

# Репликационный проект

## 1 Введение

### 1.1 Исследование

Для репликации были использованы данные исследования "The Relationship Between Social Media Use and Beliefs in Conspiracy Theories and Misinformation"(2021)<sup>1</sup>, опубликованного в журнале Political Behavior.

Исследование направлено на изучение влияния использования социальных сетей на веру в теории заговора (*conspiracy theories and desinformation*).

По результатам регрессионного анализа данных об использовании социальных сетей (на индивидуальном уровне, источник - Qualtrics), авторы приходят к выводу о том, что поддержка теорий заговора обусловлена не только получением новостей из социальных медиа, но и т. н. «конспирологическим мышлением» (*conspiracy thinking - the predisposition to interpret salient events as products of conspiracies*).

Основной вывод исследования заключается в том, что использование социальных сетей само по себе не способно оказать значительного влияния на веру в теории заговора и дезинформацию. Скорее, люди должны обладать «системой верований, восприимчивой к конспирологической информации». Чем больше вероятность увидеть заговоры во всевозможных культурных и политических событиях, тем сильнее связь между использованием социальных сетей и верой в теории заговора.

Гипотезы исследования:

- **H1:** Индивиды, которые используют социальные сети в качестве основного источника новостей и проводят в них больше времени, при прочих равных условиях будут больше верить в теории заговора (и дезинформацию), по сравнению с индивидами, которые получают новости из других источников и проводят меньше времени в социальных сетях.

---

<sup>1</sup>Enders, A.M., Uscinski, J.E., Seelig, M.I. et al. The Relationship Between Social Media Use and Beliefs in Conspiracy Theories and Misinformation. Polit Behav (2021) doi: [10.1007/s11109-021-09734-6](https://doi.org/10.1007/s11109-021-09734-6)

- **H2:** Взаимосвязь между использованием социальных сетей и верой в теории заговора слабее для тех индивидов, кто демонстрирует низкий уровень конспирологического мышления, по сравнению с теми, кто демонстрирует более высокий уровень.

## 1.2 Дизайн

В рамках репликационного проекта оценены ключевые модели статьи, позволяющие оценить эффект 1) частоты использования социальных сетей и 2) наличие модулирующего эффекта *conspiracy thinking*.

В рамках репликации частотной модели из статьи оценена OLS-модель с переменной взаимодействия: полученные оценки подтверждают гипотезы авторов исследования о взаимосвязи использования социальных сетей и верой в теории заговора. По итогам анализа данных также подтверждается гипотеза о том, что эта взаимосвязь зависит от «уровня конспирологического мышления» респондентов.

В рамках байесовского анализа оценена спецификация модели с переменными взаимодействия: оценки коэффициентов аналогичны оценкам частотной модели как по направлению взаимосвязи так и по величине эффекта.

Следующие разделы работы содержат описание данных для анализа, спецификации ключевых моделей и сравнение полученных результатов, а также основные выводы работы.

## 2 Описание данных

Данные для репликации опубликованы на портале [Political Behavior Dataverse](#) и доступны по [ссылке](#). Код для обработки данных - в формате .do.

Данные содержат 2023 наблюдения. Описательные статистики представлены в Таблице 1 в конце документа.

Зависимая переменная - количество убеждений - сконструирована на основе 15 вопросов, представленных респондентам. Среднее значение составляет 5 убеждений, а медиана — 4.

Кроме того, в исследовании использованы 3 ключевые независимые переменные:

- Социальные сети как основной источник новостей
- Частота использования социальных сетей (агрегированная переменная) Facebook, Twitter, Instagram, Reddit, Youtube (5-балльная шкала)
- «Конспирологическое мышление» (5-балльная шкала, Uscinski and Parent (2014))

В модели включен ряд контрольных переменных, в т. ч. партийная принадлежность, идеологическая самоидентификация, интерес к политике, уровень образования, возраст, доход, пол, расовая и этническая принадлежность.

### 3 Описание модели

#### 3.1 Частотный анализ

В реплицируемой статье отражены оценки для нескольких спецификаций моделей, каждая из которых была оценена в рамках проекта. Для оценки влияния использования социальных сетей и наличия «конспирологического мышления» на количество *conspiracy beliefs* использована следующая спецификация (Table 1, Model 3. Использована OLS-модель с переменными взаимодействия):

$$\begin{aligned} NumberBeliefs_i = & \beta_0 + \beta_1 SocialMediaNews_i + \beta_2 SocialMediaFrequency_i + \\ & + \beta_3 ConspiracyThinking_i + \beta_4 SMFrequency * ConspiracyThinking_i + \\ & + \beta_5 SMNews * ConspiracyThinking_i + \beta_6 X_i + \epsilon_i, \end{aligned}$$

где *NumberBeliefs<sub>i</sub>* - количество теорий, с которыми согласен индивид *i*, *SocialMediaNews<sub>i</sub>* - использование социальных медиа для чтения новостей, *SocialMediaFrequency<sub>i</sub>* - частота использования социальных медиа, *ConspiracyThinking<sub>i</sub>* - «конспирологическое мышление» (психологическая предрасположенность интерпретировать события как результат теорий заговора; операционализировано как индекс, сконструированный на основе четырех вопросов опросника). *X<sub>i</sub>* - вектор контрольных переменных (партийная принадлежность, идеологическая самоидентификация, интерес к политике, уровень образования, возраст, доход семьи,

пол, расовая и этническая принадлежность).

### 3.2 Байесовский анализ

В рамках байесовского анализа была оценена линейная модель с двумя переменными взаимодействия (спецификация отражена в разделе выше).

Так как априорная информация о распределениях неизвестна, в моделях использованы неинформативные («объективные») априорные распределения для всех параметров - подобный выбор априорных значений делает все правдоподобные значения оцениваемых параметров приблизительно равновероятными. Выбранные распределения обусловлены как спецификациями оригинальных моделей в исследовании (OLS), так и множествами значений переменных, включаемых в анализ.

Error term  $\sim Uniform(0, 1)$

Coefficients:  $\beta \sim Normal(0, 0.01)$

По результатам оценки результатов моделирования при разных значениях burn-in ( $n$  начальных значений, не учитывающихся при оценке параметров модели) и thinning (определяет порядок удалений реализаций в целях избежания проблемы автокорреляции) были определены следующие значения:

Burn-in: 50000. Thinning: 10.

Величина выборки, полученной по итогам реализации модели и семплинга, составила 200000 наблюдений. Несмотря на то, что подобные значения значительно увеличили время, необходимое для оценки параметров регрессионной модели, такой подход позволил получить достаточно хорошие результаты (диагностики модели представлены в Таблице A2 и на графиках в Приложении).

## 4 Результаты

Таблица 1 содержит результаты частотного анализа. Коэффициенты OLS-моделей аналогичны значениям, представленным авторами реплицируемой статьи. Таблица 2 содержит результаты байесовского анализа. Таблица 3 отражает оценки обеих моделей. Очевидно, что

результаты, полученные с использованием обоих способов, различаются минимально - и знаки при оценках, и величина параметров моделей практически для всех переменных одинаковы. Кроме того, на основании полученных 95% HPD-интервалов можно сделать вывод о значимости 14 из 15 параметров (кроме Age: среднее -0.0002, интервал (-0.001, 0.0002)), так как рассчитанные для них HPD-интервалы не включают 0.

Для модели байесовской регрессии также представлены ключевые диагностики (в Приложении). Таблица A2 отражает значения диагностики Гевеке для каждой переменной: несмотря на то, что среди множества значений присутствуют достаточно высокие (1.64 для  $\beta_8$ ), все значения не превышают по модулю 1.96.

Кроме того, графики в Приложении, позволяют оценить сходимость модели визуально. Очевидно, что определенное на начальных этапах количество итераций позволило достичь стационарного распределения для параметров модели.

Таблица 1: Результаты репликации моделей статьи

	<i>Dependent variable:</i>		
	Number of Conspiracy Beliefs		
	(1)	(2)	(3)
Social Media Frequency		0.878*** (0.086)	−0.753*** (0.238)
Social Media for News	0.596*** (0.169)	0.349** (0.167)	0.503 (0.591)
Conspiracy Thinking	3.028*** (0.069)	2.909*** (0.068)	1.493*** (0.203)
Partisanship	0.122** (0.057)	0.146*** (0.056)	0.137** (0.056)
Ideology	−0.011 (0.045)	0.021 (0.044)	0.028 (0.043)
Interest in Politics	0.384*** (0.070)	0.269*** (0.069)	0.260*** (0.069)
Education	0.029 (0.049)	−0.058 (0.048)	−0.064 (0.048)
Age	−0.017*** (0.005)	0.003 (0.005)	−0.0002 (0.005)
Income	0.136*** (0.042)	0.081* (0.042)	0.052 (0.041)
Female	−0.525*** (0.137)	−0.372*** (0.134)	−0.327** (0.132)
Black	0.141 (0.197)	−0.007 (0.192)	−0.021 (0.190)
Hispanic	0.157 (0.183)	0.111 (0.179)	0.151 (0.177)
SM Freq * Conspiracy			0.492*** (0.067)
SM for News * Conspiracy			−0.061 (0.168)
Constant	−5.846*** (0.440)	−8.122*** (0.483)	−3.292*** (0.811)
Observations	2,022	2,022	2,022
R <sup>2</sup>	0.534	0.557	0.569

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Таблица 2: Результаты байесовской регрессии. 95% HPD-интервалы.

	Mean	Lower Bound	Upper Bound
$\beta_1$	-3.293	-3.350	-3.236
$\beta_2$	0.504	0.463	0.545
$\beta_3$	-0.753	-0.770	-0.737
$\beta_4$	1.493	1.479	1.507
$\beta_5$	0.137	0.133	0.141
$\beta_6$	0.028	0.025	0.031
$\beta_7$	0.260	0.255	0.265
$\beta_8$	-0.064	-0.067	-0.061
$\beta_9$	-0.0002	-0.001	0.0002
$\beta_{10}$	0.052	0.049	0.054
$\beta_{11}$	-0.327	-0.336	-0.318
$\beta_{12}$	-0.021	-0.034	-0.007
$\beta_{13}$	0.151	0.138	0.163
$\beta_{14}$	0.492	0.488	0.497
$\beta_{15}$	-0.061	-0.072	-0.049
<i>Sigma</i>	0.100	0.100	0.100

Таблица 3: Сравнение моделей

Variable	Frequentist	Bayesian
intercept	-3.292 (0.811)	-3.293 (-3.350, -3.236)
socialmedia	-0.753 (0.238)	-0.753 (-0.770, -0.737)
conthink	1.493 (0.203)	1.493 (1.479, 1.507)
socmedia	0.503 (0.591)	0.504 (0.463, 0.545)
pid	0.137 (0.056)	0.137 (0.133, 0.141)
ideo	0.028 (0.043)	0.028 (0.025, 0.031)
interest	0.260 (0.069)	0.260 (0.255, 0.265)
edu	-0.064 (0.048)	-0.064 (-0.067, -0.061)
age	-0.0002 (0.005)	-0.0002 (-0.001, 0.0002)
income	0.052 (0.041)	0.052 (0.049, 0.054)
female	-0.327 (0.132)	-0.327 (-0.336, -0.318)
black	-0.021 (0.190)	-0.021 (-0.034, -0.007)
hispanic	0.151 (0.177)	0.151 (0.138, 0.163)
socialmedia*conthink	0.492 (0.067)	0.492 (0.488, 0.497)
conthink*socmedia	-0.061 (0.168)	-0.061 (-0.072, -0.049)



## 5 Заключение

По итогам реализации проекта по репликации были оценены вероятностные модели из Таблицы 2 рассматриваемой статьи и осуществлено оценивание модели байесовской регрессии. По итогам сравнения моделей можно заключить, что полученные с использованием обоих способов, различаются минимально - и знаки при оценках, и величина параметров моделей практически для всех переменных одинаковы.

Два метода, использованных в настоящей работе прежде всего направлены на поиск ответов на разные вопросы. В рамках классических/частотных методов статистики вероятность понимается как объективное свойство изучаемого феномена («объективная неопределенность»). Так, вероятности в рамках данной парадигмы напрямую связаны с частотой рассматриваемого события «в долгосрочной перспективе». В свою очередь, байесовский подход предполагает иную интерпретацию, согласно которой вероятность необъективна, а случайность - это «мера нашего незнания». Таким образом, в то время как классические методы статистики полностью основаны на данных, байесовский подход позволяет инкорпорировать априорную информацию непосредственно в анализ.

В связи с этим, на вопрос об убедительности результатов представленных моделей нельзя ответить однозначно в контексте сравнения результатов частотного и байесовского анализов - похожие результаты, полученные по итогам оценивания двух моделей, не отменяют того факта, что два подхода по своему определению отвечают на разные вопросы. В данном случае «неинформативные» априорные распределения позволяют сфокусироваться на точечных оценках параметров и минимизировать влияние субъективной составляющей, и иные значения priors (определенные на основе литературы или личных наблюдений) бы повлекли за собой другие результаты.

Так или иначе, если рассматривать применение данных методов в контексте тестируемых авторами статьи гипотез, частотный подход представляется более уместным.

## 6 Приложение

Таблица A1: Описательные статистики (ключевые переменные)

Statistic	N	Mean	St. Dev.	Min	Max
female	2,023	0.513	0.500	0	1
edu	2,023	3.358	1.504	1	6
white	2,023	0.646	0.478	0	1
black	2,023	0.151	0.358	0	1
hispanic	2,023	0.178	0.383	0	1
age	2,023	41.544	16.083	17	90
income	2,023	3.225	1.746	1	7
pid	2,023	2.784	1.416	1	5
ideo	2,022	3.885	1.782	1	7
interest	2,023	3.111	0.993	0	4
socialmedia	2,023	2.832	0.923	1.000	5.000
newspaper	2,023	0.038	0.191	0	1
socmedia	2,023	0.211	0.408	0	1
conthink	2,023	3.184	0.973	1.000	5.000
cpurpose	2,023	2.758	1.379	1	5
onepercent	2,021	3.521	1.154	1	5
deepstate	2,021	3.217	1.223	1	5
falseflag	2,023	2.140	1.271	1	5
rothschilds	2,023	2.952	1.176	1	5
jfk	2,023	3.256	1.226	1	5
replaceworkers	2,023	2.765	1.258	1	5
aliens	2,023	2.900	1.323	1	5
aids	2,022	2.446	1.300	1	5
vaccines	2,023	2.674	1.333	1	5
jewskilled	2,023	2.090	1.248	1	5
epstein	2,023	3.559	1.121	1	5
smallgroup	2,022	2.987	1.274	1	5
number	2,023	5.040	4.257	0	15

Таблица A2: Geweke Diagnostics

Variable	Value
$\beta_1$	-0.610004
$\beta_2$	1.113342
$\beta_3$	0.004268
$\beta_4$	-0.310996
$\beta_5$	0.779849
$\beta_6$	-0.572602
$\beta_7$	-0.172079
$\beta_8$	0.873788
$\beta_9$	1.639658
$\beta_{10}$	-0.393914
$\beta_{11}$	0.353768
$\beta_{12}$	0.547575
$\beta_{13}$	0.857866
$\beta_{14}$	0.335251
$\beta_{15}$	-0.983390
<i>Sigma</i>	0.796193



