

**Е.А. Федорова**

Департамент корпоративных финансов и корпоративного управления  
Финансового университета при Правительстве РФ, Москва

**И.В. Пыльцин**

НИУ ВШЭ, Факультет экономических наук, Москва

**Ю.А. Ковальчук**

МГИМО МИД России, Москва

**П.А. Дроговоз**

МГТУ им. Н.Э. Баумана, Москва

## **Новости и социальные сети российских компаний: степень влияния на рынок ценных бумаг**

**Аннотация.** Основная цель данной работы — определить периоды, в которых доходность фондового рынка России частично объясняется официальной информацией (новости РБК) и неофициальной (твиты компаний). Методология исследования включала модель VAR со скользящим окном в 50 дней. Эмпирическая база исследования состояла из 32 тысяч новостных статей и 111 тысяч твитов за 2011–2020 гг. Оценка тональности текста проводилась на основе словаря EcSentiThemeLex. Данный словарь является одним из первых словарей, позволяющих оценивать тональность экономических и финансовых текстов на русском языке по пяти тональностям. Результаты исследования показывают, что доходность рынка объясняется как экономическими показателями (цена на нефть и курс валюты), так и тональностью текстовой информации. Более того, с 2011 по 2015 г. тональность новостного фона была главенствующим информационным фактором доходности рынка ценных бумаг, а в 2017–2020 гг. доминанция перешла к настроению текстов социальных сетей. В ходе анализа был выявлен краткосрочный период влияния текстовой информации, когда внезапные всплески настроений инвесторов приводили к временному скачку цен на акции, которые быстро возвращались на уровень, близкий к исходному.

**Ключевые слова:** *текстовый анализ, твиты, новости РБК, фондовый рынок.*

Классификация JEL: C51, G11, G17.

Цитирование: **Федорова Е.А., Пыльцин И.В., Ковальчук Ю.А., Дроговоз П.А.** (2022). Новости и социальные сети российских компаний: степень влияния на рынок ценных бумаг // *Журнал Новой экономической ассоциации.* № 1 (53). С. 32–52.  
DOI: 10.31737/2221-2264-2022-53-1-2

### **Введение**

На протяжении десятилетий многих зарубежных и отечественных исследователей волнуют вопросы, связанные с объяснением факторов, влияющих на движение фондового рынка. Финансовая теория утверждает, что доходность акций объясняется рыночными механизмами эффективного финансового рынка, однако в реальности рыночная доходность формируется в результате распространения информации среди инвесторов, которые несут определенное настроение. При этом дискуссионным остается вопрос о том, как профессиональные участники финансового рынка реагируют на эту информацию и какие издания являются приоритетными для потенциальных инвесторов. На данный момент всю информацию условно можно разделить на две группы: официальную и неофициальную, среди которых можно выделить официальные новости и социальные сети. В России существуют такие новостные агрегаторы, как «РБК» или

«Ведомости», а также социальная сеть Twitter, где инвестор может изучить аналитику, экспертные оценки и узнать важные экономические факты.

Основная цель данной работы — определить периоды, в которых доходность фондового рынка России частично объясняется официальной (новости «РБК») и неофициальной (твиты<sup>1</sup> компаний) информацией.

Данный вопрос практически впервые исследуется для отечественного рынка по ряду причин:

- 1) авторы данного исследования в составе рабочей группы создали словарь, позволяющий оценивать тональность экономических и финансовых текстов на русском языке, поэтому в данной работе оцениваются и твиты, и новости «РБК» именно на русском языке<sup>2</sup>;
- 2) применение модели VAR со скользящим окном позволяет выявить периоды приоритетного влияния двух источников информации.

### 1. Обзор литературы и гипотезы исследования

Тема, связанная с выявлением факторов, влияющих на доходность фондового рынка, является довольно популярной среди отечественных и зарубежных исследователей. Первый блок исследований посвящен фундаментальным факторам, влияющим на доходность фондового рынка, и к ним разные авторы относят макроэкономические и финансовые факторы. Одним из направлений исследований является оценка влияния курса валюты на фондовый рынок; в ряде зарубежных исследований доказывается такое влияние (Tachibana, 2018; Patro et al., 2014). Что касается отечественного фондового рынка, то в ряде работ также доказано влияние курса валюты на доходность фондового рынка (Федорова, Панкратов, 2010). Следующим направлением исследования является анализ влияния цены на нефть на фондовый рынок; ряд авторов доказали такое влияние (Sevik et al., 2020; Thorbecke et al., 2019). Однако вопрос остается дискуссионным, так как в исследовании (Kilian, 2009) доказано, что высокие цены на нефть вызывают рост стоимости акций для одних отраслей и их падение для других. Также в ряде исследований доказывается влияние динамики кредитного рынка на фондовый рынок (Abbassi et al., 2016). Суммируя вышесказанное, можно отметить, что показатели кредитного сектора, цена на нефть и состояние валютного рынка влияют на движение фондового рынка. Однако помимо этих фундаментальных факторов на фондовый рынок влияет тональность различной информации, доступная инвесторам.

Поэтому следующим блоком отечественных и зарубежных исследований является оценка влияния тональности новостей из различных источников (новости, годовые отчеты, социальные сети и т.д.) на движение фондового рынка. Однако влияние информации может отличаться в зависимости от тональности, как минимум, положительная или негативная окраска новостей может по-разному влиять на котировки акций компаний. Так, (Vorjigin et al., 2020) изучили особенности распространения позитивной и негативной информации на всем рынке, были выявлены различия в распространении в зависимости от тональности: негативная информация распространяется быстрее позитивной. В отечественных источниках в ряде исследований данная тема также обсуждалась. Так, например в исследовании (Дробышевский и др., 2017), исследуется информационная

<sup>1</sup> Твит – короткое опубликованное сообщение.

<sup>2</sup> Словарь оформлен в виде библиотеки в среде R ([https://dmafanashev.github.io/rulexicon/reference/hash\\_sentiment\\_essentithemelex.html](https://dmafanashev.github.io/rulexicon/reference/hash_sentiment_essentithemelex.html)).

политика Банка России на основе анализа тональности новостных сообщений информационных агентств, а также запросов пользователей в поисковой системе Google. Выявлена степень воздействия информационных сигналов на денежный и валютный рынки. В исследовании (Федорова, Хрустова, Демин, 2020) оценивается влияние раскрытия нефинансовой информации на инвестиционную привлекательность российских компаний. В ходе проведенного авторами исследования эмпирически обосновано влияние полноты раскрытия информации о качестве корпоративного управления, экономической результативности и политике в области управления персоналом в соответствии со стандартами GRI на инвестиционную привлекательность компании на финансовом рынке.

Одним из официальных источников информации являются новости различных агентств и ведомств. В работах (Al-Maadid et al., 2020; Guan-Ru et al., 2019; Афанасьев и др., 2020) авторы доказали, что общая тональность новостей из официальных источников влияет на доходность фондового рынка. Помимо таких новостных статей, широкую популярность получили социальные сети в качестве транслятора общественного мнения и мнения инвесторов. К таким социальным сетям можно отнести Twitter, ВКонтакте или Facebook. Ряд авторов доказали, что тональность сообщений в социальных сетях также влияет на динамику фондового рынка (Oliveira et al., 2020; Nisar et al., 2018). Влияние тональности информации из новостей и твитов имеет дискуссионный характер. Данный вопрос недостаточно исследован на материалах российского фондового рынка. В рамках данного исследования было выдвинуто несколько гипотез.

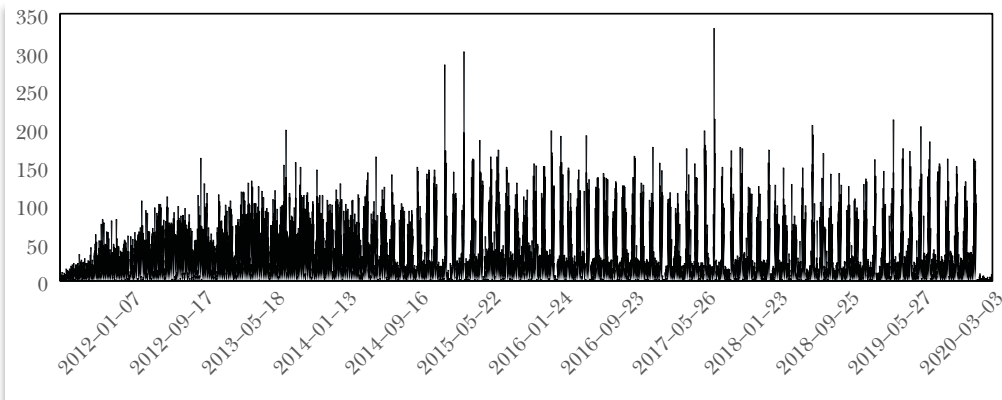
**Гипотеза Н1.** Существуют такие временные интервалы, в одних из которых доходность российского рынка ценных бумаг более чувствительна к настроению новостных финансовых статей, а в других – к тональности микроблогов Твиттера, транслирующих суждения инвесторов относительно движения рынка (Gan et al., 2020).

В реальном мире инвесторы общаются и усваивают информацию с помощью комплекса средств массовой информации и социальных сетей, что делает социальное влияние критическим фактором процесса распространения информации и ценообразования активов. Социальные сети, как известно, создают привлекающую внимание горячую тему, которая может поколебать представление инвесторов относительно будущих перспектив компании, тем самым формируя настроения инвесторов, которые в конечном итоге влияют на цены акций.

**Гипотеза Н2.** С течением времени тональность микроблогов из российских аккаунтов Твиттера становится доминирующим информационным фактором ценообразования активов.

На рис. 1 показана динамика использования социальной платформы Твиттер путем количественного анализа написанных микроблогов за каждый день, находящихся в исследуемом промежутке времени. Таким образом, можно заметить, что существует тенденция популяризации Твиттера как передатчика мыслей, страхов и идей среди инвесторов. Иными словами, общее число твитов экономической и финансовой тематики со временем увеличилось.

В качестве контрольных переменных на основе обзора литературы для отечественного рынка были выбраны следующие факторы: цена на нефть, курс валюты и ставка кредитного рынка.



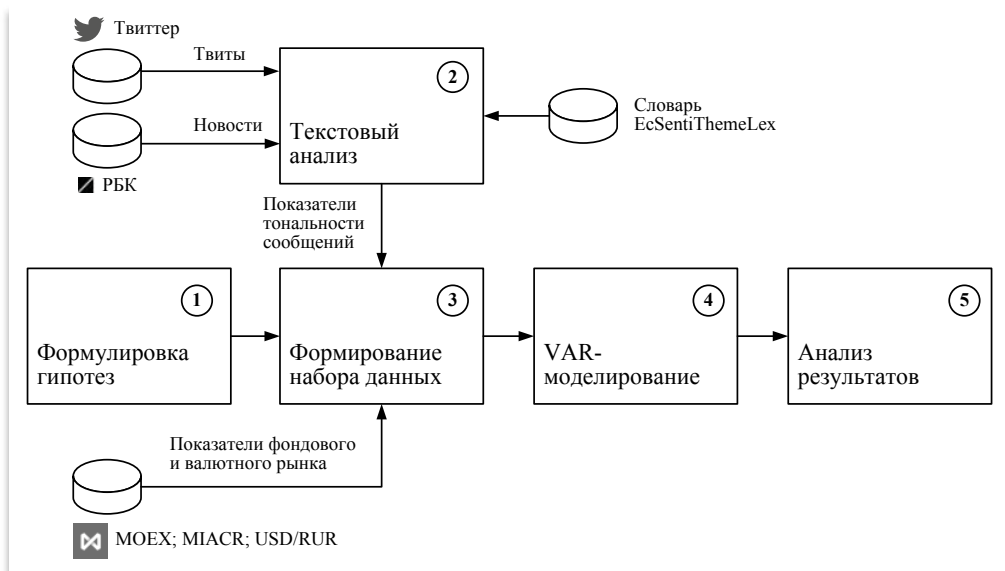
**Рис. 1**

*Динамика числа твитов за исследуемый период времени*

## 2. Методология исследования

### 2.1. Схема исследования

Схема нашего исследования представлена на рис. 2 и включает основные этапы: формулировку гипотез, текстовый анализ твитов и новостных сообщений, формирование набора данных посредством слияния показателей фондового и валютного рынка с показателями тональности сообщений, VAR-моделирование с 50-дневным скользящим окном и анализ полученных результатов. Далее будет дано подробное описание всех этапов исследования.



**Рис. 2**

*Схема исследования*

## 2.2. Текстовый анализ

В нашем исследовании одной из задач выступала оценка тональности экономических и финансовых текстов на русском языке. Авторы статьи в рамках рабочей группы разработали словарь EcSentiThemeLex, позволяющий оценивать эту тональность. Составленный словарь отличается от предыдущих по двум направлениям: 1) он является одним из первых словарей, позволяющих оценивать тональность экономических и финансовых текстов на русском языке по пяти тональностям; 2) позволяет оценить тональность и смысловую направленность текста по 12 экономическим темам (например, «макроэкономика», «монетарная политика», «фондовые и товарные рынки» и т.д.). Разработанный словарь был оформлен в виде библиотеки в среде R<sup>3</sup>, этот словарь мы использовали в данном исследовании для оценки тональности новостей. Методология разработки авторского словаря описана в работе (Федорова и др., 2020). Был проведен частотный анализ слов, построены матрицы корреляций, выделены кластеры в корпусе текстов (320 годовых отчетов российских компаний, более 500 новостей и экономических официальных документов в 2002–2019 гг., более 150 тыс. твитов компаний, входящих в индекс РТС, более 32 тыс. новостных статей «РБК» по категориям: «Финансы», «Экономика», «Деньги» и «Бизнес»). На следующем этапе каждому слову и словосочетанию присваивалось значение тональности: «-2» – сильно негативные слова; «-1» – негативные слова; «0» – нейтральные слова; «1» – позитивные слова; «2» – сильно позитивные слова. Так, слову «рекордный» была присвоена тональность «2», слову «продуктивность» – «1», слову «налог» – «0», слову «нерентабельность» – «-1», слову «кризис» – «-2». Впоследствии словарь добавлялся синонимами и переведенными словами из зарубежных финансовых словарей. Если мы сравним разработанный авторами словарь с зарубежными аналогичными словарями (Loughran & McDonald и NRC), то заметим, что в этом словаре негативных слов (1038 слов) намного больше, чем позитивных (614 слов), что соответствует общемировой практике. Для сравнения, в словаре Loughran & McDonald негативных слов также больше – 1281 против 218 слов, а в словаре NRC – соответственно 3019 против 2126 слов. Хотелось бы отметить, что к преимуществам данного исследования также можно отнести оформление разработанного словаря EcSentiThemeLex в виде программного пакета (библиотеки) rulexicon для среды моделирования R<sup>4</sup> и Python<sup>5</sup>.

Введем несколько переменных по анализу текстов, которые будем использовать в нашем исследовании. При оценке тональности текстов мы будем учитывать важность слова в контексте документа, являющегося частью массива текстовых данных. Это будет выполнено при помощи статистической меры TF-IDF (term frequency – частота слова, inverse document frequency – обратная частота документа):

$$tf(t,d) = \frac{n_t}{\sum_k n_k}, \quad idf(t,D) = \log \frac{|D|}{|\{d_i \in D | t \in d_i\}|},$$

$$TF-IDF = tf(t,d)idf(t,D),$$

где  $t$  – слово,  $n$  – число вхождений слова в текст,  $d$  – документ (или текст),  $D$  – коллекция документов (или массив документов),  $|D|$  – число документов в мас-

<sup>3</sup> <https://dmafanashev.github.io/rulexicon/index.html>

<sup>4</sup> <https://dmafanashev.github.io/rulexicon/index.html>

<sup>5</sup> <https://pypi.org/project/ecsenthemex/>

сиве, а  $|\{d_i \in D | t \in d_i\}|$  – число документов из коллекции, в которых встречается слово  $t$ . Далее, в формулах вес  $TF-IDF$  будем записывать как  $tf-idf(word)$ .

После проведения данных операций над текстом будут оценена тональность каждого текста по формуле  $S_i^j = \left( \sum_{k=1}^n S_k^{word} TF-IDF(word) \right) / n$ . Для каждого текста  $i$ , отнесенного к категории  $j = \{S, N\}$ , где  $S$  – новостные статьи, а  $N$  – микроблоги из социальных сетей, будет произведена оценка тональности путем отношения сложения тональности всех  $k$  слов текста –  $S_k^{word}$ , перемноженных на вес важности определенного слова текста –  $tf-idf(word)$  к числу слов в тексте –  $n$ .

Агрегированная дневная оценка сентиментальности для текстов новостей и микроблогов определяется по формуле  $S_i^j = \sum_i S_i^j w_i$ , где  $w_i$  – доля длины отдельного текста в общей сумме слов во всех текстах  $i$  новостной или социальной категории текстов за день  $t$ .

### 2.3. Применение VAR с 50-дневным скользящим окном

В данном исследовании мы будем применять эконометрическое моделирование на основе VAR-модели с 50-дневным скользящим окном. Введем следующие обозначения:  $r_t^{MOEX}$  – 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. К данным, рассчитанным при помощи текстового анализа, будут отнесены:  $Sent_t^S$  – агрегированная дневная тональность микроблогов;  $Sent_t^N$  – агрегированная дневная тональность новостных статей;  $v_t^S$  – ежедневная частота появления слов, обозначающих радикальные события в социальных сетях (в данном случае – в микроблогах);  $v_t^N$  – ежедневная частота появления слов, обозначающих радикальные события в новостных статьях.

Для того чтобы понять, какие фундаментальные переменные следует включать в исследование, будет построена модель:

$$\begin{cases} r_t^{MOEX} = B_{1,0} + B_{1,1}r_{t-1}^{MOEX} + B_{1,2}r_{t-1}^{MIACR} + B_{1,3}r_{t-1}^{curr} + B_{1,4}r_{t-1}^{brent} + m_{1,t}; \\ r_t^{MIACR} = B_{2,0} + B_{2,1}r_{t-1}^{MOEX} + B_{2,2}r_{t-1}^{MIACR} + B_{2,3}r_{t-1}^{curr} + B_{2,4}r_{t-1}^{brent} + m_{2,t} \\ r_{t-1}^{curr} = B_{3,0} + B_{3,1}r_{t-1}^{MOEX} + B_{3,2}r_{t-1}^{MIACR} + B_{3,3}r_{t-1}^{curr} + B_{3,4}r_{t-1}^{brent} + m_{3,t}. \end{cases} \quad (1)$$

Система уравнений (1) позволит найти ответ на вопрос, объясняется ли динамика доходности фондового рынка России колебаниями на валютном, кредитном и нефтяном рынках.

Проверка взаимного влияния твитов и новостных статей на определение доминирующего положения одного из источников в конкретные периоды изучаемого времени (H1) будет проведена методом, реализованным в работе (Gan et al., 2020), а именно методом построения VAR-моделей с 50-дневным окном исследования и получением множества коэффициентов. Данные модели будут рассмотрены в следующей спецификации:

$$\begin{cases} Sent_t^S = A_{1,0} + A_{1,1}Sent_{t-1}^S + A_{1,2}Sent_t^N + m_{1,t}; \\ Sent_t^N = A_{2,0} + A_{2,1}Sent_{t-1}^S + A_{2,2}Sent_t^N + m_{2,t}. \end{cases} \quad (2)$$

Можно выделить четыре случая для такой системы VAR:

- 1) однонаправленная причинно-следственная связь – если  $A_{1,2}$  существенно отличается от нуля,  $A_{2,1}$  не отличается от нуля;
- 2) обратная однонаправленная причинность – если  $A_{2,1}$  существенно отличается от нуля, а  $A_{1,2}$  существенно не отличается от нуля;
- 3) обратная связь, или двусторонняя причинность, – когда и  $A_{1,2}$ , и  $A_{2,1}$  существенно отличаются от нуля;
- 4) независимость – когда ни  $A_{1,2}$ , ни  $A_{2,1}$  существенно не отличаются от нуля.

Для оценки наших гипотез будут построены две модели VAR на разных промежутках времени, выделенных на основе результатов, полученных при оценке системы (2) на всем исследуемом интервале. Эконометрические модели данного типа будут представлены в виде системы уравнений:

$$\begin{aligned} r_t^{MOEX} &= B_{1,0} + B_{1,1}r_{t-1}^{MOEX} + B_{1,2}r_{t-1}^{MIACR} + B_{1,3}r_{t-1}^{curr} + B_{1,4}r_{t-1}^{brent} + B_{1,5}Sent_{t-1}^N + B_{1,6}Sent_{t-1}^S + m_{1,t}; \\ r_t^{MIACR} &= B_{2,0} + B_{2,1}r_{t-1}^{MOEX} + B_{2,2}r_{t-1}^{MIACR} + B_{2,3}r_{t-1}^{curr} + B_{2,4}r_{t-1}^{brent} + B_{2,5}Sent_{t-1}^N + B_{2,6}Sent_{t-1}^S + m_{2,t}; \quad (3) \\ r_{t-1}^{curr} &= B_{3,0} + B_{3,1}r_{t-1}^{MOEX} + B_{3,2}r_{t-1}^{MIACR} + B_{3,3}r_{t-1}^{curr} + B_{3,4}r_{t-1}^{brent} + B_{3,5}Sent_{t-1}^N + B_{3,6}Sent_{t-1}^S + m_{3,t}; \end{aligned}$$

Отметим, что эти уравнения будут проверены на двух временных периодах в соответствии с нашими выдвинутыми гипотезами. В выделенных временных периодах будут построены кривые функции импульсного отклика (impulse response function, IRF), которые демонстрируют влияние шоков настроений в твитах и новостях на доходность российского рынка ценных бумаг.

### 3. Анализ данных

В работе были собраны следующие данные: индекс MOEX, Межбанковская кредитная ставка MIACR, валютный курс рубля к доллару за период с 2011 по 2020 г.<sup>6</sup> Ежедневные новостные статьи были взяты из Интернет-ресурса агрегатора новостей «РБК»<sup>7</sup>, поскольку данный ресурс освещает не только финансовые и экономические события, то были отобраны новостные статьи по следующим категориям: «финансы», «экономика», «деньги» и «бизнес», было собрано более 32 тыс. новостных статей. Ежедневные короткие сообщения (твиты) были взяты из социальной сети Твиттер, сообщения были взяты с аккаунтов компаний, составляющих Индекс МосБиржи. Стоит отметить, что не все компании, включенные в базу расчета данного инструмента, имеют аккаунт в данной социальной сети. В итоге 27 российских компаний написали более 111 тыс. твитов за период с 2011 г. по настоящее время.

Поскольку в данном исследовании используется VAR-моделирование, то необходимо было провести тест на стационарность, по расширенному тесту Дики-Фуллера все переменные являются стационарными. В табл. 1 представлена описательная статистика исследуемых переменных.

<sup>6</sup> Данные были собраны с investing.com

<sup>7</sup> www.rbc.ru

**Таблица 1**  
Описательная статистика экономических и текстовых переменных

Переменная	Среднее	Стандартная ошибка	Медиана	Стандартное отклонение	Дисперсия выборки
$r_t^{MOEX}$	0,0111	0,0253	0,0100	1,1757	1,3823
$r_t^{MIACR}$	0,0719	0,0722	0,0000	3,3567	11,2674
$r_t^{curr}$	0,0325	0,0217	0,0302	1,0100	1,0200
$r_t^{brent}$	-0,0223	0,0417	0,0380	1,9408	3,7668
$Sent_t^N$	-0,0065	0,0002	-0,0040	0,0093	0,0001
$Sent_t^S$	-0,0375	0,0008	-0,0372	0,0359	0,0013

Заметим, что факторы, представленные в работе, распределены в одном масштабе, поэтому коэффициенты при переменных в эконометрических моделях будут демонстрировать максимально приближенный к экономическим реалиям эффект. Также стоит отметить, что в среднем в 2011–2020 гг. настроение в новостях и микроблогах в России было негативное, что демонстрируют показатели «среднее» и «медиана». Данный факт соответствует тому, что в новостном потоке преобладают статьи с негативными заголовками. В табл. 2 представлена частота появления слов в текстах новостей и микроблогов.

**Таблица 2**  
Частота появления слов в текстах новостей и микроблогов

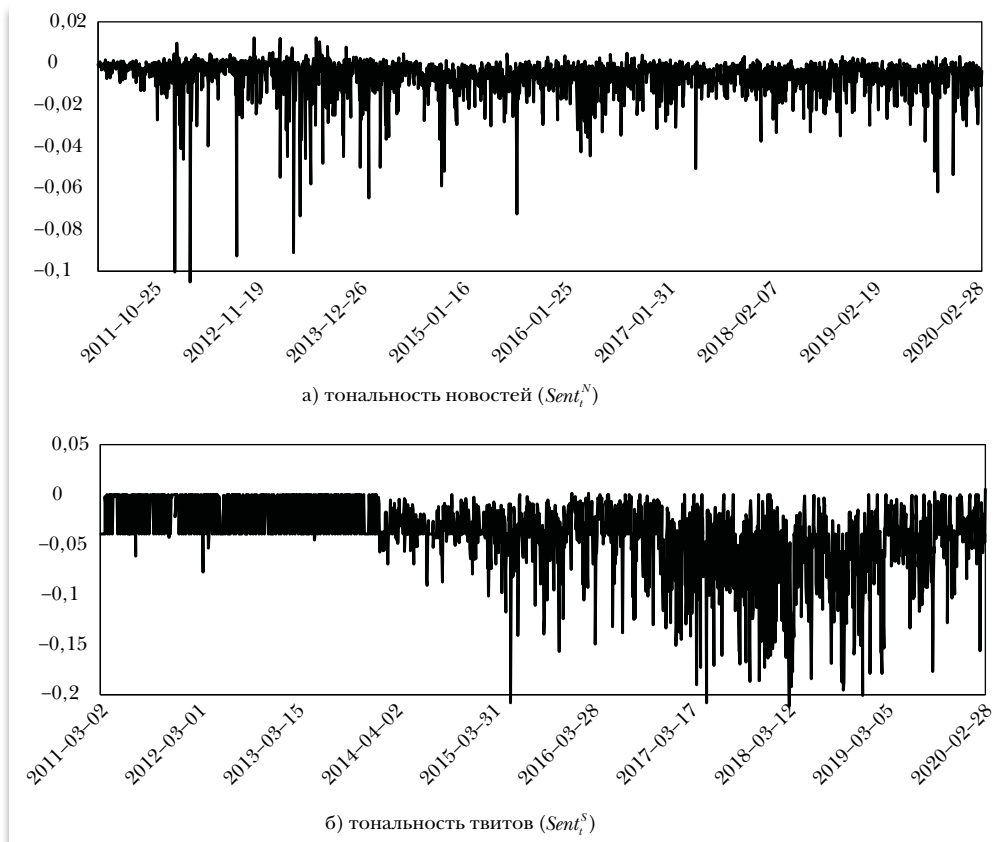
Новости		Социальные сети	
Сочетание слов (в начальной форме)	Частота	Сочетание слов (в начальной форме)	Частота
год	42 695	карта	7094
банк	34 178	банк	5615
который	26 334	система	3056
компания	25 943	информация	2753
млрд	22 866	фск	2733
руб	22 557	еэс	2618
россия	21 280	коллега	2521
рынок	20 996	который	2513
российский	17 781	компания	2494
цена	15 083	операция	2397

Окончание таблицы 2

Новости		Социальные сети	
Сочетание слов (в начальной форме)	Частота	Сочетание слов (в начальной форме)	Частота
составлять	14 807	работа	2364
рост	14 472	вопрос	2357
акция	13 910	время	2248
нефть	12 094	первый	2181
фото	11 668	индекс	2154
сообщать	10 796	россия	2142
страна	9639	получать	1991
долл	9375	принимать	1975
уровень	9315	перевод	1904

Из анализа частотности данных в табл. 2 можно понять, какие слова популярны в новостном и социальном поле. Стоит отметить, что язык новостного потока более четко публицистически направленный, более официальный, в то время как социальные сети больше содержат разговорных слов для межличностного взаимодействия на неофициальном уровне. Но как данные особенности двух диаметрально противоположных по стилю источников информации могут отразиться на тональности текстов в рамках построения эконометрической модели? Для ответа на этот вопрос были построены графики на рис. 3.

Анализируя данные на рис. 3(а, б), можно отметить, что негативная тональность превалирует на исследуемом промежутке вне зависимости от того, какой источник информации рассматривается. Причем реакция инвесторов, экономических агентов на такие негативные экономические шоки, как, например, валютный кризис в России в 2014–2015 гг. и пандемия, связанная с распространением COVID-19 (конец 2018–2019 гг.), отражается резким скачком негативных значений тональности твиттов (рис. 3б) в данных промежутках времени (с  $-0,05$  до  $-0,2$ ). А подобный, негативный, импульс в новостях проявляется значительно слабее (с  $-0,05$  до  $0,1$ ).



**Рис. 3**

Динамика тональностей за исследуемый период времени

**Таблица 3**

Оценка отдельных новостных текстов и ее сравнение с заголовками публикаций

Заголовок новостной статьи	Сентиментальная оценка
Shell вышел из сделки с «Газпромнефтью» по освоению месторождений в ЯНАО :: Бизнес :: РБК	-0,0108
Основатели ИТ-интегратора IBS заявили о разделе бизнеса :: Бизнес :: РБК	0,0136
Внешний долг России упал до минимума за 11 лет из-за ослабления рубля :: Экономика :: РБК	-0,0909
Российские банки в марте выдали рекордный объем кредитов населению :: Финансы :: РБК	0,0592
Bloomberg назвал причину изменения позиции России по нефтяной сделке :: Экономика :: РБК	-0,039
ЦБ разрешил Сбербанку занять 500 млрд :: Финансы :: РБК	0,0597
Приток иностранных инвестиций в Россию остановился :: Экономика :: РБК	0,013

Окончание таблицы 3

Заголовок новостной статьи	Сентиментальная оценка
Власти раскрыли размеры принадлежащей украинцам земли в Крыму :: Экономика :: РБК	-0,02
Fitch впервые ухудшило прогноз рейтинга Госбанка из-за коронавируса :: Финансы :: РБК	-0,0491
Алексей Кудрин – о нарастающем кризисе. Полная версия интервью :: Экономика :: РБК	-0,1
Половина работающих россиян сообщили о сокращении доходов из-за вируса :: Экономика :: РБК	-0,0482
Путин поручил рассмотреть возможность субсидирования доставки еды :: Бизнес :: РБК	0,1268
Кремль назвал способы избежать падения уровня доходов россиян :: Экономика :: РБК	0,0476
Кудрин предупредил об ударе следующей стадии кризиса по банкам :: Финансы :: РБК	-0,1429
Путин назвал «справедливую формулу» помощи бизнесу из-за вируса :: Бизнес :: РБК	0,0252
Минфин не смог продать ОФЗ на первом аукционе после обвала рынков и рубля :: Финансы :: РБК	-0,0217
Титов не исключил падения ВВП более чем на 10% в 2020 году :: Экономика :: РБК	-0,169
Сбербанк назвал категории товаров с падением трат практически до нуля :: Финансы :: РБК	-0,0192

В табл. 3 демонстрируется распространенный факт о том, что большинство событий, освещаемых новостными изданиями, – негативные. Прочитав заголовок той или иной статьи, каждый человек формирует свое отношение к тексту, еще не взглянув на него. Данные сентиментальные оценки текстов новостей подтверждают, что если в заголовке есть слово «кризис» или слово «вирус», то негативность самого заголовка перетекает в отрицательное настроение самого текста.

#### 4. Результаты исследования

В табл. 4 показано влияние динамики фундаментальных экономических переменных на доходность фондового индекса. При построении модели учитывался тот факт, что ставка межбанковского кредитования оказалась незначимой; вследствие этого в дальнейших моделях она была исключена из вектора объясняющих переменных. Результаты моделирования представлены ниже.

Одним из самых значимых показателей, влияющих на доходность фондового рынка России, является курс валюты, потому что отечественные компании, входящие в биржевой индекс, экспортируют свои товары и услуги на мировой рынок, где расчеты производятся в долларах.

Данные табл. 5 показывают, что коэффициенты при сентиментальных переменных являются положительными и значимыми на уровне 5% (подтвержде-

**Таблица 4**

Экономические переменные как факторы доходности российского рынка

Переменная	Значение коэффициента при переменной	Стандартное отклонение	t-статистика	P-value
<i>const</i>	-0,056215	0,068474	-0,821	0,412
$r_{t-1}^{MOEX}$	0,026153	0,055374	0,472	0,637
$r_{t-1}^{MIACR}$	-0,017887	0,014605	-1,225	0,221
$r_{t-1}^{curr}$	-0,268610**	0,123988	-2,166	0,030**
$r_{t-1}^{brent}$	0,102766**	0,051360	-2,001	0,045**

**Примечание.** В таблице символами «\*», «\*\*», «\*\*\*» отмечены оценки, значимые на уровне 10, 5 и 1% соответственно.

ние гипотезы H1). Значение выше нуля при данных временных рядах свидетельствует о том, что позитивная информация положительно отражается на доходности рынка, что соотносится с предыдущими исследованиями в данной области. Также стоит отметить незначимость изменения индекса МосБиржи в прошлом периоде. Это может быть следствием особенности подобранных для исследования данных. В дневном периоде инвесторы уделяют меньше внимания предыдущим значениям самого рынка, ориентируясь больше на «рыночное настроение» и макроэкономические факторы.

**Таблица 5**

Влияние совокупности сентиментальных и экономических переменных на доходность российского рынка

Переменная	Значение коэффициента при переменной	Стандартное отклонение	t-статистика	P-value
Зависимый фактор – рыночная доходность ( $r_{t-1}^{MOEX}$ )				
<i>const</i>	-0,082589**	0,040010	-2,064	0,039**
$r_{t-1}^{MOEX}$	0,016339	0,022411	0,729	0,466
$r_{t-1}^{curr}$	-0,062648**	0,028239	-2,219	0,027**
$r_{t-1}^{brent}$	0,011695	0,014876	-0,786	0,432
$Sent_{t-1}^N$	0,566471**	0,270859	2,091	0,036**
$Sent_{t-1}^S$	0,161562**	0,070044	2,307	0,021**

Продолжение таблицы 5

Переменная	Значение коэффициента при переменной	Стандартное отклонение	t-статистика	P-value
Зависимый фактор – настроение новостей ( $Sent_{t-1}^N$ )				
<i>const</i>	0,055908***	0,003171	17,629	0,000***
$r_{t-1}^{MOEX}$	0,000018	0,001776	0,010	0,992
$r_{t-1}^{curr}$	-0,000176	0,002238	-0,079	0,937
$r_{t-1}^{brent}$	0,001452	0,001179	1,232	0,218
$Sent_{t-1}^N$	0,079339***	0,021470	3,695	0,000***
$Sent_{t-1}^S$	0,009876*	0,005552	1,779	0,075*
Зависимый фактор – настроение Твиттера ( $Sent_{t-1}^S$ )				
<i>const</i>	0,183876***	0,010628	17,301	0,000***
$r_{t-1}^{MOEX}$	0,002843	0,005953	0,477	0,633
$r_{t-1}^{curr}$	0,003855	0,007501	0,514	0,607
$r_{t-1}^{brent}$	0,002022	0,003952	0,512	0,609
$Sent_{t-1}^N$	0,033689	0,071952	0,468	0,640
$Sent_{t-1}^S$	0,503984***	0,018607	27,086	0,000***
Зависимый фактор – валютный курс ( $r_{t-1}^{curr}$ )				
<i>const</i>	0,062782*	0,034458	1,822	0,068*
$r_{t-1}^{MOEX}$	-0,049506***	0,019301	-2,565	0,010***
$r_{t-1}^{curr}$	0,052097**	0,024320	2,142	0,032**
$r_{t-1}^{brent}$	-0,041071***	0,012812	-3,206	0,001***
$Sent_{t-1}^N$	-0,000921	0,233273	-0,004	0,997
$Sent_{t-1}^S$	-0,082414	0,060324	-1,366	0,172
Зависимый фактор – цена на нефть марки Brent ( $r_{t-1}^{brent}$ )				
<i>const</i>	0,115287*	0,066068	1,745	0,081*
$r_{t-1}^{MOEX}$	-0,108982***	0,037007	-2,945	0,003***

Окончание таблицы 5

Переменная	Значение коэффициента при переменной	Стандартное отклонение	<i>t</i> -статистика	P-value
$r_{t-1}^{curr}$	0,035165	0,046630	0,754	0,451
$r_{t-1}^{bent}$	-0,104840***	0,024565	-4,268	0,000***
$Sent_{t-1}^N$	-0,373238	0,447268	-0,834	0,404
$Sent_{t-1}^S$	-0,180484	0,115664	-1,560	0,119

**Примечание.** В таблице символами «\*», «\*\*», «\*\*\*» отмечены оценки, значимые на уровне 10, 5 и 1% соответственно.

Если рассматривать сентиментальные переменные – тональность новостей и твитов в качестве зависимых факторов в рамках построенной VAR-модели, то можно заметить ряд интересных фактов. Влияние рыночной доходности оказалось очень мало в абсолютных величинах и незначимо при любом разумном уровне значимости. Но наблюдается чувствительность одной сентиментальной переменной к другой. Данный факт наталкивает на мысль о том, что тональность одних источников информации может влиять на другие, а направленность влияния может меняться во времени с развитием технологий и методов коммуникации между людьми.

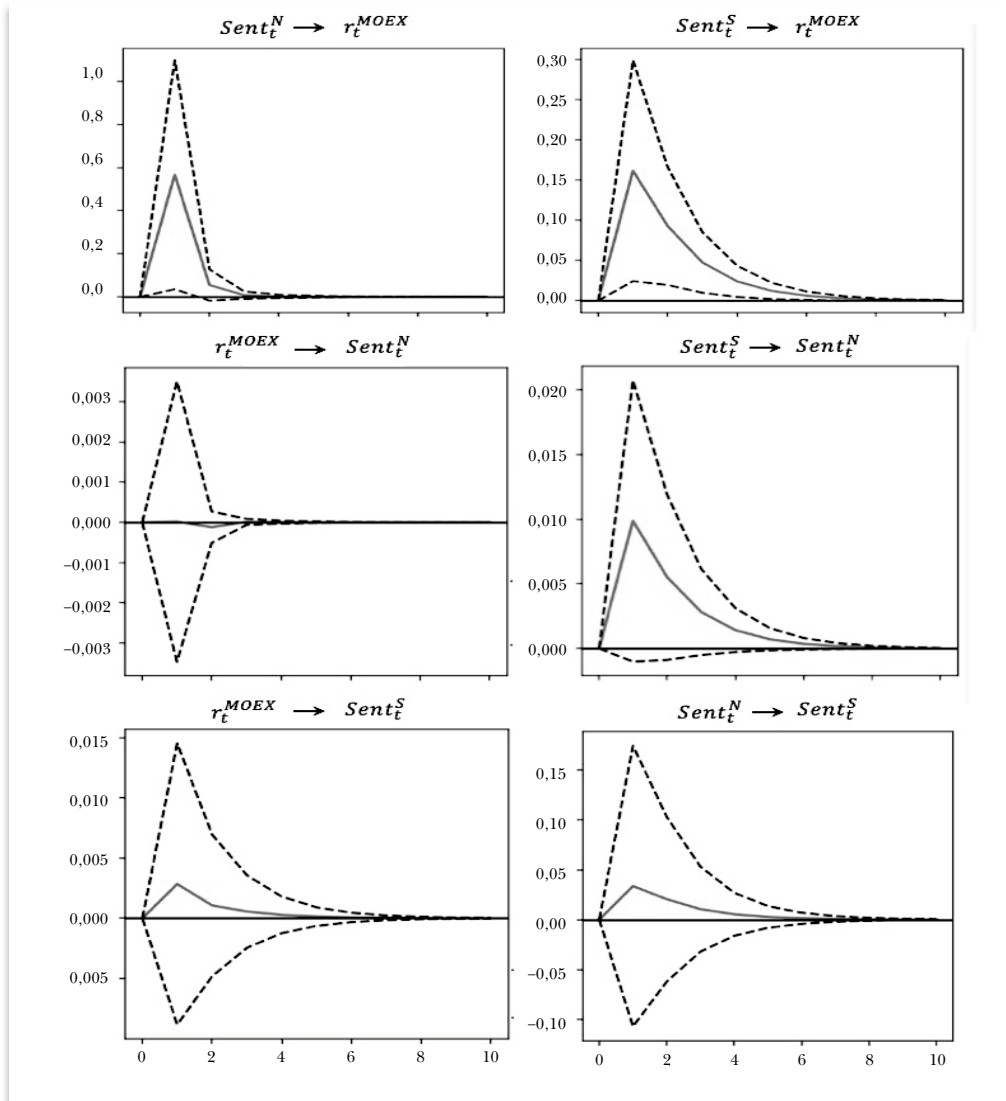
Анализ графиков функций реакций импульсов (рис. 4) показывает, что рыночная доходность положительно откликается на шоки тональностей, новостей и твитов, но с разной амплитудой (более чувствительна к шокам новостей).

Проверка гипотез осуществляется в несколько этапов. Сначала была оценено взаимное влияние сентиментальных переменных, описывающих тексты новостей и микроблогов при помощи системы (2). VAR-модель была построена на скользящем окне в 50 периодов (иными словами, каждые 50 дней, начиная с 2011 г.). Результаты представлены на рис. 5.

Полужирной линией на рис. 5 представлена динамика коэффициента влияния настроения микроблогов на новостной фон (влияние *S* на *N*), тонкой линией – динамика коэффициента влияния статей в СМИ на твиты (влияние *N* на *S*). Точками обозначены значимые коэффициенты среди всех построенных VAR-моделей. Таким образом, можно заметить, что в период с начала 2011 г. по июнь 2015 г. новости влияют на тональность микроблогов, т.е. имеется односторонняя связь между переменными. Между июнем 2015 г. и сентябрем 2017 г. наблюдается так называемый переходный, или трансформационный, период. В данном интервале времени наблюдается взаимное влияние тональностей СМИ и твитов российских компаний. И, наконец, после сентября 2017 г. тональность микроблогов становится доминантой для определения настроения инвесторов (большинство значимых коэффициентов, при отсутствии значимости у настроения новостей как объясняющей переменной).

Далее, учитывая полученные временные интервалы, а именно предпереходный (01,01,2011–01,06,2015) и постпереходный (01,09,2017–01,04,2020)

периоды, строятся регрессионные модели по системе (3) на каждом периоде. Результаты моделирования представлены в табл. 6.



**Рис. 4**

Графики функций импульсного отклика для переменной доходности российского фондового рынка и сентиментальных переменных



**Рис. 5**

Динамика коэффициентов системы (2) со скользящим окном в 50 дней и их значимость

**Таблица 6**

Построение моделей VAR на двух выявленных периодах

Переменная	Значение коэффициента при переменной	Стандартное отклонение	t-статистика	P-value
Предпереходный период (01.01.2011–01.06.2015)				
<i>const</i>	-0,105957	0,090498	-1,171	0,242
$r_{t-1}^{MOEX}$	0,032459	0,055364	0,586	0,558
$r_{t-1}^{curr}$	-0,278941**	0,123965	-2,250**	0,024
$r_{t-1}^{brent}$	0,110393**	0,051302	2,152**	0,031
$Sent_{t-1}^N$	1,316064**	0,661064	1,991**	0,046
$Sent_{t-1}^S$	-0,027817	0,386381	-0,072	0,943
Постпереходный период (01.09.2017– 01.04.2020)				
<i>const</i>	-0,064794	0,072472	-0,894	0,371

Окончание таблицы 6

Переменная	Значение коэффициента при переменной	Стандартное отклонение	t-статистика	P-value
$r_{t-1}^{MOEX}$	-0,036155	0,041807	-0,865	0,387
$r_{t-1}^{curr}$	-0,172067***	0,058348	-2,949	0,003
$r_{t-1}^{brent}$	0,030073	0,020848	1,442	0,149
$Sent_{t-1}^N$	0,182116	0,486358	0,374	0,708
$Sent_{t-1}^S$	0,189476**	0,085992	2,203	0,028

**Примечание.** В таблице символами «\*», «\*\*», «\*\*\*» отмечены оценки, значимые на уровне 10, 5 и 1% соответственно.

В первую очередь нас интересуют коэффициенты первого уравнения в пред- и постпереходный периоды. Заметим, что в первом коэффициент при сентиментальной новостной переменной значим на уровне 5% и имеет положительный знак. Этот факт свидетельствует о том, что положительная информация способствует росту цены на активы, торгуемые на российском фондовом рынке. Иначе говоря, настроение инвесторов задает положительную тенденцию движения рынка. В то время как коэффициент при сентиментальной переменной, посчитанный на текстах микроблогов, не значим на любом разумном уровне. Это сигнализирует о том, что только новости в предпереходном периоде являются фактором ценообразования активов. Если рассматривать постпереходный период, то наблюдается обратная картина: здесь значим коэффициент при социальных сентиментах и статистически равен нулю коэффициент при новостных сентиментах, что показывает доминацию социального влияния на ценообразование активов в постпереходном периоде. Данные рассуждения приводят нас к выводу о том, что выдвинутая ранее гипотеза H2 подтверждается.

Стоит обратить внимание еще на один факт. Значение коэффициентов при сентиментальных переменных без учета значимости в предпереходном периоде при новостных сентиментах по модулю выше, чем при социальных сентиментах, а в постпереходном (более позднем) – ситуация противоположная.

Для модели, представленной выше, были построены функции импульса (рис. 6).

На рис. 6 представлены графики реакции временного ряда доходности российского рынка на шоки объясняющих переменных. Следуя данной аналогии, можно отметить, что все представленные шоки стремятся к нулевому значению через несколько периодов (или дней), что указывает на краткосрочное влияние шока и стремление рынка к равновесному состоянию. Обратим внимание на шоки сентиментальных переменных. В предпереходный период более существенная реакция на шоки новостной тональности, а в постпереходный – на шоки тональности социальных сетей. Краткосрочность таких шоков указывает на то, что внезапные всплески настроений инвесторов приводят к временному скачку цен на акции, которые скоро возвращаются на близкий к исходному уровень.

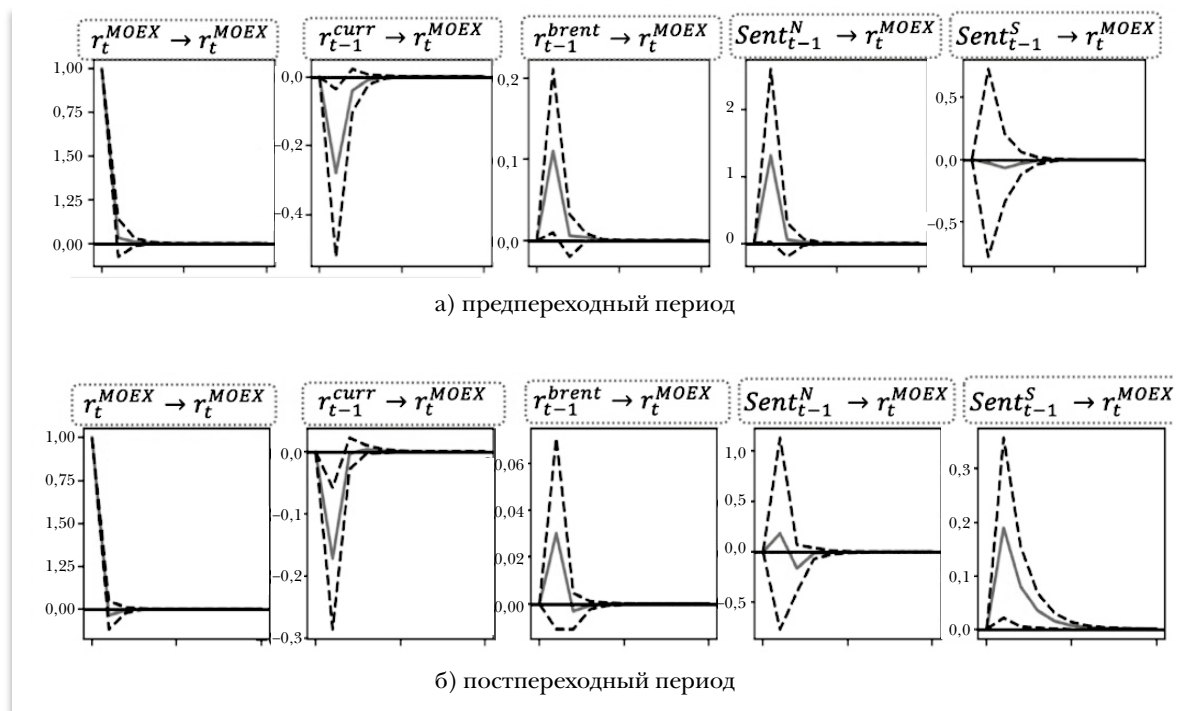


Рис. 6

Функции реакции доходности рынка ценных бумаг России на шоки экономических и сенсительных переменных

### Заключение

Сенсительные переменные, а именно тональность новостей и тональность социальных сетей (микроблогов), могут быть факторами ценообразования активов наравне с такими экономическими показателями, как цена на нефть марки Brent и валютный курс USD/RUB. Более того, были установлены следующие периоды:

- 1) одностороннее влияние тональности новостей на социальное настроение (предпереходный период);
- 2) взаимовлияние сенсительности новостей и микроблогов (переходный период);
- 3) одностороннее влияние тональности социальных сетей на новости.

Данные периоды позволили показать, что изначально новости были доминирующим транслятором инвестиционных настроений и убеждений, которые позитивно отражались на доходности фондового рынка. Но с течением времени микроблоги из российских аккаунтов Твиттера стали все сильнее и сильнее влиять на ценообразование активов и на данных последних двух лет перехватили инициативу у новостных изданий.

Ограничения в исследовании состоят в следующих проблемах:

- 1) мало компаний в России, входящих в Индекс МосБиржи, имеют аккаунты в социальных сетях и, более того, предоставляют там информацию не о своих коммерческих предложениях для конечных потребителей (например, сезонные скидки);

- 2) слабая мощность (число слов и словоформ) разработанного нами словаря, которая в дальнейшем будет увеличена, трудности во взаимодействии с другими социальными сетями в России, например «ВКонтакте», в силу норм законодательства и отсутствие утилит для доступа к текстовым данным;
- 3) новости «РБК» могут содержать оценочные мнения журналистов и поэтому условно относиться к официальным источникам.

Однако данное исследование является практически одним из первых, в котором сравнивают одновременно текстовый анализ социальных сетей и новостей.

#### ЛИТЕРАТУРА / REFERENCES

- Афанасьев Д.О., Федорова Е.А., Демин И.С., Рогов О.Ю.** (2020). О влиянии зарубежных СМИ на российский фондовый рынок: текстовый анализ // *Экономика и математические методы*. Т. 56. № 1. С. 153–165 [**Afanasyev D.O., Fedorova E.A., Demin I.S., Rogov O.Yu.** (2020). On the influence of foreign media on the Russian stock market: Text analysis. *Economics and Mathematical Methods*, 56, 1, 153–165 (in Russian).]
- Дробышевский С.М., Трунин П.В., Божечкова А.В., Горюнов Е.Л., Петрова Д.А.** (2017). Анализ информационной политики Банка России // *Вопросы экономики*. № 10. С. 88–110. [**Drobyshevsky S.M., Trunin P.V., Bozhechkova A.V., Gorunov E.L., Petrova D.A.** (2017). Analysis of the Bank of Russia information policy. *Voprosy Ekonomiki*, 10, 88–110 (in Russian).]
- Федорова Е.А., Афанасьев Д.О., Демин И.С., Пыльцин И.В., Нерсесян Р.Г., Лазарев А.М.** (2020). Разработка тонально-тематического словаря EcSentiThemeLex для анализа экономических текстов на русском языке // *Прикладная информатика*. Т. 15. № 6. С. 58–77. DOI: 10,37791/2687-0649-2020-15-6-58-77 [**Fedorova E.A., Afanasyev D.O., Demin I.S., Pyl'tsin I.V., Nersesjan R.G., Lazarev A.M.** (2020). Development of the EcSentiThemeLex tonal and thematic dictionary for the analysis of economic texts in Russian. *Applied Informatics*, 15, 6, 58–77. DOI: 10,37791/2687-0649-2020-15-6-58-77 (in Russian).]
- Федорова Е.А., Панкратов К.А.** (2010). Влияние макроэкономических факторов на фондовый рынок России // *Проблемы прогнозирования*. № 2 (119). С. 78–83. [**Fedorova E.A., Pankratov K.A.** (2010). Influence of macroeconomic factors on the Russian stock market. *Studies on Russian Economic Development*, 2 (119), 78–83 (in Russian).]
- Федорова Е.А., Хрустова Л.Е., Демин И.С.** (2020). Влияние качества раскрытия нефинансовой информации российскими компаниями на их инвестиционную привлекательность // *Российский журнал менеджмента*. Т. 18. № 1. С. 51–72. [**Fedorova E.A., Khrustova L.E., Demin I.S.** (2020). Completeness of non-financial disclosure by Russian companies: The influence on investment attractiveness. *Russian Management Journal*, 18, 1, 51–72 (in Russian).]
- Abbassi P., Iyer R., Peydró J.-L., Tous F.R.** (2016). 'Securities Trading by Banks and Credit Supply: Micro-Evidence from the Crisis'. *Journal of Financial Economics*, 121 (3), 569–594.
- Borjigin S., Gao T., Sun Y., An B.** (2020). For evil news rides fast, while good news baits later? A network based analysis in Chinese stock market. *Physica A: Statistical Mechanics and Its Applications*, 551 (August), 124593.
- Cevik N.K., Cevik E.I., Diboğlu S.** (2020). Oil prices, stock market returns and volatility spillovers: Evidence from Turkey. *Journal of Policy Modeling*, 42 (3), 597–614.
- Gan B., Alexeev V., Bird R., Yeung D.** (2020). Sensitivity to sentiment: News vs social media.

- International Review of Financial Analysis*, 67 (January), 101390.
- Guan-Ru W.G., Hou T., Lin J.** (2019). Can economic news predict Taiwan stock market returns? *Asia Pacific Management Review*, 24 (1), 54–59.
- Kilian L.** (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99 (3), 1053–1069.
- Maadid A., Caporale M., Spagnolo F., Spagnolo N.** (2020). The impact of business and political news on the GCC stock markets. *Research in International Business and Finance*, 52, 101102.
- Nisar T.M., Yeung M.** (2018). Twitter as a tool for forecasting stock market movements: A short-window event study. *The Journal of Finance and Data Science*, 4 (2), 101–119.
- 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*, 73 (May), 125–144.
- Patro D.K., Wald J., Wu Y.** (2014). Currency devaluation and stock market response: An empirical analysis. *Journal of International Money and Finance*, 40, 79–94.
- Tachibana M.** (2018). Relationship between stock and currency markets conditional on the US stock returns: A vine copula approach. *J. of Multi. Fin. Manag.*, 46, 75–106.
- Thorbecke W.** (2019). How oil prices affect east and Southeast Asian economies: Evidence from financial markets and implications for energy security. *Energy Policy*, 128 (May), 628–638.

Поступила в редакцию 20.05.2019

Received 20.05.2019

**E.A. Fedorova**

Department of Corporate Finance and Corporate Governance, Financial University under the Government of the Russian Federation, Moscow, Russia

**I.V. Pyltsin**

NRU-HSE, Department of Economic Sciences, Moscow, Russia

**Yu.A. Kovalchuk**

MGIMO MFA of Russia, Moscow, Russia

**P.A. Drogovoz**

Bauman Moscow State Technical University, Moscow, Russia

## **News and social networks of Russian companies: Degree of influence on the securities market**

**Abstract.** The main purpose of this work is to determine the periods in which the profitability of the Russian stock market is partially explained by official information (RBC news) and unofficial (tweets of companies). The methodology of the study included a VAR model with a sliding window of 50 days. The empirical base of the study included 32 thousand news articles and 111 thousand tweets for 2011–2020. The assessment of the tonality of the text was carried out on the basis of the EcSentiThemeLex dictionary, this dictionary is one of the first dictionaries that allow assessing the tonality of economic and financial texts in Russian by five tonalities. The results of the study show that the profitability of the market is explained both by economic indicators (the price of oil and the exchange rate) and by the tone of textual information. Moreover, from 2011 to 2015, the tone of the news background was the dominant information factor in the profitability of the securities market, and in 2017–2020 – dominance has shifted to the mood of social media texts. A short-term period of influence of textual information was also revealed, sudden bursts of investor

sentiment lead to a temporary jump in stock prices, which quickly return to close the initial level.

**Keywords:** *text analysis, tweets, RBC news, stock market.*

JEL Classification: C51, G11, G17.

For reference: **Fedorova E.A., Pylytsin I.V., Kovalchuk Yu.A., Drogozov P.A.** (2022). News and social networks of Russian companies: Degree of influence on the securities market. *Journal of the New Economic Association*, 1 (53), 32–52. DOI: 10,31737/2221-2264-2022-53-1-2