

ЭФФЕКТ СПЛОЧЕНИЯ В (РУССКОЯЗЫЧНЫХ)  
СОЦИАЛЬНЫХ СЕТЯХ:  
ПРОБЛЕМЫ ИЗМЕРЕНИЯ

Анкудинов Иван Андреевич

Москва, 2023

# Сегодня в повестке

## 1. Введение

- Проблема
- Формальности
- Теория

## 2. Результаты

- Данные
- Методы
- Спецификации
- Обсуждение

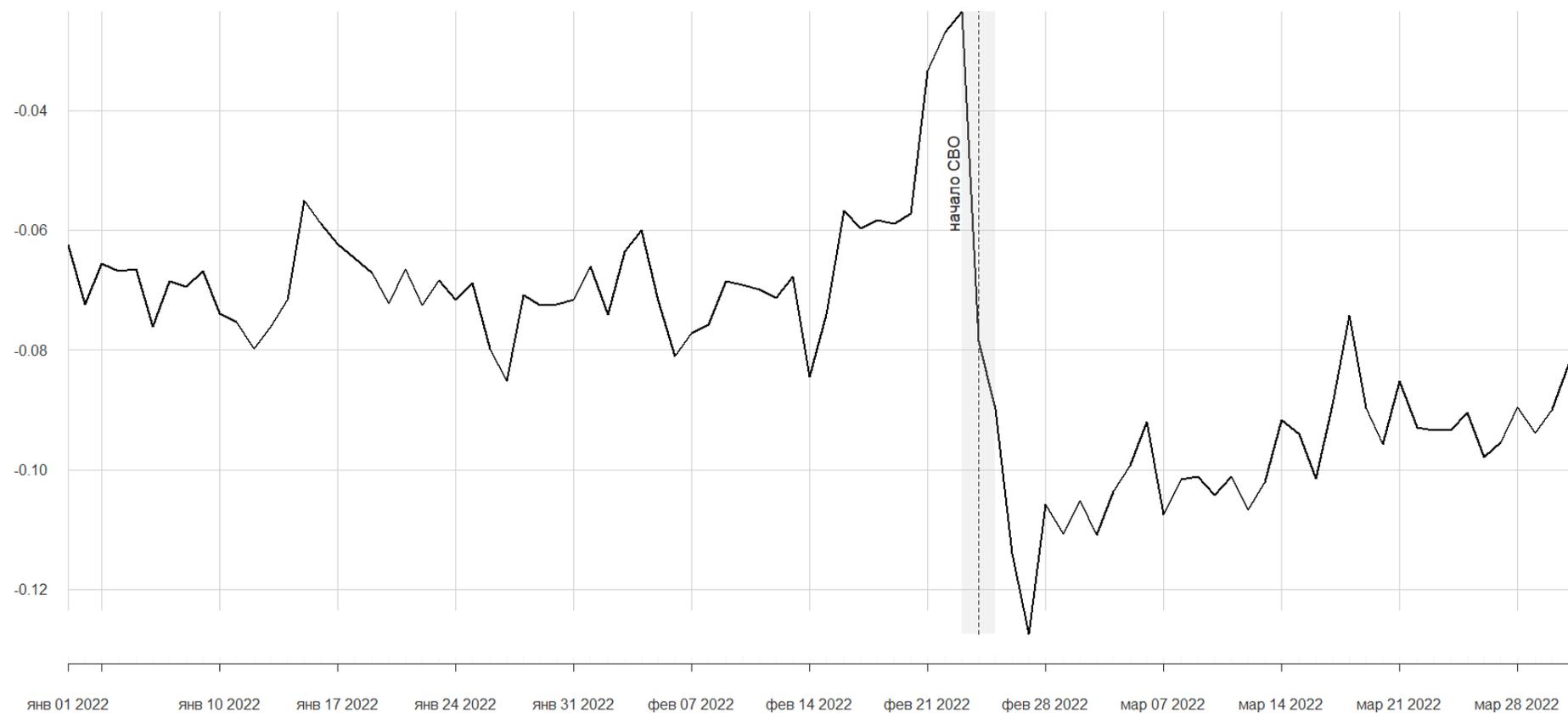
## 3. Выведение

# Проблема

- После 24 февраля россияне испытали патриотический подъем, по масштабу сравнимый с «Крымским» [ВЦИОМ, 2022; ФОМ, 2022]
- Вопрос о «просачивании» этого эффекта за пределы цифр поллстеров остается открытым (массовые митинги? акции? публичные заявления?)
- Рунет, невероятно смещенный и нерепрезентативный, дает возможность оценить динамику патриотического дискурса по цифровым следам
- С одной стороны, социальные сети склонны «перепредставлять» оппозиционную аудиторию [Reuter, Szakonyi, 2015]
- С другой, информационные потоки стали жестче контролироваться, а самыми популярными наполнителями новостных лент стали Z-каналы [Urman, Makhortykh, 2022]

# Проблема

2022-01-01 / 2022-04-01



# Формальности

- **Вопрос:** насколько (количественно, качественно) изменилось наполнение патриотических публикаций после 24 февраля 2022 г.?
- **Объект:** «патриотические» публикации
- **Предмет:** динамика их содержания после начала СВО
- **Цель:** подтвердить или опровергнуть «нулевую» гипотезу об отсутствии просачивания на материале социальных сетей

# Теория

- Россия – на самом деле, неплохо представленная в Интернете страна [89% в глобальной сети, 73% в социальных – WebCанаре]
- Тем не менее, пользователей Рунета лучше рассматривать не как выборочную, а как самостоятельную генеральную совокупность [см. Smetanin, 2020 и разбор внутри]
- Распространение среди них лоялистских настроений сходно с распространением протестных, но с меньшим «коэффициентом переноса» из онлайн в офлайн [Spaiser, 2017]
- Однако наша **основная гипотеза** заключается в том, что действие сильных разнонаправленных потоков информации (условно говоря, проправительственного и антиправительственного) привело к **нивелированию эффекта сплочения**, и какого-либо значимого «перекоса» в пользу позитивных упоминаний за спецоперацией не последовало

# Данные

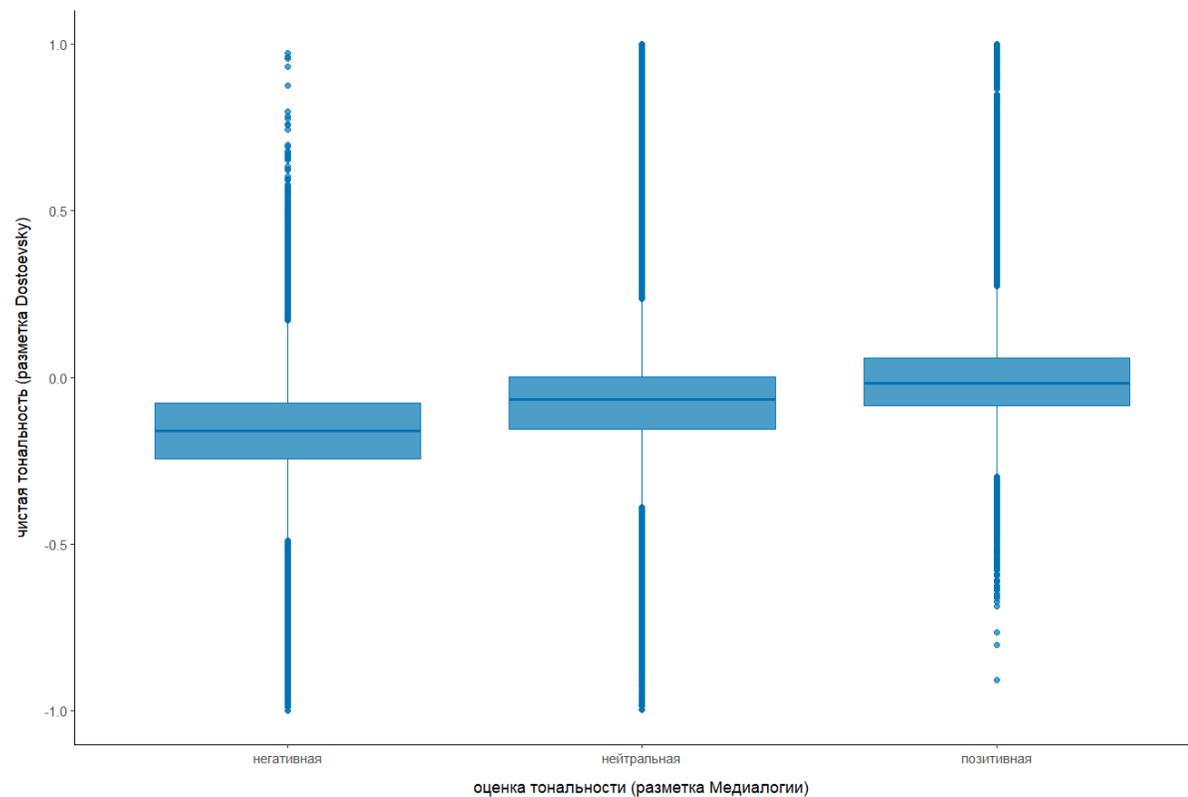
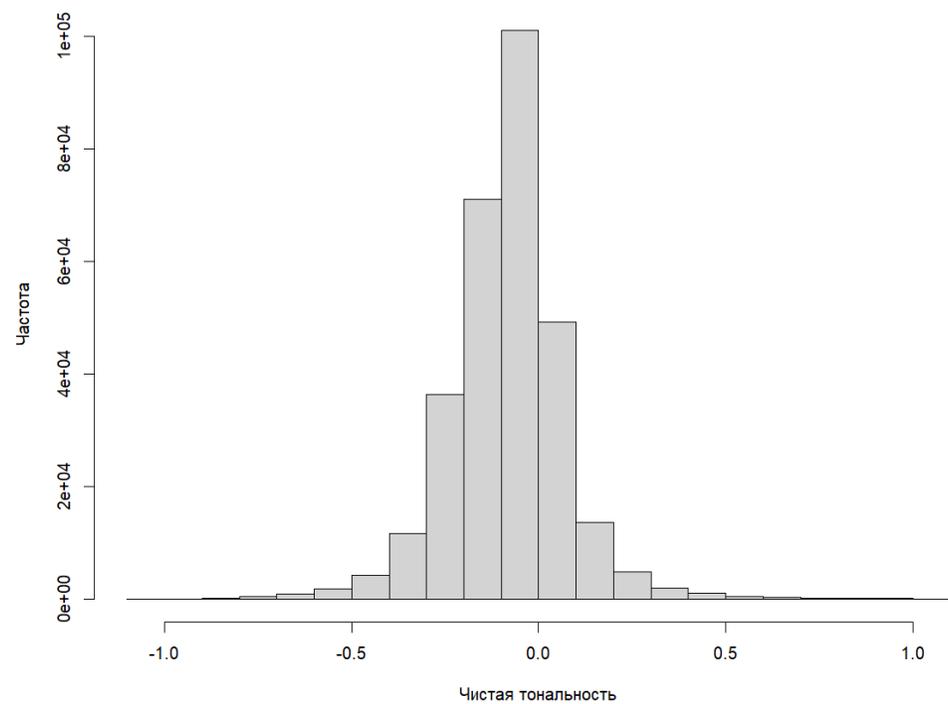
- Период: 01.01.2022 – 01.04.2022. Выглядят примерно так:

№	Дата	Тип	Сообщение	Автораспознанный текст	Площадка	Тип площадки	Ссылка	Автор	Профиль автора	...	Страна	Регион	Город	Теги	Спам
1	01.04.2022 23:59	Пост	На тему низкоклонства. Константин Симонов.\n...	NaN	ВКонтакте	Соцсеть	<a href="http://vk.com/wall489949567_14536">http://vk.com/wall489949567_14536</a>	Артур Назаров	<a href="http://vk.com/id489949567">http://vk.com/id489949567</a>	...	Россия	Новосибирская	Новосибирск	NaN	NaN
2	01.04.2022 23:59	Пост	<a href="https://ru.wikipedia.org/wiki/Гоголь,_Николай_...">https://ru.wikipedia.org/wiki/Гоголь,_Николай_...</a>	NaN	Одноклассники	Соцсеть	<a href="http://ok.ru/profile/541751803575/statuses/154...">http://ok.ru/profile/541751803575/statuses/154...</a>	Анна Сурис	<a href="http://ok.ru/profile/541751803575">http://ok.ru/profile/541751803575</a>	...	Россия	Санкт-Петербург	Санкт-Петербург	NaN	NaN
3	01.04.2022 23:58	Пост	СПАСИБО ЗА ПЕСНЮ,СЕРГЕЙ ИСТИННЫЙ ПАТРИОТ СВОЕЙ...	NaN	ВКонтакте	Соцсеть	<a href="http://vk.com/wall167830298_5210">http://vk.com/wall167830298_5210</a>	Валентина Беляева	<a href="http://vk.com/id167830298">http://vk.com/id167830298</a>	...	Россия	Вологодская	Бекетовская	NaN	NaN
4	01.04.2022 23:57	Пост	Британская компания Shell, вероятно, не сможет...	NaN	ВКонтакте	Соцсеть	<a href="http://vk.com/wall46074126_1047">http://vk.com/wall46074126_1047</a>	Елена Евграфова	<a href="http://vk.com/id46074126">http://vk.com/id46074126</a>	...	Россия	Москва	Москва	NaN	NaN
5	01.04.2022 23:57	Пост	Войну выигрывают патриоты а не те у кого недви...	NaN	Одноклассники	Соцсеть	<a href="http://ok.ru/profile/526754119856/statuses/154...">http://ok.ru/profile/526754119856/statuses/154...</a>	Анатолий Зиновьев	<a href="http://ok.ru/profile/526754119856">http://ok.ru/profile/526754119856</a>	...	Россия	Москва	Москва	NaN	NaN

- Выгружены с помощью «Медиалогии» по словарю синонимов (15 производных и эрратизмов слова «патриот» + близкие слова)
- 300 809 записей (всего – 3,1 млн), 65 529 уникальных авторов
- Очищены от спама, дополнены оценками тональности Dostoevsky

# Данные

- В разрезе зависимых переменных:



# Методы

- Предварительно: машинная оценка тональности
- Регрессионный анализ: «кусочная», LSDV-, LDV-, PPO-модели
- Тематическое моделирование: ?
  - Есть вопросы
  - Перебор числа тем до получения осмысленного результата? Использование расчетных метрик вроде perplexity / predictive likelihood / topic coherence?
  - Сопоставление относительных весов тем между периодами? Грубое сравнение «топа-10»? Изменения в распределении по авторам?
- Другие способы обработки больших текстов?

# Спецификации

- Общее уравнение для моделей «экзогенного шока»:

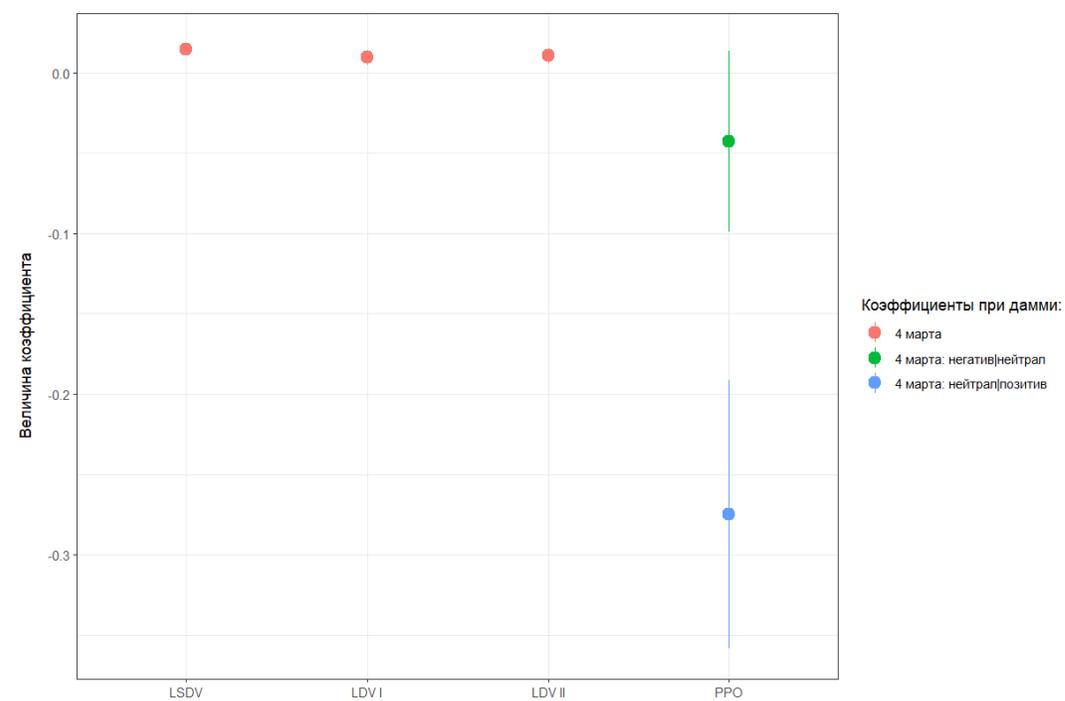
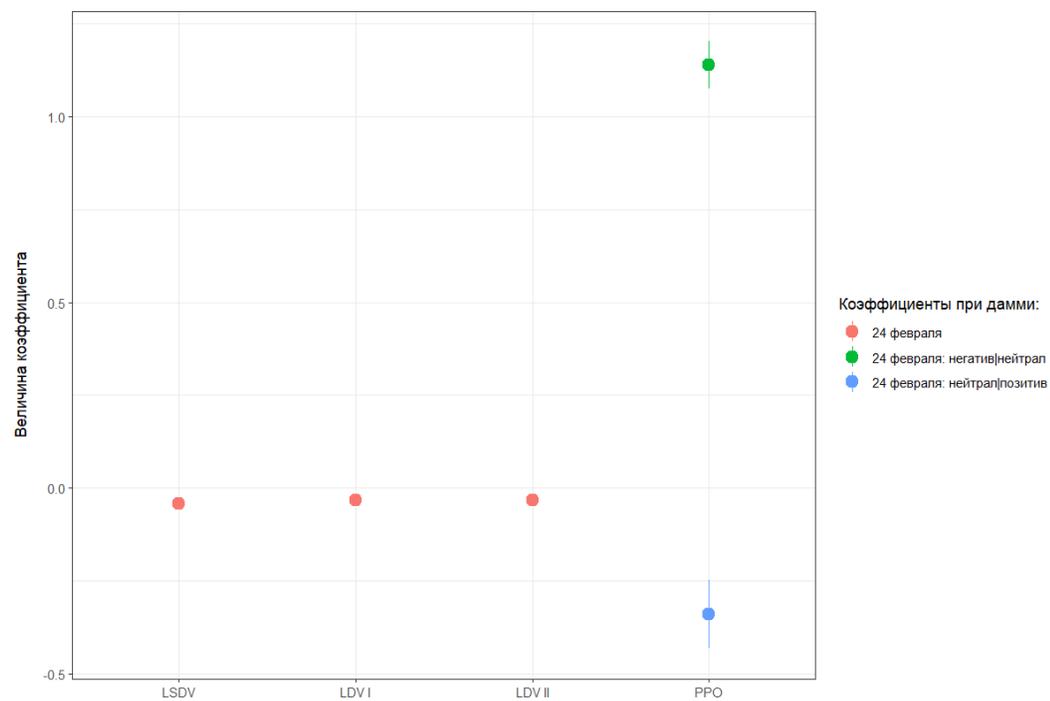
$$Tone\_of\_Publication_{it} = \beta_0 + \beta_1 Time_{it} + \beta_2 After\_24F_{it} + \beta_3 Time_{it} \times After\_24F_{it} + \beta_4 After\_4M_{it} + \beta_5 Time_{it} \times After\_4M_{it} + \varepsilon_{it}$$

- Общее уравнение для последующих моделей с FE:

$$Tone\_of\_Publication_{it} = \beta_0 + \beta_1 Platform_{it} + \beta_2 Author\_Subscribers_{it} + \beta_3 Author\_Sex_{it} + \beta_4 Author\_Age_{it} + \beta_5 Author\_Age^2_{it} + \beta_6 Author\_Education_{it} + \beta_7 Number\_of\_Duplicates_{it} + \beta_8 Publication\_Status_{it} + \beta_9 After\_24F_{it} + \beta_{10} After\_4M_{it} + FE_{region} + \varepsilon_{it}$$

- В процессе работы меняется как его правая часть (там появляется лагированный отклик  $\rightarrow$  LDV), так и левая (вместо *Tone\_of\_Publication* используется категориальная оценка тональности от «Медиалогии»  $\rightarrow$  PPO)
- LDV ограничивается двумя лагами, PPO позволяет варьировать только эффекту ключевых предикторов – *After24M* и *After4M*

# Результаты



# Результаты [см. python-код]

# Обсуждение

## ■ Выводы:

- В абсолютном выражении – маленькие, с завышенной из-за большого числа наблюдений значимостью, но все же отрицательные оценки эффекта
- Очень далеко от ожидаемого эффекта сплочения
- Скорее всего, в сети действительно имел место нивелирующий процесс: за быстротечным «сплочением» последовала масса негативных комментариев, перекрывших потенциал дальнейшего роста

## ■ Ограничения:

- Патриотический дискурс воспринимается буквально как разговоры о патриотизме
- Узкая интерпретация, основанная преимущественно на количественных показателях
- Возможный недоучет особенностей Интернет-дискуссии (эмоции, ирония, эхо-камеры...)
- Нехватка данных из Telegram (очень специфическая среда)

Спасибо за внимание!

Иван Анкудинов • ВШЭ • [iaankudinov@edu.hse.ru](mailto:iaankudinov@edu.hse.ru)

# Приложение

- *Tone\_of\_Publication* – «чистая тональность» публикации  
 $P_{Dostoevsky}[i\text{-я публикация} = \text{«позитивная»}] - P_{Dostoevsky}[i\text{-я публикация} = \text{«негативная»}]$
- *Time* – временной период (день)
- *After\_24F* – дамми-переменная, принимающая значение «1» в каждый период после 24 февраля
- *After\_4M* – дамми-переменная, принимающая значение «1» в каждый период после 4 марта
- *Platform* – площадка, на которой было размещено сообщение (Facebook, Instagram, Twitter, «ВКонтакте», «Одноклассники»)\*

\* первые две сети принадлежат Meta, признанной российским судом экстремистской организацией

# Приложение

- *Author\_Subscribers* – аудитория автора сообщения (число подписчиков)
- *Author\_Sex* – пол автора (мужчина, женщина)
- *Author\_Education* – уровень образования автора (неоконченное, среднее, высшее)
- *Number\_of\_Duplicates* – число дубликатов сообщения (постов, полностью идентичных данному)
- *Publication\_Status* – статус сообщения на момент выгрузки (не менялось, отредактировано, удалено, восстановлено)
- *FE* – фиксированные эффекты (фиктивные переменные для k-1 регионов)