяют выстроить логику дальнейшего исследования в направлении регрессионного анализа в разрезе каждого эмитента, учитывая единственный фактор сентимента – согласованность мнений частных инвесторов на платформе «Тинькофф Пульс».

Методология и результаты проверки гипотезы о влиянии различных метрик сентимента на биржевые характеристики акций в разрезе каждого эмитента

Если детально рассмотреть график распределения средних значений индекса согласованности мнений (*AG*) (Рис. 3.4), то можно увидеть, что не наблюдается сильных колебаний, кроме начала периода анализа с декабря 2019 г. по февраль 2020 г. Наше объяснение заключается в том, что платформа «Тинькофф Пульс» в этот период имела не очень высокую популярность среди розничных игроков на фондовом рынке, так как она была основана осенью 2019 г.

Увеличение аудитории платформы позволило снизить степень согласованности мнений среди частных инвесторов, что также подтверждает предположение о наличии различных групп пользователей на изучаемой онлайн-платформе.

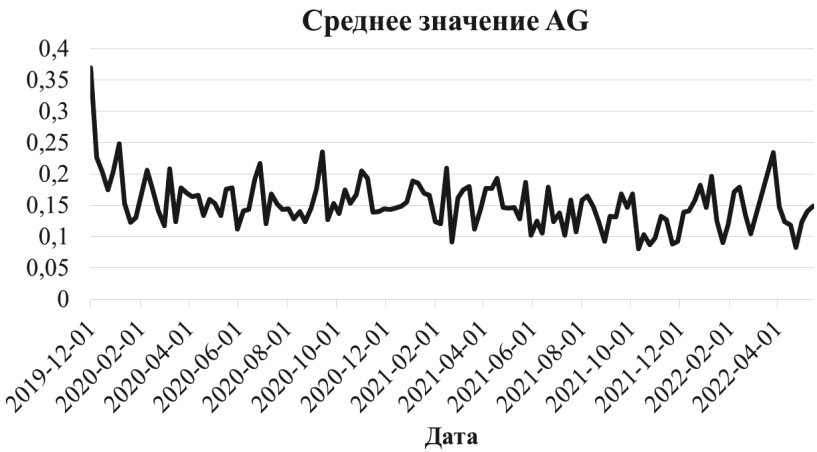

Рис. 3.4. Динамика средних значений сентимент-метрики AG (индекс согласованности мнений)

Однако, если рассмотреть график панельных данных по индексу согласованности мнений частных инвесторов (Рис. 3.5), то увидим часто встречающиеся максимальные значения данного индекса.

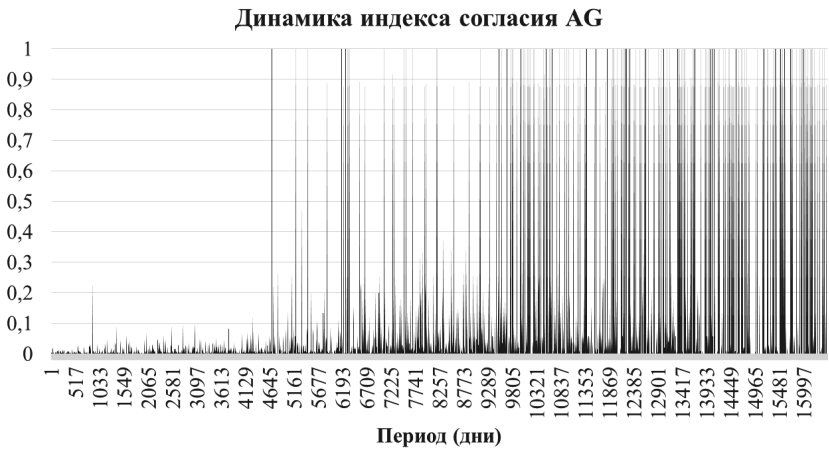

Рис. 3.5. Динамика значений сентимент-метрики AG (индекс согласованности мнений) по всем периодам наблюдения в выборке

В выборке присутствуют компании, по акциям которых была отмечена наибольшая сила негативного или позитивного сентимента среди участников онлайн-платформы. Указанные факты позволяют перейти к следующему этапу анализа сентимента пользователей платформы «Тинькофф Пульс» по каждой компании-эмитенту акций в отдельности (37), (38).

$$\begin{aligned} \mathrm{Re}\,turn_{i,t} &= \alpha + \beta_1 IMOEX_t + \beta_2 Sh_TV_{i,t} + \beta_3 LnRVI_t + \beta_4 Delt_RUONIA_t + \\ &+ \beta_5 AG_{i,t} + \beta_6 B_scaled_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{37}$$

$$\begin{aligned} Sh_TV_{i,t} &= \alpha + \beta_1 IMOEX_t + \beta_2 \,\mathrm{Re}\,turn_{i,t} + \beta_3 LnRVI_t + \beta_4 Delt_RUONIA_t + \\ &+ \beta_5 AG_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{38}$$

Стоит отметить, что при изучении влияния сентимента на ценовую реакцию акций российских компаний-эмитентов наблюдалась достаточно слабая и неустойчивая сила объяснения динамики доходности по акциям эмитентов за счет сентимент-метрики превышения позитивных сообщений над негативными (показатель *B*), что позволило исключить данный параметр из моделей (36), (37).

Полученные результаты моделирования (Приложение Б, Табл. Б.2) отражают ситуацию, когда более распространенным вариантом объяснения ценовой реакции акций компаний-эмитентов на российском фондовом рынке является использование нормированной метрики сентимента превышения позитивного настроения частных инвесторов над негативным (*B_scaled*).

Также стоит отметить, что в большинстве случаев индекс согласованности мнений (*AG*) имеет положительное воздействие в объяснении различий доходности акций компаний-эмитентов. Для 8 из 19 компаний, доходность акций которых статистически значимо объясняется данным индексом, обнаружен отрицательный эффект. Можно предположить, что влияние сентимента частных инвесторов на ценовую реакцию акций российских эмитентов неоднозначно. Особенность методики расчета метрики AG заключается в том, что она показывает степень согласованности мнений от 0 до 1, но не учитывает тон мнений инвесторов (когда все комментарии негативны или все комментарии позитивны, метрика принимает одно и то же значение). Мы предполагаем, что если коэффициент регрессии при переменной AG по определенной компании имеет отрицательный знак, значит, скорее всего, преобладают негативные комментарии, и чем выше AG, тем ниже доходность. Наоборот, по другим компаниям преобладают оптимистичные высказывания, и чем выше значение AG, тем выше доходность акции. Таким образом, некоторые компании могут обсуждаться частными инвесторами больше в негативном ключе, тогда как другие – преимущественно в позитивном ключе.

Заметим, что индекс согласованности мнений (*AG*) является наиболее значимым фактором в объяснении различий ликвидности акций российских компаний-эмитентов (Приложение Б, Табл. Б.3). Наблюдается схожая ситуация, когда по различным активам присутствует неоднозначная реакция относительного объема торгов в зависимости от уровня согласованности мнений частных инвесторов. Стоит отметить, что, несмотря на статистическую значимость данной сентимент-метрики для большей части компаний, она имеет достаточно близкое к нулю значение коэффициента регрессии.

Данный случай может описываться наличием в выборке компаний, чьи акции не пользуются особой популярностью среди инвесторов, что обосновывается низкой долей в структуре общего объема торгов на российском фондовом рынке. Так, среди компаний с наиболее крупным объемом торгов акциями наблюдается отрицательный эффект в объяснении ликвидности. Выявленный факт может свидетельствовать о наличии согласованного негативного мнения по наиболее торгуемым компаниям на российском фондовом рынке, когда большее значение индекса согласованности (*AG*) демотивирует совершать биржевые сделки. Для поиска объяснений такого влияния сентимента следует учитывать долю объема торгов, которую формируют физические лица на российском фондовом рынке. Примером может служить то, что Мосбиржа публикует месячные данные по объему торгов среди частных инвесторов как суммарно по фондовому рынку, так и с выделением топ-10 компаний, акции которых были наиболее популярны среди непрофессиональных участников.

Методология и результаты проверки гипотезы о влиянии метрик сентимента на биржевые характеристики популярных акций

Анализ влияния сентимента розничных участников рынка позволит определить, существуют ли определенные взаимосвязи между биржевыми характеристиками популярных акций и сентиментом частных инвесторов. В качестве контрольных переменных будут использоваться показатели: а) доли объема торгов среди физических лиц к совокупному объему торгов по рынку акций (*Sh_Retail*), б) доли объема торгов по каждой компании в выборке Топ-10 наиболее популярных акций среди непрофессиональных инвесторов к совокупному объему торгов акциями данной компании (*Retail_asset*). Таким образом, можно определить, как увеличение или уменьшение популярности акций среди непрофессиональных игроков может влиять на ценовую реакцию и ликвидность популярных акций (входящих в Топ-10 среди физических лиц по данным МосБиржи).

Для следующего этапа расчетов (учет популярности акций) собраны данные в месячном разрезе. В Табл. 3.7 представлена описательная статистика по показателям, которые используются в приведенных ниже регрессионных моделях по панельным данным (39), (40). В выборку нашего исследования входят топ-10 компаний с наиболее популярными среди физических лиц акциями по данным МосБиржи (список таких компаний меняется ежемесячно).

Таблица 3.7

Описательная статистика переменных в помесячном разрезе

Переменная	*Среднее*	*Медиана*	*Стандартное отклонение*	*Минимум*	*Максимум*
Returns	-0,0048	0,0015	0,0822	-0,2840	0,2380
IMOEX	-0,0009	0,0123	0,0547	-0,0977	0,1270
Sh_TV	0,0562	0,0002	0,1970	0,0000	0,8800
Sh_Retail	0,4130	0,4140	0,0187	0,3800	0,4640
LnRVI	0,0657	0,0387	0,1940	-0,2440	0,5280
Delta_RUONIA	0,1000	0,1300	0,3870	-0,8100	0,8500
Retail_Asset	10,0000	7,5000	6,0000	3,8000	32,2000
AG	0,0070	0,0047	0,0073	0,0000	0,0598
VA	-0,0008	-0,0015	0,0152	-0,0468	0,0314
Deposits	135,5000	121,6000	597,6000	-881,1000	1696,0000
Assets	137,1000	143,2000	129,7000	-167,1000	519,0000

$$Return_{i,t} = \alpha + \beta_1 IMOEX_t + \beta_2 Sh_TV_{i,t} + \beta_3 Sh_Retail_t + \beta_4 Retail_Asset_{i,t} + \\ + \beta_5 LnRVI_t + \beta_6 Delt_RUONIA_t + \beta_k SENT_{i,t} + \varepsilon_{i,t} \quad (39)$$

$$Sh_TV_{i,t} = \alpha + \beta_1 IMOEX_t + \beta_2 Return_{i,t} + \beta_3 Sh_Retail_t + \beta_4 Retail_Asset_{i,t} + \\ + \beta_5 LnRVI_t + \beta_6 Delt_RUONIA_t + \beta_k SENT_{i,t} + \varepsilon_{i,t} \quad (40)$$

где Sh_Retail_t – доля объема торгов среди частных инвесторов в совокупном объеме торгов по рынку акций за месяц t, $Retail_Asset_{i,t}$ – доля объема торгов акциями среди частных инвесторов в общем объеме торгов акциями компании i за месяц t.

Однако, следует упомянуть о том, что для ряда сентимент-метрик была обнаружена проблема отсутствия стационарности во временных рядах. Поэтому далее в нашем исследовании мы фокусируемся на двух метриках: индексе согласованности мнений (AG) и разности относительных долей позитивных сообщений за прошедший и текущий моменты наблюдений (VA).

Полученные результаты (Приложение Б, Табл. Б.4) свидетельствуют о наличии значимой на 5% уровне взаимосвязи между объемом торгов по наиболее популярным среди физических лиц акциям и индексом согласованности мнений. Наблюдается значимое отрицательное влияние показателя *Retail_Asset* на долю объема торгов акциями той или иной компании в выборке с достаточно высоким уровнем коэффициента регрессии. Мы предполагаем, что увеличение доли непрофессиональных участников негативно влияет на ликвидность актива по причине избыточного внимания розничных инвесторов. Избыток внимания может вызвать избыточную волатильность актива или непредсказуемое его ценовое движение за счет присутствия манипулятивных действий на различных онлайн-платформах.

Также наблюдается отрицательный эффект воздействия индекса согласованности мнений на ликвидность. Эта зависимость может быть объяснена наличием избыточного отрицательного сентимента: большее количество негативных сообщений способно оказывать эффект на другие группы инвесторов, побуждая их удерживать активы в моменты пессимизма (проявление эффекта «якорения»). Однако, не следует забывать о низком уровне коэффициента детерминации в объяснении доли объема торгов «народными» (т.е. наиболее популярными среди физических лиц) акциями. Данный факт не позволяет с достаточной уверенностью полагать, что полученные результаты могут описывать любое движение объясняемой переменной за счет отобранных регрессоров.

Иная ситуация наблюдается в части объяснения ценовой реакции топ-10 популярных акций. Несмотря на достаточно высокий уровень коэффициента детерминации, не было выявлено статистически значимых взаимосвязей между доходностью акций и сентиментом частных инвесторов, а также долей частных инвесторов в объеме торгов.

Методология и результаты проверки гипотезы о влиянии метрик сентимента на биржевые характеристики акций с учетом макроэкономических факторов

На основании полученных выше результатов следует предположить, что на объясняемые биржевые характеристики могут оказывать влияние другие экзогенные факторы, связанные с состоянием сбережений населения и ди-

намикой объема средств домохозяйств, выделяемых на покупку акций компаний, паев и акций инвестиционных фондов.

Стоит отметить, что рост популярности фондового рынка среди физических лиц имел тенденцию до событий весны 2020 г. Так, на Рис. 3.6 наблюдается плавный и относительно небыстрый рост торгового оборота среди частных инвесторов с января 2019 г. по февраль 2020 г. С марта по май 2020 г. наблюдалась взрывная динамика как общего, так и относительного объема торгов среди розничных инвесторов, которая объясняется ростом популярности фондового рынка среди населения.

Начиная с весны 2020 г., не было зафиксировано отрицательной динамики торгового оборота акций среди сегмента непрофессиональных участников фондового рынка. Это может говорить о закреплении части инвесторов, которые активно торгуют акциями на бирже для получения дополнительного дохода.

Активность физических лиц обусловлена множеством макроэкономических факторов, которые влияют на инвестиционные стратегии. Одними из ключевых факторов являются объем денежных средств на депозитах физических лиц и средств, вложенных в активы фондового рынка (Приложение Б, Рис. Б.1).

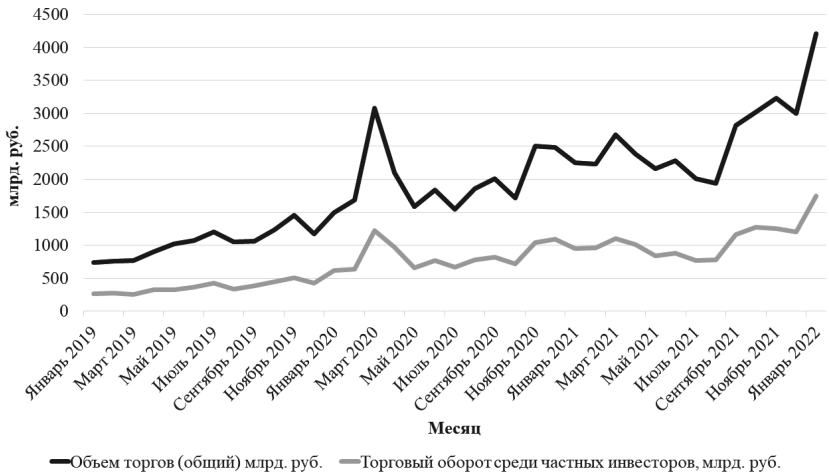

Рис. 3.6. Динамика объема торгов на Мосбирже с выделением частных инвесторов

С одной стороны, закономерностью можно считать превышение норм среднемесячного потребления со стороны домохозяйств в декабре-январе, что связано с праздничными днями в России. Это мотивирует многие домохозяйства снимать свои сбережения с депозитов для компенсации убытков. С другой стороны, не наблюдается резких всплесков в эти периоды по

движению средств на счетах частных инвесторов, чьи деньги вложены в акции и паи. Предполагается, что розничный сегмент участников фондового рынка не заинтересован продавать свои активы для компенсации собственных убытков. Это говорит о присутствии рационального подхода среди частных инвесторов, которые не реагируют импульсивно в случае недостатка собственных средств.

Макроэкономические оценки могут оказать влияние на биржевые характеристики популярных среди частных инвесторов акций. Для тестирования эффекта макроэкономических параметров используются регрессионные уравнения (41) и (42):

$$\begin{aligned} \mathrm{Re}\,turn_{i,t} &= \alpha + \beta_1 IMOEX_t + \beta_2 Sh_TV_{i,t} + \beta_3 Sh_\mathrm{Re}\,tail_t + \beta_4\,\mathrm{Re}\,tail_Asset_{i,t} + \\ &+ \beta_5 LnRVI_t + \beta_6 Delt_RUONIA_t + \beta_7 Deposits_t + \beta_8 Assets_t + \beta_k SENT_{i,t} + \varepsilon_{i,t} \end{aligned} \quad (41)$$

$$\begin{aligned} Sh_TV_{i,t} &= \alpha + \beta_1 IMOEX_t + \beta_2 Return_{i,t} + \beta_3 Sh_Retail_t + \beta_4 Retail_Asset_{i,t} + \\ &+ \beta_5 LnRVI_t + \beta_6 Delt_RUONIA_t + \beta_7 Deposits_t + \beta_8 Assets_i + \beta_k SENT_{i,t} + \varepsilon_{i,t} \end{aligned} \quad (42)$$

где $Deposits_t$ – изменение объема денежных средств на депозитах домохозяйств за месяц *t*, $Assets_t$ – динамика стоимости портфеля акций компаний, паев и акций инвестиционных фондов среди российских домохозяйств за месяц *t*.

Полученные результаты (Табл. 3.8) позволяют подтвердить ранее полученные выводы о неоднозначном влиянии доли объема торгов среди частных инвесторов по наиболее популярным у таких инвесторов акциям.

Таблица 3.8

Анализ влияния сентимента на доходность и объем торгов по популярным среди частных инвесторов акциям с учетом макроэкономических факторов

Переменная	*Зависимая переменная – $Return_{i,t}$*		*Переменная*	*Зависимая переменная – $Sh_TV_{i,t}$*	
	(1)	*(2)*		*(1)*	*(2)*
const	0,1120	0,1144	const	-0,1756	-0,1752
IMOEX	0,9277***	0,9270***	IMOEX	0,0352	0,0351
Sh_TV	-0,0229	-0,0229	Return	-0,2021	-0,2021
Sh_Retail	-0,2622	-0,2693	Sh_Retail	0,7651	0,7640
Retail_Asset	-0,0004	-0,0004	Retail_Asset	-0,0065***	-0,0065***
LnRVI	0,0007	0,0012	LnRVI	-0,0047	-0,0046
Delt_RUONIA	0,0010	0,0012	Delt_RUONIA	0,0063	0,0063
Deposits	0,0001	0,0001	Deposits	0,0001	0,0001
Assets	0,0001	0,0001	Assets	0,0001	0,0001
AG	0,5667	0,5951	AG	-2,5139**	-2,5096**
VA		0,1089	VA		0,0160
AIC	-629,95	-628,10	AIC	-94,72	-92,72

Переменная	*Зависимая переменная – $Return_{i,t}$*		*Переменная*	*Зависимая переменная – $Sh_TV_{i,t}$*	
	(1)	*(2)*		*(1)*	*(2)*
BIC	-594,89	-589,54	BIC	-59,67	-54,17
R^2	0,3807	0,3810	R^2	0,0467	0,0467
Obs	246	246	Obs	246	246

Введение дополнительных макроэкономических контрольных переменных не внесло особых изменений в объяснительную силу сентимент-метрик, однако сами макроэкономические регрессоры не оказывают значимого влияния на ценовую реакцию акций и относительный объем торгов по ним. Робастность моделей (38)-(39) подтверждается текущими результатами, что еще раз доказывает важность учета внимания частных инвесторов. Ликвидность значимо объясняется активностью непрофессиональных участников рынка.

Методология и результаты проверки гипотезы о влиянии метрик сентимента на биржевые характеристики акций с учетом динамики фондового рынка

Динамика рынка в целом может являться значимой в объяснении биржевых характеристик. Фондовый рынок может описываться двумя периодами - роста и падения, когда возникает положительное или отрицательное изменение общерыночного состояния. В качестве контрольной переменной мы выбираем дамми роста рынка (*Growth*), принимающую значение 1 в случае, если наблюдается положительное изменение индекса Мосбиржи, и 0 в ином случае. Следовательно, требуется протестировать модифицированные регрессионные модели, построенные на основе недельных данных, по изучению влияния сентимента частных инвесторов на ценовую реакцию акций российских компаний-эмитентов и объема торгов акциями (43), (44).

$$\begin{aligned} \mathrm{Re}\,turn_{i,t} &= \alpha + \beta_1 IMOEX_t + \beta_2 Sh_TV_{i,t} + \beta_3 LnRVI_t + \beta_4 Delt_RUONIA_t + \\ &+ \beta_5 Growth_t + \beta_k SENT_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{43}$$

$$\begin{aligned} Sh_TV_{i,t} &= \alpha + \beta_1 IMOEX_t + \beta_2 Return_{i,t} + \beta_3 LnRVI_t + \beta_4 Delt_RUONIA_t + \\ &+ \beta_5 Growth_t + \beta_k SENT_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{44}$$

Результаты тестирования моделей (42), (43) представлены в Приложении Б (Табл. Б.5). Новый регрессор был проверен на целесообразность в использовании в моделях при помощи теста Чоу, который не отверг гипотезу об использовании данного параметра на уровне p-value, равном 0,0001 и 0,0269 для случаев объяснения доходности и ликвидности, соответственно.

Табл. Б.5 (Приложение Б) подтверждает ранее полученные выводы по моделям (34) и (35). Индекс согласованности мнений с отрицательным знаком связан с доходностью акций эмитентов в периоды растущего рынка.

Дамми-переменная динамики рынка подтвердила свою значимость в регрессионном моделировании, что позволило увеличить коэффициент детерминации примерно на 1%. Разнонаправленное влияние индекса согласованности мнений частных инвесторов (*AG*) и относительного объема торгов требует изучения влияния этих факторов в разрезе двух состояний фондового рынка – падения и роста.

Изучение динамики объема торгов по акциям российских компаний эмитентов (Приложение Б, Табл. Б.6) показало отсутствие значимых различий в результатах с предыдущим этапом тестирования сентимент-метрик по объяснению различий в ликвидности (Табл. 3.6).

Разделим выборку исследования на две части: период роста и падения рынка. Это позволит понять, как настроение частных инвесторов на онлайн-платформе «Тинькофф Пульс» может объяснять ценовую реакцию и объем торгов акциями российских эмитентов в различных условиях: на растущем и падающем рынке. Мы строим модели (34) и (35) для двух состояний рынка.

Табл. Б.7 (Приложение Б) показывает результаты эмпирической проверки моделей (34) и (35) для растущего рынка. Индекс согласованности мнений значимо отрицательно влияет на доходность активов. Это может говорить о том, что увеличение неопределенности в отношении ряда активов ведет к падению цен на них в моменты общего роста фондового рынка. Инвесторы негативно оценивают те активы, по которым наблюдается расхождение мнений в моменты, когда рынок растет.

Что касается ликвидности акций эмитентов на растущем российском фондовом рынке, то чем выше разногласие среди частных инвесторов, тем ниже ликвидность (Приложение Б, Табл. Б.8). Неоднозначность мнений различных групп частных инвесторов по отдельным акциям компаний-эмитентов – сигнал для инвесторов о целесообразности «удерживать» актив, не совершать сделки.

На падающем фондовом рынке наблюдается ситуация, схожая с растущим рынком (Приложение Б, Табл. Б.9): чем ниже уровень согласованности мнений частных инвесторов, тем ниже доходность акций. При этом наблюдается отрицательное и значимое на 10% уровне влияние изменения ставки RUONIA на текущую недельную доходность акций российских компаний-эмитентов. Также наблюдается отрицательное влияние показателя изменения доли положительных сообщений на платформе «Тинькофф Пульс» (VA) на доходность акций. Это может говорить о том, что частные инвесторы не доверяют нарастанию оптимистичных мнений и ожидают, что оптимизм может смениться пессимизмом. Это открывает возможность для отдельных групп частных инвесторов не ждать и продавать свои активы на фондовом рынке, когда другие ожидают дальнейший рост одобрения других участников на рынке.

На падающем фондовом рынке (Приложение Б, Табл. Б.10), как и на растущем, мы отмечаем, что чем выше разногласие среди частных инвесторов, тем ниже ликвидность. Влияние фактора *AG* значимо на 1% уровне.

Наконец, протестируем модели (34) и (35) для двух различных подвыборок, которые отличаются по составу акций. Данный метод основывается на выявленных различиях в поведении частных инвесторов (Рис. 3.2) в зависимости от уровня «индекса страха» и степени согласованности мнений по каждому тикеру. Это позволит протестировать влияние дивергенции мнений розничных инвесторов на биржевые характеристики акций, по которым наблюдается положительная / отрицательная корреляция. В случае определения положительной корреляции подразумевается, что разногласие мнений частных инвесторов может быть минимальным в случае роста волатильности на фондовом рынке (все инвесторы придерживаются одной точки зрения в условиях повышенного риска). И наоборот, может фиксироваться рост разногласий непрофессиональных участников рынка в условиях, когда рынок находится в менее рискованном состоянии – снижение волатильности активов рынка. Однако может быть и ситуация отрицательной корреляции индекса согласованности мнений и «индекса страха». В этом случае увеличение уровня волатильности («индекса страха» RVI) может приводить к росту разногласий среди розничных инвесторов, а падение волатильности - сопровождаться ростом показателя согласованности мнений непрофессиональных участников рынка. Значения корреляции между показателями *AG* и *LnRVI* представлены в Приложении Б (Табл. Б.11).

Если корреляция *AG* и *LnRVI* положительна (Приложение Б, Табл. Б.11), то не наблюдается отличительных особенностей поведения сентимент-метрики AG в объяснении доходности акций эмитентов: влияние отрицательно и значимо на уровне 1%. При одновременном увеличении индекса RVI и индекса согласованности мнений частных инвесторов AG доходность рассматриваемых акций снижается. Мы объясняем это тем, что на рынке могут присутствовать различные группы инвесторов, некоторые из которых оценивают подобную ситуацию как сигнал к продаже активов во время роста страха (индекс RVI) и снижения разрыва мнений на онлайн-платформах.

В моделях, построенных с целью изучения ликвидности акций, статистическая значимость индекса согласованности AG в объяснении относительных объемов торгов акций снижается, но сохраняет отрицательный знак (Приложение Б, Табл. Б.11). При одновременном увеличении индекса RVI и индекса согласованности мнений инвесторов ликвидность на рынке снижается.

По выборке акций, доходность которых отрицательно коррелирована с относительным изменением индекса RVI (Приложение Б, Табл. Б.12), тоже наблюдается отрицательное и значимое на 5% уровне влияние индекса согласованности мнений частных инвесторов на доходность. Напротив, влияние индекса страха RVI положительно, но статистически не значимо. Таким образом, если возникает увеличение индекса RVI и меньший разрыв мнений, влияние на доходность акций не однозначно. В случае возникновения больших волнений на рынке и увеличения консенсуса среди частных

инвесторов, значимым фактором остается только индекс согласованности мнений.

Мы предполагаем, что отрицательная зависимость доходности акций и их ликвидности от уровня согласованности мнений частных инвесторов (Приложение Б, Табл. Б.12) может возникать по причине наличия участников рынка, которые могут зарабатывать на разности мнений.

Можно сделать вывод о том, что в случае возникновения роста / падения «индекса страха» RVI согласованность мнений частных инвесторов является важным фактором в определении движения доходности акций российских компаний-эмитентов. Во всех случаях индекс согласованности AG отрицательно объясняет недельную доходность рассматриваемых активов. Предположение такого поведения цены может описываться наличием большего количества отрицательных сообщений на онлайн-платформах или присутствием групп инвесторов, которые идут против рынка, когда мнение других участников становится однозначным.

3.4. АНАЛИЗ ВЛИЯНИЯ СЕНТИМЕНТА ЧАСТНЫХ ИНВЕСТОРОВ НА ВОЛАТИЛЬНОСТЬ ДОХОДНОСТИ АКЦИЙ РОССИЙСКИХ КОМПАНИЙ

Начиная с конца XX века, проявился тренд к усилению глобализации (Wei et al., 2022), увеличился интерес инвесторов к развивающимся рынкам (Muguto & Muzindutsi, 2022). В последние годы, в связи с пандемией COVID-19 и введением санкций в отношении ряда развивающихся стран (Россия, Беларусь, Иран) фиксируется и противоположный тренд к изоляции и протекционизму. Исследование детерминант волатильности доходности на развивающихся рынках представляет научный и практический интерес: повышение волатильности сопряжено с повышенным риском инвестирования и может привести к процессам «бегства в качество» на развитые рынки капитала (Talwar et al., 2021). Волатильность является одним из ключевых показателей, принимаемых инвесторами во внимание при формировании портфелей на фондовом рынке (Muguto & Muzindutsi, 2022), и оказывает существенное влияние на будущую ценовую динамику рынка под воздействием шоков (Nguyen & Nguyen, 2019).

В работе Chousa et al. (2016) показано, что инвесторы обращают внимание не только на фундаментальные показатели, в т.ч. рыночные мультипликаторы, но и на настроения других инвесторов. Открытым исследовательским вопросом является влияние сентимента участников фондового рынка на биржевые характеристики активов, в т.ч. волатильность доходности, на развивающихся рынках (Fariska et al., 2021; Behrendt & Schmidt, 2018). В первых работах по данной теме (Barber & Odean, 2008; Tetlock, 2007) в качестве прокси сентимента использовались косвенные метрики – объем торгов и аномальная доходность, а также количество упоминаний компаний-эмитентов акций в новостной ленте и тональность аналитических колонок в профильных журналах. При использовании косвенных метрик, таких как объем торгов, возникает проблема потенциальной эндоген-

ности. Анализ новостной ленты также имеет недостатки: во-первых, подписчики деловых изданий не обязательно являются инвесторами, во-вторых, неизвестно, были ли прочитаны инвесторами заголовки и тексты статей. Начиная с 2010-х гг., появляются исследования, в которых сентимент выявляется на основе социальных сетей, например, Twitter (Oliveira et al., 2013; Oliveira et al., 2017; Al-Nasseri & Ali, 2018). Преимуществом таких работ является построение индексов и метрик сентимента по обсуждениям компаний частными инвесторами. Все участники обсуждений совершают сделки и открыто выражают свое мнение о перспективах роста и снижения цен акций. В ряде работ представлены доказательства взаимосвязи между настроениями инвесторов, измеряемыми на основе поисковых запросов в Интернете, и волатильностью доходности акций во время пандемии коронавируса (Tripathi & Pandey, 2021).

Мы отмечаем, что предыдущие работы преимущественно фокусируются на влиянии сентимента инвесторов на доходность акций и на рынке США (Oliveira et al., 2013; Oliveira et al., 2017; Al-Nasseri & Ali, 2018), тогда как в нашей работе впервые исследовано влияние сентимента на волатильность доходности на развивающемся российском рынке акций. Актуальность нашего исследования обусловлена тем, что в России стремительно растет количество частных инвесторов. Из Рис. В.1 Приложения В видно, что количество активных счетов частных инвесторов с января 2014 г. по ноябрь 2019 г. выросло в 7,5 раз. В то же время, этот рост не привел к открытию новых счетов юридическими лицами (Рис. В.2). В-третьих, в России активно развиваются social-trading платформы.

В качестве источников сообщений инвесторов мы рассматриваем новый популярный мессенджер Telegram, включающий каналы и чаты с инвестиционной тематикой, и старейший российский профессиональный биржевой форум mfd.ru. Выбранные нами платформы Telegram и mfd.ru охватывают разные возрастные группы инвесторов, что позволяет сопоставить влияние их сентимента на биржевые характеристики акций и выявить отличия. Важное нововведение нашей работы - уникальные данные и высокий охват большей части сентимента частных инвесторов. Количество пользователей форума MFD и отслеживаемых Telegram-каналов в сумме составляет 50% от активных счетов на Московской бирже (Теплова и др., 2022).

Еще одно преимущество нашей работы – учет сленга и неформальных выражений инвесторов, широко применяемых в социальных сетях, что привело к необходимости отказа от готовых словарей и пакетов[1]. Нами вручную были размечены свыше 30 тысяч сообщений инвесторов, после чего для классификации сообщений по тональности использовались и сопоставлялись по метрикам качества методы машинного обучения (дерево решений, случайный лес) и модель на основе трансформеров и CNN, предложенная Yang et al. (2019).

[1] Нами создан авторский словарь, учитывающий сленг частных инвесторов, патент 2020622801. URL: https://www1.fips.ru/registers-doc-view/fips_servlet?DB=DB&DocNumber=2020622801&TypeFile=html

3.4.1. Обзор предыдущих работ

Подавляющее большинство работ о влиянии сентимента на биржевые характеристики акций посвящено рынку США (Tetlock, 2007; Oliveira et al., 2013; Al-Nasseri & Ali, 2018; Huang et al. 2019), тогда как акции компаний развивающихся рынков гораздо реже являются объектом исследования (Renault, 2017). В нашей работе впервые исследовано влияние сентимента на российском рынке акций. Во многих работах тестируются только модели влияния сентимента на доходность акций (Baker & Wurgler, 2006; Renault, 2017; Huang et al., 2019), тогда как мы рассматриваем в качестве зависимой переменной волатильность доходности. Поскольку ряд эмпирических исследований показывает, что сентимент инвесторов может объяснять доходность акций (Renault, 2017; Huang et al., 2019), можно предположить, что сентимент может объяснять и второй момент цены – волатильность (Gong et al., 2022).

Gong et al. (2022) приводят два теоретических обоснования того, что на основе сентимента инвесторов можно прогнозировать волатильность доходности акций. Во-первых, сделки, совершаемые иррациональными инвесторами, приводят к аномалиям ценообразования вопреки действиям рациональных инвесторов (арбитражеров). На рынках широко распространенным является эффект толпы, когда одни иррациональные инвесторы повторяют действия других. Во-вторых, существуют эффекты как недостаточной, так и чрезмерной реакции цен акций на новости. Недостаточная реакция обусловлена консерватизмом: частные инвесторы не меняют своих убеждений относительно перспектив изменений цен финансовых активов даже при появлении новых фактов. Чрезмерная реакция обусловлена тем, что частные инвесторы находят мнимые закономерности в случайных последовательностях. Это свидетельствует о неэффективности рынка: цены на акции могут не отражать полную информацию. Таким образом, настроения инвесторов можно рассматривать как фактор риска на финансовом рынке.

Gong et al. (2022) строят авторский композитный индекс сентимента NISI по рынку акций в целом на основе таких базисных метрик как: дисконт / премия закрытых инвестиционных фондов, оборот фондового рынка, количество IPO, средняя доходность IPO в первый день, количество новых счетов инвесторов, индекс потребительского доверия, разность количества компаний с растущими в цене акциями и количества компаний с падающими в цене акциями, отношение количества компаний с растущими в цене акциями к количеству компаний с падающими в цене акциями. Построение индекса NISI осуществляется с применением метода PLS. Авторы анализируют рынок Китая за 2003-2021 гг. Разработанный авторами композитный индекс NISI сопоставляется с другими индексами сентимента, предложенными в предыдущих работах (Табл. 3.9), по качеству прогнозирования реализованной волатильности. Авторы приходят к выводу, что для краткосрочного временного горизонта (один месяц) NISI - единственная метрика

сентимента инвесторов, которая хорошо работает как в кризисные периоды, так и в периоды стабильности. Композитный индикатор сентимента на фондовом рынке Китая, предложенный Yi and Mao (2009), и индекс сентимента инвесторов, предложенный Wei et al. (2014), не являются значимыми предикторами волатильности для периода 2003-2021 гг. в целом, но они обладают значимой прогностической способностью в отдельные подпериоды. Для более долгосрочных периодов (3, 6 и 12 месяцев) авторский индикатор NISI, процентное отношение цены открытия к цене закрытия предыдущего дня по индексу SSEC и индикатор технического анализа, предложенный Liu and Pan (2020), обладают высокой прогностической способностью. При этом уровень значимости NISI наибольший среди рассматриваемых метрик.

Hsu and Tang (2022) выявляют влияние настроений инвесторов на неожиданную условную волатильность фондовых рынков под контролем качества государственного управления. Неожиданная условная волатильность определяется как разница между фактической волатильностью и ожидаемой волатильностью, если бы пандемии Covid-19 не произошло. Для оценки ожидаемой волатильности применяются модели GARCH(1,1), EGARCH(1,1) и GJR-GARCH(1,1). Настроения инвесторов фиксируются по поисковым запросам в сети Интернет по тематике пандемии COVID-19. Авторы подчеркивают, что Интернет представляет собой жизненно важный канал для получения информации, что особенно актуально во время пандемии COVID-19, когда многие люди были вынуждены оставаться дома. В выборку исследования входит 12 фондовых рынков: 7 развитых и 5 развивающихся, в т.ч. РФ, Китай, Индия. Исследование проводится на уровне рынка в целом, без рассмотрения динамики цен акций отдельных компаний.

Hsu and Tang (2022) выявили положительную корреляцию между индексом количества поисковых запросов в Google по тематике COVID-19 и неожиданной волатильностью. Более пессимистические настроения инвесторов, связанные с COVID-19, во время пандемии приводят к более высокой неопределенности на фондовом рынке. Спецификой китайского рынка является то, что среднее значение неожиданной волатильности отрицательно. Авторы объясняют это тем, что индивидуальные инвесторы, доминирующие на китайском рынке, склонны больше внимания уделять хорошим новостям и меньше реагировать на плохие новости, что приводит к положительной зависимости между доходностью акций и волатильностью. Это отличает Китай от большинства фондовых рынков развитых стран, где основными инвесторами являются институциональные инвесторы, и часто наблюдается отрицательная связь между доходностью акций и волатильностью. Также авторы получили результат о том, что неожиданная условная волатильность ниже, когда выше качество государственного управления, метриками которого являются: защита прав миноритарных акционеров, эффективность государственного управления, защита прав инвесторов.

Fariska et al. (2021) констатируют увеличение роли социальных сетей как источника информации в период пандемии COVID-19 и ставят целью выявить объясняющую роль настроений инвесторов в микроблогах на биржевые характеристики акций компаний Индонезии. Индивидуальные инвесторы имеют ограниченный доступ к информации о компаниях, а социальные сети могут стать для них источниками информации. Определение тональности сообщений инвесторов в сети Twitter осуществляется с применением интеллектуального анализа текста и наивной байесовской классификации сообщений на положительные, отрицательные и нейтральные. Период анализа – 2019-2020 гг., в выборку исследования вошли сообщения по 68 компаниям. Для выявления краткосрочных и долгосрочных взаимосвязей между сентиментом инвесторов и волатильностью авторы применяют модели векторной авторегрессии, функции импульсного отклика и тест Грейнджера на причинность.

Fariska et al. (2021) выявили, что настроения инвесторов в микроблогах оказывают значительное влияние на доходность и волатильность акций. Показано, что настроения инвесторов с лагом в 1 и 2 дня оказывают значимое влияние на волатильность доходности акций, причем связь отрицательная. Выявлена и обратная связь: волатильность с лагом 1 и 2 дня положительно влияет на сентимент. Таким образом, связь между волатильностью и сентиментом является двунаправленной. Связь между сентиментом и доходностью акций является также двунаправленной, но при этом положительной. Реакция на шоки является конвергентной: инвесторы, ведущие микроблоги, сами осуществляют мониторинг новостей. Авторы отмечают консерватизм частных инвесторов: они медленно меняют свои взгляды и предпочтения при появлении новой информации.

Behrendt and Schmidt (2018) исследовали внутридневную динамику волатильности доходности акций под влиянием настроений инвесторов и их активности в социальной сети Twitter. В выборку включены компании, входившие в промышленный индекс Доу Джонса в период 2015-2017 гг. С применением двумерной модели VAR выявлено статистически значимое влияние сентимента инвесторов и активности обсуждений по компании в сети Twitter на внутридневную волатильность акций данной компании. Но авторы отмечают, что коэффициенты в регрессионной модели имеют малую величину, т.е. влияние на волатильность невелико. Авторы приходят к выводу, что и сентимент инвесторов, и активность обсуждения по компании являются причинами по Грейнджеру для волатильности. В то же время, результаты декомпозиции дисперсии ошибки прогноза показывают, что включение экзогенной информации на базе Twitter в модели прогнозирования не влияет существенно на эффективность прогноза. Наконец, авторы строят HAR-модель по панельным данным с включением лагированных данных по настроениям инвесторов и активности обсуждений, и эффективность прогнозов вне выборки не улучшается. Авторы заключают, что высокочастотная информация в сети Twitter не представляет большой практической значимости для инвесторов в целях прогнозирования волатильности

по «голубым фишкам», но для компаний, характеризующихся меньшей ликвидностью акций, такая информация может представлять интерес.

Chen et al. (2020) исследуют динамическую взаимосвязь между настроениями инвесторов и реализованной волатильностью фондового рынка Китая с применением метода thermal optimal path. Авторы строят интегральный индекс сентимента посредством метода главных компонент на основе статистических метрик, отражающих динамику фондового рынка (количество новых брокерских счетов и т.д.), и количества поисковых запросов по темам, связанным с фондовым рынком. Авторы получили вывод о том, что когда фондовый индекс сильно колеблется, настроения инвесторов опережают реализованную волатильность с лагом 1-2 месяца. Так, всплеск оптимизма инвесторов указывает на шок фондового рынка в следующем месяце. Наконец, прогностическая сила авторского индекса сентимента инвесторов проверяется с использованием метрики RMSE и коэффициента неравенства Theil. Предложенный авторами индекс сентимента, учитывающий количество запросов по темам, связанным с фондовом рынком, позволяет получить лучшие метрики качества прогнозных моделей, чем ряд других экзогенных переменных.

Большинство авторов рассматривают только один источник сообщений для построения прокси переменных сентимента, не деля источники на классы - форумы, социальные сети, мессенджеры (Tetlock, 2007; Oliveira et al., 2013; Renault, 2017; Al-Nasseri & Ali, 2018). В русле работы Тепловой и др. (2022), мы рассматриваем и сопоставляем объясняющую силу индексов сентимента по двум источникам: новому популярному мессенджеру Telegram и одному из старейших биржевых форумов mfd.ru.

Мы отмечаем, что большинство работ, посвященных анализу влияния сентимента инвесторов на волатильность, фокусируются на рынке в целом и не рассматривают отдельные компании (Chen et al., 2020; Hsu & Tang, 2022; Gong et al., 2022). В отличие от них, мы исследуем влияние сентимента на волатильность доходности акций отдельных компаний.

Одним из недостатков многих предыдущих работ по рынку США мы видим то, что выводы строятся на основе выборки данных, которая не охватывает большую часть сентимента инвесторов (Oliveira et al., 2017; Huang et al., 2019; Behrendt & Schmidt, 2018). Обсуждение крупнейшего развитого рынка ведется инвесторами из различных стран, на разных языках. Наша работа посвящена развивающемуся российскому рынку, охват сентимента частных инвесторов составляет более 50%.

В русле современных работ Renault (2017), Al-Nasseri and Ali (2018), Теплова и др. (2022), мы используем методы машинного обучения для классификации сообщений инвесторов по тональности. Для проверки гипотез о влиянии сентимента инвесторов на волатильность доходности акций мы тестируем регрессионные модели с учетом эффектов нелинейности. В большинстве предыдущих работ для тестирования также применялись регрессионные модели, но эффекты нелинейности не учитывались (Табл. 3.9).

Таблица 3.9

Обзор предыдущих работ по теме исследования

Работа	*Прокси-переменные сентимента и внимания инвесторов*	*Рынок, объект исследования, период тестирования*	*Методы анализа тональности и построения индекса сентимента / внимания инвесторов*	*Методы тестирования гипотез*	*Влияние сентимента / внимания инвесторов на волатильность*
Gong et al. (2022)	Новый композитный авторский индикатор сентимента (New investor sentiment index, NISI, рассчитанный на основе базовых метрик), композитный индикатор сентимента на фондовом рынке Китая CICSI (предложенный Yi & Mao, 2009), индекс сентимента инвесторов ISI (предложенный Wei et al., 2014), отношение количества растущих и падающих в цене акций к отношению объема растущих и падающих в цене акций (ARMS индикатор), процентное отношение цены открытия к цене закрытия предыдущего дня по индексу	Китай, индекс Shanghai Stock Exchange Composite Index (SSEC), 2003-2021 гг.	Partial Least Squares (PLS) для расчета авторского индекса NISI на основе базовых метрик сентимента	ARDL модель	Авторский индекс NISI обладает большей прогнозной силой по сравнению с рядом других индексов сентимента. NISI обладает высокой прогностической силой в некризисные периоды, тогда как альтернативные индикаторы CICSI и ISI хорошо работают только в кризисные периоды

Работа	*Прокси-переменные сентимента и внимания инвесторов*	*Рынок, объект исследования, период тестирования*	*Методы анализа тональности и построения индекса сентимента / внимания инвесторов*	*Методы тестирования гипотез*	*Влияние сентимента / внимания инвесторов на волатильность*
	SSEC (Aboody et al., 2018), индекс неопределенности экономической политики (Huang & Luk, 2020), индикатор технического анализа, предложенный Liu and Pan (2020), объем торгов				
Hsu and Tang (2022)	Количество поисковых запросов в Google по тематике пандемии COVID-19	12 рынков (7 развитых и 5 развивающихся), фондовые индексы, 2010-2020 гг.	Тональность не выявлялась, фиксировалось количество запросов в Google	Многоуровневая модель со смешанными эффектами (multilevel mixed effect model)	Количество запросов по тематике COVID-19 значимо положительно влияет на неожиданную волатильность на фондовом рынке
Fariska et al. (2021)	Индекс сентимента по заданной компании, рассчитанный на основе доли позитивных, нейтральных и негативных сообщений	Индонезия, акции 68 компаний, 2019-2020 гг.	Наивный байесовский классификатор для разделения сообщений на три класса по тональности	VAR модель, тест Грейнджера на причинность	Связь между сентиментом инвесторов и волатильностью является двунаправленной, при этом сентимент

Работа	*Прокси-переменные сентимента и внимания инвесторов*	*Рынок, объект исследования, период тестирования*	*Методы анализа тональности и построения индекса сентимента / внимания инвесторов*	*Методы тестирования гипотез*	*Влияние сентимента / внимания инвесторов на волатильность*
					значимо отрицательно влияет на волатильность с лагом 1 и 2 дня
Behrendt and Schmidt (2018)	Метрики, рассчитываемые Bloomberg: количество сообщений и сентимент инвесторов в сети Twitter по выбранной компании за указанный период времени	США, акции компаний, входящих в индекс DIJA, 2015-2017 гг., внутридневные данные	Алгоритм Bloomberg	Модель векторной авторегрессии (VAR), модель гетероскедастической авторегрессии (HAR) по панельным данным, тест Грейнджера	Настроения инвесторов и их активность в сети Twitter значимо влияют на волатильность доходности акций, но размер этого эффекта не очень велик. Чем меньше объем торгов по акциям компании, тем больше влияние активности обсуждений и сентимента на волатильность

Работа	*Прокси-переменные сентимента и внимания инвесторов*	*Рынок, объект исследования, период тестирования*	*Методы анализа тональности и построения индекса сентимента / внимания инвесторов*	*Методы тестирования гипотез*	*Влияние сентимента / внимания инвесторов на волатильность*
Chen et al. (2020)	Количество новых брокерских счетов, маржинальный баланс, коэффициент оборачиваемости, чистый приток капитала инвесторов в акции, количество поисковых запросов по фондовому рынку в сервисе Baidu	Китай, индекс фондового рынка, 2013-2019 гг.	Метод главных компонент для расчета авторского индекса на основе базовых метрик сентимента и внимания инвесторов	Thermal optimal path (TOP) method	Если рынок нестабилен и фондовый индекс достигает высокого уровня, сентимент инвесторов опережает реализованную волатильность. Если же рынок движется в горизонтальном коридоре, сентимент инвесторов следует за реализованной волатильностью с запаздыванием
Oliveira et al. (2017)	Тональность и количество сообщений в Twitter, индикаторы тональности обзоров в Datastream	США, индексы широкого рынка и отдельных секторов, портфели акций,	Применение словарей для анализа тональности текста, фильтр Калмана	Регрессионный анализ и методы машинного обучения	Сентимент частных инвесторов позволяет

Работа	*Прокси-переменные сентимента и внимания инвесторов*	*Рынок, объект исследования, период тестирования*	*Методы анализа тональности и построения индекса сентимента / внимания инвесторов*	*Методы тестирования гипотез*	*Влияние сентимента / внимания инвесторов на волатильность*
		факторы модели Fama-French, 2012-2015 гг.			спрогнозировать волатильность доходности индекса DIJA

3.4.2. Гипотезы

Основываясь на предыдущих работах, мы выдвигаем следующие гипотезы:

Гипотеза 1. Сентимент частных инвесторов значимо и нелинейно влияет на волатильность доходности акций.

Мы рассматриваем пессимистические настроения инвесторов как фактор риска и предполагаем, что чем сильнее негативная тональность сообщений инвесторов, тем выше волатильность. Maitra and Dash (2017) по рынку Индии показали, что чем более позитивны настроения инвесторов, тем ниже волатильность. Ho et al. (2013) также выявили отрицательную корреляцию между настроениями инвесторами в Twitter и индексом волатильности VIX по рынку США. Мы также предполагаем, что чем интенсивнее идут обсуждения компаний в социальных сетях, тем выше волатильность доходности. Antweiler and Frank (2004) по рынку США выявили, что увеличение количества обсуждений на форумах по 45 компаниям, входившим в индекс Dow Jones Industrial Average, приводит к росту волатильности по этим акциям. Audrino et al. (2020) и Behrendt and Schmidt (2018) выявили значимое влияние метрик сентимента и активности обсуждений в социальных сетях на волатильность на рынке акций США. Поскольку наш индикатор Hype является интегральной мерой позитивной тональности сообщений и интенсивности обсуждений, мы предполагаем нелинейное значимое влияние данного фактора на волатильность.

Гипотеза 2. Пессимистические настроения инвесторов сильнее влияют на волатильность доходности акций, чем оптимистические.

Как показано Caporale et al. (2016) по рынкам восьми развитых стран, негативные новости оказывают большее влияние на доходность акций, чем позитивные. В работе Phuong et al. (2015) аналогичный эффект выявлен для волатильности доходности акций на вьетнамском фондовом рынке. В работе Chen et al. (2020) по рынку Китая показано, что влияние сентимента на волатильность зависит от текущей динамики фондового рынка. В связи с этим, мы предполагаем асимметрию влияния сентимента на волатильность доходности акций на российском рынке.

Гипотеза 3. Количество поисковых запросов в сети Интернет значимо и нелинейно влияет на волатильность доходности акций.

В работе Tripathi and Pandey (2021) выявлено значимое влияние настроений инвесторов, измеряемых на основе поисковых запросов в Интернете, на волатильность доходности акций во время пандемии коронавируса. Мы предполагаем, что если количество запросов относительно мало или, напротив, велико, то волатильность доходности снижается. Повышенный интерес к компании свидетельствует о негативных или позитивных событиях, в результате которых на рынке формируется устойчивый тренд, и волатильность доходности при этом снижается. Пониженный интерес (малое количество поисковых запросов) свидетельствует о том, что значимых со-

бытий не происходит, нет предпосылок к существенным колебаниям цен, и они остаются относительно стабильными.

3.4.3. Методология

Применение модифицированной нейросети, подготовка и разметка данных

Данная часть исследования выполнена в рамках работы Тепловой и др. (2022). Для построения индекса сентимента требовалось обрабатывать большие массивы сообщений (высказываний) в профильных социальных сетях, посвященных акциям и инвестированию в них на российском рынке. Мы отмечаем, что участники социальных сетей в переписке выражают свои мысли самыми разными способами, переделывая названия компаний и глаголы, используя жаргонизмы и искаженные английские термины, эмодзи. В связи с этим, широко применяемый метод выявления тональности слов мы сочли для нашей работы неприемлемым. Оригинальность нашего исследования состоит в том, что единицей обработки в нейронной сети является не слово, а сообщение в целом. Проблему структурирования неструктурированной информации (текста) мы решаем с применением методов искусственного интеллекта.

На этапе предварительной обработки данных мы применили процедуры лемматизации и стемминга для приведения слов к нормальной форме и удаления окончаний и стоп-слов. В результате такой обработки распределение количества повторов слов стало более равномерным, уменьшилась доля выбросов. Далее нами был составлен оригинальный русскоязычный словарь с официальными (включая тикеры) и неформальными названиями компаний. В качестве последних на форумах часто используются жаргонизмы, фамилии CEO и основных владельцев, названия материнских компаний, названия основных продуктов и услуг.

Следующим этапом для построения моделей машинного обучения является преобразование сообщений в числовые векторы. Для этого применим алгоритмы OneHotEncoding и TF-IDF Encoding. OneHotEncoding создает бинарные переменные в количестве, равном количеству слов в выборке данных, что позволяет каждое сообщение представить в виде вектора из нулей и единиц. TF-IDF Encoding сопоставляет сообщению вектор, элементы которого характеризуют частоту употребления каждого слова в данном сообщении и во всей выборке сообщений (Sparck, 1972). Это позволяет уменьшить значимость слов, не несущих практического смысла.

Для решения поставленной задачи использовался метод нейронных сетей. Опишем процесс присвоения сообщению вероятности принадлежности соответствующему классу сентимента. Мы различаем три класса эмоциональных окрасок: негативные (для выделения медвежьего тренда и рекомендаций продавать), нейтральные (для выделения сообщений, из которых непонятно, как поведут себя участники обсуждения) и позитивные (для вы-

деления прогноза участников на рост цены и рекомендации, зачастую завуалированной, к открытию позиции на покупку).

Класс задач, в котором первичный набор данных содержит обучающую выборку (наши сообщения) и метки классов (соответствующую сообщениям тональность), называется «обучение с учителем». Для решения проблемы классификации нами была выбрана модель, описанная в статье Wang et al. (2018) и представляющая собой нейронную сеть, которая находит закономерности между словами, имея информацию об априорном семантическом взаимоотношении слов.

Чтобы оценить преимущества нейронной сети (модели Wang et al., 2018), мы реализовали на нашей выборке сообщений стандартные методы машинного обучения в качестве бенчмарков. Цель такого сопоставления - оценить, насколько собранные и вручную размеченные данные информативны. Результаты показаны в Табл. 3.10.

Таблица 3.10

Сопоставление результатов обработки текста с применением различных моделей ИИ[1]

Метод	*One-HotEncoding*						*TF-IDF*					
	Acc (Train)	*Acc (Test)*	*F1 (Train)*	*F1 (Test)*	*F1 weigh*	*F1 weigh*	*Acc (Train)*	*Acc (Test)*	*F1 (Train)*	*F1 (Test)*	*F1 weigh*	*F1 weigh*
SGD	0,81	0,73	0,71	0,56	0,81	0,73	0,77	0,73	0,63	0,56	0,78	0,73
Случайный лес	0,98	0,74	0,98	0,52	0,98	0,76	0,99	0,74	0,98	0,51	0,99	0,78
Дерево решений	0,79	0,55	0,74	0,46	0,77	0,51	0,89	0,61	0,86	0,47	0,89	0,59
Константная модель	0,75	0,61	0,29	0,25	0,64	0,46						

В Табл. 3.10 показаны результаты по тестированию моделей на обучающей и тестируемой выборках (train и test, соответственно) по двум типам метрик: Accuracy и F1-мера. Accuracy представляет собой показатель средней точности распознавания класса и используется для качественной интерпретации результатов, демонстрируя вероятность угадывания правильного сентимента в соответствующей выборке. Однако, сами модели отбираются по другой метрике - F1, которая учитывает качественные особенности набора данных, влияющие на процесс обучения, включая проблему несбалансированности классов.

Из Табл. 3.10 следует, что особенности наших данных позволяют обучить модель, предсказывающую верный класс сообщения, в 72,8% случаях.

[1] Расчеты проведены Д.В. Бучко под руководством Т.В. Тепловой.

F1-мера определила нижнюю границу эффективности для обучаемой модели (SGD + OHE). Как следствие, был сделан следующий вывод: тестирующий набор данных может позволить моделировать связь между сообщениями и их сентиментом с точностью не меньше, чем 72,8% и F1-мерой, равной 0,564.

Далее для анализа данных мы применили модель на основе трансформеров и CNN, предложенную (Yang et al., 2019). Так как One-Hot Encoding может привести к переобучению и не очень хорошо справляется с неформальной, часто меняющейся лексикой, в итоговой модели мы используем процедуру CNN Embedding.

На втором шаге после CNN Embedding мы применили полносвязную нейронную сеть с тремя слоями. В первом слое - 128 нейронов, во втором - 64. В обоих слоях в качестве функции активации используется гиперболический тангенс. Также используется dropout (исключение определенного процента случайных нейронов) с параметром 0,5 после каждого слоя для предотвращения переобучения. Последний слой состоит из трех нейронов и имеет функцию активации softmax. Модель обучалась 20 эпох с коэффициентом скорости обучения (learning rate) 0,0015. CNN-based-модель на тестовой выборке показала F1-меру 57,7% и точность (Accuracy), равную 69,4%.

Построение метрик сентимента частных инвесторов

В нашей работе анализируется широкий спектр метрик сентимента частных инвесторов (Табл. 3.11), выбор метрик основан на работе Тепловой и др. (2022). Авторские метрики, рассчитанные с применением методов машинного обучения для классификации сообщений, сопоставляются с дополнительными метриками.

Разбив все сообщения пользователей на классы с позитивной и негативной тональностью, мы строим индикатор Hype[1] (Табл. 3.11). Он учитывает как долю позитивных сообщений, так и рост активности обсуждений инвесторов по акциям компании. Рассчитывается по формуле (45):

$$HYPE_T = \frac{MSG_T^{POSITIVE}}{MSG_T^{ALL}} \cdot \left(\frac{MSG_T^{ALL}}{MSG_{T-1}^{ALL}} - 1 \right) \tag{45}$$

где $MSG_T^{POSITIVE}$ - количество позитивных сообщений по каждой компании, MSG_T^{ALL} – суммарное количество позитивных и негативных сообщений по каждой компании. Предложенный Hype индикатор принимает значения от 0 (все сообщения негативные, или количество сообщений снизилось / не изменилось) до 100 (все сообщения позитивные, и количество сообщение выросло в 10 и более раз).

Наряду с индикатором Hype мы рассматриваем дополнительные метрики сентимента, традиционно применяемые в научных работах (Табл. 3.11).

[1] Индикатор Hype разработан А.Ф. Томтосовым под руководством Т.В. Тепловой.

Таблица 3.11

Прокси-переменные сентимента частных инвесторов

Прокси сентимента	*Описание*	*Ожидаемое направление влияния на волатильность*
Авторские метрики сентимента, построенные на основе разметки тональности с применением нейросети		
Mes_pos	Количество сообщений позитивной тональности по акциям данной компании	-
Mes_neutr	Количество сообщений нейтральной тональности по акциям данной компании	+ / -
Mes_neg	Количество сообщений негативной тональности по акциям данной компании	+
Hype	Авторский индикатор, одновременно учитывающий и рост количества сообщений позитивной тональности, и рост общества количества сообщений. Рассчитывается по формуле (44)	+ / -
Дополнительные прокси сентимента		
Yandex_abs	Количество запросов в Яндексе Метрике, рассчитывается для каждой акции отдельно на основе запросов в виде «*Название компании* + акции»	+
Yandex_relat	Относительное количество запросов в Яндекс Метрике, определяемое как отношение абсолютного количества запросов по компании к совокупному количеству поисковых запросов	+
Num_mes	Количество сообщений по каждой компании в разрезе источников MFD и Telegram без разметки тональности сообщений	+
TR_Indicator	Индикатор Refinitiv Credit Text Mining News Component, оценивающий кредитный риск компании на основе публикаций, брокерских отчетов, записей конференц-звонков с аналитиками о компании. Измеряется по шкале от 1 (максимальный риск) до 100 (минимальный риск)	+

Источник: Теплова и др. (2022)

Выбор контрольных переменных

Мы рассматриваем контрольные переменные, представленные в Табл. 3.12. Выбор переменных основан на работе Тепловой и др. (2022).

Таблица 3.12

Контрольные переменные

Контрольная переменная	*Описание*	*Ожидаемое направление влияния на волатиль-ность*	*Научные работы о влиянии показателя на биржевые характеристики акций*
Показатели ликвидности акций			
TradeVolFF	Объем торгов с учетом free-float, %	+	Brown et al. (2009)
TradeVolFF_2	Квадрат переменной TradeVolFF	+	
FreeFloat	Доля акций в свободном обращении, %	+	
NzTradeDaysShare	Доля торговых дней с ненулевым объемом торгов за месяц, %	+	
Финансовые показатели эмитента			
ROA	Отдача на активы эмитента	-	Анкудинов и Лебедев (2014); Nurazi et al. (2015)
Leverage	Леверидж, рассчитанный как отношение активов к собственному капиталу	+	
EBITDAtoAssets	Отношение EBITDA к активам	+	
NetDebttoEBITDA	Коэффициент чистого долга, рассчитывается как отношение Net Debt к EBITDA	+	
EBITtoInterest	Коэффициент покрытия EBIT процентных расходов	+	
Показатели корпоративного управления			
BoardSize	Количество членов Совета директоров	-	Andreou et al. (2016); Rostami et al. (2016); Ali et al. (2017)
IndepDirPercent	Доля независимых директоров в Совете	-	
Макроэкономические показатели			
GDPgrowth	Темп роста ВВП в РФ, квартал-к-кварталу, %	-	Celebi and Honig (2019); Salamat et al. (2016); Debata and Mahakud
RVI_change	Изменение индекса волатильности RVI на	+	

Контрольная переменная	*Описание*	*Ожидаемое направление влияния на волатиль-ность*	*Научные работы о влиянии показателя на биржевые характеристики акций*
	российском рынке за месяц, %		(2018); Khan and Rehman (2020)
Gov10Y	Ставка доходности к погашению по 10-летним ОФЗ, %	+	
Gov10Y_2	Квадрат переменной Gov10Y	+	
PMI_US	Индекс деловой активности PMI в сфере услуг в США на начало месяца	-	
d14, d15, … , d18	Дамми-переменные временных периодов: 2014, 2015, …, 2018 г. Соответственно	+ / -	

Тестирование эконометрических моделей

Мы тестируем следующие эконометрические модели со случайными эффектами (46):

$$y_{it} = \beta_1 \cdot Sentiment_{it} + \beta_2 \cdot Sentiment_{it}^2 + \sum \beta_{3,k} \cdot x_{it,k} + u_{it} \qquad (46)$$

где i – номер компании, t – месяц наблюдения, y_{it} - зависимая переменная, $Sentiment_{it}$ - одна из метрик сентимента (включаются в модели по одной, Табл. 3.11), $x_{it,k}$ - контрольные переменные (Табл. 3.12), $\beta_1, \beta_2, \beta_{3,k}$ - коэффициенты регрессии, $u_{it,k}$ - индивидуальные эффекты.

Зависимые переменные и показатели сентимента берутся в помесячной разбивке, как и ряд макроэкономических показателей (например, доходность к погашению по 10-летним ОФЗ (среднее значение по месяцу) или индекс деловой активности) и показателей корпоративного управления. Поквартально взяты некоторые контрольные переменные: финансовые показатели компаний–эмитентов акций и годовой темп роста ВВП. Переменные включаются в модели с учетом проблемы мультиколлинеарности.

3.4.4. Данные

Выборка нашего исследования основана на работе Тепловой и др. (2022) и состоит из 60 компаний, временной период с 1 января 2014 по 31 марта 2020 г. Выборка включает: а) 20 компаний высоколиквидных «голубых

фишек», входящих в индекс Мосбиржи и имеющих кросс-листинг на биржах LSE / NYSE / Франкфурт, среднедневной объем торгов выше 100 млн руб.; б) 20 компаний с акциями средней ликвидности: среднедневной объем торгов от 10 до 100 млн руб., все компании находятся в отраслевых индексах и хоть один раз были включены в индекс Мосбиржи; в) 20 компаний с низколиквидными акциями, которые не входят в индекс Мосбиржи и не имеют кросс-листинга на других биржах.

По всем компаниям были собраны данные о волатильности доходности, метрики сентимента, финансовые и нефундаментальные показатели. Источник данных о биржевых характеристиках, финансовых и нефундаментальных показателях, макроэкономических переменных: Eikon Refinitiv.

Вручную нами было размечено 32 тысячи сообщений. На Рис. 3.7 показано распределение сообщений по классам после ручной разметки. Сформированный корпус был исходно несбалансированным (4061 негативное, 24 249 нейтральных и 4266 позитивных сообщений), что препятствовало обучению, поэтому мы использовали веса.

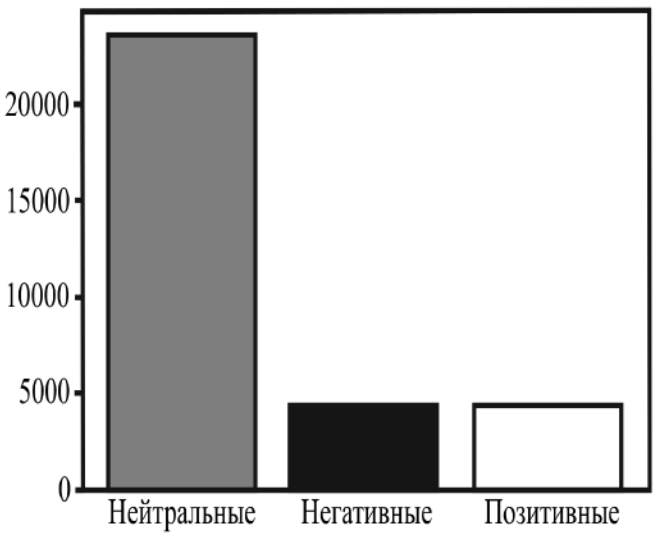

Рис. 3.7. Распределение сообщений по классам после ручной разметки

В Табл. 3.13 представлена описательная статистика по выборке.

Таблица 3.13

Описательная статистика по выборке

Переменная	*Количество наблюдений*	*Среднее*	*Стандартное отклонение*	*Минимум*	*Максимум*
Зависимая переменная					
volatility	5187	0.021	0.014	0.005	0.08
Объясняющие переменные сентимента					
ln_tele_Hype	2339	1.389	1.311	0	4.805
ln_mfd_Hype	4320	1.799	0.877	0	5.0
ln_Hype	4499	1.816	0.857	0	5.0
ln_Yandex_abs	1440	8.276	1.965	3.829	13.511
ln_Yandex_ relat	1440	3.846	1.956	-0.728	9,038
ratingTR	4859	6.693	1.093	0.526	10.0
ln_all_mes	4575	4.915	2.355	0	10.643
ln_all_neg_mes	4575	3.427	2.192	0	8.888
ln_all_ntr_mes	4575	4.450	2.281	0	10.211
ln_all_pos_mes	4572	3.408	2.074	0	8.38
Контрольные переменные					
TradeVolFF	5185	6.682	1.858	0	12.843
TradeVol withoutFF	5157	7.24	2.089	0	13.267
FreeFloat	5039	36.478	25.564	0	100
NzTradeDays Share	5187	96.848	12.92	0	100
ROA	5082	1.24	2.576	-7.421	7.58
BoardSize	4995	9.858	3.012	0	19
IndepDir Percent	4389	29.443	20.736	0	100
GDPgrowth	5307	0.069	0.283	-0.704	0.61
RVI_change	4636	3.124	20.219	-32.74	86.48
Gov10Y	5307	8.515	1.620	6.136	14.38
PMI_US	5307	54.415	2.802	39.8	61

3.4.5. Результаты эконометрических расчетов

В Табл. 3.14 показано влияние различных факторов на волатильность доходности акций российских компаний. Гипотеза 1 о нелинейном значимом влиянии сентимента инвесторов на волатильность доходности акций подтвердилась частично. Авторский индикатор сентимента Hype (интегральный по двум платформам – mfd и Telegram) оказывает положительное и значимое на 1% уровне влияние на волатильность. Вывод о значимом вли-

янии позитивной тональности и общего количества сообщений – двух компонент индикатора Hype - совпадает с результатами Audrino et al. (2020) по рынку США. В то же время, наш вывод расходится с результатами работ Maitra and Dash (2017) и Ho et al. (2013), в которых по рынкам Индии и США соответственно выявлено, что чем более позитивны настроения инвесторов, тем ниже волатильность. Это расхождение мы объясняем, во-первых, тем, что наш индикатор Hype является комплексной мерой позитивной тональности сообщений и роста общего количества сообщений. Во-вторых, если количество сообщений в социальных сетях растет, и их тональность чрезмерно позитивна, часть инвесторов не следует за толпой и не доверяет высокой оценке перспектив компании, в результате растет волатильность доходности.

Влияние индикатора Hype, рассчитанного на основе тональности сообщений в мессенджере Telegram, положительно (как и влияние интегрального индикатора по двум платформам) и значимо на 5% уровне. Влияние индикатора Hype, рассчитанного на основе тональности сообщений на форуме MFD, нелинейно и описывается параболой ветвями вверх. Волатильность доходности акций повышается как при высоком значении индикатора (что соответствует выводам по интегральному показателю Hype по двум платформам и показателю Hype по мессенджеру Telegram), так и при низком значении индикатора. В отличие от мессенджера Telegram, низкая активность обсуждений и преимущественно негативная тональность сообщений на популярном инвестиционном форуме MFD также приводят к увеличению волатильности доходности акций. Это свидетельствует о различиях в восприятии инвесторами сообщений на двух платформах.

Подтвердилась Гипотеза 3 нашего исследования: зависимость волатильности от количества запросов по компании в Yandex описывается параболой ветвями вниз, причем и количество запросов, и его квадрат значимы на 1% уровне (Табл. 3.14). Если количество запросов относительно мало или, напротив, велико, то волатильность доходности снижается. Мы объясняем это тем, что повышенный интерес к компании свидетельствует о негативных или позитивных событиях, в результате которых на рынке формируется устойчивый тренд, и волатильность доходности при этом снижается. Пониженный интерес (малое количество поисковых запросов) свидетельствует о том, что значимых событий не происходит, нет предпосылок к существенным колебаниям цен, и они остаются относительно стабильными.

Мы отмечаем, что оценка кредитного риска компании на основе сентимента (*tr_indicator*), разработанная Refinitiv, статистически не значима во всех моделях (Табл. 3.14). В то же время, ее влияние устойчиво: чем выше кредитный риск с точки зрения отчетов и рекомендаций аналитиков, тем ниже волатильность. Мы объясняем это тем, что высокий кредитный риск сопровождается устойчивым трендом к снижению цен, и волатильность не высока.

Таблица 3.14

Влияние прокси сентимента на волатильность

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
Wald	290.1	70.59	415.58	87.33	119.83	316.5	307.57	326,2	336,83
Num obs	3836	2142	4135	1239	1239	3976	3686	3976	3976
ln_Hype			0,0029*** (0,0006)						
ln_mfd_Hype	-0,0028** (0,0012)								
ln_mfd_Hype_2	0,0011*** (0,0003)								
ln_tele_Hype		0,0007** (0,0003)							
LN_Yandex_Abs				0,0186*** (0,0042)					
LN_Yandex_Abs_2				-0,0009*** (0,0002)					
LN_Yandex_Rel					0,0108*** (0,0021)				
LN_Yandex_Rel_2					-0,0009*** (0,0002)				
ln_all_mes						0,0018*** (0,0003)			
ln_all_neg_mes							0,0006** (0,0003)		
ln_all_neg_mes_2							0,0003* (0,0002)		

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
ln_all_neutr_mes								0,0018*** (0,0003)	
ln_all_pos_mes									0,0021*** (0,0004)
TR_indicator						0,0005 (0,0004)	0,0007 (0,0005)	0,0005 (0,0004)	0,0005 (0,0004)
TradeVolFF	0,0045*** (0,0005)						-0,0014 (0,0022)		
TradeVolFF_2							0,0005*** (0,0002)		
FreeFloat		-0,0001 (0,0001)	-0,0001 (0,0001)	-0,0001 (0,0001)	-0,0001 (0,0001)	-0,0001 (0,0001)		-0,0001 (0,0001)	-0,0001 (0,0001)
NzTradeDaysShare		0,0003*** (0,0001)	0,0003*** (0,0001)	0,0001 (0,0001)	0,0008 (0,001)	0,0001*** (0,01)		0,0001*** (0,01)	0,0001*** (0,01)
NzTradeDaysShare_2			-0,0001 (0,0001)		-0,0001 (0,0001)				
ROA_lag	0,0002 (0,0002)						-0,0001 (0,0002)		
ROA		-0,0009*** (0,0003)	-0,0004*** (0,0002)	-0,0016*** (0,0003)	-0,0016*** (0,0003)	-0,0005*** (0,0002)		-0,0005*** (0,0002)	-0,0005*** (0,0002)
BoardSize	-0,0009*** (0,0003)	-0,0008*** (0,0002)	-0,0006*** (0,0002)	-0,0018*** (0,0004)	-0,0018*** (0,0004)	-0,0008*** (0,0002)	-0,0008*** (0,0002)	-0,0008*** (0,0002)	-0,0008*** (0,0002)
IndepDirPercent	-0,0001*** (0,01)						-0,0001* (0,01)		
GDP growth	-0,0001 (0,001)						0,0007 (0,0008)		
RVI_change		0,0001***	0,0001***	0,0001***	0,0001***	0,0001***		0,0001***	0,0001***

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
		(0,0001)	(0,0001)	(0,0001)	(0,0001)	(0,0001)		(0,0001)	(0,0001)
Gov10Y	0,0024*** (0,0004)						-0,001 (0,0012)		
Gov10Y_2							0,0002*** (0,0001)		
PMI_US		-0,0007*** (0,0001)	-0,0005*** (0,0001)	-0,0003*** (0,0002)	-0,0003*** (0,0002)	-0,0004*** (0,0001)		-0,0004*** (0,0001)	-0,0004*** (0,0001)
_cons	-0,0235*** (0,0054)	0,0349*** (0,0105)	0,0283*** (0,0057)	-0,0424** (0,0188)	-0,0044 (0,0351)				
d14	0,0071*** (0,0017)		0,0106*** (0,0011)			0,0125*** (0,0013)	0,0079*** (0,0018)	0,0125*** (0,0013)	0,0125*** (0,0012)
d15	0,0023 (0,0016)		0,0117*** (0,0009)			0,0121*** (0,001)	0,0037* (0,0019)	0,0122*** (0,0011)	0,012*** (0,001)
d16	0,0033*** (0,001)		0,0052*** (0,0009)			0,0053*** (0,0008)	0,0047*** (0,0012)	0,0053*** (0,0008)	0,0051*** (0,0008)
d17	0,0048*** (0,001)	0,0067*** (0,0013)	0,0056*** (0,001)			0,0058*** (0,0008)	0,0059*** (0,0011)	0,0059*** (0,0008)	0,0057*** (0,0008)
d18	0,0038*** (0,0008)	0,0048*** (0,001)	0,004*** (0,001)	0,0067*** (0,0012)	0,0064*** (0,0011)	0,0042*** (0,0009)	0,0041*** (0,0009)	0,0042*** (0,0009)	0,0042*** (0,0009)

Значима тональность обсуждений компании (Табл. 3.14). При росте совокупного количества сообщений (*ln_all_mes*) волатильность повышается, что соответствует выводам Antweiler and Frank (2004) по рынку США. Аналогичный эффект роста волатильности наблюдается в случае увеличения количества позитивных (*ln_all_pos_mes*), негативных (*ln_all_neg_mes*) и нейтральных (*ln_all_neutr_mes*) сообщений в отдельности. Все переменные значимы на 1-5% уровне. Отметим специфику влияния количества негативных сообщений: для них зависимость является квадратичной, причем влияние квадрата переменной также значимо и положительно. Поэтому мы делаем вывод, что инвесторы сильнее реагируют на негативные сообщения, с ростом их количества волатильность доходности растет в большей степени по сравнению с количеством позитивных или нейтральных сообщений. Таким образом, подтвердилась Гипотеза 2 нашего исследования. Данный вывод соответствует результатам Phuong et al. (2015) о более сильном влиянии негативных новостей на волатильность доходности акций на вьетнамском фондовом рынке.

Влияние объема торгов (*TradeVolFF*) на волатильность доходности значимо и положительно (см. модель (1) в Табл. 3.14). Более того, относительно объема торгов наблюдаются эффекты нелинейности: чем выше квадрат объема торгов, тем выше волатильность (см. модель (7) в Табл. 3.14). Значимо влияние количества дней с ненулевым объемом торгов (*NzTradeDaysShare*): чем выше торговая активность, тем выше волатильность доходности, но по данной переменной нелинейность влияния не выявлена.

Чем выше отдача на активы, тем ниже волатильность доходности акций: фактор ROA по прошлому кварталу значим на 1% уровне во всех моделях (Табл. 3.14). Значимы факторы корпоративного управления. Чем больше размер совета директоров и чем больше доля независимых директоров, тем ниже волатильность доходности акций компании (Табл. 3.14).

Чем выше доходность к погашению по 10-летним государственным облигациям США (*Gov10Y*), а, следовательно, выше уровень риска на мировых финансовых рынках, тем выше волатильность доходности акций российского рынка. Мы выявили эффекты нелинейности: влияние квадрата доходности по 10-летним облигациям США тоже значимо положительно (модель (7), Табл. 3.14). Темп роста ВВП не значим, влияние данного фактора разнонаправленно (Табл. 3.14). Значимы дамми временных периодов: в 2014, 2017 и 2018 гг. фиксировался рост волатильности доходности акций российского рынка (Табл. 3.14). Мы отмечаем устойчивое положительное и значимое на 1% уровне влияние изменения индекса волатильности на российском рынке (*RVI_change*), что соответствует ожиданиям. Чем выше индекс деловой активности США (*PMI_US*), т.е. чем более благоприятна макроэкономическая ситуация в США, тем ниже волатильность акций на российском рынке.

Поскольку в нашей выборке данные по инвестиционному форуму MFD доступны с 2014 г., а по мессенджеру Telegram – только с 2017 г., мы про-

верили устойчивость полученных результатов на подвыборке с 2017 по I квартал 2020 гг. (Приложение В, Табл. В.1).

Большинство выявленных ранее зависимостей сохраняется (Приложение В, Табл. В.1). Индикатор Hype (интегральный по двум платформам) значимо положительно на 10% уровне влияет на волатильность доходности акций. Индикатор Hype по мессенджеру Telegram оказывают положительное и значимое на 5% уровне влияние на волатильность доходности, тогда как Hype по форуму MFD влияет нелинейно (парабола ветвями вверх).

Рентабельность активов, размер совета директоров и процент независимых директоров значимо отрицательно на волатильность, что соответствует ранее полученным выводам. Чем выше уровень долговой нагрузки (EBITDA to Assets, Assets to Equity), тем ниже волатильность доходности. Мы объясняем это тем, что инвесторы уделяют внимание долговой нагрузке, и при ее чрезмерном увеличении формируется устойчивый ценовой тренд, что снижает волатильность. Чем выше торговая активность (объем торгов и число дней с ненулевым объемом торгов), тем выше волатильность доходности: влияние фактора значимо на 1% уровне.

С ростом доходности по государственным облигациям США (прокси риска на глобальном рынке) растет волатильность доходности акций российских компаний. Напротив, чем выше индекс деловой активности PMI по рынку США, тем ниже волатильность доходности на российском рынке акций. Значимо положительно влияние индекса волатильности RVI по российскому рынку.

3.4.6. Выводы

В нашей работе впервые раскрывается влияние настроений частных инвесторов в социальных сетях на волатильность доходности акций компаний российского рынка. Выборка сообщений построена на уникальных данных, охватывающих более чем 40% владельцев активных брокерских счетов за 2014-2020 гг.

В работе предложен Hype индикатор, построенный на основе ручной разметки 32 тыс. сообщений и применения нейронной сети на основе CNN и трансформеров. Индикатор строится по двум платформам – широко используемому, несмотря на официальный запрет (действовавший до июня 2020 г.), мессенджеру Telegram и самому популярному в России инвестиционному форуму mfd.ru. Мы выявляем отличия влияния индикатора Hype по платформам Telegram и MFD. Наряду с индикатором Hype, мы рассматриваем широкий спектр метрик сентимента: количество позитивных, нейтральных и негативных сообщений частных инвесторов по платформам Telegram и MFD, количество поисковых запросов в Yandex, предложенный Refinitiv индикатор кредитного риска, построенный по сентименту брокеров и аналитиков.

Тестируя гипотезы о влиянии настроений инвесторов с применением методов регрессионного анализа, мы выявили взаимосвязи между метриками

сентимента частных инвесторов и волатильностью доходности акций с учетом эффектов нелинейности.

Предложенный нами индикатор сентимента Hype оказывает положительное и значимое на 1% уровне влияние на волатильность доходности акций. Мы заключаем, что если количество сообщений в социальных сетях велико, и их тональность чрезмерно позитивна, часть инвесторов не следует за толпой и не доверяет высокой оценке перспектив компании, в результате растет волатильность доходности. Мы отмечаем различия по платформам Telegram и MFD: влияние Hype индикатора, построенного по мессенджеру Telegram, линейно, положительно и значимо на 5% уровне, тогда как влияние индикатора, построенного по форуму MFD, значимо, но не линейно. Волатильность доходности акций повышается как при чрезмерно высоком, так и при низком значении индикатора Hype по обсуждениям на форуме MFD, т.е. малое количество обсуждений и их негативная окраска тоже приводят к увеличению волатильности доходности акций.

Значима тональность обсуждений компании: при росте общего количества сообщений, а также количества позитивных, негативных и нейтральных сообщений в отдельности волатильность повышается. Мы отмечаем, что инвесторы сильнее реагируют на негативные сообщения, с ростом их количества волатильность доходности растет в большей степени (зависимость является квадратичной).

Влияние оценки кредитного риска компании на основе сентимента аналитиков от Refinitiv устойчиво, хотя статистическая значимость невелика: чем выше кредитный риск, тем ниже волатильность. Мы объясняем это тем, что высокий кредитный риск сопровождается формированием устойчивого ценового тренда, и волатильность невысока.

По большинству контрольных переменных получены ожидаемые выводы. Рентабельность активов компании-эмитента значимо отрицательно влияет на волатильность. Значимы факторы корпоративного управления: чем больше размер совета директоров и чем выше доля независимых директоров, тем ниже волатильность доходности. Чем выше доходность десятилетних государственных облигаций США (прокси риска на мировых финансовых рынках), тем выше волатильность доходности акций российского рынка. Напротив, индекс деловой активности PMI по рынку США значимо отрицательно влияет на волатильность доходности акций российского рынка. Торговая активность положительно связана с волатильностью. Тестирование регрессий на подвыборке за 2017-2020 гг. показало устойчивость полученных результатов.

Для будущих исследований выглядит перспективным переход к дневным или даже внутридневным данным, чтобы понять, влияние каких факторов первично – фундаментальных факторов или реакции частных инвесторах в социальных сетях. В разрезе данных целесообразно разделить сообщения в зависимости от популярности автора (блогера), количества просмотров и реакций («лайков»).

Глава 4
КРИПТОАКТИВЫ И КРИПТОВАЛЮТЫ КАК ОБЪЕКТ ИССЛЕДОВАНИЙ ВЛИЯНИЯ СЕНТИМЕНТА

Некоторые исследователи и известные инвесторы рассматривают криптоактивы и криптовалюты исключительно как «хайповые истории», что позволяет предположить, что сентимент участников рынка для этих инвестиционных активов будет играть крайне важную роль. В данной главе мы попытаемся провести анализ этой гипотезы. Кроме того, к концу 2021 г. стало понятно, что шумиха в СМИ и инвестиционных сообществах вокруг криптовалют намного опередила реальный прогресс и практическую ценность этих специфических активов. Возможно, следующий этап позволит сформировать фундаментальную ценность этих специфических активов.

4.1. ОБЗОР РЫНКА КРИПТОВАЛЮТ ПО СОСТОЯНИЮ НА КОНЕЦ 2021 Г. – СЕРЕДИНУ 2022 Г.

В 2021-2022 гг. инвесторы проявляли яркий интерес к криптоиндустрии, и последняя демонстрировала высокую волатильность из-за новостного фона (Рис. 4.1). За 2021 г. общая рыночная капитализация отрасли криптовалют выросла более чем в три раза, достигнув исторического максимума в $3 трлн долл. и снизившись к концу года до уровня $2,4 трлн. долл.

Рис. 4.1. Капитализация и объем торгов спот топ-30 монет за 2021 г. Источник: CoinGecko Report 2021.

Биткоин достиг максимальной стоимости $69 045 в ноябре 2021 г., но к концу года скорректировался до уровня $45 000-$48 000, что, тем не менее,

представляет собой двукратный рост за год (Рис. 4.2, источник: Yahoo Finance). Ценовая динамика пары ETH-USD в 2021 г. представлена в Приложении Г на Рис. Г.1.

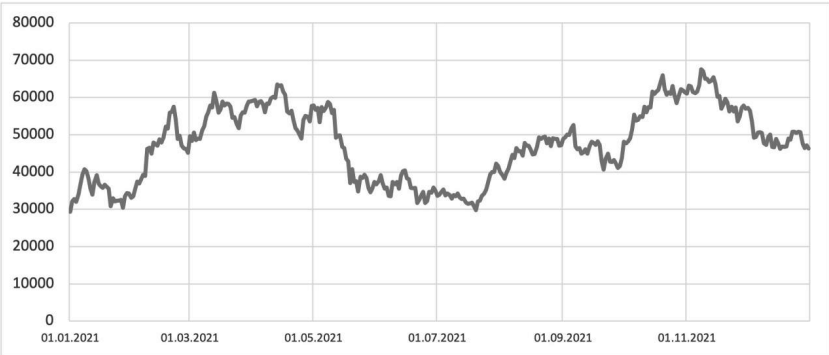

Рис. 4.2. Ценовая динамика пары BTC-USD в 2021 г.
Источник: Yahoo Finance

В 2021 г. биткоин опередил все основные классы активов по доходности инвестирования, продемонстрировав рост на 62%. Рынки в основном находились в режиме принятия риска, поскольку интерес сместился с золота и казначейских облигаций на акции и товары.

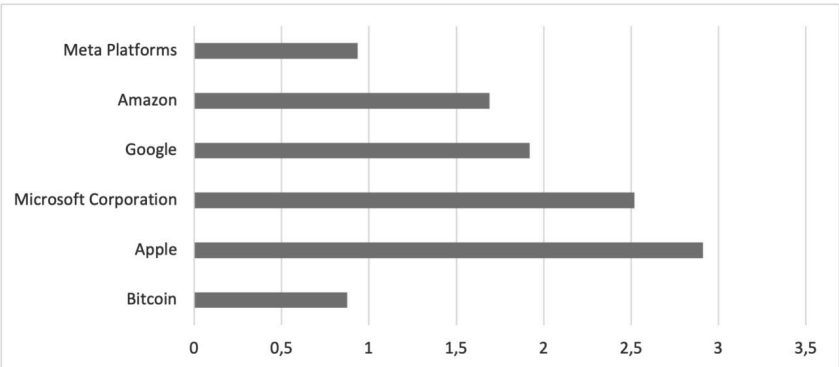

Рис. 4.3. Рыночная капитализация Bitcoin и топ-5 компаний S&P 500
Источник: Yahoo Finance

В 2021 г. корреляция биткоина с основными классами активов последовательно снижалась, но еще в 2020 г. биткоин имел значимую корреляцию с такими активами, как золото и акции. Сопоставление годовой доходности Bitcoin и иных классов активов показано в Табл. 4.1.

Таблица 4.1

Сопоставление годовой доходности Bitcoin и иных классов активов (в %, год к году)

Актив	*Доходность за 2021 г., %*	*Доходность за 2022 г., %*
BTC	62	281
Нефть	58	-21
S&P 500	29	15
NASDAQ	23	42
DXY (US Dollar Index)	6	-7
Золото	-6	24
TLT (Treasury Bonds)	6	13

Для Ethereum, так же как и для Bitcoin, 2021 г. был крайне удачным. За 2021 г. ETH обновил исторические максимумы дважды. В первый раз 11 мая, когда он пробил отметку $4000 и достиг $4183. После коррекции стоимость ETH продолжала расти и 9 ноября достигла нового максимума $4815. ETH закончил 2021 г. с ценой $3715: его годовой рост составил впечатляющие 403%. Ценовая динамика пары ETH-USD в 2021 г. показана в Приложении Г на Рис. Г.1.

Однако, доминирующее положение биткоина на рынке ослабло, упав к концу 2021 г. до уровня ниже 40% (Приложение Г, Рис. Г.2; источник: CoinGecko Report 2021). Большая часть внимания пользователей переключилась на Ethereum, его многочисленные форки и L2-чейны, которые заявляют о решении проблем с масштабируемостью, а также новые L1-чейны, такие как Solana и Terra.

2022 г. стал проблемным не только для основных классов активов, но и для криптоиндустрии, опровергнув гипотезу о наличии защитных свойств (Рис. 4.4). Общий объем рынка криптовалют на середину 2022 г. оценивался в $900 млрд долл., что представляет собой падение более чем на 50% по сравнению с I кварталом 2022 г. и примерно на 70% с момента достижения пика рынка в ноябре 2021 г. Цена и BTC, и ETH продемонстрировала аналогичные по размеру падения, опустившись ниже своих предыдущих максимумов бычьего цикла 2018 г.

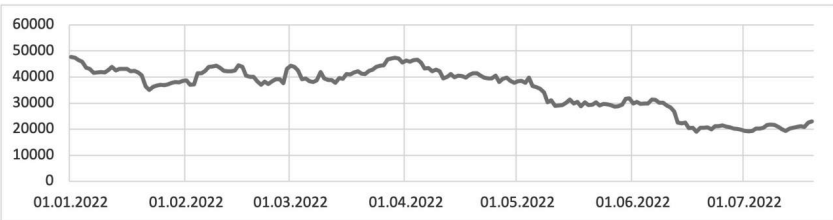

Рис. 4.4. Ценовая динамика пары BTC-USD в первом полугодии 2022 г.
Источник: Yahoo Finance

Важные роли в падении рынка криптовалют в 2022 г. сыграли геополитическая напряженность, ужесточение денежно-кредитной политики ЦБ развитых стран и сжатие свободной денежной ликвидности. Имели место и громкие обвалы на этом рынке. Самым крупным из них является обвал стейблкоина Terra (UST) и связанной с ней криптовалюты Luna. На пике своей стоимости LUNA и UST достигали рыночной стоимости в $40B и $18B соответственно. Однако при возникновении проблем практически вся стоимость UST была разрушена в течение недели из-за цунами панического оттока средств.

Серьезные ценовые просадки по всем криптоактивам также вызвали каскадную ликвидацию позиций с кредитным плечом. Еженедельные многомиллиардные ликвидации на CEX стали обычным явлением, поскольку многие трейдеры неоднократно оказывались в минусе. Наиболее заметной была компания Three Arrows Capital, которой в настоящее время грозит банкротство и ликвидация.

BTC находился в нисходящем тренде на протяжении всего II квартала 2022 г. (Табл. 4.2). Впервые в истории существования BTC девять недель подряд закрытие торгов происходило в минусе. Bitcoin пробил множество уровней поддержки, включая основные на уровнях $30 тыс. и $20 тыс. Краткосрочно BTC достиг годового минимума в $17 760, но впоследствии вернулся на уровни выше $20 тыс. Биткоин оказался худшим по доходности активом по сравнению с основными классами активов. Также в течение 2022 г. наблюдается тенденция к росту корреляции между биткоином и фондовым рынком, которая снижалась в 2021 г.

Таблица 4.2

Сопоставление годовой доходности Bitcoin и иных классов активов (в %, год к году)

Актив	*Доходность за II квартал 2022 г., %*	*Доходность за II квартал 2021 г., %*
BTC	-55,8	-40
Нефть	7	22
S&P 500	-17	8
NASDAQ	-23	9
DXY (US Dollar Index)	7	-1
Золото	-6	3
TLT (Treasury Bonds)	-13	6

За II квартал 2022 г. ETH зафиксировал 11 еженедельных закрытий в красной зоне, что также произошло впервые за все время (Приложение Г, Рис. Г.3). ETH ненадолго опускался ниже $900, но достаточно быстро восстановился выше $1 000.

Несмотря на то, что многие журналисты СМИ в очередной раз успели «похоронить» криптовалюты и цифровые активы, присутствуют опреде-

ленные положительные моменты по развитию этого сегмента инвестиционного рынка. Во-первых, растет осведомленность широкого круга инвесторов. Всего три-четыре года назад во время предыдущего бычьего цикла криптовалюты были представлены преимущественно блокчейнами BTC и ETH и покупались исключительно для спекуляций. Сегодня на рынке представлено большое количество альтчейнов со своими собственными экосистемами, которые решают ряд проблем традиционных блокчейнов. Сектор DeFi значительно увеличил свои позиции в отрасли и занял соответствующую нишу. Энтузиасты NFT и цифровых активов сформировали свои собственные сообщества внутри криптовалют, которые развивают не только сегменты Game и Metaverse, но и проекты, несущие практическую пользу и относящиеся к сегменту Utility. Во-вторых, схлопывание пузыря позволяет индустрии избавиться от многих неэффективных проектов и финансовых пирамид, накопленных во время бычьего рынка.

4.2. РЫНОК NFT В 2018–2022 ГГ.[1]

Одним из наиболее ранних популярных NFT является коллекция CryptoKitties, которая представляет собой набор художественных изображений, представляющих виртуальных кошек из одноименной игры на Ethereum, позволяющей игрокам покупать, собирать, разводить и продавать их (Team, 2021). В декабре 2017 г. волна популярности CryptoKitties захлестнула Интернет (Wong, 2017). Однако, несмотря на огромную популярность CryptoKitties, которую многие считают главным примером иррациональности (Tepper, 2017), эта коллекция оставалась единственным популярным примером NFT в течение почти двух лет. В июле 2020 г. рынок NFT начал расти и привлек к себе огромное внимание в марте 2021 г., когда дизайнер и художник Майк Винкельман, известный как Beeple, продал NFT авторские права на свою работу «Everydays: Первые 5000 дней» за 69,3 млн долл. на аукционе Christie's (Team, 2021; Phillips, 2021).

Затем последовало еще несколько рекордных продаж (Phillips, 2021; Howcrof, 2021), в том числе привлёкшая значительное внимание аудитории продажа основателем Twitter Джеком Дорси NFT первой записи в социальной сети за 2,9 млн долларов.

С момента рекордных продаж Christie's продал еще более 100 произведений искусства, для которых выпущен NFT. Помимо этого, аукционный дом также стоял за четырьмя из пяти самых дорогих NFT, проданных в 2021 г. Согласно отчету Bloomberg, в декабре 2021 г. аукционный дом заявил, что за 2021 г. продал NFT на сумму 150 млн долл. Компания заявила, что новый формат помог ей достичь общего объема

[1] Данный параграф подготовлен Куркиным А.В. на основе собственной ВКР, выполненной под руководством Т.В. Тепловой.

чистых продаж в 2021 г. на сумму 7,1 млрд долл., что стало самым высоким показателем за последние пять лет (Howcroft, 2022).

Значительный успех NFT-проектов привлек в данную сферу большое количество частных инвесторов. К июлю-августу 2021 г. на рынке NFT ряд коллекций консолидировались до статуса "голубых фишек", что привело к резкому росту спроса на них, вызвав значительный приток денег на рынок NFT. Так, объем сделок за этот период составил 10,7 млрд долл. (Рис. 7). Данный период получил название «NFT-лето».

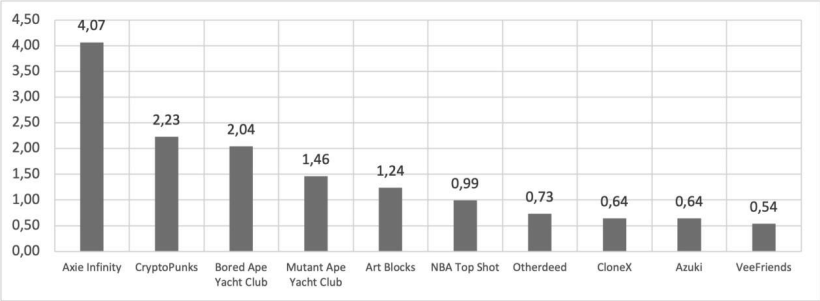

Рис. 4.5. Рейтинг коллекций NFT по объему продаж на конец 2021 г. (млрд долл.).

Источник: CoinGecko.com

Подобная активность привлекла большое внимание брендов по всему миру. Популярные коллекции NFT и дорогие продажи объектов цифровой собственности становятся высоко обсуждаемыми темами, что привело к возникновению спроса на покупку NFT в рекламных целях. Так, в 2021 г. в выпуске или приобретении NFT-объектов участвовали такие бренды как Visa, Nike, Budweiser, Coca-Cola, Burberry, Louis Vuitton, Gucci, журнал Times и многие другие. Покупка NFT брендами является альтернативной формой рекламной кампании, поскольку общественное внимание, возникающее в результате сделки, также привлекается и к компании-покупателю NFT-актива, что благотворно сказывается на продажах (CoinGecko.com, 2022).

Объем продаж NFT в 2021 г. составил 24,9 млрд долл. по сравнению с 94,9 млн долл. в 2020 г., что означает увеличение месячного объема торгов год к году на 5 500% (DappRadar.com, 2022). Однако, оценки величины рынка NFT значительно отличаются у разных поставщиков данных. Транзакции, которые происходят «вне блокчейна», такие, как крупные продажи произведений искусства NFT в аукционных домах, часто не фиксируются записями в блокчейне, и, соответственно, их труднее оценить. Так портал CryptoSlam, отслеживающий записи о NFT с нескольких блокчейнов, сообщил, что общая сумма продаж в 2021 г. составила 18,3 млрд долларов (Приложение Г, Рис. Г.3). Ресурс NonFungible.com, который отслеживает только блокчейн Ethereum, оценивает объем продаж в 2021 г. в 15,7 млрд

долл. По данным DappRadar, в 2021 г. около 28,6 млн уникальных блокчейн-кошельков участвовали в торговле NFT, по сравнению с примерно 545 000 в 2020 г.

88% годового объема торгов 2021 г. пришлось всего на три торговые площадки: OpenSea (61%), Axie Infinity (17%) и CryptoPunks (10%). Остальные маркетплейсы в среднем занимают от 1% до 1.5% доли рынка (источник: CoinGecko, 2022).

Что касается распределения продаж по крупнейшим блокчейнам, то Ethereum, Ronin и Solana были явными лидерами по торговой активности NFT с совокупной долей рынка более 90%. Однако, несмотря на то, что самые дорогие NFT продаются за десятки миллионов долларов, а средняя стоимость предметов из популярных коллекций, как, например, CryptoPunks, выросла примерно со 100 000 долл. в июле до почти 500 000 долл. в ноябре, медианный ценовой диапазон для рынка ценовых активов составляет от 100 до 1000 долл., а значительная часть NFT продается за сумму менее 10 долл. (Howcroft, 2022).

С данной информацией коррелирует усредненный портрет инвестора в NFT. Согласно отчету портала Chainalysis на долю розничных покупателей, которые потратили менее 10 000 долл. на NFT, пришлось 80% всех транзакций на рынке NFT в 2021 г. 19% транзакций NFT были совершены крупными частными инвесторами, которые потратили от 10 000 до 100 000 долл. на цифровые токены. Только менее 1% покупок NFT было совершено институциональными инвесторами, которые потратили более 100 000 долл. Таким образом, основной движущей силой рынка NFT в настоящий момент являются розничные неквалифицированные инвесторы. Распределение стоимости продажи NFT представлено в Приложении Г на Рис. Г.5.

Тем не менее, к декабрю 2021 г. показатели рынка NFT начали снижаться, и эта тенденция продолжилась в первом квартале 2022 г. Среди возможных причин данного снижения необходимо учесть рост уровня инфляции, ужесточение политики ФРС США, рост геополитической напряженности в мире. По состоянию на апрель 2022 г., большинство индикаторов, представленных в Табл. 4.3, показывают «медвежий» тренд, однако значительное снижение количества покупок NFT было почти полностью компенсировано ростом их стоимости. Таким образом, в настоящий момент целесообразнее говорить об охлаждении и стабилизации рынка NFT после периода бурного роста, нежели чем о его крахе (Nonfungible.com, 2022).

Таблица 4.3

Основные показатели рынка NFT-активов в IV квартале 2021 г. – I квартале 2022 гг.

Показатель	*IV кв. 2021 г.*	*I кв. 2022 г.*	*Изменение (в %)*
Совокупный объем сделок (млрд. долл. США)	14,54	16,45	+13,25%
Количество продаж (млн	14,04	7,44	-46,96%

Показатель	*IV кв. 2021 г.*	*I кв. 2022 г.*	*Изменение (в %)*
транзакций)			
Средняя цена продажи (долл.)	587	1057	+80,1%
Количество покупателей, млн уникальных адресов	1,696	1,172	-30,91%
Количество продавцов, млн уникальных адресов	0,967	0,816	-15,61%
Количество активных кошельков, млн	1,948	1,455	-29,34%
Количество активных смарт-контрактов	1271	1404	+10,46%

Два показателя демонстрируют позитивную динамику:

1) количество смарт-контрактов увеличилось на 10%, что означает увеличение количества проектов, которые создаются и являются активными, а, следовательно, что рынок продолжает развиваться;

2) средняя цена NFT увеличилась на 80% до 1057 долл. Во многом значительное увеличение средней цены продажи произошло в результате успешности продаж проектов-голубых фишек, таких как Bored Ape Yacht Club, CryptoPunks, NBA Top Shot.

Таким образом, результаты I квартала 2022 г. значительно хуже результатов IV квартала 2021 г., однако совокупные продажи на 8 млрд долл. во много раз превосходят результаты рынка до наступления «NFT-лета». Также, несмотря на то, что количество активных кошельков уменьшается, количество покупателей по-прежнему превышает количество продавцов, что означает, что интерес к покупке NFT все еще присутствует.

4.3. МЕТРИКИ СЕНТИМЕНТА ДЛЯ КРИПТОВАЛЮТ И ЦИФРОВЫХ АКТИВОВ

Одной из самый популярных и используемых инвесторами метрик сентимента крипторынка является «Индекс страха и жадности» (Fear & Greed Index). Традиционно этот индекс помогает инвесторам и трейдерам анализировать рынок биткоина и криптовалют с точки зрения настроений инвесторов.

Основная предпосылка для создания данного индекса заключается в том, что преобладание на рынке «страха» может указывать на недооценку / заниженную стоимость биткоина и остальных криптовалют, что может представлять хорошую возможность для покупки. Обратная ситуация также применима: преобладание на рынке «жадности» может указывать на то, что цена биткоина слишком высока по сравнению с его внутренней стоимостью, и это может быть хорошим моментом для продажи.

Индекс страха и жадности может быть полезен для инвесторов, поскольку предоставляет возможность оценить преобладающие настроения на

рынке и совершать выгодные покупки, когда рынок характеризуется состоянием страха, и выгодные продажи, когда рынок характеризуется состоянием жадности. Также данный индекс позволяет инвестору отделить собственные эмоции от остального рынка и избежать иррациональной реакции на резкие движения цен и волатильность рынка.

Индекс страха и жадности может принимать значения в диапазоне от 0 до 100. Если показатель находится на отметках ниже 45 пунктов и снижается, то это свидетельствует о крайнем страхе и может указывать на то, что многие инвесторы в биткоин ($BTC) и криптовалюты слишком обеспокоены дальнейшим падением цены. В этот момент страх перед криптовалютой высок, и инвесторы, возможно, не могут мыслить рационально и могут совершать ошибки, вызванные паникой. Если показатель превысил 55 пунктов и продолжает подниматься, то это может свидетельствовать о крайней жадности и указывать на то, что многие инвесторы в биткоин и криптовалюты слишком воодушевлены тем, что цены в будущем будут расти гораздо больше. В связи с этим, на рынке может преобладать страх за упущенные возможности (FOMO - Fear of missing out), и инвесторы готовы совершать покупки по завышенным ценам, что может представлять хорошую возможность для продажи. Динамика индекса представлена на Рис. 4.6.

Рис. 4.6. Динамика индекса Fear & Greed для рынка криптовалют в 2018–2022 гг.

Исходные данные для индекса страха и жадности генерируются каждый день из следующих источников:

а) Волатильность (25%). Измеряется текущая волатильность и максимальные просадки биткоина, и происходит их сравнение с соответствующими средними значениями за последние 30 дней и 90 дней. Необычный рост волатильности является признаком нарастания страха на рынке.

б) Рыночный импульс / объем (25%). Измеряется текущий объем и рыночный импульс в сравнении со средними значениями за последние 30/90

дней. Как правило, когда наблюдаются высокие объемы покупок на растущем рынке на ежедневной основе, можно сделать вывод, что рынок ведет себя слишком жадно / «по-бычьи».

в) Социальные медиа (15%). Для данного сегмента происходит сбор и подсчет сообщений по различным хэштегам для каждой криптовалюты и анализ количества социальных взаимодействий, которые они получают в определенные временные промежутки. Необычно высокая скорость взаимодействия приводит к росту общественного интереса к криптовалюте и соответствует жадному поведению рынка.

г) Опросы (15%). Совместно с strawpoll.com авторы индекса страха и жадности проводят еженедельные опросы относительно мнений о будущем движении цен на криптовалюты. Обычно в каждом опросе участвуют 2 000-3 000 человек, в связи с чем представление о настроениях группы криптоинвесторов может быть достаточно репрезентативным. В настоящее время опросы приостановлены.

д) Рыночное доминирование Bitcoin (10%). Авторы считают, что рост доминирования биткоина вызван страхом (и, следовательно, сокращением) слишком спекулятивных инвестиций в альткоины, поскольку биткоин становится в некотором роде аналогом доллара и выполняет защитные функции на рынке криптовалют. С другой стороны, когда доминирование биткоина сокращается, люди становятся более жадными, инвестируя в более рискованные альткоины. Напротив, больший интерес к альткоинам может свидетельствовать о бычьем / жадном поведении в отношении этой конкретной монеты.

е) Тренды (10%). Для расчёта используются данные Google Trends для различных поисковых запросов, связанных с биткоином.

Широко известно и семейство индексов сентимента для крипторынка от компании Sentix:

а) *Sentix Bitcoin Sentiment Index*. Данный индекс показывает месячные ожидания инвесторов для рынка биткоина и отражает эмоции жадности и страха участников рынка. Экстремальные негативные настроения обычно являются признаком последующего роста цен. Высокий оптимизм может быть предупреждающим сигналом о предстоящей консолидации и охлаждении рынка. Расхождение настроений, в основном, указывает на поворотные моменты на рынке и смену тренда.

б) *Bitcoin Overconfidence index*. Индекс чрезмерной уверенности показывает вероятность, что в данный момент времени ценовое движение на рынке будет воспринято инвесторами как тренд. Чем выше эта вероятность, тем более восприимчивыми становятся инвесторы к переоценке своих собственных способностей к прогнозированию. В случае такой «самоуверенности» вероятность разворота тренда возрастает, так как инвесторы склонны наращивать чрезмерные позиции. Экстремальные значения индекса чрезмерной уверенности (больше +7 или меньше -7) сигнализируют о рискованном рынке, на котором инвесторам следует сократить свои риски.

в) *Bitcoin Time differential index*[1]. Для рынка биткоина данный индекс рассчитывается как разница между краткосрочными настроениями и (среднесрочным) восприятием ценности актива. Высокие значения индикатора являются признаком позитивных настроений, которые не поддерживаются среднесрочной убежденностью инвесторов, что является фактором стресса для рынка. В свою очередь, низкие значения возникают, когда есть опасения на краткосрочной основе, но уверенность и росте стоимости в среднесрочной перспективе, что обычно благоприятно для рынка.

г) *Strategic bias on bitcoin*[2]. Индекс отражает стратегические взгляды участников рынка, а также их базовые убеждения и восприятие стоимости биткоина на горизонте 6 месяцев. Поскольку данный индекс отражает общую готовность инвесторов покупать или продавать, его не следует интерпретировать как сигнал к действию, он обычно опережает рынок на несколько недель. Поскольку индекс показывает, в основном, долгосрочные убеждения инвесторов и их восприятие ценности криптовалюты, он является индикатором коллективного мнения инвесторов.

Динамика индексов семейства Sentix представлена в Приложении Г на Рис. Г.7-Г.10.

Однако, основной отрицательной стороной перечисленных выше индексов сентимента является их концентрация только на биткоине (преобладающая доля биткоина в расчётах). Такие сегменты рынка как NFT и De-Fi демонстрируют слабую связь с поведением биткоина. Наиболее частым способом обойти данные ограничения в академической литературе выступает использование данных сервиса Google Trends.

Согласно официальным данным компании Google, сервис Google Trends делит частоту запроса на общее количество поисковых запросов в конкретном регионе за выбранное время. В соответствии с этим он определяет уровень интереса по шкале до 100 баллов. Оценку 100 получает максимальное количество за период, остальные ранжируются в зависимости от отношения к этому количеству. Оценка 0 означает, что запрос не популярный. На Рис. 4.7 приведена динамика популярности запросов, по ключевым словам «Crypto», «Bitcoin», «Ethereum» и «NFT».

Можно предположить, что использование Google Trends в качестве барометра настроений со временем будет становиться все менее эффективным. Это связано с тем, что все больше людей знакомятся с криптоиндустрией и знают, что такое биткоин и криптовалюты, необходимость в поиске информации о них уменьшается. Чем больше инвесторов будет ориентироваться на Google Trends при совершении сделок, тем меньше возможность для каждого инвестора воспользоваться ситуацией «страха» или «жадности» на рынке[3].

---

[1] https://www.crypto-sentiment.com/

[2] https://www.crypto-sentiment.com/

[3] https://bitcoinist.com/google-trend-sentiment-breaks-down-will-bitcoin-follow/

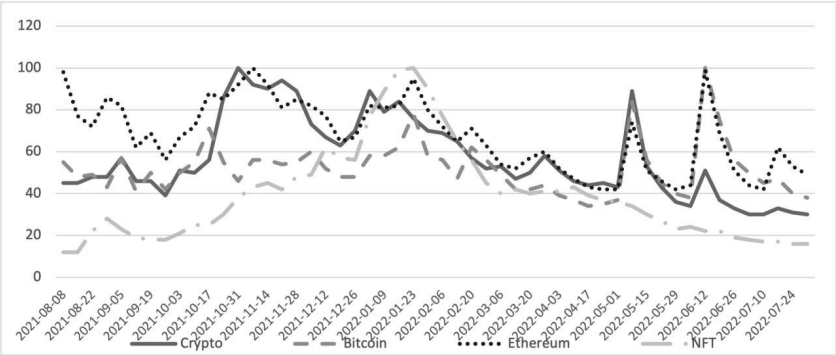

Рис. 4.7. Динамика популярности запросов по ключевым словам, связанным с криптоактивами (Google Trends)

Одной из потенциальных возможностей для обхода ограничений и недостатков Google Trends является использование данных платформы LunarCRUSH. LunarCRUSH предоставляет доступ к более чем 200 миллионам собранных социальных постов, которые классифицируются по криптовалютам с помощью машинного обучения и отображаются вместе с уникальными сведениями. LunarCRUSH собирает более 150 000 социальных постов с 20 000 ссылками в день. LunarCRUSH поддерживает более 2000 криптовалют. LunarCRUSH позволяет сравнивать различные метрики по токенам, чтобы определить, какие криптовалюты показывают лучшие результаты. Платформа собирает данные о наиболее влиятельных авторах, пишущих о NFT, их вовлеченности, частоте и степени влиянии на те или иные криптовалюты. Платформа рассчитывает ряд сентимент метрик, среди которых среднее значение сентимента, доля «бычьих» и «медвежьих» постов, социальная вовлеченность и некоторые другие. Плюсы использования данного набора данных состоят в использовании для методов машинного обучения текстовых данных, специфичных для криптосообщества. Самым важным для нашего исследования является метрика «рейтинг сентимента» (Sentiment Ranking). Этот показатель представляет собой среднее значение по шкале от 1 до 5 или от 0 до 100 (Relative Sentiment) настроения всех людей, упоминавших токен в публично доступных записях в Интернете, чтобы представить общее настроение инвесторов о данном токене. Записи ранжируются следующим образом в зависимости от тональности сообщения:

5/(75-100) = Сильно выраженный бычий тренд;
4/(55-74) = Бычий тренд;
3 /(45 – 55)= Нейтральный;
2 /(25-45) = Медвежий тренд;
1 / (0-24) = Сильно выраженный медвежий тренд.

В Приложении Г на Рис. Г.11-Г.16 представлены данные о количестве ежедневных записей в социальной сети Twitter (Twitter Posts), упоминаний в новостных источниках (News) и о рейтинге сентимента (Relative Sentiment) для криптовалют Bitcoin и Ethereum. Выбор Bitcoin и Ethereum обусловлен длительностью истории наблюдений и популярностью, однако подобная статистика может быть получена также для большого количества криптовалют (альткоинов) и ряда NFT проектов.

4.4. АВТОРСКИЕ ИНДЕКСЫ СЕНТИМЕНТА ДЛЯ РЫНКА КРИПТОВАЛЮТ И NFT

Авторами исследования «The cryptocurrency uncertainty index» был разработан новый индекс неопределенности криптовалют (UCRY – cryptocurrency uncertainty index), основанный на исследовании новостных источников. Разработанный индекс UCRY отражает два типа неопределенности: неопределенность цен криптовалют (UCRY Price) и неопределенность политики относительно криптовалют (UCRY Policy). Для сентимент-анализа авторы использовали материалы, содержащиеся в базе данных LexisNexis News & Business. Это отличает исследование от большинства предыдущих, которые опирались почти исключительно на крупные газеты (см., например, (Rice, 2020)). Обоснованием для использования более широкого спектра источников, включая, но не ограничиваясь, новостными лентами и стенограммами новостей СМИ, было признание «социального» аспекта криптовалют. Будучи новым явлением, криптовалюты стали предметом широкого обсуждения не только в традиционных, но и в альтернативных и социальных СМИ. Для построения индекса авторами отобраны переменные, указанные в Табл. 4.4.

Таблица 4.4

Объясняющие переменные для индексов UCRY Policy и UCRY Price

Переменная	*Описание*
UCRY Policy	Значение индекса неопределенности криптовалютой политики
Global EPU	Всемирный индекс неопределенности экономической политики
VIX	Индекс волатильности CBOE
Bitcoin	Цена криптовалюты Bitcoin
USFS	US Financial System stress index – индекс стресса финансовой системы США
USEPU	US economic policy uncertainty index - индекс неопределенности экономической политики США
Gold	Цена унции золота
UCRY Price	Значение индекса неопределенности цен криптовалют

Конструирование индекса

Индекс неопределенности политики в отношении криптовалют (The cryptocurrency policy uncertainty index – UCRY Policy) рассчитывается следующим образом (47):

$$UCRY\ Policy_t = \left(\frac{N_{1t} - \mu_1}{\delta_1} \right) + 100 \quad (47)$$

где $UCRY\ Policy_t$ - значение индекса неопределенности криптовалютной политики на недельном интервале в период с декабря 2013 г. по февраль 2021 г., N_{1t} - количество новостных статей на тему государственного регулирования криптовалют, которые были опубликованы за рассматриваемый недельный интервал и доступны в базе данных LexisNexis News & Business, μ_1 - среднее количество статей, δ_1 - стандартное отклонение.

Индекс неопределенности цен криптовалют (The cryptocurrency price uncertainty index – UCRY Price) рассчитывается аналогичным образом (48):

$$UCRY\ Price_t = \left(\frac{N_{2t} - \mu_2}{\delta_2} \right) + 100 \quad (48)$$

где $UCRY\ Price_t$ - значение индекса неопределенности цен криптовалют на недельном интервале в период с декабря 2013 г. по февраль 2021 г., N_{2t} - количество новостных статей на тему стоимости и изменения цен криптовалют, которые были опубликованы за рассматриваемый недельный интервал и доступны в базе данных LexisNexis News & Business, μ_2 - среднее количество статей, δ_2 - стандартное отклонение.

На Рис. 4.8 показаны еженедельные значения индексов неопределенности UCRY Policy и UCRY Price на основе 726,9 млн новостных статей, собранных за период 2013–2021 гг., на котором авторы также выделили основные изменения в соответствии с событиями на рынке криптовалют и в смежных областях экономики. Историческая декомпозиция индекса UCRY представлена в Приложении Г, Рис. Г.17. Вклад шоков политики UCRY в историческую декомпозицию индекса UCRY показан светло-голубым цветом, а вклад шоков цены UCRY - оранжевым. Эти шоки в определенной степени совпадают с ожиданиями общественности. Например, голосование по Brexit, победа Дональда Трампа на президентских выборах в США в 2016 г., запрет ICO в Китае, пузырь BTC, взлет DeFi и другие события положительно повлияли на индексы UCRY Policy и UCRY Price.

Авторы отразили не только положительные для криптовалют события, но также и отрицательные. Так, были показаны некоторые из крупнейших хакерских атак на криптовалютные биржи, такие как атаки на биржи Bitfinex, MintPal, Crispy и Dao. Они произошли с апреля 2014 г. по декабрь 2020 г., и авторами было доказано, что индексы неопределенности цен и политики UCRY отреагировали негативно на эти события. Корректировки

фискальной политики способствовали небольшим сдвигам в индексе UCRY Policy, однако значимость этих событий может возрасти в будущем. Декомпозиция также показала, что индексы UCRY отражают неопределенность, которая может быть более отчетливо отнесена к крупным событиям на рынке криптовалют по сравнению с VIX, EPU и глобальным индексом EPU. Хотя цена биткоина, индексы неопределенности политики UCRY и цен UCRY сильно коррелируют, эти индексы, как показывает разложение, отражают неопределенность, выходящую за рамки изменения цены биткоина. Наконец, кризис COVID-19 увеличил оба индекса неопределенности UCRY Policy и Price, поэтому их можно использовать в качестве эффективной меры неопределенности во время пандемии.

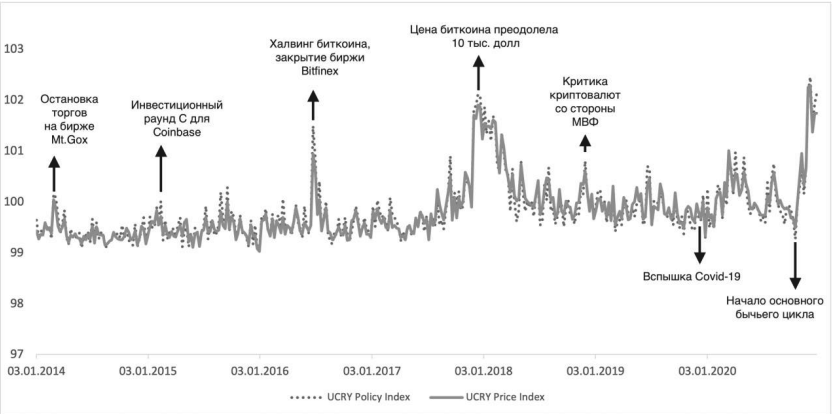

Рис. 4.8. Аннотированные расчетные значения индексов UCRY Policy и UCRY Price

В связи с тем, что ряд криптовалют используют энергоемкую процедуру майнинга для выполнения условия Proof–of–work, в публичной среде часто высказывается озабоченность по поводу воздействия криптовалют на окружающую среду. Также остается спорным вопрос о том, как внимание к экологии и озабоченность общественности негативно влияют на цены криптовалют. Ответы на данные вопросы пытаются найти авторы работы Wang et al. (2022). Основной целью работы является введение индекса экологического внимания к криптовалютам (Index of Cryptocurrency Environmental Attention - ICEA), который призван отразить относительную степень обсуждения в СМИ экологического воздействия криптовалют. Для исследования авторы собрали большое количество новостных сообщений, связанных с экологическими проблемами и критикой использования криптовалют, - 778,2 млн новостных статей из базы данных LexisNexis News & Business, что можно рассматривать как Big Data, и провели анализ этого обширного набора данных с использованием различных количественных

методов, таких как векторная модель коррекции ошибок (VECM) и структурная VECM (SVECM). Применение данных моделей полезно для характеристики динамических отношений между ICEA и экономической деятельностью.

В качестве объясняющих переменных были выбраны показатели неопределенности политики относительно криптовалют (UCRY), неопределенности цен криптовалют (UCRY), глобальный индекс экономической неопределенности (Global EPU), индекс волатильности VIX, стоимость нефти Brent, цена криптовалюты Bitcoin, индекс температуры GTU (global temperature uncertainty) и индекс производственной активности IP (industrial production index).

Методология построения индекса

Чтобы получить соответствующие новостные материалы, авторы выполнили следующие запросы в LexisNexis News & Business. Поисковая строка выглядит следующим образом:

("cryptocurrency" или "bitcoin" или "Ethereum") и atl1("energy" или "energy consumption" или "energy footprint" или "carbon footprint" или "environment" или "environmental" или "environmental impact" или "climate change")

К полученным необработанным данным впоследствии был применен процесс стандартизации и нормализации. Индекс рассчитывается по формуле (49):

$$ICEA_t = \left(\frac{N_{1t} - \mu_1}{\delta_1} \right) + 100 \qquad (49)$$

где $ICEA_t$ - значение индекса за неделю t в период 30.12.2013 г. по 02.05.2021 г.; N_{1t} - количество новостных статей, соответствующих поисковому запросу, опубликованных за рассматриваемый недельный интервал и доступных в базе данных LexisNexis News & Business, μ_1 - среднее количество этих статей, δ_1 - стандартное отклонение.

На Рис. Г.18 в Приложении Г представлены результирующие данные для значений индекса, а также выделен вклад каждой из переменной неопределенности. Помимо этого, авторы выделили основные события и факторы, связанные с изменениями в сфере криптовалют и экологической устойчивости.

Разложение дисперсии прогноза на индивидуальные вклады от экзогенных шоков показало, что на начальном этапе наблюдений переменная неопределенности политики по отношению к криптовалютам (UCRY Policy) вносила наибольший вклад в колебания сконструированного индекса ICEA (19,17%), в то время как переменные цены биткоина и неопределенности цен на криптовалюты (Price UCRY) вносили всего 4,71% и 0,187% соответственно. Эти результаты могут являться убедительным доказательством того, что экологические проблемы зародились в сфере политики и регулиро-

вания и до недавнего времени не интересовали инвесторов в криптовалюты, которых привлек этот класс активов из-за быстрого роста цен. Историческая декомпозиция ICEA показала более высокие связи между вниманием к экологии, ценой биткоина, переменной Policy UCRY и Price UCRY во время ключевых событий, которые значительно изменили цены цифровых активов, например, кибератаки на криптовалютные биржи, кризис COVID-19, бумы ICO и DeFi и периоды, похожие на пузырь биткоина. Поэтому можно сделать вывод, что общий рост внимания к экологическим последствиям использования криптовалют усилит колебания цен. Наконец, регрессионная модель OLS с панельной выборкой показывает, что ICEA положительно влияет на цену биткоина, цену Ethereum и индексы UCRY Policy и UCRY Price.

Работа Wang et al. (2022) посвящена широко обсуждаемому в последние годы проекту выпуска криптовалют центральными банками. К концу 2020 г. центральные банки по всему миру начали массово объявлять об интересе к развитию цифровых валют центральных банков (ЦВЦБ), и из точечных инициатив отдельных стран цифровые валюты стали общим трендом в развитии индустрии.

CBDC (ЦВЦБ) - это цифровые валюты, основанные на государственном кредите, что позволяет снижать их риски. Многие регуляторы и исследователи рассматривают CBDC как национально эмитированный «стейблкоин» (stablecoin) и считают, что такие валюты могут сбалансировать банковскую систему (Sissoko, 2020) и положительно влиять на финансовую стабильность (Larina & Akimov, 2020; Copeland, 2020; McLaughlin, 2021; Buckley et al., 2021). Однако, в настоящее время отсутствуют исследования взаимосвязи CBDC с финансовыми рынками. Одной из возможных причин этого исследовательского пробела является отсутствие переменных, которые можно было бы использовать как прокси временного ряда, связанного с CBDC.

В работе Wang et al. (2022) оценивается влияние новостей о CBDC на финансовые рынки с использованием более 660 млн новостных статей, собранных из базы данных LexisNexis News & Business. В частности, авторы вводят два новых показателя неопределенности и внимания к CBDC. Индекс неопределенности CBDC и индекс внимания CBDC были построены и доступны для периода с января 2015 г. по июнь 2021 г. В качестве объясняющих переменных авторы используют ранее описываемые индексы поведения криптовалютных рынков (например, индексы UCRY, ICEA и Bitcoin), другие популярные метрики неопределенности (например, индекс волатильности VIX и USEPU), индексы фондовых рынков (например, FTSE All-World Index), банковского сектора (например, MSCI World Bank Index), рынки облигаций (например, FTSE World Government Bond Index), значения обменных курсов основных пар валют (например, EUR/USD, GBP/USD, RUB/USD, JPY/USD и CNY/USD), а также цену золота.

Для сбора данных для построения индексов CBDCAI и CBDCUI Wang et al. (2022) произвели поиск и выгрузку из базы данных LexisNexis News &

Business, используя комбинации ключевых слов, относящихся к теме CBDC. Были выявлены наиболее распространенные валюты, и рассмотрено, какие официальные неанглийские термины для «цифровых валют центрального банка» используются в странах-эмитентах данных валют. Таким образом, поисковые термины были переведены на китайский, японский, русский языки, чтобы обеспечить полный охват историй в основных странах, которые лидируют в развитии CBDC. Учитывая, что испанский, португальский, французский и немецкий являются основными языками в странах ЕС, авторы также перевели термин «цифровая валюта центрального банка» на эти четыре языка. Кроме того, поскольку CBDC является разновидностью цифровой валюты, а некоторые страны оценивают CBDC как инструмент противодействия криптовалютам, авторы включили термин «цифровая валюта» в качестве еще одной ключевой фразы для поиска. Поскольку USD, EUR, GBP, CHF, RUB, JPY и CNY также могут быть использованы в контексте CBDC, авторы заменили ключевые слова «валюта» или «деньги» на официальные названия этих валют. Компиляция этих ключевых поисковых терминов вместе составили поисковую строку для использования в LexisNexis News & Business.

Авторы отмечают, что одним из недостатков построения индекса на основе любого литературного архива является то, что статьи поступают в архив и покидают его, поэтому общий объем статей может варьироваться по источникам публикаций и времени. Поэтому процедуры стандартизации и нормализации должны проводиться в соответствии с исходными данными подсчета, так как это позволяет сортировать данные по одной и той же шкале. Например, CBDCUI оценивает наблюдаемое значение количества новостных статей за каждую неделю по количеству статей, соответствующих поисковой строке за ту же неделю. Затем ряд стандартизируется, чтобы получить набор данных временного ряда в качестве исходного индекса. Наконец, исходный индекс нормализуется путем добавления среднего значения 100 для устранения потенциальных отрицательных значений. Таким образом, для конструирования индекса используется ранее описанная формула.

Еженедельные индексы CBDCUI и CBDCAI представлены на Рис. 4.9 и показывают, какие события вызывали скачки индексов. Значения индексов получены на основе 663 млн новостей, собранных с января 2015 г. по июнь 2021 г. Согласно (Turrin, 2021), Эквадор был первой страной, запустившей CBDC в феврале 2015 г. для продвижения дедолларизации экономики. Именно поэтому январь 2015 г. был выбран в качестве начала периода наблюдения. График доказывает, что рост значений индекса могут вызвать не только новые события, связанные с CBDC, но также значительные события, связанные с криптовалютами.

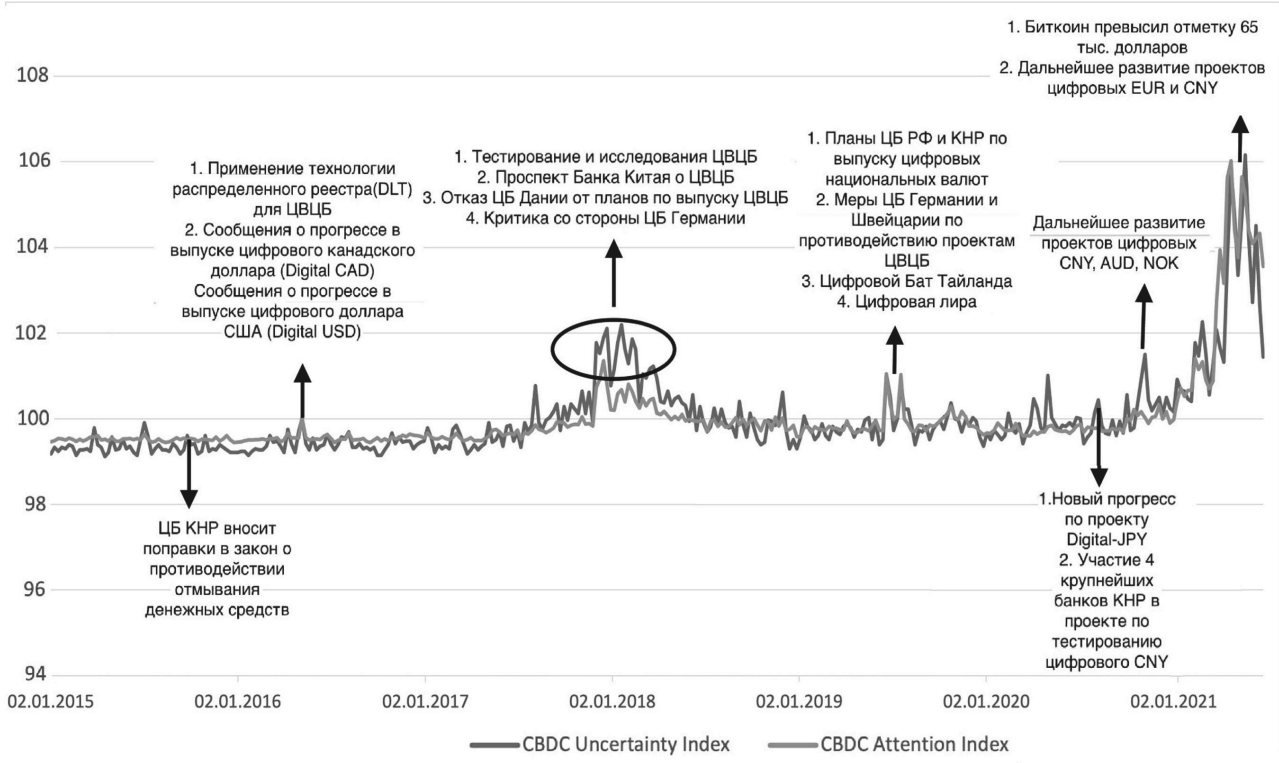

Рис. 4.9. Аннотированные расчетные значения индексов CBDCUI и CBDCAI

Эмпирические результаты показывают, что индексы CBDC оказывают значимое отрицательное влияние на волатильность индексов MSCI World Banks Index, USEPU и FTSE All-World Index.

Однако, индексы CBDC оказывают значимое положительное влияние на волатильность UCRY Policy, UCRY Price, ICEA и Bitcoin (криптовалютные рынки), FTSE World Government Bond Index (рынки облигаций), EUR/USD, GBP/USD, RUB/USD, JPY/USD и CNY/USD (валютные рынки), а также VIX и цену золота.

Более того, волатильность финансовых переменных более чувствительна к CBDCUI по сравнению с реакцией на шоки CBDCAI, что подчеркивает важность неопределенности CBDC в этой взаимосвязанной системе.

Кумулятивные положительные и отрицательные эффекты возмущений CBDCUI на финансовые переменные больше, чем эффекты возмущений CBDCAI.

Эти результаты показывают, что неопределенность вокруг новостей о CBDC играет более важную роль, чем просто внимание к этим новым цифровым активам, что говорит о том, что введение CBDC может принести значительные изменения в экономику.

В работе Wang (2022) изучается связь переливов волатильности между рынком NFT-активов и традиционными финансовыми рынками.

Для изучения данной зависимости в работе предлагается новый прямой показатель внимания инвесторов к рынку NFT: индекс внимания к NFT токена (NFTsAI - NFT attention index) основанный на 590 млн. новостных записей из базы данных LexisNexis News & Business Database.

Авторы демонстрируют результаты эмпирического анализа с помощью модели волатильности TVP-VAR. Основываясь на предыдущих исследованиях в данной сфере, авторы выдвигают две гипотезы:

H1: Индекс NFTsAI имеет связь с рынками NFT.

H2: Интенсивность и величина эффекта от перетока волатильности с других финансовых рынков на рынок NFT, выше, чем с рынка NFT на другие финансовые рынки.

Для сбора базы текстов использовалось ПО LexisNexis News & Business, отметим, что этот мультирегиональный и многоязычный источник охватывает последние новостные статьи, архивы публикаций и блоги.

Итоговый вид поискового запроса, используемого авторами, представлен в Табл. 4.5.

Таблица 4.5

Поисковый запрос по тематике NFT и смежным темам

Тема 1	*Тема 2*	*Тема 3*	*Тема 4*	*Тема 5*
«Non-fungible tokens» «NFTs»	«digitable art» «crypto art» «cryptocurrency art»	«digital collectibles» «crypto collectibles»	«digital identity» «IdToken» «token unique»	«CryptoKittles» «WCK» «CryptoPunks» «Axie Infinity»

Тема 1	*Тема 2*	*Тема 3*	*Тема 4*	*Тема 5*
	«artwork tokenised» «digital image licensing»	«cryptocurrency collectibles»	«unique digital property»	«Bored Ape Yacht Club» «The Sandbox» «Art Blocks» «nonfungible. com»

Источник: Wang (2022)

В качестве отправной точки для сбора новостных данных для построения индекса NFTsAI было выбрано 01.01.2017 г., поскольку, согласно данным платформы nonfungible.com, в 2017 г. было выпущено много коллекций NFT, ставших впоследствии популярными (например, Curio Card, CryptoPunks, Moon Cats и Decentraland).

В качестве объясняющих переменных была выбрана средняя цена коллекций Decentraland и CryptoPunks, как двух наиболее ликвидных и известных активов NFT, которые могли бы выступить в качестве прокси переменных рынка NFT.

В качестве переменных рынка DeFi выбраны Chainlink (LINK) и Maker (MKR). Выбор осуществлялся на основании объема торгов и наличия длительной истории торгов.

В качестве основной переменной для представления рынка криптовалют был выбран Bitcoin (BTC), поскольку он является самой популярной и ликвидной криптовалютой (биткоин имеет самую высокую цену, объем продаж и рыночную капитализацию).

Помимо этого, биткоин обычно рассматривается как прокси показатель для всего рынка криптовалют.

В качестве еще одного криптовалютного актива автор использует цену криптовалюты Ethereum.

BGCI Bloomberg Galaxy Crypto Index предназначен для оценки показателей крупнейших криптовалют, торгуемых в долларах США (Umar& Gubareva, 2020).

Поскольку одна из гипотез исследования состоит в установлении взаимосвязи между рынком NFT-активов и финансовыми рынками, в исследовании рассматриваются инструменты традиционных финансовых рынков:

- рынок акций (FTSE All-World Index),
- рынок облигаций (FTSE World Government Bond Index и PIMCO Corporate & Income Strategy Fund),
- товарный рынок (Invesco DB Commodity Index Tracking Fund [DBC]),
- валютный рынок (U.S. Dollar Index)
- и рынок золота (COMEX Gold).

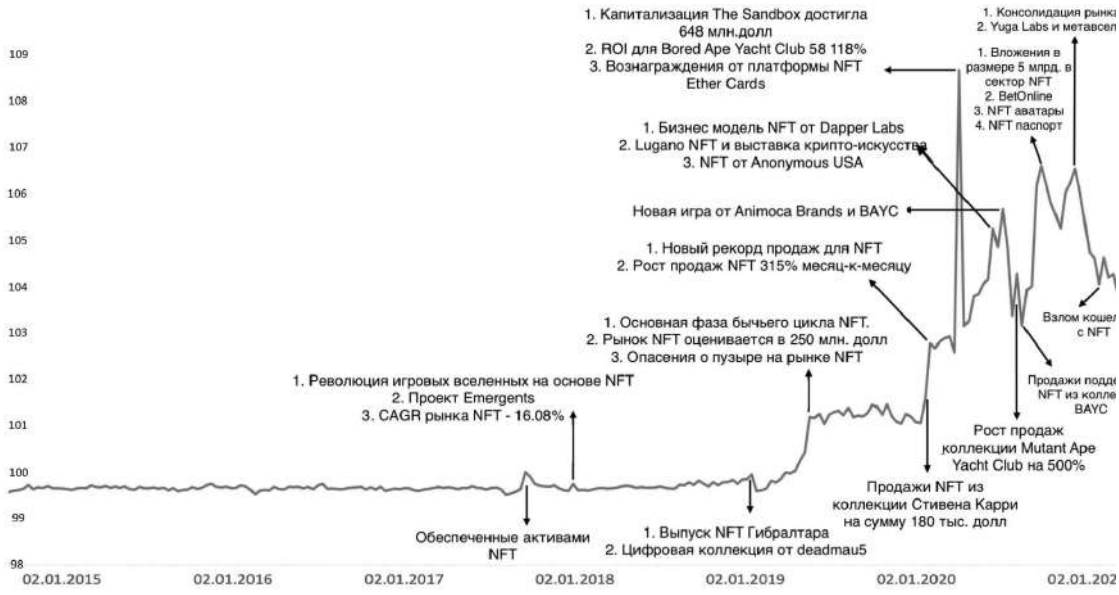

Рис. 4.10. Аннотированные расчетные значения индекса NFTsAI

Индекс внимания для рынка NFT рассчитывается по аналогичной другим работам формуле (50):

$$NFTsAI_t = \left(\frac{N_{1t} - \mu_1}{\delta_1} \right) + 100 \qquad (50)$$

где $NFTsAI_t$ - значение индекса внимания к NFT в неделю t в интервале с января 2017 г. по май 2022 г., N_t - количество новостных статей по тематике NFT, которые были опубликованы за рассматриваемый недельный интервал и доступны в базе данных LexisNexis News & Business, μ - среднее количество статей, отражающих внимание к рынку NFT в период с 26 декабря 2016 г. по 5 июня 2022 г., σ – рассчитанное значение стандартного отклонения. Добавление значения 100 призвано устранить потенциальное негативное воздействие, вызванное тем, что общий объем статей варьируется в зависимости от источников публикации и временного периода.

После построения и декомпозиции индекса (Рис. Г.19 Приложения Г) были сделаны следующие выводы:

1) Динамика цен CryptoPunks и Decentraland коррелирует с NFTsAI. Чем выше внимание к NFT, тем выше волатильность активов NFT. NFTsAI может служить косвенным индикатором процессов, происходящих на рынке NFT. Таким образом, H1 подтверждается.

2) Исторические вариации значений NFTsAI в достаточной степени подтверждают гипотезу авторов. Положительные новости, касающиеся рынков NFT, вызывают положительный шок для исторических вариаций NFTsAI, а отрицательные новости, касающиеся рынков NFT, вносят отрицательный шок в изменения NFTsAI.

3) Анализ динамики NFTsAI показывает, что рынки NFT-активов обладают относительно независимой и изолированной природой волатильности по сравнению с другими финансовыми рынками. Другими словами, рынки NFT в меньшей степени подвержены влиянию криптовалют, DeFi, акций, облигаций, товаров, валютных пар и золота, что может играть положительную роль в диверсификации инвестиций.

4.5. ЭМПИРИЧЕСКИЙ АНАЛИЗ ФАКТОРОВ, ВЛИЯЮЩИХ НА РЫНОК NFT-АКТИВОВ

Гипотезы исследования

В нашей авторской работе мы эмпирически протестировали несколько гипотез (за основу взята ВКР А. Куркина на магистерской программе ФРФИ под руководством Т.В. Тепловой на ФЭН НИУ ВШЭ):

Гипотеза 1. Рост цен на криптовалюты и, соответственно, повышение их доходности увеличивает интерес инвесторов к цифровым активам, что отражается в положительной ценовой динамике на рынке NFT.

Поскольку NFT является близкородственным для криптовалют сегментом активов, то целесообразно предположить, что прошлая доходность таких криптовалют, как Ethereum и Bitcoin, будет привлекать внимание к рынку цифровых активов в целом, что будет отражаться в положительной динамике доходности NFT (Wang, 2022; Ante, 2022).

Гипотеза 2. Рост доходности криптовалют отражается в росте доходности NFT рынка не мгновенно, а с определенным лагом, который может достигать нескольких дней.

Предшествующие работы по данной тематике (Pinto-Gutiérrez, 2022; Dowling, 2021) на недельных данных выявили, что эффект от роста стоимости криптовалют отражается на рынке NFT с лагом в одну неделю.

Гипотеза 3. Существует зависимость между индексами сентимента инвесторов и индексом рынка NFT-активов, что отражается в росте цен на NFT-активы по мере роста «бычьих» настроений среди инвесторов, и наоборот.

На текущем этапе развития рынка цифровых активов NFT является крайне волатильным и, в значительной мере, спекулятивным активом. Как отмечалось ранее, большинство сделок на рынке NFT совершается частными неквалифицированными инвесторами. Так, до 80% всех продаж NFT-рынка были осуществлены в ценовом сегменте стоимостью до 10 000 долларов частными инвесторами, и только 20% продаж приходится на долю крупных частных инвесторов и институциональных инвесторов, причем доля последних по оценкам не превышает 1–2%. Иными словами, можно говорить о высокой степени иррациональности, присущей рынку NFT. На основании предшествующих работ (Kraaijeveld & Smedt, 2020), посвященных рынку криптовалют, для которого также характерна высокая доля розничных инвесторов, можно сделать предположение о том, что на поведение инвесторов рынка NFT влияют не фундаментальные факторы, а общее настроение на финансовых рынках, то есть в благоприятный период будут преобладать «бычьи» настроения инвесторов, что отразится в положительной динамике рынка NFT, и, наоборот, в периоды «медвежьего» рынка доходность NFT-рынка будет снижаться.

Гипотеза 4. Ценовые колебания на товарных рынках, таких как золото и нефть, отражаются на динамике цен NFT-активов.

Поскольку большинство блокчейнов, на которых выпущены NFT, основано на концепции Proof-of-Work, они являются очень энергоемкими, что, в свою очередь, должно ставить их в зависимость от цены на энергоресурсы, в частности, на нефть. Иными словами, чем ниже стоимость нефти, тем менее затратно поддержание работы системы и, соответственно, тем больше доходность.

Золото часто воспринимается инвесторами как один из главных защитных активов во времена нестабильности рынка и высокой волатильности. Более высокая биржевая стоимость золота сигнализирует о преобладании медвежьих настроений на рынке и выходе из рисковых и альтернативных активов, к которым относятся криптовалюты и NFT.

Гипотеза 5. Существует взаимосвязь между присутствием избыточной рыночной ликвидности и ростом стоимости NFT-активов.

NFT, как и криптовалюты, может выступать в качестве поглотителя избыточной ликвидности на рынке. Можно предположить наличие взаимосвязи между увеличением ликвидности на рынке и ростом стоимости NFT-активов. (Corbet et al., 2022). В качестве прокси-показателя ликвидности была выбрана переменная изменения величины баланса федеральной резервной системы США.

Описание данных

Для анализа и проверки гипотез был собран набор данных, включающий показатели доходности NFT-рынка, криптовалют, финансовых индексов, а также фьючерсов на нефть и золото. Данные для исследования были получены с помощью терминалов Refinitiv и Bloomberg (до марта 2022 г.), а также сервиса Yahoo Finance (с апреля по июнь 2022 г.). Показатели сентимента были получены со страниц авторов данных сентимент-индексов. Данные содержат временные ряды с 01.07.2017 г. до 30.06.2022 г. Детальное описание данных представлено в Табл. 4.6.

Таблица 4.6

Описание исследуемых переменных

Переменная	*Описание*
Average_Price	Средняя стоимость проданных NFT-активов за недельный интервал на основании данных сервиса https://nonfungible.com/
Криптовалюты	
BTC	Стоимость криптовалюты Bitcoin (BTC) в долл. США
ETH	Стоимость криптовалюты Ethereum (ETH) в долл. США
Финансовые индексы	
VIX	Значения индекс волатильности Чикагской фондовой биржи (CBOE)
Nasdaq	Значения индекса высокотехнологичных компаний Nasdaq
Fed_Bal_change	Изменение баланса Федерального резерва США в процентах: $\ln\left(\frac{Fed_balance_t}{Fed_balance_{t-1}}\right)\cdot 100\%$
Ресурсы	
Brent	Стоимость фьючерсных контрактов на нефть марки Brent
Gold	Стоимость фьючерсных контрактов на унцию золота
Индексы сентимента	
NFTsAI	Индекс внимания инвесторов к рынку NFT-активов. Индекс отнормирован около значения 100 для предотвращения отрицательных значений. Значение индекса, равное 100, показывает нейтральное отношение инвесторов, значения индекса выше 100 говорят о росте интереса инвесторов к NFT-активам, значения индекса ниже 100 говорят о

Переменная	*Описание*
	снижении интереса инвесторов к NFT-активам
UCRY_Policy	Индекс неопределенности политики относительно криптовалют. Индекс отнормирован около значения 100 для предотвращения отрицательных значений. Значение индекса, равное 100, показывает нейтральное отношение со стороны органов власти и инвесторов, значения индекса выше 100 говорят о положительных настроениях инвесторов и снижении неопределенности, значения индекса ниже 100 говорят о росте неопределенности мнений относительно криптовалют и ухудшении настроений инвесторов

Графики динамики временных рядов рассматриваемых переменных представлены на Рис. Г.20 Приложения Г. Можно отметить существенный скачок в стоимости как криптовалют, так и рынка традиционных финансовых инструментов в 2020–2021 гг. На Рис. Г.21 представлена корреляционная матрица для рассматриваемых переменных.

Табл. 4.7 содержит описательную статистику для рассматриваемых переменных. Для проверки стационарности временных рядов был использован расширенный тест Дики-Фуллера (ADF). Проверка временных рядов на стационарность имеет важное значение, поскольку нестационарные данные могут привести к ложным результатам регрессии.

Таблица 4.7

Описательная статистика рассматриваемых переменных

	Среднее	*Медиана*	*Минимум*	*Максимум*	*Стандартное отклонение*
Average_USD	361.37	40.835	1.7700	3014.2	601.43
BTC	19124.0	9714.4	1929.8	65467	17564.0
ETH	1197.8	480.22	85.262	4626.4	1260.1
Brent	66.411	65.370	21.440	122.01	18.632
Gold	1563.3	1522.8	1176.5	2010.1	260.23
Nasdaq	9966.0	8571.5	6216.5	16057.0	2988.2
VIX	19.621	17.385	9.1400	66.040	8.5688
NFTsAI	100.65	99.677	99.508	108.67	1.8977
UCRY Policy	101.61	100.23	99.271	114.65	2.7976
Fed_bal_ change	0.00266	0.00048	-0.01268	0.11827	0.012765

Результаты, представленные в Табл. 4.8, показывают, что только для некоторых из исследуемых рядов на основании теста ADF можно сделать вывод о нестационарности. Для решения проблемы нестационарности было решено использовать значения первой разности для ряда переменных. По-

вторное тестирование переменных с помощью ADF – теста говорит о стационарности временных рядов.

Таблица 4.8

Результаты ADF-теста для исследуемых переменных

	ADF тест	*ADF, первая разность*
Average_USD	-4.0436*	-4.0436*
BTC	-1.4470	-13.9740*
ETH	-1.8289	-14.8995*
Brent	-0.7505	-16.0817*
Gold	-2.2171	-18.7467*
Nasdaq	-1.0474	-16.4553*
VIX	-4.3658*	-4.3658*
NFTsAI	-1.9496	-16.2476*
UCRY Policy	-3.3742	-12.8448*
Fed_bal_change	-6.0078*	-6.0078*

Методология исследования

Методология проверки гипотез базируется на VAR-модели (Pinto-Gutiérrez, 2022), в указанной работе исследуются факторы, которые могли бы объяснить увеличение внимания инвесторов к NFT-активам. По мнению авторов, рост цен на криптовалюты в 2020–2021 гг. мог сыграть главную роль во всплеске популярности цифровых активов и привлечь внимание к NFT и, как следствие, привести к взрывному росту цен на NFT-активы. В качестве основного метода исследования авторами использовались VAR-модели с экзогенными переменными. Эти экзогенные переменные включают экономические факторы, которые также могут определять внимание инвесторов к рынку NFT-активов. Например, авторы используют такие переменные, как индекс волатильности CBOE (VIX), цена на золото и доходность индекса S&P 500. Авторы также оценивают уровень внимания к Bitcoin и Ethereum. Модель VAR основывается на уравнениях (51), (52):

$$NFTattention_t = \alpha + \sum_{j=1}^{p} \beta' NFTattention_{t-j} + \sum_{j=1}^{p} \gamma' Cryptoreturn_{t-j} + \delta' Z_{t-1} + \mu_t \qquad (51)$$

$$Cryptoreturn_t = \alpha + \sum_{j=1}^{p} \beta' NFTattention_{t-j} + \sum_{j=1}^{p} \gamma' Cryptoreturn_{t-j} + \delta' Z_{t-1} + \mu_t \qquad (52)$$

где α - константа, β - вектор коэффициентов для первой эндогенной переменной (коэффициента внимания инвесторов к NFT), γ - вектор коэффициентов для второй эндогенной переменной (еженедельная доходность Bitcoin или Ethereum). Вектор Z*t* представляет собой экзогенные контрольные переменные, δ - вектор коэффициентов на эти контрольные переменные, μ – случайная ошибка. В уравнениях (50) и (51) величина *p* обозначает

количество лагов. Pinto-Gutiérrez (2022) определяют оптимальное число лагов, используя несколько информационных критериев, включая информационный критерий Акаике (AIC).

В контексте нашей исследовательской работы предлагается следующая трансформация данной модели:

1) В качестве основной зависимой переменной выбрана переменная средней стоимости NFT-активов (Average_Price). Для исследования была выбрана средняя стоимость NFT-активов (для рынка в целом), а не совокупная стоимость продаж NFT-активов за день в связи с тем, что при использовании совокупного объема продаж будет наблюдаться большой скос в сторону единичных, но крупных продаж, таких как, например, активов из коллекций CryptoPunks или Bored Ape Yacht Club. Причины роста стоимости данных коллекций, как правило, мало связаны с рынком NFT в целом, а обусловлены внутренними успехами проекта. Использование индикатора среднедневной стоимости NFT позволить избежать выбросов в данных.

2) В качестве стоимости криптоактивов будет выступать стоимость криптовалют Bitcoin (BTC) и Ethereum (ETH).

3) В качестве показателя сентимента инвесторов будут использованы индексы NFTsAI и UCRY Policy, представленные в работах Lucey et al. (2021), Wang (2022).

4) В качестве экзогенных переменных для модели VAR будут использованы значения индекса технологических компаний, цены фьючерсов на золото (Gold) и нефть (Brent), значения индекса волатильности VIX, а также изменение баланса ФРС США на еженедельной основе (*Fed_bal_change*), как прокси-переменная избыточной рыночной ликвидности.

Таким образом, в данной работе для исследования взаимосвязи между средней стоимостью NFT и другими инвестиционными активами использована VAR регрессия с экзогенными переменными, решающая задачу структурного анализа зависимостей между несколькими переменными.

Мы моделируем векторы переменных, зависящие от собственных лагов и лагов эндогенных переменных, а также показателей эндогенных переменных.

Поскольку в работе планируется оценить изолированно влияние каждого из индексов сентимента, а также криптовалют Bitcoin и Ethereum, построены четыре спецификации базовой модели (53)-(56).

Модель 1. Индекс NFTsAI. Криптовалюта BTC. Порядок лага 2

$$\begin{bmatrix} Average_USD \\ BTC \\ Nasdaq \\ NFTsAI \end{bmatrix} = A_0 + A_1 \cdot \begin{bmatrix} Average_USD_{t-1} \\ BTC_{t-1} \\ Nasdaq_{t-1} \\ NFTsAI_{t-1} \end{bmatrix} + A_2 \cdot \begin{bmatrix} Average_USD_{t-2} \\ BTC_{t-2} \\ Nasdaq_{t-2} \\ NFTsAI_{t-2} \end{bmatrix} + A_3 Gold +$$

$$+ A_4 Brent + A_5 VIX + A_6 Fed_bal_change + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \end{bmatrix} \quad (53)$$

где A_0 - константа регрессии, A_1-A_2 - коэффициенты - матрицы 4*4, *Gold*, *Brent*, *VIX*, *Fed_bal_change* – экзогенные переменные, A_3-A_6 - коэффициенты при экзогенных переменных, $\varepsilon_{1,t},...,\varepsilon_{4,t}$ - случайные ошибки регрессии.

Модель 2. Индекс NFTsAI. Криптовалюта ETH. Порядок лага 2

$$\begin{bmatrix} Average_USD \\ ETH \\ Nasdaq \\ NFTsAI \end{bmatrix} = A_0 + A_1 \cdot \begin{bmatrix} Average_USD_{t-1} \\ ETH_{t=1} \\ Nasdaq_{t-1} \\ NFTsAI_{t-1} \end{bmatrix} + A_2 \cdot \begin{bmatrix} Average_USD_{t-2} \\ ETH_{t-2} \\ Nasdaq_{t-2} \\ NFTsAI_{t-2} \end{bmatrix} + A_3 Gold +$$

$$+ A_4 Brent + A_5 VIX + A_6 Fed_bal_change + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \end{bmatrix} \quad (54)$$

Модель 3. Индекс UCRY Policy. Криптовалюта BTC. Порядок лага 2

$$\begin{bmatrix} Average_USD \\ BTC \\ Nasdaq \\ UCRY\ Policy \end{bmatrix} = A_0 + A_1 \cdot \begin{bmatrix} Average_USD_{t-1} \\ BTC_{t-1} \\ Nasdaq_{t-1} \\ UCRY\ Policy_{t-1} \end{bmatrix} + A_2 \cdot \begin{bmatrix} Average_USD_{t-2} \\ BTC_{t-2} \\ Nasdaq_{t-2} \\ UCRY\ Policy_{t-2} \end{bmatrix} + A_3 Gold +$$

$$+ A_4 Brent + A_5 VIX + A_6 Fed_bal_change + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \end{bmatrix} \quad (55)$$

Модель 4. Индекс UCRY Policy. Криптовалюта ETH. Порядок лага 2

$$\begin{bmatrix} Average_USD \\ ETH \\ Nasdaq \\ UCRY\ Policy \end{bmatrix} = A_0 + A_1 \cdot \begin{bmatrix} Average_USD_{t-1} \\ ETH_{t-1} \\ Nasdaq_{t-1} \\ UCRY\ Policy_{t-1} \end{bmatrix} + .A_2 \cdot \begin{bmatrix} Average_USD_{t-2} \\ ETH_{t-2} \\ Nasdaq_{t-2} \\ UCRY\ Policy_{t-2} \end{bmatrix} + A_3 Gold +$$

$$+ A_4 Brent + A_5 VIX + A_6 Fed_bal_change + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \end{bmatrix} \quad (56)$$

Результаты регрессионного анализа

Полученные оценки по двум моделям показаны в Табл. 4.9. В модели 1 для зависимой переменной *Average_USD* значения ее лаговых переменных значимы на 99% уровне доверия до второго лага. Значимы оценки для стоимости криптовалюты Bitcoin (первый и второй лаги) и для индекса высокотехнологичных компаний Nasdaq (второй лаг). Что касается индекса внимания инвесторов к рынку NFT (*NFTsAI*), то для обоих лагов регрессии получены статистически значимые оценки. В модели 2 для зависимой переменной Ethereum практически все эндогенные переменные, кроме переменной индекса Nasdaq, при регрессионном анализе оказались незначимы.

Таблица 4.9

Результаты VAR-регрессий для модели 1–2

VAR регрессия	*Модель 1 (BTC)*		*Модель 2 (ETH)*	
	Коэффициент	*t-статистика*	*Коэффициент*	*t-статистика*
Const	−54.4230	−1.198	−45.3821	−0.8685
AverageUSD_1	0.555937***	9.150	0.591753***	9.440
AverageUSD_2	0.358110***	5.841	0.260091***	4.117
d_BTC_1/ d_ETH_1	0.0461619***	7.796	−0.0900164	−1.231
d_BTC_2/ d_ETH_2	0.0223279***	3.386	−0.0454779	−0.6269
d_NFTsAI_1	−49.0264*	−1.728	−7.48273	−0.2319
d_NFTsAI_2	−83.3255***	−2.942	−48.2010	−1.506
d_Nasdaq_1	−0.0124250	−0.2307	0.118829**	1.979
d_Nasdaq_2	−0.0912611*	−1.677	−0.0142057	−0.2321
d_Brent	−2.95968	−0.6340	−5.30056	−0.9922
d_Gold	0.270919	0.5825	0.229423	0.4303
Fed_bal_change	−764.659	−0.5276	96.1861	0.05796
VIX	0.280548	0.1103	−1.35852	−0.4644
Time	0.661405**	2.056	0.990962***	2.705

VAR регрессия	*Модель 1 (BTC)*		*Модель 2 (ETH)*	
	Коэффициент	*t-статистика*	*Коэффициент*	*t-статистика*
Кол-во наблюдений	260		260	
R^2	0.858942		0.814798	
R^2 adj	0.851396		0.804891	

На Рис. 4.11 и 4.12 представлены импульсные отклики зависимой переменной *Average_USD* на шоки эндогенных переменных для моделей 1 и 2 соответственно. Говоря о результатах, полученных с помощью функции импульсного отклика, было обнаружено, что однократное повышение цен на BTC в результате шока стандартного отклонения положительно влияет на продажи NFT. Шоковое значение для индекса Nasdaq, наоборот, оказывает отрицательное воздействие. Данный эффект, вероятно, связан с перетоком денежных средств из области цифровых активов в область высокотехнологичных компаний в случае роста последней. Удивительно, но для цены ETH характерен обратный эффект. Для индекса NFTsAI также характерно отрицательное влияние на среднюю стоимость NFT-активов, причем наибольшее влияние характерно для недель 3–5.

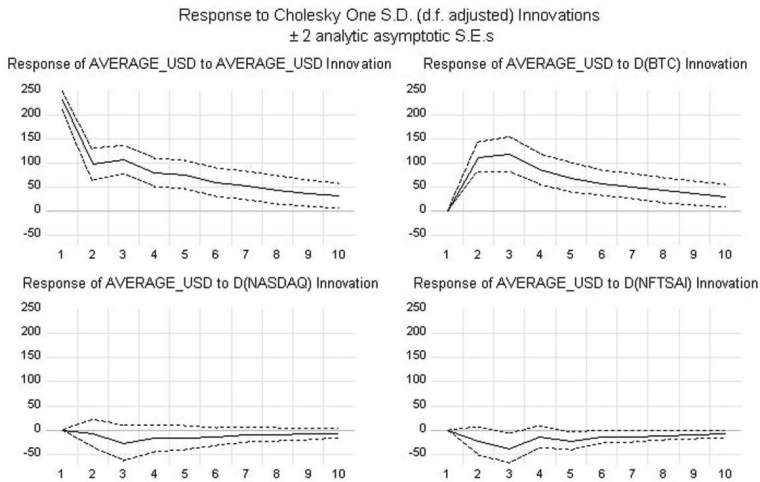

Рис. 4.11. Модель 1. Импульсный отклик переменной Average_USD на шоки эндогенных переменных

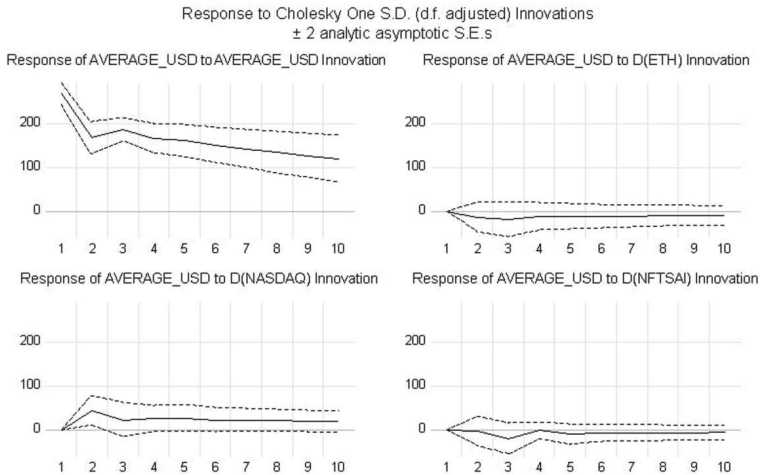

Рис. 4.12. Модель 2. Импульсный отклик переменной Average_USD на шоки эндогенных переменных

Результаты регрессионного анализа для моделей 3 и 4 представлены в Табл. 4.10. В модели 3 для зависимой переменной *Average_USD* значения ее собственных лаговых переменных значимы на 99% уровне доверия (первый и второй лаги). Значимы оценки для стоимости криптовалюты Bitcoin (первый и второй лаги) и индекса высокотехнологичных компаний Nasdaq (второй лаг). Что касается индекса внимания инвесторов к рынку NFT (*NFTsAI*), то оценки для данной переменной значимы для второго лага. В модели 4 для зависимой переменной *Ethereum* практически все эндогенные переменные, кроме индекса Nasdaq, при регрессионном анализе оказались незначимы.

Таблица 4.10

Результаты VAR-регрессий для модели 3–4

VAR регрессия	*Модель 1 (BTC)*		*Модель 2 (ETH)*	
	Коэффициент	*t-статистика*	*Коэффициент*	*t-статистика*
Const	−59.6919	−1.308	−53.3484	−1.033
AverageUSD_1	0.550437***	9.030	0.573993***	9.222
AverageUSD_2	0.359395***	5.838	0.282333***	4.488
d_Nasdaq_1	−0.0377741	−0.6920	0.0816265	1.355
d_Nasdaq_2	−0.100668*	−1.839	−0.0370150	−0.6078
d_BTC_1/ d_ETH_1	0.0449497***	7.442	−0.0843237	−1.167
d_BTC_2/	0.0164896**	2.461	−0.0245309	−0.3430

VAR регрессия	*Модель 1 (BTC)*		*Модель 2 (ETH)*	
d_ETH_2				
d_UCRYPolicy_1	−14.4650	−1.009	−33.3386**	−2.092
d_UCRYPolicy_2	−38.9391***	−2.643	−40.5853**	−2.481
VIX	0.787635	0.3079	−0.612301	−0.2114
Fed_bal_change	−1055.72	−0.7234	−361.247	−0.2197
d_Brent	−1.38227	−0.2945	−3.36816	−0.6364
d_Gold	0.174349	0.3733	0.234757	0.4457
Time	0.651423**	2.018	0.958671***	2.648
Кол-во наблюдений	260		260	
R^2	0.857995		0.819157	
R^2 adj	0.850398		0.809482	

Результаты функции импульсного отклика переменной *Average_USD* на шок переменной *UCRY Policy* свидетельствуют о первоначальном незначительном отрицательном воздействии (Рис. 4.13, 4.14). После недели 3 имеет место незначительное отрицательное воздействие на среднюю стоимость NFT-активов. Как и в моделях 1-2, шок цены биткоина оказывает значительное положительное воздействие на среднюю стоимость NFT на дополнительную величину до 100-120 долл. США. Шок цены *Ethereum*, напротив, не оказывает практически никакого влияния.

В Приложении Г на Рис. Г.22-Г.29 представлены импульсные отклики всех переменных для моделей 1-4.

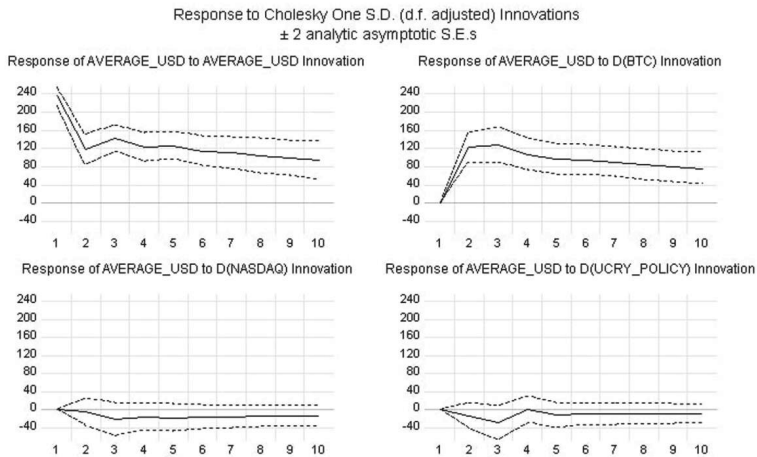

Рис. 4.13. Модель 3. Импульсный отклик переменной Average_USD на шоки эндогенных переменных

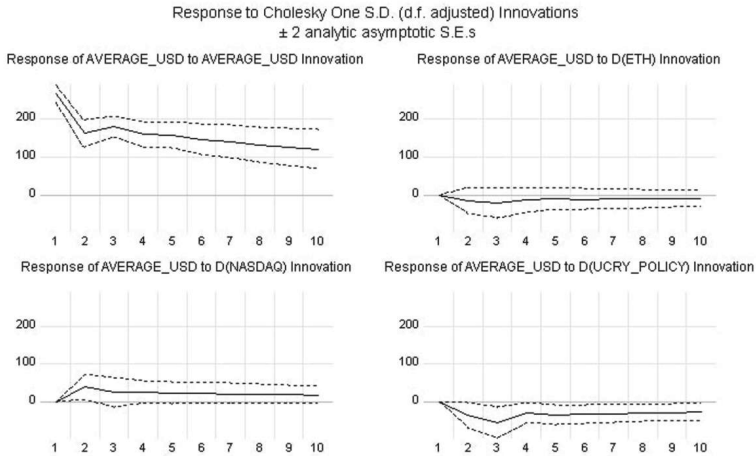

Рис. 4.14. Модель 4. Импульсный отклик переменной Average_USD на шоки эндогенных переменных

На основании проведенных тестов можно сделать следующие выводы:

1) Поскольку все переменные или их разности стационарны, а обратные корни оцененных моделей VAR по модулю меньше единицы и находятся внутри единичного круга, следовательно, все оцененные VAR-модели стационарны.

2) С помощью коррелограммы остатков и LM-теста было оценено наличие сериальной корреляции остатков определенного порядка VAR-регрессии. В результате тестов было подтверждено отсутствие сериальной корреляции.

Гипотеза 1 частично подтверждена. На основании моделей 1, 3, статистической значимости оценок переменной *Bitcoin*, а также функции импульсного отклика можно сделать вывод о наличии существенной взаимосвязи между ростом доходности криптовалюты Bitcoin и ростом доходности рынка NFT. Данная гипотеза подтверждается только для криптовалюты Bitcoin, для криптовалюты Ethereum она отклоняется в связи с незначимостью коэффициентов.

Гипотеза 2 о том, что рост доходности криптовалют отражается в росте доходности NFT рынка не мгновенно, а с определенным лагом, который может достигать нескольких дней, не отвергается. На основании моделей 1, 3 для переменной стоимости криптовалюты *Bitcoin* можно отметить наличие статистически значимых оценок для лагов 1 и 2 данной переменной. Данный порядок лагов согласуется с выводами исследований Pinto-Gutiérrez et al. (2022), Dowling (2021), в которых авторы на основании не-

дельных данных выявили влияние фактора роста биткоина с лагом в одну неделю.

Гипотеза 3 о зависимости между индексами сентимента инвесторов и индексом рынка NFT-активов не отвергается. Для всех моделей, кроме модели 2, можно отметить статистическую значимость оценок для лаговых переменных сентимента *NFTsAI* и *UCRY Policy*. Стоит отметить, что для переменной *UCRY Policy* характерна значимость оценок дальнего (второго) лага. Можно говорить о том, что изменение настроений инвесторов, как и изменение доходности криптовалюты биткоин, отражается на рынке NFT с определенным запозданием, что в контексте исследуемых данных составляет срок одну или две недели. Результаты моделирования функции импульсного отклика говорят о достаточно малом влиянии переменных сентимента инвесторов крипторынка, что согласуется с выводами работы Pinto-Gutiérrez et al. (2022).

Гипотеза 4 о влиянии активов товарного рынка (нефти и золота) на цены NFT-активов не подтверждена. Также не нашла подтверждения гипотеза 5 о влиянии прокси переменной рыночной ликвидности в виде баланса ФРС США. Мы считаем, что гипотеза 5 требует дальнейшего изучения с целью поиска более подходящей прокси избыточной рыночной ликвидности.

Ограничения исследования

Важно также подчеркнуть основные ограничения данного исследования:

1) Данные о NFT-активах были получены с площадок-поставщиков данных, таких как https://nonfungible.com/, https://cryptoslam.io/ и других, а не непосредственно с блокчейнов Ethereum / Solana / Polygon, поэтому существует вероятность пропуска ряда независимых / малых площадок-провайдеров информации о NFT.

2) Несмотря на то, что определенный объем NFT-активов был создан в 2017–2018 гг., подавляющее большинство NFT-коллекций было создано в 2020–2021 гг. В связи с этим, для периода 2017–2019 гг. наблюдается меньшая ликвидность и большее колебание цен, что могло привести к ухудшению результатов моделирования.

3) В нашем анализе не учитывалась информация о создателе (цифрового) объекта, для которого выпущен NFT. Несмотря на то, что часть коллекций набрала популярность по причине популярности их создателей, во многих случаях личность создателя недоступна или вовсе не существует (например, для изображений, созданных генеративными нейросетями). Во избежание несбалансированности в данных было принято решение не учитывать данный параметр.

Выводы

Быстрый рост рынка NFT-активов привлек внимание инвесторов, заинтересованных в альтернативных формах инвестиций. Несмотря на то, что рынок NFT-активов рассматривается многими инвесторами как новая возможность для получения высокой доходности или диверсификации порт-

феля, экономика рынка NFT и его инвестиционные показатели в силу новизны данной сферы еще в значительной мере не исследованы. Подобно традиционным финансовым рынкам, рынок NFT-активов нестабилен и подвержен иррациональному поведению инвесторов. Однако, в отличие от традиционных финансовых рынков, NFT рынок крайне неоднороден и в значительной мере состоит из частных инвесторов. Только 1% активов торгуется на сумму более 1500 долларов, а 75% активов - на сумму менее 15 долларов, при этом доля покупок частных инвесторов с объемом капитала до 10 000 долларов составляет более 80% от общего объема покупок. Однако, привлекают внимание к данному сектору активов единичные продажи на суммы в десятки миллионов долларов. В связи с этим, важно оценить финансовые показатели этого нового класса активов, выявить факторы, вызывающие спрос на данные активы, а также роль сентимента инвесторов в данном процессе.

В результате исследования была подтверждена гипотеза о наличии связи между ростом доходности криптовалют и ростом доходности рынка NFT, однако, данное утверждение оказалось верно только для криптовалюты Bitcoin. Аналогично более ранним исследованиям были получены данные о том, что взаимодействие между биткоином и рынком NFT происходит не мгновенно, а с недельным лагом. Аналогично гипотезе Dowling (2021), было обнаружено, что ценообразование BTC и ETH влияет на рынок NFT, в то время как рынок NFT не оказывает существенного влияния на ценообразование криптовалют. Таким образом, меньший рынок NFT следует за рынком криптовалют. Данное предположение выглядит правдоподобно, поскольку криптовалюты являются инструментом для покупки и торговли NFT. Падение стоимости криптовалюты означает снижение покупательной способности, что, вероятно, приведет к снижению NFT рынка. Наоборот, когда криптовалюты дорожают, инвесторы, как правило, ищут новые или альтернативные инвестиционные возможности. Однако, в ходе исследования не обнаружено существенной связи между стоимостью Ethereum и NFT, данный эффект обнаружен для BTC.

Гипотезы о наличии связи между индексами настроений инвесторов и спросом на NFT-активы также подтвердились. По аналогии с передачей импульса от криптовалюты Bitcoin, изменение настроений инвесторов отражается на рынке с лагом (до одной–двух недель).

Гипотеза о зависимости между спросом на NFT-активы и динамикой цен на ресурсы, такие как нефть и золото, не была подтверждена. Гипотезу о влиянии избыточной рыночной ликвидности на рост средней стоимости NFT-активов также подтвердить не удалось.

Таким образом, в данной работе была предпринята попытка исследования факторов, влияющих на спрос инвесторов на NFT-активы. Данная работа вносит вклад в научную литературу по NFT, посвященную влиянию различных факторов на NFT-активы. Результаты данного исследования имеют практическое значение для исследователей и инвесторов, которые хотят понять эту быстроразвивающуюся отрасль как часть новой цифровой

экономики. В будущих исследованиях по данной теме необходимо уделить большее внимание рассмотрению влияния сентимента инвесторов в социальных сетях, а также выявить роль таких характеристик актива, как личность создателя актива и его медийная популярность.

Заключение

Во многих ранее опубликованных работах отмечается нерациональность поведения частных инвесторов (Schmeling, 2006; Kuo et al., 2015; Adachi et al., 2017; De Bortoli et al., 2019). В настоящее время на многих развитых и развивающихся рынках, включая США и Китай, Россию, доля частных инвесторов растет. Наша коллективная монография впервые поднимает вопросы роли частных инвесторов на рынке акций в РФ, анализирует статистические данные и складывающиеся тренды. Мы рассматриваем как значимый фактор, влияющий на поведение биржевых характеристик инвестиционных активов, общерыночный сентимент и внимание (как результат заметности актива) и сентимент по отдельному активу. Мы подчеркиваем значимость не просто новостной ленты, а ее интерпретации в инвестиционных сообществах, трактовки этих новостей и событий, степень переживания по ним.

Проблемой в количественной оценке влияния сентимента на биржевые характеристики акций или иных активов мы видим построение прокси-переменных моделей для измерения ненаблюдаемых явно настроений. В нашей монографии мы систематизируем и классифицируем метрики сентимента, предложенные в ранее проведенных исследованиях, выделяем метрики внимания к активам (заметности активов) и настроения (тональности восприятия событий на рынке, новостной ленты, действий других участников), метрики общерыночного сентимента и сентимента в отношении отдельных акций, метрики по различным каналам (новостная лента и СМИ, социальные сети, брокерские каналы в мессенджерах и на Youtube). Кроме того, нами предложены новые авторские метрики сентимента, которые могут быть полезны для построения портфелей и торговых стратегий. Мы уделяем внимание влиянию метрик сентимента на биржевые характеристики активов и проводим ряд собственных эмпирических исследований по российскому рынку.

Мы отмечаем, что роль частных инвесторов на российском рынке меняется и растет значимость инвестиционных сообществ, «авторитета мнений». Ковид-кризис изменил многие процессы в экономике, политике и в поведении инвесторов. С 2019 г. по октябрь 2021 г. общее число счетов клиентов брокеров на Московской бирже выросло с 3,9 млн до 15,3 млн, или в 3,9 раза, а число счетов активных клиентов брокеров на Санкт-Петербургской бирже – с 87 тыс. до 825 тыс., или в 9,5 раза (Абрамов и др., 2021). Еще одной поворотной точкой стало введение санкций против России, контрсанкций со стороны России и отток нерезидентов. До февраля 2022 г. во владении активами и в объемах торгов высока была доля иностранных инвесторов, более 50% оборота на рынке акций приходилось на долю институциональных инвесторов, по многим бумагам – нерезидентов. После начала СВО доля частных инвесторов в объеме торгов акциями превысила 75%. Значимую роль играют физические лица и в торгах высокодо-

ходными облигациями (до 100% объема торгов по большинству выпусков ВДО). Риски притока розничных инвесторов на фондовый рынок состоят в том, что инвесторы могут кооперироваться и совершать похожие сделки, что может привести к росту волатильности и манипулированию ценами.

Несколько теоретических конструкций претендуют на объяснение поведения частных инвесторов, которые отходят от рационального выбора. В нашей книге систематизированы как модельные конструкции, так и их эмпирические проверки. Caplin et al. (2011) подчеркивает важность учета поведенческого искажения статуса-кво (status-quo bias). Авторы показывают, что задача поиска индивидом оптимального решения может быть приостановлена при появлении удовлетворяющего его предпочтениям «первого» решения. Также на выбор инвестора может влиять эффект эмоциональной привязки (anchoring bias), «якорения». Для частного инвестора та или иная акция может иметь дополнительную субъективную ценность - эмоциональную привязку, как например, с ярыми последователями Илона Маска и покупкой акций Тесла.

«Эффект внимания» как объясняющий парадокс в поведении цен впервые был предложен Cherry (1953), позднее идеи наличия фильтров в обработке информации и доказательства ограниченной рациональности частных инвесторов развил Kahneman (1973). В более поздних работах было показано, что повышенное внимание инвесторов к активу влечет рост волатильности его доходности (Aouadi et al., 2013). Существует несколько теоретических концепций, в рамках которых анализируется эффект внимания. «Теория осведомленности инвестора» (the investor recognition hypothesis), разработанная Merton (1987), утверждает, что акции компаний, привлекающие меньше внимания, должны демонстрировать более высокую ожидаемую доходность. Отсутствие внимания рассматривается как принятие инвестором большего идиосинкратического (уникального) риска. Альтернативная «теория внимания» (attention theory) была предложена позднее Barber and Odean (2007). Инвесторы предпочитают покупать акции, которые в большей степени привлекают их внимание, прокси которого могут выступать аномальные объемы торгов, экстремальные уровни доходности или большой новостной фон. Для инвесторов приобретает важность не только само содержание рыночной информации, но и способы ее представления (Hirshleifer et al., 2004; Huberman & Regev, 2001). Еще одно объяснение поведения доходности - «теория выдающихся событий» (salience theory, теория заметных событий) Bordalo et al. (2012). Индивиды, принимающие решения, преувеличивают вероятность заметных событий, если они знают о потенциальной возможности их наступления.

Открытый вопрос – выбор адекватной метрики внимания инвесторов (заметности актива). В ранних работах по данной тематике исследовались, в том числе, расходы на рекламу компании-эмитента (Grullon et al., 2004; Chemmanur & Yan, 2009; Lou, 2014), новости и их заголовки (Barber & Odean, 2007; Yuan, 2015), объем торгов по ценной бумаге (Gervais et al., 2001; Barber & Odean, 2007; Hou et al., 2009), охват медиа (media coverage)

(Engelberg & Parsons, 2011; Drake et al., 2012; Yuan, 2015), число выходящих аналитических отчетов (Drake et al., 2012).[1] Однако, все перечисленные прокси имеют свои недостатки. Например, новостной фон и охват медиа могут быть измерены только в момент выхода новостей и не могут гарантировать того, что значительное число инвесторов будет следить за этими источниками информации. Увеличение объема торгов может быть результатом манипуляций на рынке, осуществляемых группой инвесторов (Tantaopas et al., 2016). Яркий пример сомнений в охвате пользователей – тяжба Илона Маска в процессе покупки Твиттер с менеджментом компании (оспаривалась сумма покупки в 44 млрд долл. США).

Мы отмечаем интересные выводы по ряду работ, посвященных эффектам нейминга и алфавитизма. На российском рынке, при сравнительно небольшом числе публичных эмитентов с ликвидными акциями, нейминг ярче проявляется на рынке облигаций. Внятное название, особенно указывающее на реальное производство (завод, мясокомбинат, с/х предприятие) с высокой вероятностью привлечет инвестора, а значит, позволит привлечь долговые деньги в компании и быстрее, и дешевле. В академических исследованиях в качестве зависимых переменных рассматривают относительную рыночную оценку компании (Xing et al., 2018), доходность инвестирования на разных временных горизонтах (Head et al., 2009, Baer et al., 2020), показатели ликвидности (часто используется коэффициент Амихуда, как в работе Green et al., 2013). Теплова и Микова (2019) в авторском исследовании с Е. Глинской демонстрируют, что эффекты нейминга выявляются и на развивающихся рынках капитала (рынки акций Индии, Бразилии, Китая). В нашей книге представлены результаты эмпирического тестирования нейминга на бирже Гонконга (расчеты Д.А. Ерохина под руководством Т.В. Тепловой). Кумулятивная доходность авторского портфеля, основанного на выборе «умных тикеров», превысила доходность бенчмарка на 12,75%. Коэффициент Шарпа авторского портфеля выше, чем аналогичный показатель для рыночного индекса.

Отдельный пласт работ, связанный с эффектом внимания – копирование состава портфеля, который инвесторы могут наблюдать по крупным инвестиционным фондам (хедж-фондам) или гуру-инвесторам. В этот пласт работ попадают и различные стратегии автоследования. В контексте этого направления интересна работа Brown and Schwarz (2020), где показано, что наиболее интересным с точки зрения копирования набора активов являются стратегии, которые используют информацию по хедж-фондам.

Начиная с 2010-х гг., появились работы, в которых оценка внимания инвесторов базируется на частоте поисковых запросов в Google (Da et al., 2011) или Wikipedia (Behrendt et al., 2020). Behrendt et al. (2020) показали, что рыночным инвесторам для принятия инвестиционного решения может

[1] Данный обзор выполнен студентов факультета экономических наук НИУ ВШЭ Мажовым М.А. под руководством Т.В. Тепловой.

быть достаточно краткого описания деятельности компании, если на следующем шаге они не будут проводить анализ финансовой отчетности.

В нашей работе проведено демо-исследование влияния метрик заметности на доходность портфелей по компаниям российского рынка (автор – студент ФЭН ВШЭ Мажов М.). При помощи метода главных компонент построен авторский индекс внимания по компаниям российского рынка и проведено сравнение с популярным показателем внимания - частотой поисковых запросов (SVI Google). Результаты портфельного анализа показали, что доходность крайних портфелей существенно отличается и согласуется с логикой, заложенной в эффект внимания, однако серединные портфели выборки ведут себя неоднозначно. Нами выявлено, что для российского фондового рынка наибольшее значение имеет фактор размера. Для крупных компаний наблюдается монотонное снижение как средних, так и накопленных доходностей при уменьшении двух построенных индексов внимания. Показатель SVI более релевантен для малых компаний и нетипичных портфелей. Коэффициент альфа Йенсена для большинства портфелей положителен, при этом отрицательные значения коэффициента характерны для портфелей, относящихся к группе с низким показателем внимания. Таким образом, отбор акций в портфель по прокси-внимания позволяет получить повышенную доходность.

Мы различаем общерыночные метрики настроений и метрики сентимента по отдельным ценным бумагам. Примером общерыночных метрик являются различные индексы, которые уже зарекомендовали себя для предсказания макроэкономической динамики и косвенно - поведения биржевых индексов акций (Ding et al., 2014; Sun et al., 2016; Dimpfl & Kleiman, 2019; Smales, 2017; Long & ManyiZhao, 2020). Недостатком таких индексов, как прокси сентимент-метрик, является ограниченность данных, которые используются в методике их формирования. Как правило, такие индексы рассчитываются при помощи статистических данных или результатов опроса инвесторов.

Только с середины 2000-х гг. появляются работы, где метрики сентимента выстраиваются на базе обработки текстовых сообщений, включая такие объемные базы, как переписка в мессенджерах, т.е. на основе коммуникаций инвесторов между собой. В работе Baker and Wurgler (2006) показано, что авторские сентимент-метрики по-разному проявляются в поведении доходности акций, принадлежащих к разным инвестиционным стилям (акции роста, акции стоимости, акции компаний различного размера) и разным рынкам (важны межстрановые различия). Schmeling (2009) по прокси-сентимента на основе опросов потребительского поведения выявил отрицательную связь между месячной доходностью страновых индексов и прошлыми настроениями потребителей на горизонте 1 и 24 месяца. Gao and Xie (2020) по китайскому рынку показали, что на дневном горизонте общерыночный индекс настроений надежно предсказывает избыточную доходность и аномальные объемы торгов. Важный вывод ранее проведенных исследований по влиянию сегодняшнего сентимента на будущую доход-

ность – отрицательная зависимость объясняется не только «стадностью» поведения инвесторов, но и ошибками при необходимости принятия быстрых решений, по сути - сверхреакцией, например, на плохие новости.

Глава 3 нашей монографии посвящена разработке метрик сентимента на основе текстового анализа и тестированию гипотез о наличии преимуществ этих метрик как в объяснении поведения отдельных активов, так и в прогнозировании их динамики. Первые работы измеряли сентимент на основе новостной ленты. Tetlock (2007) приходит к выводу, что высокий пессимизм в аналитических колонках в Wall Street Journal прогнозирует будущий разворот на падающем рынке акций США, а экстремальные значения пессимизма и оптимизма приводят к росту объема торгов. На наш взгляд, такой подход содержит критические допущения. Во-первых, авторы не раскрывают вопрос о том, являются ли все подписчики указанных изданий инвесторами на фондовом рынке США, и какова доля среди них частных инвесторов. Во-вторых, крайне затруднительно проверить реакцию подписчиков на опубликованные тексты, остаются открытыми вопросы о том, были ли в действительности прочитаны тексты и заголовки, и совершали ли подписчики сделки с акциями за период исследования в соответствии с тональностью статьи (Теплова и др., 2022).

С 2010-х гг. появились работы, посвященные анализу сентимента инвесторов на базе социальных сетей (StockTwits, Twitter). Yang et al. (2020) показали, что социальные медиа имеют более тесную связь с динамикой акций, чем традиционные СМИ. До 2020-х гг. подавляющее большинство работ фокусировалось на рынке США и на англоязычных обсуждениях (Oliveira et al., 2016; Al-Nasseri & Ali, 2018), в последние годы появились работы по китайским акциям (Dong & Gil-Bazo, 2020, Yang et al., 2020) и другим рынкам капитала.

Если текст в газетах относится к деловому письму и сообщения можно разметить на позитивные, нейтральные и негативные группы с помощью готовых словарей и пакетов, то на площадках свободного общения в основном используется неформальный стиль, много сленговых выражений. Для неанглоязычных платформ это представляет отдельный этап текстовой обработки. По развивающимся рынкам многие исследования основываются на создании уникальных баз текстовых данных, их классификации и формировании авторских индексов. Такие работы отличаются использованием различных моделей на основе искусственного интеллекта по автоматической разметке текстовых данных, для каждого сообщения определяется класс тональности (Renault, 2017; Al-Nasseri & Ali, 2018; Wang et al., 2021; Bari & Agah, 2020; Ko & Chang, 2021). Популярные методы включают SVM, наивный Байес и словарь оценочной лексики (Jing et al., 2021), нейронные сети с полносвязными и LSTM слоями (Karlemstrand & Leckstrem, 2021). Не по всем работам есть консенсус относительно влияния сентимента на динамику цен акций: так, если в работе Wang et al. (2021) показано, что положительный сентимент частных инвесторов порождает положительную доходность акций на следующий день, то Kim and Kim

(2014) не выявили значимость влияния обсуждений на доходность. Консенсус выводов связан с волатильностью: обсуждения в инвестиционном сообществе увеличивают торговую активность и волатильность акций (Lee et al., 2002; Huang et al. 2020). Behrendt and Schmidt (2018) ставят вопрос о горизонте, на котором можно построить выигрышную стратегию по рынку США, и заключают, что только короткие интервалы (внутридневные таймфреймы в пять минут) обеспечивают выигрышность стратегий.

Мы считаем, что неоднозначность ранее полученных результатов эмпирических исследований в области влияния сентимента может быть объяснена тем, что многие исследования строятся на базе одной платформы, одного канала коммуникации, который не покрывает мнения всех инвесторов в рассматриваемые активы (Oliveira et al., 2013; Renault, 2017; Al-Nasseri & Ali, 2018). Мы считаем важным учет дивергенции мнений инвесторов как по различным каналам (Ren et al., 2021), так и по одному выбранному для исследования каналу (Shen et al., 2022).

В нашей работе впервые протестирован ряд гипотез относительно влияния сентимента инвесторов на биржевые характеристики акций на российском рынке. В первой части нашего эмпирического исследования выявлено влияние сентимента инвесторов, фиксируемого по текстам платформы «Тинькофф Пульс», на доходность и ликвидность акций российских компаний-эмитентов. Мы рассчитываем оригинальную метрику согласованности мнений частных инвесторов. С применением регрессионного анализа по панельным данным получен вывод о том, что если все инвесторы соглашаются, что нарастает бычье или медвежье настроение, то доходность акций снижается, заработать вместе с толпой на недельном таймфрейме не удастся. С ростом согласованности мнений инвесторов ликвидность снижается: консенсус мнений не мотивирует участников торгов к сделкам.

При проведении анализа влияния сентимента по отдельным акциям наблюдается неоднозначность влияния: для 8 компаний из 19 выявлено отрицательное направление влияния на доходность акций компаний-эмитентов, тогда как для остальных - отрицательное. Мы объясняем это тем, что непрофессиональные участники фондового рынка могут чаще негативно обсуждать некоторые акции. Для показателей ликвидности наблюдается схожая ситуация, когда по различным активам присутствует неоднозначная реакция относительного объема торгов в зависимости от уровня согласованности мнений частных инвесторов. По выборке наиболее популярных у частных инвесторов акций наблюдается отрицательный эффект в объяснении ликвидности. Большее значение уровня согласованности мнений демотивирует инвесторов совершать биржевые сделки.

Нами проведено исследование влияния согласованности мнений инвесторов во время роста или падения цен на российском рынке акций. Получен вывод о том, что увеличение неопределенности во мнениях ведет к падению доходности и снижению ликвидности в моменты роста фондового рынка. То же характерно и для периодов снижения рынка.

Во второй части нашего эмпирического исследования мы раскрыли влияние настроений частных инвесторов в социальных сетях на волатильность доходности акций компаний российского рынка. Выборка сообщений построена на уникальных данных, охватывающих более чем 40% владельцев активных брокерских счетов за 2014-2020 гг. В нашей работе используется Hype индикатор, построенный на основе ручной разметки 32 тыс. сообщений инвесторов и применении нейронной сети на основе CNN и трансформеров. Мы выявили отличия влияния индикатора Hype по платформам Telegram и MFD. Наряду с авторским индикатором Hype, мы рассматриваем широкий спектр метрик сентимента.

Индикатор сентимента Hype оказывает положительное и значимое на 1% уровне влияние на волатильность доходности акций. Мы заключаем, что если количество сообщений в социальных сетях велико, и их тональность чрезмерно позитивна, часть инвесторов не следует за толпой и не доверяет высокой оценке перспектив компании, в результате растет волатильность доходности. Влияние индикатора Hype по форуму MFD нелинейно. Значима тональность обсуждений компании: при росте общего количества сообщений, а также количества позитивных, негативных и нейтральных сообщений в отдельности волатильность повышается. Мы отмечаем, что инвесторы сильнее реагируют на негативные сообщения, с ростом количества таких сообщений волатильность доходности растет в большей степени (зависимость является квадратичной).

Быстрый рост рынка NFT-активов привлек внимание частных инвесторов, заинтересованных в альтернативных формах инвестиций. В результате эмпирического исследования мы подтвердили гипотезу о наличии связи между ростом доходности криптовалюты Bitcoin и ростом доходности рынка NFT (с недельным лагом). Показано, что ценообразование BTC и ETH влияет на рынок NFT, в то время как рынок NFT не оказывает существенного влияния на ценообразование криптовалют. Таким образом, меньший рынок NFT следует за рынком криптовалют. Гипотезы о наличии связи между индексами настроений инвесторов и спросом на NFT-активы также подтвердились. Не подтвердилась гипотеза о зависимости между спросом на NFT-активы и динамикой цен на ресурсы, такие как нефть и золото.

Список литературы

Абрамов, А. Е., Косырев, А. Г., Радыгин, А. Д., Чернова, М. И. Поведение частных инвесторов на фондовых рынках России и США // Экономическое развитие России. - 2021. - Т. 28, № 12. - С. 54-59.

Афанасьев, Д. О., Федорова, Е.А., Рогов, О. Ю. О влиянии тональности новостей в международных СМИ на рыночный курс российского рубля: текстовый анализ // Экономический журнал Высшей школы экономики. – 2019. – Т. 23, № 2. - С. 264-289.

Теплова, Т. В., Микова, Е. С. Инвестиции на рыночных неэффективностяхи поведенческих искажениях. – М.: ИНФРА-М, 2019.

Теплова, Т.В., Соколова, Т.В., Томтосов, А.Ф., Бучко, Д.В., Никулин, Д.Д. Сентимент частных инвесторов в объяснении различий в биржевых характеристиках акций российского рынка // Журнал Новой экономической ассоциации. - 2022. – Т. 53, № 1. - С. 53-84.

Федорова, Е. А., Афанасьев, Д. О., Дёмин, И. С., Пыльцин, И. В., Нерсесян, Р. Г., Лазарев, А. М. Разработка тонально-тематического словаря EcSentiThemeLex для анализа экономических текстов на русском языке // Прикладная информатика. – 2020. – Т. 15, №6. – С. 58-77.

Федорова, Е. А., Осетров, Р. А., Демин, И. С., Федоров, Ф. Ю., Хрустова, Л. Е. Влияние тональности писем CEO на финансовые показатели компании // Российский журнал менеджмента. – 2017. – Т. 15, № 4. – С. 441-462.

Федорова, Е. А., Пыльцин, И. В., Ковальчук, Ю. А., Дроговоз, П. А. Новости и социальные сети российских компаний: степень влияния на рынок ценных бумаг // Журнал Новой экономической ассоциации. – 2022. – Т. 1, № 53. – С. 32-52.

Adachi, Y., Masuda, M., Takeda, F. (2017). Google search intensity and its relationship to the returns and liquidity of Japanese startup stocks // Pacific-Basin Finance Journal. – Vol. 46. – P. 243-257.

Adebiyi, A. A., Adewumi, A. O., Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction // Journal of Applied Mathematics. – Vol. 2014. – P. 1-7.

Afanasyev, D. O., Fedorova, E., Ledyaeva, S. (2021). Strength of words: Donald Trump's tweets, sanctions and Russia's ruble // Journal of Economic Behavior & Organization. – Vol. 184. – P. 253-277.

Aharon, D. Y. (2021). Uncertainty, fear and herding behavior: Evidence from size-ranked portfolios // Journal of Behavioral Finance. – Vol. 22, № 3. - С. 320-337.

Alanyali, M., Moat, H., Preis, T. (2013). Quantifying the relationship between financial news and the stock market // Scientific Reports. – Vol. 3. - 3578.

Alhaj-Yaseen, Y. S., Yau, S. K. (2018). Herding tendency among investors with heterogeneous information: Evidence from China’s equity markets // Journal of Multinational Financial Management. – Vol. 47. – P. 60-75.

Al-Nasseri, A., Ali, F. M. (2018). What does investors' online divergence of opinion tell us about stock returns and trading volume? // Journal of Business Research. – Vol. 86. – P. 166-178.

Al-Nasseri, A., Ali, F. M., Tucker, A. (2021). Investor sentiment and the dispersion of stock returns: Evidence based on the social network of investors // International Review of Financial Analysis. – Vol. 78. - 101910.

Anderson, A. G., Larkin, Y. (2019). Does noninformative text affect investor behavior? // Financial Management. – Vol. 48, №1. – P. 257-289.

Anderson, A., Brockman, P. (2018). An examination of 13F filings // Journal of Financial Research. – Vol. 41, №3. – P. 295-324.

Antweiler, W., Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards // The Journal of Finance. – Vol. 59, №3. – P. 1259-1294.

Aouadi, A., Arouri, M., Teulon, F. (2013). Investor attention and stock market activity: Evidence from France // Economic Modelling. – Vol. 35. - P. 674-681.

Aragon, G. O., Hertzel, M., Shi, Z. (2013). Why do hedge funds avoid disclosure? Evidence from confidential 13F filings // Journal of Financial and Quantitative Analysis. – Vol. 48, №5. – P. 1499-1518.

Asur, S., Huberman, B. A. (2010). Predicting the future with social media. In: 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology. IEEE. - Vol. 1. - P. 492-499.

Baer, N., Barry, E., Smith, G. (2020). The name game: The importance of resourcefulness, ruses, and recall in stock ticker symbols // The Quarterly Review of Economics and Finance. – Vol. 76. – P. 410-413.

Baker, M., Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns // The Journal of Finance. – Vol. 61, №4. – P. 1645-1680.

Bank, M., Larch, M., Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks // Financial Markets and Portfolio Management. – Vol. 25, №3. – P. 239-264.

Barber, B. M., Odean, T. (2011). The Behavior of Individual Investors. URL: http://ssrn.com/abstract=1872211 (дата доступа: 01.11.2022).

Barber, B. M., Lee, Y. T., Liu, Y. J., Odean, T. (2009). Just how much do individual investors lose by trading? // The Review of Financial Studies. – Vol. 22, №2. – P. 609-632.

Barber, B. M., Odean, T., Zhu, N. (2008). Do retail trades move markets? // The Review of Financial Studies. – Vol. 22, №1. – P. 151-186.

Barber, B. M., Odean, T., Zhu, N. (2009). Systematic noise // Journal of Financial Markets. – Vol. 12, №4. – P. 547-569.

Barber, B. M., Odean, T. (2006). All that glitters: The effect of attention on the buying behavior of individual and institutional investors. Working paper. URL: http://ssrn.com/abstract=460660 (дата доступа: 01.11.2022).

Bari, O. A., Agah, A. (2020). Ensembles of text and time-series models for automatic generation of financial trading signals from social media content // Journal of Intelligent Systems. – Vol. 29, №1. – P. 753-772.

Basak, S. (2005). Asset pricing with heterogeneous beliefs // Journal of Banking & Finance. – Vol. 29, №11. – P. 2849-2881.

Behrendt, S., Schmidt, A. (2018). The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility // Journal of Banking & Finance. – Vol. 96. – P. 355-367.

Behrendt, S., Peter, F. J., Zimmermann, D. J. (2020). An encyclopedia for stock markets? Wikipedia searches and stock returns // International Review of Financial Analysis. – Vol. 72. - 101563.

Benos, E., Jochec, M. (2013). Patriotic name bias and stock returns // Journal of Financial Markets. – Vol. 16, №3. - P. 550-570.

Bernile, G., Lyandres, E. (2011). Understanding investor sentiment: The case of soccer // Financial Management. – Vol. 40, №2. – P. 357-380.

Blankespoor, E., deHaan, E., Zhu, C. (2018). Capital market effects of media synthesis and dissemination: Evidence from robo-journalism // Review of Accounting Studies. – Vol. 23, №2. - P. 1-36.

Bollen, J., Mao, H., Zeng, X. (2011). Twitter mood predicts the stock market // Journal of computational science. – Vol. 2, №1. – P. 1-8.

Bordalo, P., Gennaioli, N., Shleifer, A. (2012). Salience theory of choice under risk // The Quarterly journal of economics. – Vol. 127, №3. – P. 1243-1285.

Borovkova, S., Tsiamas, I. (2019). An ensemble of LSTM neural networks for high-frequency stock market classification // Journal of Forecasting. – Vol. 38, №6. – P. 600-619.

Bossman, A., Teplova, T., Umar, Z. (2022). Do local and world COVID-19 media coverage drive stock markets? Time-frequency analysis of BRICS // Complexity. - Vol. 2022. - 2249581.

Brown, G. W., Cliff, M. T. (2005). Investor sentiment and asset valuation // The Journal of Business. – Vol. 78, №2. – P. 405-440.

Brown, S., Schwarz, C. (2020). Do Market Participants Care about Portfolio Disclosure? Evidence from Hedge Funds' 13F Filings. Working Paper. Available at http://ssrn.com/abstract=1683628.

Caplin, A., Dean, M., Martin, D. (2011). Search and satisficing // American Economic Review. – Vol. 101, №7. – P. 2899-2922.

Carhart, M. M. (1997). On persistence in mutual fund performance // The Journal of Finance. – Vol. 52, №1. – P. 57-82.

Chakrabarty, B., Seetharaman, A., Swanson, Z., Wang, X. (2018). Management risk incentives and the readability of corporate disclosures // Financial Management. - Vol. 47, №3. – P. 583-616.

Charoenrook, A. (2003). Change in consumer sentiment and aggregate stock market returns. The Owen Graduate School of Management. Vanderbilt University. Working Paper.

Chemmanur, T., Yan, A. (2009). Product market advertising and new equity issues // Journal of Financial Economics. – Vol. 92, №1. – P. 40-65.

Chen, H., De, P., Hu, Y. J., Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media // The Review of Financial Studies. – Vol. 27, №5. – P. 1367-1403.

Chen, S., Ge, L. (2019). Exploring the attention mechanism in LSTM-based Hong Kong stock price movement prediction // Quantitative Finance. – Vol. 19, №9. – P. 1507–1515.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears // The Journal of the Acoustical Society of America. – Vol. 25, №5. – P. 975-979.

Choi, K. H., Yoon, S. M. (2020). Investor sentiment and herding behavior in the Korean stock market // International Journal of Financial Studies. – Vol. 8, №2. – P. 34.

Chong, E., Han, C., Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies // Expert Systems with Applications. – Vol. 83. – P. 187–205.

Chui, A. C., Titman, S., Wei, K. J. (2010). Individualism and momentum around the world // The Journal of Finance. – Vol. 65, №1. – P. 361-392.

Corredor, P., Ferrer, E., Santamaria, R. (2015). The impact of investor sentiment on stock returns in emerging markets: The case of Central European Markets // Eastern European Economics. – Vol. 53, №4. – P. 328-355.

Corwin, S. A., Coughenour, J. F. (2008). Limited attention and the allocation of effort in securities trading // The Journal of Finance. – Vol. 63, №6. – P. 3031-3067.

Da, Z., Engelberg, J., Gao, P. (2011). In search of attention // The Journal of Finance. – Vol. 66, №5. – P. 1461-1499.

Dai, Z., Chang, X. (2021). Forecasting stock market volatility: Can the risk aversion measure exert. An important role? // North American Journal of Economics and Finance. – Vol. 58. - 101510.

Das, S. R., Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web // Management Science. – Vol. 53, №9. – P. 1375-1388.

De Bortoli, D., da Costa, J. N., Goulart, M., Campara, J. (2019). Personality traits and investor profile analysis: A behavioral finance study // PloS One. – Vol. 14, №3. - e0214062.

De Long J. B., Shleifer A., Summers L. H., Waldmann, R. J. (1990). Noise trader risk in financial markets // Journal of Political Economy. – Vol. 98, №4. – P. 703–738.

De Vries, L., Gensler, S., Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing // Journal of Interactive Marketing. – Vol. 26, №2. – P. 83-91.

DellaVigna, S., Pollet, J. M. (2009). Capital budgeting vs. market timing: An evaluation using demographics. No. w15184. National Bureau of Economic Research.

Demirer, R., Gkillas, K., Gupta, R., Pierdzioch, C. (2021). Risk aversion and the predictability of crude oil market volatility: A forecasting experiment with random forests // Journal of Operational Resource Society. – P. 1–13.

Dimpfl, T., Kleiman, V. (2019). Investor pessimism and the German stock market: Exploring Google search queries // German Economic Review. – Vol. 20, №1. – P. 1-28.

Ding, C. G., Wang, H. J., Lee, M. C., Hung, W. C., & Lin, C. P. (2014). How does the change in investor sentiment over time affect stock returns? // Emerging Markets Finance and Trade. – Vol. 50, №2. – P. 144-158.

Doellman, T. W., Itzkowitz, J., Itzkowitz, J., & Sardarli, S. H. (2019). Alphabeticity bias in 401 (k) investing // Financial Review. – Vol. 54, №4. – P. 643-677.

Dong, H., Gil-Bazo, J. (2020). Sentiment stocks // International Review of Financial Analysis. – Vol. 72. - 101573.

Drake, M. S., Roulstone, D. T., Thornock, J. R. (2012). Investor information demand: Evidence from Google searches around earnings announcements // Journal of Accounting Research. – Vol. 50, №4. – P. 1001-1040.

Eickhoff, M., Muntermann, J. (2016). Stock analysts vs. the crowd: Mutual prediction and the drivers of crowd wisdom // Information & Management. – Vol. 53, №7. – P. 835-845.

Engelberg, J. E., Parsons, C. A. (2011). The causal impact of media in financial markets // The Journal of Finance. – Vol. 66, №1. – P. 67-97.

Fama, E. F., French, K. R. (1993). Common risk factors in the returns on stocks and bonds // Journal of Financial Economics. – Vol. 33, №1. – P. 3-56.

Fama, E. F., French, K. R. (1996). Multifactor explanations of asset pricing anomalies // The Journal of Finance. – Vol. 51, №1. – P. 55-84.

Fama, E. F., French, K. R. (2015). A five-factor asset pricing model // Journal of Financial Economics. – Vol. 116, №1. – P. 1-22.

Fehle, F., Tsyplakov, S., Zdorovtsov, V. (2005). Can companies influence investor behaviour through advertising? Super bowl commercials and stock returns // European Financial Management. – Vol. 11, №5. – P. 625-647.

Forgosh, D., Wittler, I. (2018). Stock market prediction of FAANG stocks using Twitter. Hood College. URL: https://danielforgosh.com/papers/dataMiningFinal.pdf (дата доступа 01.11.2022).

Frame, A., Brachotte, G. (2015). Le tweet stratégique: Use of Twitter as a PR tool by French politicians // Public Relations Review. – Vol. 41, №2. – P. 278-287.

Fu, J., Zhou, Q., Liu, Y., Wu, X. (2019). Predicting stock market crises using daily stock market valuation and investor sentiment indicators // The North American Journal of Economics and Finance. – Vol. 51. - 100905.

Gao B., Xie J. (2020). Forecasting excess returns and abnormal trading volume using investor sentiment: Evidence from Chinese stock index futures market // Emerging Markets Finance and Trade. – Vol. 56, №3. – P. 593-612.

Gao, B., Yang, C. (2018). Investor trading behavior and sentiment in futures markets // Emerging Markets Finance and Trade. – Vol. 54, №3. – P. 707-720.

Gervais, S., Kaniel, R., Mingelgrin, D. H. (2001). The high-volume return premium // The Journal of Finance. – Vol. 56, №3. – P. 877-919.

Giannini, R., Irvine, P., Shu, T. (2019). The convergence and divergence of investors' opinions around earnings news: Evidence from a social network // Journal of Financial Markets. – Vol. 42. – P. 94-120.

Gibbons, M. R., Ross, S. A., Shanken, J. (1989). A test of the efficiency of a given portfolio. Econometrica // Journal of the Econometric Society. – P. 1121-1152.

Göçken, M., Özçalıcı, M., Boru, A., Dosdoğru, A. T. (2016). Integrating metaheuristics and artificial neural networks for improved stock price prediction // Expert Systems with Applications. – Vol. 44. – P. 320-331.

Goddard, J., Kita, A., Wang, Q. (2015). Investor attention and FX market volatility // Journal of International Financial Markets, Institutions and Money. – Vol. 38, September. – P. 79-96.

Green, T. C., Jame, R. (2013). Company name fluency, investor recognition, and firm value // Journal of Financial Economics. – Vol. 109, №3. – P. 813-834.

Groth, S. S., Muntermann, J. (2011). An intraday market risk management approach based on textual analysis // Decision Support Systems. – Vol. 50, №4. – P. 680–691.

Grullon, G., Kanatas, G., Weston, J. P. (2004). Advertising, breadth of ownership, and liquidity // The Review of Financial Studies. – Vol. 17, №2. – P. 439-461.

Han, X., Li, Y. (2017). Can investor sentiment be a momentum time-series predictor? Evidence from China // Journal of Empirical Finance. – Vol. 42. – P. 212-239.

Harris, M., Raviv, A. (1993). Differences of opinion make a horse race // The Review of Financial Studies. – Vol. 6, №3. – P. 473-506.

Harrison, J. M., Kreps, D. M. (1978). Speculative investor behavior in a stock market with heterogeneous expectations // The Quarterly Journal of Economics. – Vol. 92, №2. – P. 323-336.

He, Z., He, L., Wen, F. (2019). Risk compensation and market returns: The role of investor sentiment in the stock market // Emerging Markets Finance and Trade. – Vol. 55, №3. – P. 704-718.

Head, A., Smith, G., Wilson, J. (2009). Would a stock by any other ticker smell as sweet? // The Quarterly Review of Economics and Finance. – Vol. 49, №2. – P. 551-561.

Hillebrand, E., Medeiros, M. C. (2010). The benefits of bagging for forecast models of realized volatility // Econometric Reviews. – Vol. 29, №5–6. – P. 571–593.

Hirshleifer, D., Hou, K., Teoh, S. H., Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets? // Journal of Accounting and Economics. – Vol. 38. – P. 297-331.

Hofstede, G. (2001). Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations. Thousand Oaks, CA: Sage.

Hong, H., Stein, J. C. (2007). Disagreement and the stock market // Journal of Economic perspectives. – Vol. 21, №2. – P. 109-128.

Hou, K., Xiong, W., Peng, L. (2009). A Tale of Two Anomalies: The Implications of Investor Attention for Price and Earnings Momentum // SSRN Electronic Journal. doi: 10.2139/ssrn.976394.

Hu, C., Li, B. G. (2021). Chinese lexicography and stock trading // International Review of Economics & Finance. – Vol. 73. – P. 44-59.

Hu, T., Tripathi, A. (2016, June). Impact of social media and news media on financial markets. In Thirty Seventh International Conference on Information Systems, Dublin. URL: http://ssrn.com/abstract=2796906 (дата доступа 01.11.2022).

Huang, C., Wen, S., Yang, X., Cao, J., Yang, X. (2022). Measurement of individual investor sentiment and its application: evidence from Chinese stock message board // Emerging Markets Finance and Trade. – Vol. 58, №3. – P. 681-691.

Huang, C.-J., Liao, J.-J., Yang, D.-X., Chang, T.-Y., Luo, Y.-C. (2010). Realization of a news dissemination agent based on weighted association rules and text mining techniques // Expert Systems with Applications. – Vol. 37, №9. – P. 6409–6413.

Huang, D., Jiang, F., Tu, J., Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns // The Review of Financial Studies. – Vol. 28, №3. – P. 791-837.

Huberman, G., Regev, T. (2001). Contagious speculation and a cure for cancer: A nonevent that made stock prices soar // The Journal of Finance. – Vol. 56, №1. – P. 387-396.

Hvidkjaer, S. (2008). Small trades and the cross-section of stock returns // The Review of Financial Studies. – Vol. 21, №3. – P. 1123-1151.

Itzkowitz, J., Itzkowitz, J. (2017). Name-based behavioral biases: are expert investors immune? // Journal of Behavioral Finance. – Vol. 18, №2. – P. 180-188.

Jackson, A. (2004). The aggregate behaviour of individual investors. Working paper. URL: http://ssrn.com/abstract=536942 (дата доступа 01.11.2022).

Jacobs, H., Hillert, A. (2016). Alphabetic bias, investor recognition, and trading behavior // Review of Finance. – Vol. 20, №2. – P. 693-723.

Jaggi, M., Mandal, P., Narang, S., Naseem, U., Khushi, M. (2021). Text mining of stocktwits data for predicting stock prices // Applied System Innovation. – Vol. 4, №1. – 13.

Jiang, Y., Mo, B., Nie, H. (2018). Does investor sentiment dynamically impact stock returns from different investor horizons? Evidence from the US stock market using a multi-scale method // Applied Economics Letters. – Vol. 25, №7. – P. 472-476.

Jin, F., Self, N., Saraf, P., Butler, P., Wang, W., Ramakrishnan, N. (2013). Forexforeteller: Currency trend modeling using news articles. ACM SIGKDD

International Conference on Knowledge Discovery and Data Mining. – P. 1470–1473.

Jin, X., Shen, Y., Yu, B. (2021). Stock name length and high visibility premium // Finance Research Letters. – Vol. 39. - 101595.

Jing, N., Wu, Z., Wang, H. (2021). A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction // Expert Systems with Applications. – Vol. 178. - 115019.

Joseph, K., Wintoki, M. B., Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search // International Journal of Forecasting, 27, №4. – P. 1116-1127.

Junqu´e de Fortuny, E., De Smedt, T., Martens, D., Daelemans, W. (2014). Evaluating and understanding text-based stock price prediction models // Information Processing & Management. – Vol. 50, №2. – P. 426–441.

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

Kara, Y., Boyacioglu, M. A., Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange // Expert systems with Applications. – Vol. 38, №5. – P. 5311-5319.

Karlemstrand, R., Leckström, E. (2021). Using Twitter attribute information to predict stock prices. arXiv preprint arXiv:2105.01402.

Kazem, A., Sharifi, E., Hussain, F. K., Saberi, M., Hussain, O. K. (2013). Support vector regression with chaos-based firefly algorithm for stock market price forecasting // Applied Soft Computing. – Vol. 13, №2. – P. 947–958.

Kim, S. H., Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns // Journal of Economic Behavior & Organization. – Vol. 107. – P. 708-729.

Klemola, A. (2020). Internet search-based investor sentiment and value premium // Finance Research Letters. – Vol. 33. - 101224.

Ko, C. R., Chang, H. T. (2021). LSTM-based sentiment analysis for stock price forecast // PeerJ Computer Science. – Vol. 7. - e408.

Kordonis, J., Symeonidis, S., Arampatzis, A. (2016, November). Stock price forecasting via sentiment analysis on Twitter. In: Proceedings of the 20th Pan-Hellenic Conference on Informatics. - P. 1-6.

Kumar, A., Lee, C. M. (2006). Retail investor sentiment and return comovements // The Journal of Finance. – Vol. 61, №5. – P. 2451-2486.

Kuo, W. Y., Lin, T. C., Zhao, J. (2015). Cognitive limitation and investment performance: Evidence from limit order clustering // The Review of Financial Studies, 28, №3. – P. 838-875.

Lee, C. M., Shleifer, A., Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle // The Journal of Finance. – Vol. 46, №1. – P. 75-109.

Lee, E. J., Lee, H. Y., Choi, S. (2020). Is the message the medium? How politicians' Twitter blunders affect perceived authenticity of Twitter communication // Computers in Human Behavior. – Vol. 104. - 106188.

Lee, J., Lim, Y. S. (2016). Gendered campaign tweets: the cases of Hillary Clinton and Donald Trump // Public Relations Review. – Vol. 42, №5. – P. 849-855.

Lee, W. Y., Jiang, C. X., Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment // Journal of Banking & Finance. – Vol. 26, №12. – P. 2277-2299.

Lemmon, M., Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence // The Review of Financial Studies. – Vol. 19, №4. – P. 1499-1529.

Li, F. (2010). The Information content of forward-looking statements in corporate filingsA Naïve Bayesian machine learning approach // Journal of Accounting Research. – Vol. 48, №5. – P. 1049–1102.

Li, K., Wang, T., Cheung, Y. L., Jiang, P. (2011). Privatization and risk sharing: Evidence from the split share structure reform in China // The Review of Financial Studies. – Vol. 24, №7. – P. 2499-2525.

Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., Chen, Y. (2014). The effect of news and public mood on stock movements // Information Sciences. – Vol. 278. – P. 826-840.

Li, T., van Dalen, J., van Rees, P. J. (2018). More than just noise? Examining the information content of stock microblogs on financial markets // Journal of Information Technology. – Vol. 33, №1. – P. 50-69.

Li, X., Xie, H., Chen, L., Wang, J., Deng, X. (2014). News impact on stock price return via sentiment analysis // Knowledge-Based Systems. – Vol. 69, №1. – P. 14–23.

Li, Y., Bu, H., Li, J., Wu, J. (2020). The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning // International Journal of Forecasting. – Vol. 36, №4. – P. 1541-1562.

Liang C., Li Y., Tang L., Wei Y. (2020). Which sentiment index is more informative to forecast stock market volatility? Evidence from China // International Review of Financial Analysis. – Vol. 71. – P. 1-11.

Liu, L. X., Sherman, A. E., Zhang, Y. (2014). The long-run role of the media: Evidence from initial public offerings // Management Science. – Vol. 60, №8. – P. 1945-1964.

Long W., Zhao M., Tang Y. (2020). Can the Chinese volatility index reflect investor sentiment? // International Review of Financial Analysis. – Vol. 73. – 101612.

Lou, D. (2014). Attracting investor attention through advertising // The Review of Financial Studies. – Vol. 27, №6. – P. 1797-1829.

Loughran, T., McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks // The Journal of Finance. – Vol. 66, №1. – P. 35-65.

Ma, Y., Duan, Q., Wu, H. (2021). Does a stock's name affect its return? Evidence from the Chinese stock market during the China–US trade conflict // Finance Research Letters. – Vol. 40. - 101733.

Mackintosh P. (2020) Who counts as a retail investor? NASDAQ Stock Exchange website. URL: https://www.nasdaq.com/articles/who-counts-as-a-retail-investor-2020-12-17 (дата доступа: 01.11.2022).

McGurk, Z., Nowak, A., Hall, J. C. (2020). Stock returns and investor sentiment: textual analysis and social media // Journal of Economics and Finance. – Vol. 44, №3. – P. 458-485.

Menkhoff, L., Rebitzky, R. R. (2008). Investor sentiment in the US-dollar: Longer-term, non-linear orientation on PPP // Journal of Empirical Finance. – Vol. 15, №3. – P. 455-467.

Merton, R. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information // Journal of Finance. – Vol. 42, №3. – P. 483-510.

Meta, R. (2015). Behavioral finance: The psychology of investing. Unpublished working paper. University of Torino.

Michaux, E. (2019). Twitter and its relationship with returns and trading volume of European stocks. Louvain School of Management, Université catholique de Louvain.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion // The Journal of Finance. – Vol. 32, №4. – P. 1151-1168.

Neal, R., Wheatley, S. M. (1998). Do measures of investor sentiment predict returns? // Journal of Financial and Quantitative Analysis, 33, №4. – P. 523-547.

Nishimura, Y., Sun, B. (2021). President's Tweets, US-China economic conflict and stock market Volatility: Evidence from China and G5 countries // The North American Journal of Economics and Finance. – Vol. 58. - 101506.

Oliveira, N., Cortez, P., Areal, N. (2017). The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices // Expert Systems with Applications. – Vol. 73. – P. 125-144.

Piñeiro-Chousa, J., López-Cabarcos, M. Á., Caby, J., Šević, A. (2021). The influence of investor sentiment on the green bond market // Technological Forecasting and Social Change. – Vol. 162. - 120351.

Qiu, L., Welch, I. (2004). Investor sentiment measures. Working paper, Brown University. URL: http://ssrn.com/abstract=589641 (дата доступа: 01.11.2022).

Qiu, M., Song, Y., Akagi, F. (2016). Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market // Chaos, Solitons, Fractals. – Vol. 85. – P. 1–7.

Reboredo J.C., Ugolini A. (2018). The impact of Twitter sentiment on renewable energy stocks // Energy Economics. – Vol. 76. - P. 153-169.

Reis P.N., Pinho C. (2020). A new European investor sentiment index (EURsent) and its return and volatility predictability // Journal of Behavioral and Experimental Finance. - Vol. 27. - 100373.

Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation // European Journal of Operational Research. – Vol. 291, №3. – P. 906-917.

Ren, J., Dong, H., Padmanabhan, B., Nickerson, J. V. (2021). How does social media sentiment impact mass media sentiment? A study of news in the financial

markets // Journal of the Association for Information Science and Technology. – Vol. 72, №9. – P. 1183-1197.

Renault, T. (2017). Intraday online investor sentiment and return patterns in the US stock market // Journal of Banking & Finance. – Vol. 84. – P. 25-40.

Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., Kim, H. C. (2021). Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions // Electronics. – Vol. 10, №21. - 2717.

Scharfstein, D. S., Stein, J. C. (1990). Herd behavior and investment // The American Economic Review. – Vol. 80, №3. - P. 465-479.

Schmeling, M. (2007). Institutional and individual sentiment: smart money and noise trader risk? // International Journal of Forecasting. – Vol. 23, №1. – P. 127-145.

Schmeling, M. (2009). Investor sentiment and stock returns: Some international evidence // Journal of Empirical Finance. – Vol. 16, №3. – P. 394-408.

Seasholes, M. S., Wu, G. (2007). Predictable behavior, profits, and attention // Journal of Empirical Finance. – Vol. 14, №5. – P. 590-610.

Seong, N., Nam, K. (2021). Predicting stock movements based on financial news with segmentation // Expert Systems with Applications. – Vol. 164. - 113988.

Shanghai Stock Exchange Market and Infrastructure Report (2020). Shanghai Stock Exchange website. URL: http://english.sse.com.cn/news/publications/sseinfrastructure/c/5271425.pdf (дата доступа 01.11.2022).

Shen, J., Najand, M., Dong, F., He, W. (2017). News and social media emotions in the commodity market // Review of Behavioral Finance. – Vol. 9, №2. – P. 148-168.

Shen, S., Xia, L., Shuai, Y., Gao, D. (2022). Measuring news media sentiment using big data for Chinese stock markets // Pacific-Basin Finance Journal. – Vol. 74. - 101810.

Singal, V., Tayal, J. (2020). Risky short positions and investor sentiment: evidence from the weekend effect in futures markets // Journal of Futures Markets. – Vol. 40, №3. – P. 479-500.

Sirignano, J., Cont, R. (2019). Universal features of price formation in financial markets: Perspectives from deep learning // Quantitative Finance. – Vol. 9, №9. – P. 1449–1459.

Smales L.A. (2021). Investor attention and global market returns during the COVID-19 crisis // International Review of Financial Analysis. – Vol. 73. – P. 1-14.

Smales, L. A. (2017). The importance of fear: investor sentiment and stock market returns // Applied Economics. – Vol. 49, №34. – P. 3395-3421.

Song Q., Liu A., Yang S.Y. (2017). Stock portfolio selection using learning-to-rank algorithms with news sentiment // Neurocomputing. – Vol. 264, № 15. – P. 20-28.

Song Q., Liu A., Yang S.Y., Deane A., Datta K. (2015). An extreme firm-specific news sentiment asymmetry based trading strategy. Proceedings of the 2015 IEEE Symposium Series of Computational Intelligence. - P. 898-904.

Sprenger, T. O., Tumasjan, A., Sandner, P. G., Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs // European Financial Management. – Vol. 20, №5. – P. 926-957.

Stambaugh, R. F., Yu, J., Yuan, Y. (2015). Arbitrage asymmetry and the idiosyncratic volatility puzzle // The Journal of Finance. – Vol. 70, №5. – P. 1903-1948.

Sun, L., Najand, M., Shen, J. (2016). Stock return predictability and investor sentiment: A high-frequency perspective // Journal of Banking & Finance. – Vol. 73. – P. 147-164.

Swaminathan, B. (1996). Time-varying expected small firm returns and closed-end fund discounts // The Review of Financial Studies. – Vol. 9, №3. – P. 845-887.

Takeda, F., Wakao, T. (2014). Google search intensity and its relationship with returns and trading volume of Japanese stocks // Pacific-Basin Finance Journal. – Vol. 27. – P. 1-18.

Takeda, F., Yamazaki, H. (2006). Stock price reactions to public TV programs on listed Japanese companies // Economics Bulletin. – Vol. 13, №7. – P. 1-7.

Tantaopas, P., Padungsaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in Asia-Pacific stock markets // Pacific-Basin Finance Journal. – Vol. 38. – P. 107-124.

Tashiro, D., Matsushima, H., Izumi, K., Sakaji, H. (2019). Encoding of high-frequency order information and prediction of short-term stock price by deep learning // Quantitative Finance. – Vol. 19, №9. – P. 1499–1506.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market // The Journal of Finance. – Vol. 62, №3. – P. 1139-1168.

Tokyo Stock exchange monthly trading volume & value by type of investor (2020). Tokyo Stock Exchange website. URL: https://www.jpx.co.jp/english/markets/statistics-equities/monthly/b5b4pj0000040gq6-att/09_t-kabu2102.pdf (дата доступа 01.11.2022).

Travers, J. R., Olivier, D. C. (1978). Pronounceability and statistical «Englishness» as determinants of letter identification // The American Journal of Psychology. - P. 523-538.

Tsai I. C. (2017). Diffusion of optimistic and pessimistic investor sentiment: An empirical study of an emerging market // International Review of Economics & Finance. – Vol. 47. – P. 22-34.

Umar, Z., Bossman, A., Sun-Yong, C., Teplova T. (2023). The relationship between global risk aversion and returns from safe-haven assets // Financial Research Letters. In Press.

Umar, Z., Gubareva, M., Yousaf, I., Ali, S. (2021). A tale of company fundamentals vs sentiment driven pricing: The case of GameStop // Journal of Behavioral and Experimental Finance. - Vol. 30. - 100501.

Wang, H., Xue, L., Du, W., Wang, F., Li, P., Chen, L., Ma, H. (2021). The effect of online investor sentiment on stock movements: an LSTM approach. In: International Conference on Intelligence Science. Springer, Cham. – P. 1-14.

Wang, Q., Xu, W., Zheng, H. (2018). Combining the wisdom of crowds and technical analysis for financial market prediction using deep random subspace ensembles // Neurocomputing. – Vol. 299. – P. 51–61.

Wang, R., Li, J. (2021). The influence and predictive powers of mixed-frequency individual stock sentiment on stock returns // The North American Journal of Economics and Finance. – Vol. 58. - 101522.

Wang, Y., Luo, J., Niemi, R., Li, Y. (2016). To follow or not to follow: Analyzing the growth patterns of the Trumpists on Twitter. The Workshops of the Tenth International AAAI Conference on Web and Social Media News and Public Opinion: Technical Report WS-16-18.

Wanyun, C., Jie, L. (2013). Investors' bullish sentiment of social media and stock market indices // Journal of Management. – Vol. 5. - 012.

Wu, X., Xie, H., Zhang, H. (2022). Time-varying risk aversion and renminbi exchange rate. Volatility: Evidence from CARR-MIDAS model // North American Journal of Economics and Finance. – Vol. 61. - 101703.

Xing, F. Z., Cambria, E., & Welsch, R. E. (2018). Intelligent asset allocation via market sentiment views. // IEEE ComputatioNal iNtelligeNCe magaziNe. – Vol. 13, №4. – P. 25-34.

Xu, Y., Liang, C., Li, Y., Huynh, T. L. (2022). News sentiment and stock return: Evidence from managers' news coverages // Finance Research Letters. – Vol. 48. - 102959.

Yan, D., Zhou, G., Zhao, X., Tian, Y., Yang, F. (2016). Predicting stock using microblog moods // China Communications. – Vol. 13, №8. – P. 244-257.

Yang, S. Y., Mo, S. Y. K., Liu, A. (2015). Twitter financial community sentiment and its predictive relationship to stock market movement // Quantitative Finance. – Vol. 15, №10. – P. 1637-1656.

Yang, X., Zhu, Y., Cheng, T. Y. (2020). How the individual investors took on big data: The effect of panic from the internet stock message boards on stock price crash // Pacific-Basin Finance Journal. – Vol. 59. - 101245.

Yao, C. Z., Sun, B. Y., Lin, J. N. (2017). A study of correlation between investor sentiment and stock market based on Copula model // Kybernetes. – Vol. 46. – P. 550–571.

Yu, Y., Duan, W., Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach // Decision Support Systems. – Vol. 55, №4. – P. 919-926.

Yuan, Y. (2015). Market-wide attention, trading, and stock returns // Journal of Financial Economics. – Vol. 116, № 3. – P. 548-564.

Zhang, X., Li, A., & Pan, R. (2016). Stock trend prediction based on a new status box method and AdaBoost probabilistic support vector machine // Applied Soft Computing. – Vol. 49. – P. 385–398.

Zhang, Y. X., Liu, X. H., Wang, W. J., Liu, Y. J. (2020, August). A study of relationship between investor sentiment and stock price: Realization of investor

sentiment classification based on Bayesian model. In: 2020 International Symposium on Computer Engineering and Intelligent Communications (ISCEIC). IEEE. – P. 34-37.

Zhou, S., Kan, P., Huang, Q., Silbernagel, J. (2021). A guided latent Dirichlet allocation approach to investigate real-time latent topics of Twitter data during Hurricane Laura // Journal of Information Science. Published Online.

Zhu, S., Liu, Q., Wang, Y., Wei, Y., Wei, G. (2019). Which fear index matters for predicting US stock market volatilities: Text-counts or option based measurement? // Physica A: Statistical Mechanics and its Applications. – Vol. 536. - 122567.

Zou, H., Hastie, T. (2005). Regularization and variable selection via the elastic net // Journal of the Royal Statistical Society: Series B (Statistical Methodology). – Vol. 67, №2. – P. 301-320.