



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 2 за 2019 г.



ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

Подлазов А.В.

Формальные методы
выявления масштабных
электоральных
фальсификаций на
материале федеральных
выборов 1999-2018 гг.

Рекомендуемая форма библиографической ссылки: Подлазов А.В. Формальные методы выявления масштабных электоральных фальсификаций на материале федеральных выборов 1999-2018 гг. // Препринты ИПМ им. М.В.Келдыша. 2019. № 2. 28 с. doi:[10.20948/prepr-2019-2](https://doi.org/10.20948/prepr-2019-2)
URL: <http://library.keldysh.ru/preprint.asp?id=2019-2>

**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.Келдыша
Российской академии наук**

А.В. Подлазов

**Формальные методы выявления
масштабных электоральных
фальсификаций на материале
федеральных выборов 1999-2018 гг.**

Москва — 2019

А.В. Подлазов

Формальные методы выявления масштабных электоральных фальсификаций на материале федеральных выборов 1999-2018 гг.

Предложены 4 различных теста для обнаружения масштабных электоральных фальсификаций на уровне субъектов федерации. Эти тесты основываются только на официальных данных о результатах выборов и не используют каких-либо предположений о поведении избирателей или эталонных результатов.

Ключевые слова: электоральная статистика, результаты выборов, фальсификация результатов, психологически привлекательные числа, жадное голосование, сгустки результатов, недействительные бюллетени

A.V. Podlazov

Formal methods for detection of large-scale electoral fraud
on the material of the federal elections 1999-2018

I offer 4 different tests for the detection of large-scale electoral fraud at the level of subjects of the federation. These tests bases only on official data on election results and do not use any assumptions about voter behavior or reference results.

Keywords: electoral statistics, election results, falsification of results, psychologically attractive numbers, greedy voting, clots of results, invalid ballots

Работа выполнена при поддержке РФФИ (проект 18-01-00619-а).

Содержание

Введение	3
Круглые числа.....	4
Идея метода.....	4
Анализ по субъектам федерации	5
Динамика фальсификаций.....	7
Жадное голосование.....	9
Простота хуже воровства	9
Пороги жадного голосования.....	11
Целевые установки.....	15
Равновероятность последних цифр	15
Сгустки явки и результатов.....	17
Недействительные бюллетени	22
Выводы	25

ВВЕДЕНИЕ

Выборы органов власти являются крупнейшим из возможных соцопросов, наиболее полно представляющим взгляды граждан. К сожалению, в России результаты выборов систематически фальсифицируются, что не позволяет опираться на них в социологических исследованиях без предварительной реконструкции. А для ее успешного осуществления необходимо понимание как того, как в отечественных условиях устроены подлинные результаты, так и того, в каких регионах нашей страны результаты фальсифицированы столь сильно, что исчезает основа для реконструкции. Попыткой ответить на первый вопрос стала работа [1], рассматривающая практически достоверные результаты нескольких последних выборов в Москве. А данная работа посвящена второму вопросу.

Проверка на наличие фальсификаций осуществляется на уровне субъектов федерации (в административных границах, существовавших на дату голосования) по причине существенной разницы между ними как в масштабах фальсификаций, так и в предпочитаемых способах их осуществления [2]. Подчеркнем, что здесь речь идет только о массовых фальсификациях, процедура выявления которых не должна предполагать наличия участков (лет), где (когда) результаты подсчитаны честно, на чем основываются все методы реконструкции [2,3,4,5,6,7,8,9,10,11,12,13,14]. Кроме того, процедура не должна опираться на какие-либо предположения об электоральном поведении вроде независимости результата кандидатов/партий от явки или гауссовости распределения их поддержки. Применимость в случае России этих предположений сложно обосновать, поскольку, как показывает иностранный опыт, в общем случае они могут и не выполняться [11,13,15,16]. Наконец, мы здесь отказываемся от методов анализа, привлекающих такие дополнительные нерегулярные данные, как сообщения наблюдателей [17] или данные «электронных урн» [7], поскольку покрытие ими избирательных участков нерепрезентативно.

Для анализа используются официальные (публикуемые ЦИК РФ) результаты 10 выборов федерального уровня, состоявшихся в России за последние 20 лет, – выборов в Государственную Думу в 1999, 2003, 2007, 2011 и 2016 гг. и выборов Президента в 2000, 2004, 2008, 2012 и 2018 гг. К сожалению, существующие базы с детализацией до избирательных участков неполны. Так, по 1999 и 2000 гг. доступные данные охватывают лишь 41% и 97% списочного состава избирателей соответственно. Поэтому к результатам анализа за эти годы следует относиться с определенной осторожностью. Кроме того, в данных за 2003 г. зафиксированы единичные случаи невыполнения контрольных соотношений или превышения числом проголосовавших избирателей их списочной численности. Также для этого года в нескольких субъектах возникли расхождения между сводными результатами вышестоящей избирательной комиссии и суммой результатов по подчиненным ей нижестоящим, интегрально достигающие нескольких тысяч человек, что в масштабах страны представляется не-

принципиальным. Начиная с 2004 г. пропусков и разночтений в данных не выявлено.

Далее представлен набор формальных методов выявления массовых фальсификаций и оценки их интегрального размера. По степени обоснованности эти методы лежат в диапазоне от вполне строгих до полуэмпирических. В силу различий в способах фальсификации, обнаружить их все каким-либо одним тестом вряд ли возможно. Однако сочетание тестов различной природы в целом успешно справляется с этой задачей, причем в подавляющем большинстве случаев фальсифицированные результаты проваливают сразу несколько тестов, что косвенно подтверждает надежность каждого из них.

Предваряя описание тестов, заметим, что при поиске собственно фальсификаций мы ориентируемся не на результаты выборов как таковые, а лишь на аномальное поведение тех или иных электоральных характеристик. А вот в качестве интегрального размера массовых фальсификаций выступает именно число голосов, вброшенных за кандидата/партию власти и/или переброшенных ему/ей от других кандидатов/партий.

КРУГЛЫЕ ЧИСЛА

Идея метода

Самой тяжелой формой фальсификации является выдумывание результатов, когда числа в избирательных протоколах не находятся в какой-либо связи с содержимым урн. При столь грубом подходе в «результатах» преобладают *психологически привлекательные числа* [18]. Для целочисленных электоральных характеристик таковыми являются круглые числа, а для процентных – значения без десятых долей (т.е. круглым оказывается количество промилле).

У целочисленной случайной величины, разброс которой измеряется многими десятками и даже сотнями единиц, последняя цифра должна практически равновероятно принимать все возможные значения. Если же частота цифры «0» в конце числа превышает 1/10, можно предполагать наличие выдуманных результатов. Для проверки этого предположения необходимо проверить статистическую гипотезу о естественном возникновении превышения. И если ее *уровень значимости* – вероятность отвергнуть данную гипотезу, когда она верна, – окажется мал, то наличие фальсификаций следует считать подтвержденным.

Регистрация на избирательном участке некоторой электоральной характеристики представляет собой испытание Бернулли с вероятностями *успеха* (круглое число) $p = 1/10$ и *неудачи* (некруглое число) $q = 1 - p = 9/10$. Количество успехов описывается биномиальным распределением, для которого вероятность зарегистрировать не менее k успехов в n испытаниях $P_{k,n} = \sum_{i=k}^n C_n^i p^i q^{n-i} = I_p(k, n - k + 1)$, где $I_p(x, y) = B_p(x, y) / B(x, y)$ – регуляризованная неполная бета-функция. Впрочем, непосредственное суммирование

здесь предпочтительнее, поскольку оно и проще, и быстрее алгоритмов вычисления неполной бета-функции, сохраняющих работоспособность при больших аргументах. Величина $\alpha = P_{k,n}$ определяет значимость проверяемой гипотезы, т.е. вероятность того, что, хотя частота k/n и может казаться аномально большой, ее подъем над уровнем p есть лишь результат вероятных флуктуаций.

Поскольку в случае масштабных фальсификаций уровень значимости α чрезвычайно мал, далее для удобства вместо него всюду приводится его *десятичный показатель* $pa = -\lg \alpha$, увеличение которого на единицу соответствует уменьшению на порядок вероятности отсутствия фальсификаций.

Учитывая, что Российская Федерация состояла в разные годы из 84÷90 субъектов, на конкретных выборах для любой электоральной характеристики в среднем должен встретиться 1 субъект с $pa \approx 1,92 \dots 1,95$. Таким образом, значения pa начиная примерно с 2 следует считать *подозрительными*, с 3 – *исключительными*, а с 4 – *невероятными*. Если же анализировать весь массив из 10 федеральных выборов (с учетом неполноты данных за 1999-2000 гг. и объединением зарубежных участков в отдельный виртуальный субъект всего получается 865 отдельных результатов выборов), то указанные пороги следует увеличить на 1, т.е. до 3, 4 и 5 соответственно.

Анализ по субъектам федерации

Чаще всего объектом фальсификации становятся следующие электоральные характеристики:

- количество избирателей, *принявших участие в выборах* (получивших избирательные бюллетени);
- *явка избирателей* – доля зарегистрированных избирателей, принявших участие в выборах;
- *результат власти* – доля избирателей, поддержавших партию/кандидата власти, измеряемая от числа принявших участие в голосовании (опустивших бюллетень в урну).

Две последние величины, являющиеся дробями, при анализе превращаются в целые числа путем умножения на 1 000 и округления до ближайшего целого. Чтобы гарантировать значительную ширину разброса рассматриваемых характеристик, учитываются только те участки, на которых участие в выборах приняли не менее 100 человек. Это, кроме того, позволяет сделать пренебрежимо малой вероятность концентрации явки и результата на дробях с малыми знаменателями. Наконец, для явки исключаются из рассмотрения участки, где участие в выборах приняли все без исключения зарегистрированные избиратели, т.к. это может быть следствием отсутствия заранее составленных списков избирателей, что делает явку фиктивной величиной.

В табл. 1, 2 и 3 приведены данные по всем субъектам, где указанные характеристики хотя бы раз оказывались круглыми на подозрительно большом числе участков ($pa \geq 3$). При этом исключительные значения значимости ($pa \geq 4$) выделены жирным шрифтом, а невероятные ($pa \geq 5$) – еще и красным

цветом. Отсортированы субъекты федерации по убыванию максимального из зафиксированных в них значений ра.

Таблица. 1. Показатель ра для участия избирателей в выборах

Субъект федерации	1999	2000	2003	2004	2007	2008	2011	2012	2016	2018
Дагестан	1,4	17,8	34,8	26,3	16,7	31,8	103,0	27,6	20,5	8,1
Татарстан	0,8	0,6	0,2	1,8	8,6	23,5	1,2	2,0	2,3	4,3
Чечня	—	—	13,3	15,2	1,4	0,2	0,3	0,0	0,6	0,4
Башкортостан	0,9	3,9	0,3	2,9	11,4	7,9	3,2	4,2	2,0	1,9
Северная Осетия	0,9	1,1	1,5	4,4	0,5	3,6	1,7	2,9	8,5	7,0
Краснодарский кр.	1,1	1,8	1,6	1,0	1,8	6,9	0,7	1,1	3,4	4,0
Кемеровская обл.	0,8	0,9	0,1	2,4	4,1	7,0	0,4	1,6	5,7	2,9
Крым	—	—	—	—	—	—	—	—	0,4	5,2
Карачаево-Черкесия	1,1	1,4	0,7	3,9	4,0	5,0	0,6	2,4	1,1	3,8
Тюменская обл.	0,2	0,3	0,3	1,1	0,9	2,0	1,4	0,1	5,0	1,6
Хабаровский кр.	0,5	0,4	0,1	0,2	0,8	1,1	0,8	1,2	0,3	4,9
Кабардино-Балкария	1,0	2,3	4,5	1,5	0,8	2,8	2,1	2,5	2,5	0,5
Липецкая обл.	0,6	1,2	0,3	0,3	0,3	5,0	1,0	1,1	1,5	1,6
Самарская обл.	0,2	0,3	0,2	0,9	0,1	4,4	0,1	0,0	0,6	0,7
Пермская обл.	1,2	0,6	3,8	2,7	—	—	—	—	—	—
Ярославская обл.	0,8	0,4	1,9	0,2	3,7	1,5	0,8	0,1	0,7	0,2
Мордовия	0,5	0,4	0,2	0,6	3,2	2,0	1,6	0,9	0,8	0,2
Алтайский кр.	0,2	0,0	0,2	0,1	0,2	1,3	0,8	0,5	2,2	3,2
Ставропольский кр.	0,3	0,1	0,2	1,0	0,4	0,2	0,8	2,8	0,7	3,1
Пензенская обл.	0,3	0,6	0,4	1,7	3,0	1,3	0,3	0,6	0,3	0,4

Таблица. 2. Показатель ра для общей явки

Субъект федерации	1999	2000	2003	2004	2007	2008	2011	2012	2016	2018
Татарстан	0,9	0,2	0,3	0,5	0,8	6,5	11,8	16,3	49,2	27,3
Дагестан	1,5	1,0	1,1	17,2	5,8	8,8	1,4	8,3	20,3	30,2
Краснодарский кр.	0,4	0,1	0,7	0,6	1,5	3,1	1,5	1,9	1,7	26,7
Башкортостан	0,9	0,6	0,0	17,7	8,9	18,1	11,9	11,7	23,7	18,5
Кабардино-Балкария	0,7	2,0	5,3	4,6	4,6	22,1	0,2	9,7	1,3	0,4
Кемеровская обл.	0,2	0,4	0,4	3,1	2,7	6,4	0,9	4,1	17,4	17,0
Карачаево-Черкесия	0,4	0,7	0,1	0,5	1,1	9,7	3,0	2,0	1,0	3,9
Саратовская обл.	0,4	0,3	0,9	0,4	1,2	1,6	0,8	1,1	4,1	7,8
Северная Осетия	0,1	0,5	0,7	4,4	3,6	7,8	0,1	0,4	4,6	2,5
Мордовия	0,3	0,0	1,1	1,8	3,8	2,0	2,2	1,8	6,9	1,3
Ямало-Ненецкий авт.окр.	0,1	0,0	0,0	0,6	0,3	2,0	0,2	6,9	2,1	0,5
Москва	0,2	0,5	0,3	2,1	0,5	5,2	6,0	0,5	1,5	0,3
Ставропольский кр.	0,1	0,2	0,2	0,2	1,1	0,2	0,1	0,1	0,5	5,3
Ростовская обл.	0,0	0,0	0,1	0,9	1,0	2,5	0,2	2,8	4,7	2,5
Ингушетия	0,3	0,0	0,5	0,4	1,6	0,9	0,2	4,2	0,5	0,3
Чечня	—	—	3,9	4,0	0,0	1,0	0,0	0,0	0,1	0,4
Московская обл.	0,5	0,4	0,4	0,8	3,1	3,6	0,9	0,7	0,7	2,5
Санкт-Петербург	0,4	0,3	0,4	0,7	0,5	3,3	0,1	0,7	0,5	0,4

Таблица. 3. Показатель $\rho\alpha$ для результата власти

Субъект федерации	1999	2000	2003	2004	2007	2008	2011	2012	2016	2018
Башкортостан	0,1	0,5	0,2	77,4	16,1	43,5	9,6	7,5	24,0	12,7
Татарстан	3,1	1,1	8,9	31,8	25,7	59,2	31,2	28,6	48,8	21,4
Дагестан	0,0	5,9	11,4	31,1	12,2	13,9	20,4	17,0	37,6	19,3
Кабардино-Балкария	0,7	2,5	3,2	4,7	7,4	26,7	4,9	1,9	1,9	0,2
Кемеровская обл.	0,2	0,5	0,1	1,3	3,4	16,1	3,2	7,0	3,2	6,7
Краснодарский кр.	0,3	0,2	0,2	0,4	0,0	12,7	3,3	3,4	1,1	5,0
Мордовия	0,2	0,2	2,4	12,2	12,4	10,1	12,2	1,4	10,9	2,2
Москва	0,0	0,1	0,1	0,5	0,4	1,3	10,1	0,3	0,9	0,4
Карачаево-Черкесия	0,1	0,2	1,2	0,3	9,9	8,0	1,0	0,2	0,0	1,8
Ростовская обл.	0,5	0,9	0,2	2,6	4,8	9,3	1,6	0,6	3,8	0,8
Чечня	—	—	3,2	5,6	0,8	0,5	0,0	8,6	0,0	0,7
Саратовская обл.	0,1	2,3	0,5	6,7	0,1	4,1	2,4	8,4	1,4	2,2
Ингушетия	1,2	0,3	0,7	0,9	6,5	0,2	0,3	0,4	0,6	0,3
Северная Осетия	0,4	0,9	0,0	1,6	2,2	2,7	0,3	0,8	5,2	0,6
Ставропольский кр.	0,1	0,2	0,2	0,9	0,7	0,1	2,6	1,4	1,8	4,6
Тюменская обл.	0,3	0,0	1,6	2,7	2,6	4,2	3,0	1,3	3,1	4,5
Тыва	0,1	0,8	0,1	0,2	1,3	2,5	1,6	2,7	0,3	4,3
Кировская обл.	0,9	0,0	0,7	1,8	0,3	3,2	0,8	0,0	0,1	2,2

Подозрительные значения показателя для участия, общей явки и результата власти суммарно зафиксированы соответственно 45, 51 и 65 раз, в т.ч. исключительные – 33, 42 и 54 раза и невероятные – 21, 33 и 46 раз. Возрастание чисел в этих рядах коррелирует с предполагаемой значимостью для фальсификаторов соответствующих электоральных характеристик (наиболее важен результат власти, чуть меньше внимания уделяется величине явки, а количество избирателей уже второстепенно).

Всего в таблицы попали 29 субъектов федерации, из них 11 (Башкортостан, Дагестан, Кабардино-Балкария, Карачаево-Черкесия, Мордовия, Северная Осетия, Татарстан, Чечня, а также Краснодарский и Ставропольский кр. и Кемеровская обл.) – во все три таблицы, а ещё 5 (Ингушетия, Москва, а также Ростовская, Саратовская и Тюменская обл.) – сразу в две.

Хотя бы по разу значение $\rho\alpha \geq 5$ продемонстрировали 17 субъектов (Башкортостан, Дагестан, Ингушетия, Кабардино-Балкария, Карачаево-Черкесия, Крым, Мордовия, Северная Осетия, Татарстан, Чечня, Москва, а также Ямало-Ненецкий авт.окр., Ставропольский и Краснодарский кр., Кемеровская, Ростовская и Саратовская обл.), а $\rho\alpha \geq 4$ – 22 субъекта (к предыдущему списку добавляются Тыва, Хабаровский кр., Липецкая, Самарская и Тюменская обл.).

Динамика фальсификаций

Как можно видеть, фальсификации данного типа имеют тенденцию концентрироваться в одних и тех же субъектах. Вместе с тем то, что некоторый субъект прошел тест на частоту круглых чисел, еще не означает, что в нем результаты выборов не выдумывались. Например, возможна ситуация, когда в разных избира-

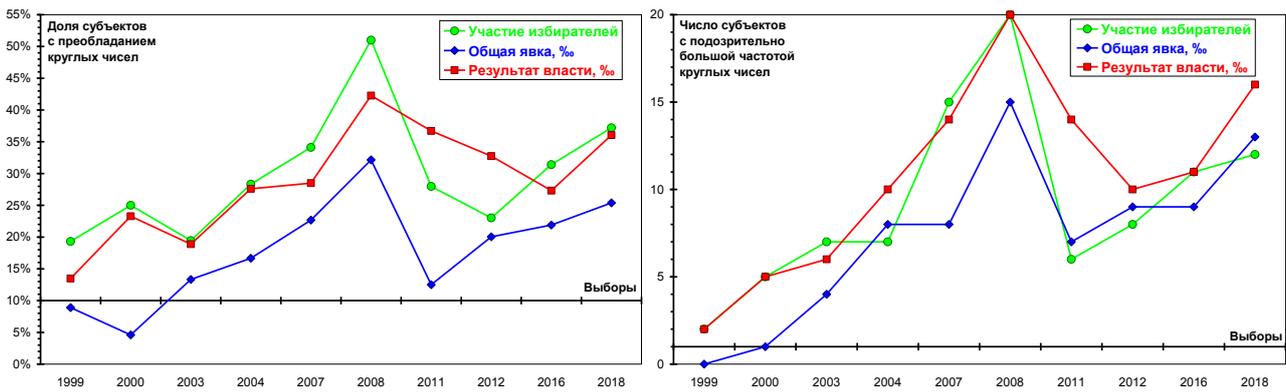


Рис. 1. Динамика интегрального выдумывания результатов

Слева – процент субъектов, где для рассматриваемой электоральной характеристики наиболее частой последней цифрой является «0». Справа – число подозрительных субъектов. Для первой меры естественный уровень – 10%, для второй – 1.

тельных комиссиях субъекта сосредотачиваются на выдумывании различных электоральных характеристик, что сдерживает рост частоты круглых результатов. По этой причине целесообразно обратиться к их интегральным мерам.

В качестве одной такой меры выбрано число субъектов, по участкам которых наиболее часто в конце рассматриваемых электоральных характеристик стоит «0» (если при этом другие цифры встречаются с той же частотой, то субъект учитывается с пропорционально меньшим весом), отнесенное к общему числу субъектов в рассматриваемом году N . Другой мерой круглости служит число подозрительных субъектов с $N \cdot \alpha \geq 1$ (напомним, что в среднем ежегодно

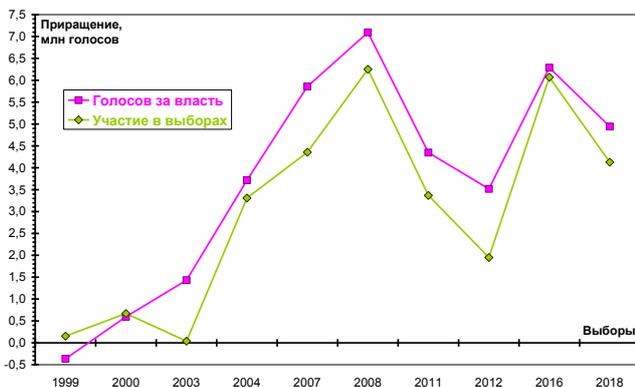


Рис. 2. Динамика выдуманных голосов

Обусловленное фальсификациями увеличение числа голосов, поданных за власть, всегда больше их вброса (увеличения количества избирателей, принявших участие в выборах), т.к. включает в себя и переброс голосов от других партий/кандидатов. Тем самым, инверсия графиков в 1999 г. показывает уровень погрешности данного метода оценки масштабов фальсификации, составляющий около полумиллиона голосов.

должен встречаться 1 субъект с такой значимостью гипотезы о естественном возникновении избытка круглых чисел). Изменение этих мер во времени, представленное на рис. 1, показывает, что по распространенности фальсификаций данного типа президентские выборы 2008 г. держат абсолютное лидерство, а президентские выборы 2018 г. никогда не опускаются ниже третьего места. Распространенное в обществе ощущение рекордных фальсификаций на парламентских выборах 2011 г. ошибочно. По-видимому, оно обусловлено тем, что тогда результат власти массово фальсифицировался даже в Москве, чего не наблюдалось в иные годы (см. табл. 3).

Зная вероятность того, что в субъекте имели место фальсификации

данного типа, можно оценить их интегральный эффект для страны в целом, если суммировать результаты по субъектам с весами, равными минимуму из вероятностей, рассчитанных для трех рассматриваемых электоральных характеристик. Это дает оценки общей явки и результата власти, какими бы они были при честном подсчете голосов. Распространение этих значений на всю страну, в свою очередь, позволяет оценить эффект от фальсификаций данного типа. Как можно видеть из рис. 2, по числу выдуманных голосов выборы 2008 г. сохраняют первенство, тогда как выборы 2018 г. съезжают на четвертое место, пропуская вперед предшествовавшие им парламентские выборы 2016 и 2007 гг.

ЖАДНОЕ ГОЛОСОВАНИЕ

Простота хуже воровства

Кроме выдумывания из головы числа избирателей, принявших участие в выборах, существует еще один механизм фальсификации, приводящий к преобладанию в данных об участии круглых чисел. Количество избирательных бюллетеней, полученных участковой избирательной комиссией, зачастую само оканчивается на «0» (скорее всего, бюллетени просто отсчитывают десятками). И если фальсификаторы, ничтоже сумняшеся, приравнивают количество пришедших избирателей количеству имеющихся бюллетеней, то это тоже может приводить к появлению круглых чисел.

Введем представление о *жадном голосовании*, понимая под ним ситуацию, когда количество избирателей, принявших участие в выборах, в точности совпадает с числом бюллетеней, полученных участковой избирательной комиссией, но при этом оказывается меньше числа избирателей, зарегистрированных в ее списках. Последнее ограничение необходимо, чтобы исключить «герметичные» участки, образованные в труднодоступных местах, где и в самом деле могут проголосовать все зарегистрированные избиратели, что требует получения такого же количества бюллетеней. Поэтому вполне возможно, что при полной явке и впрямь были использованы все бюллетени (это весьма типично для Приморского кр., Камчатской, Магаданской, Мурманской и Сахалинской обл. и в меньшей степени для некоторых других субъектов). Однако крайне подозрительно, если это происходит при неполной явке. В таком случае нельзя исключать ситуации грубого нарушения прав избирателей, которые не смогли проголосовать из-за нехватки бюллетеней, что, однако, никак не отражается и не может быть отражено в избирательных протоколах. А если протоколы недостоверны, то уже весьма вероятно, что их недостоверность возникла не в силу неудачного стечения обстоятельств, а в силу злого умысла.

В табл. 4 приведены данные о доле избирательных участков, где все бюллетени были использованы при неполной явке, для субъектов, где эта доля хотя бы однажды превысила 5%.

Таблица. 4. Доля участков с жадным голосованием

Субъект федерации	1999	2000	2003	2004	2007	2008	2011	2012	2016	2018
Карачаево-Черкесия	0%	0%	4%	4%	27%	18%	2%	2%	0%	6%
Чечня	—	—	24%	27%	0%	0%	0%	0%	0%	0%
Дагестан	3%	9%	6%	6%	14%	24%	11%	11%	7%	4%
Ингушетия	10%	0%	2%	2%	21%	0%	1%	1%	1%	1%
Агинский Бур. авт.окр.	0%	0%	0%	0%	18%	—	—	—	—	—
Усть-Ордынск. Бур. авт.окр.	0%	3%	1%	1%	18%	—	—	—	—	—
Татарстан	2%	2%	4%	5%	14%	16%	1%	1%	2%	3%
Мордовия	0%	3%	4%	6%	15%	9%	1%	1%	1%	4%
Кабардино-Балкария	0%	1%	4%	4%	13%	1%	0%	0%	1%	0%
Тыва	0%	1%	2%	2%	12%	7%	4%	4%	2%	7%
Башкортостан	1%	3%	5%	7%	9%	7%	0%	0%	0%	1%
Ямало-Ненецкий авт.окр.	2%	1%	1%	1%	4%	8%	2%	2%	0%	0%
Марий Эл	0%	0%	0%	0%	2%	1%	0%	0%	8%	7%
Северная Осетия	0%	1%	6%	7%	1%	2%	0%	0%	3%	3%
Чукотский авт.окр.	0%	0%	5%	6%	2%	0%	0%	0%	0%	0%
Коми	0%	0%	0%	0%	1%	2%	5%	5%	6%	6%
Липецкая обл.	0%	0%	0%	0%	1%	6%	0%	0%	1%	3%
Томская обл.	0%	0%	1%	1%	0%	1%	0%	0%	3%	6%
Нижегородская обл.	0%	0%	0%	0%	0%	1%	0%	0%	3%	5%
Владимирская обл.	0%	0%	0%	0%	0%	0%	0%	0%	3%	5%
Ивановская обл.	0%	0%	0%	0%	0%	0%	0%	0%	5%	5%

Таблица. 5. Доля избирателей на участках с жадным голосованием

Субъект федерации	1999	2000	2003	2004	2007	2008	2011	2012	2016	2018
Ингушетия	9%	0%	1%	1%	26%	0%	0%	0%	1%	1%
Карачаево-Черкесия	0%	0%	3%	3%	25%	13%	1%	1%	0%	6%
Агинский Бур. авт.окр.	0%	0%	0%	0%	24%	—	—	—	—	—
Чечня	—	—	19%	21%	0%	0%	0%	0%	0%	0%
Дагестан	1%	9%	4%	4%	13%	18%	7%	7%	7%	3%
Мордовия	0%	1%	3%	3%	13%	6%	1%	1%	1%	2%
Марий Эл	0%	0%	0%	0%	1%	0%	0%	0%	12%	11%
Томская обл.	0%	0%	0%	0%	0%	1%	0%	0%	4%	11%
Татарстан	1%	1%	1%	2%	7%	10%	0%	0%	1%	1%
Тыва	0%	0%	1%	1%	10%	5%	3%	3%	2%	3%
Усть-Ордынск. Бур. авт.окр.	0%	1%	0%	0%	10%	—	—	—	—	—
Коми	0%	0%	0%	0%	1%	0%	7%	7%	9%	8%
Кабардино-Балкария	0%	1%	3%	3%	9%	0%	0%	0%	0%	0%
Владимирская обл.	0%	0%	0%	0%	0%	0%	0%	0%	4%	8%
Ивановская обл.	0%	0%	0%	0%	0%	0%	0%	0%	7%	7%
Ямало-Ненецкий авт.окр.	1%	1%	1%	1%	3%	7%	2%	2%	0%	0%
Еврейская авт.обл.	0%	0%	0%	0%	0%	1%	0%	0%	3%	7%
Северная Осетия	0%	0%	6%	7%	0%	2%	0%	0%	4%	4%
Нижегородская обл.	0%	0%	0%	0%	0%	0%	0%	0%	3%	6%
Башкортостан	0%	1%	3%	3%	6%	6%	0%	0%	0%	0%

В табл. 5 аналогичные данные приведены для доли зарегистрированных на таких участках избирателей. Выделение жирным шрифтом и синим цветом выполнено по порогам в 10% и 15% соответственно.

Выбор указанных порогов отсечения и выделения является в значительной степени условным, поскольку в отличие от уровней значимости тут нет очевидных критериев того, что следует считать лишь подозрительным, а что – уже невероятным. Вместе с тем по относительному расположению субъектов в таблицах можно сделать определенные выводы о предпочитаемом способе фальсификаций. Так, например, в Карачаево-Черкесии, немного не дотянувшей до невероятных значений ρ для участия, и в Ингушетии, вообще не попавшей в табл. 1, скорее, ориентируются на количество полученных бюллетеней. С другой стороны, в Татарстане, Башкортостане и Северной Осетии, характеризующихся средневисоким уровнем жадного голосования, и тем более в Краснодарском кр. и Кемеровской обл., не попавших в табл. 4 и 5, участие избирателей именно выдумывают. Наконец, для Дагестана и Чечни, одинаково «успешных» во всех рейтингах, по-видимому, характерно комбинирование обоих способов фальсификации.

Пороги жадного голосования

Жадное голосование было зафиксировано хотя бы на 1 участке субъекта в 634 случаях из 865. И хотя его естественное возникновение нельзя полностью исключить, оно представляется экзотикой. В принципе, возможно, что на каком-то участке и в самом деле пришли ровно столько избирателей, сколько им было получено бюллетеней. Или, что более правдоподобно, предвидя угрозу нехватки бюллетеней, участковая избирательная комиссия уже по ходу голосования потребовала их дополнительное количество, которое, будучи небольшим числом, имеет больше шансов совпасть с числом лишних избирателей. Наконец, нельзя исключать существования двух разных численностей избирателей: официальной, попадающей в протоколы, и реальной, в соответствии с которой отсчитываются бюллетени в расчете на полную явку. Такая ситуация, тоже формально являющаяся нарушением, может, однако, не быть результатом злого умысла в случае избирательных участков, находящихся в труднодоступных местах.

Несмотря на то что жадное голосование на конкретном участке почти наверняка означает фальсификацию на нем результатов выборов, ее масштаб, как правило, сравнительно невелик. И ограничивает его как раз нехватка избирательных бюллетеней. Там, где их имеется в достатке, фальсификаторы, дойдя до желаемых результата власти и/или общей явки, могут и остановиться, тогда как дефицит бюллетеней вынуждает выгребать их до последнего. Однако то, что фальсификаторы имеют возможность и причины жадничать, в лучшем случае говорит об отсутствии эффективного контроля над процедурой голосования и подведением его итогов, а в худшем – о прямых установках на достижение высоких значений общей явки и/или результата власти. Иначе говоря, жадное голосование на одних участках субъекта федерации, скорее всего, свидетельствует о фальсификациях и на многих других участках, более изобильных бюл-

летениями. И чем выше распространенность жадного голосования, тем весомее это свидетельство.

Как уже было сказано, здесь нет прямых критериев того, «сколько орехов – куча». Тем не менее, с помощью инструментария ранг-размерного анализа [19] можно установить косвенные ориентиры для определения порогов жадности и их интерпретации.

Пусть *жадность* субъекта g – доля его участков, где на данных выборах при неполной явке были использованы все полученные бюллетени. Упорядочим сводные данные по всем субъектам и выборам по убыванию жадности и назовем *рангом* значения g его номер r в таком списке. Поскольку нумерация – процедура, выполняемая с точностью до произвольной аддитивной постоянной, присвоим самому жадному субъекту номер r_0 , величину которого определим далее.

Для субъектов, где фальсификации носят массовый характер, можно предполагать приближенное выполнение зависимости $g(r) \sim 1/r$. Эта формула в точности соответствует распределению величины, обратной случайной величине, равномерно распределенной на некотором интервале $(a; b)$, причем можно показать, что $r_0 = 1 + n \cdot a / (b - a)$, где n – объем анализируемой выборки. Для распределений, отличных от равномерного, гиперболическая зависимость обратных величин от их ранга сохраняется, но лишь в области сравнительно небольших r , а величина r_0 при этом становится подгоночным параметром.

Пусть субъект включает N избирательных участков, из которых на G участках голосование жадное. Тогда распределение субъектов по жадности $g = G/N$ определяется соотношением флуктуаций числителя и знаменателя. Для массовых фальсификаций числитель G ограничивается лишь организационными возможностями и аппетитами фальсификаторов, достигая десятков и даже сотен. При этом существенными становятся вариации знаменателя, что и приводит к гиперболической зависимости жадности от ранга. А для немассовых фальсификаций гиперболический вид зависимости нарушается, поскольку доминируют уже вариации числителя, который измеряется лишь единицами.

Анализ, подтверждающий эти рассуждения, представлен на рис. 3. Произведение $g \cdot r$ не зависит от r в области малых рангов, логарифмически убывает в области средних и обрушивается в области больших. Зависимость в области малых и средних рангов в полулогарифмическом масштабе аппроксимирована непрерывной кусочно-линейной функцией с константным первым участком. При обработке данные учитываются с весом $1/r$ (это обеспечивает одинаковую представительность всех частей оси абсцисс). Параметры аппроксимации – величина r_0 , а также точки перелома и завершения обработки данных – подбираются так, чтобы максимизировать произведение дисперсии ординаты на первом участке и коэффициента детерминации на втором.

Результат анализа позволяет установить на шкале жадности пороги $g_1 = 0,31\%$ и $g_2 = 1,53\%$, разделяющие фальсификации на *незначительные*, *умеренные* и *масштабные*. При $g < g_1$ последствия фальсификаций, даже если они имели место, не удается отличить от естественного стечения обстоятельств, приведше-

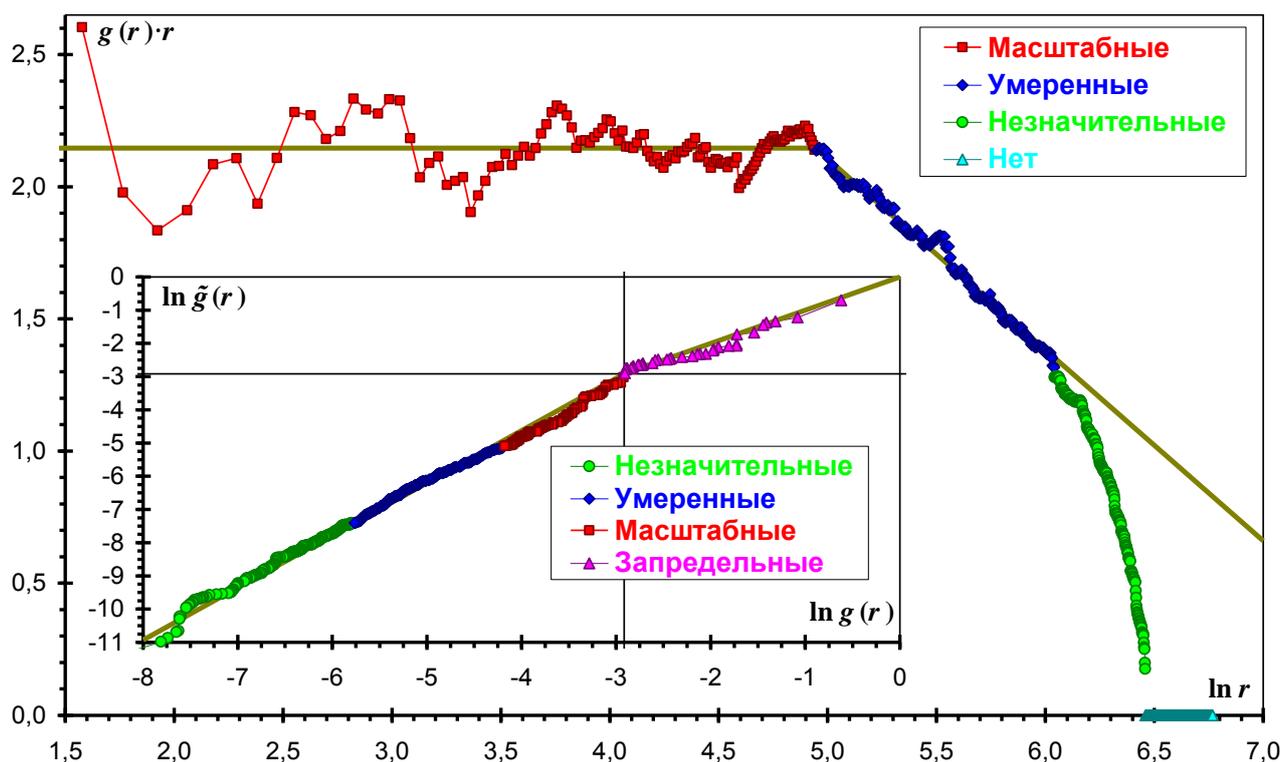


Рис. 3. Распределение субъектов федерации по жадности голосования

Наилучшее приближение достигается при $r_0 = 4,84$ и выделении 136 случаев масштабных фальсификаций и 280 – умеренных. При этом в 218 случаях фальсификации считаются незначительными, а еще в 231 – не зафиксированы.

На врезке – зависимость ранговых долей избирателей и участков, затронутых жадным голосованием. В 38 случаях самого жадного голосования, для которых точки ложатся на диагональ, фальсификации перекалифицированы из масштабных в запредельные. В остальных случаях ранговые жадности связаны формулой $\tilde{g}/g_0 = (g/g_0)^\gamma$, где $g_0 \approx -2,915\dots$ и $\gamma \approx 1,576\dots$, т.е. распределение доли избирателей оказывается более широким ($\gamma > 1$).

го к исчерпанию бюллетеней при неполной явке. А при $g_1 < g < g_2$, хотя отличия уже вполне заметны, доказать их неслучайную природу вряд ли возможно.

Проведение аналогичного анализа для доли избирателей на участках с жадным голосованием \tilde{g} представляется нецелесообразным, поскольку здесь возникает дополнительный фактор, связанный с численностью избирателей, из-за чего качество аппроксимации значительно ухудшается. Однако можно сопоставить распределения величин g и \tilde{g} друг с другом, как это сделано на врезке к рис. 3, где точки параметризованы их рангом (при этом абсцисса и ордината каждой точки могут относиться к разным субъектам и разным выборам).

Зависимость аппроксимируется в двойном логарифмическом масштабе непрерывной кусочно-линейной функцией, правая часть которой лежит на диагонали. Прохождение графика через начало координат обусловлено тем, что для гипотетического абсолютно жадного голосования, очевидно, $g = \tilde{g} = 1$. А единственный наклон правой части графика отражает независимость вероятности жадного голосования от численности избирателей. Обычно уровень фальсификаций

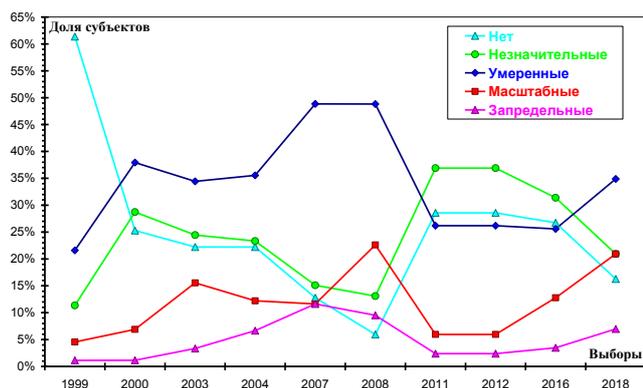


Рис. 4. Динамика распространности жадного голосования

Суммарная распространенность масштабных и запредельных фальсификаций нарастает в 1999-2008 и 2011-18 гг. В целом похожее поведение демонстрируют и умеренные фальсификации (что говорит о том, что они – всё-таки фальсификации, а не стечение обстоятельств). А вот доля субъектов, где жадных фальсификаций не было вовсе или они были незначительны, меняется прямо противоположным образом, достигая исторического минимума (19%) в 2008 г.

Доля субъектов без фальсификаций аномально высока в 1999 г., для которого нельзя исключить избирательной неполноты доступных данных.

выше на маленьких участках (где хуже контроль), чем на больших [2]. Поэтому утрата зависимости от размера участка, происходящая при $g, \tilde{g} > g_2 = 5,4\%$, означает даже не масштабный, а уже *запредельный* уровень фальсификаций. При меньшей жадности голосования распределение субъектов по доле избирателей на участках с жадным голосованием ожидаемо становится шире распределения по доле самих участков, а диапазон изменения величины \tilde{g} – больше чем g .

В свете сказанного критерии, использованные при составлении табл. 4 и 5, следует счесть чересчур мягкими (фактически отслеживались только запредельные и самые масштабные случаи фальсификаций). Более адекватно изменение распространенности фальсификаций данного типа отражает рис. 4. И хотя общий вывод о пиковой жадности выборов 2007-08 и 2018 гг. можно сделать и из таблиц, ее общий уровень они не отражают.

Последовательно исключая из рассмотрения субъекты с наибольшей долей участков, охваченных жадным голосованием, можно приблизиться к истинному положению дел, как это показано на рис. 5. Однако нужно помнить, что, уменьшая систематическую погрешность сокращением выборки, мы неизбежно увеличиваем погрешность стохастическую. Сложно понять, где здесь лежит золотая середина, при которой уровень фальсификаций уже достаточно

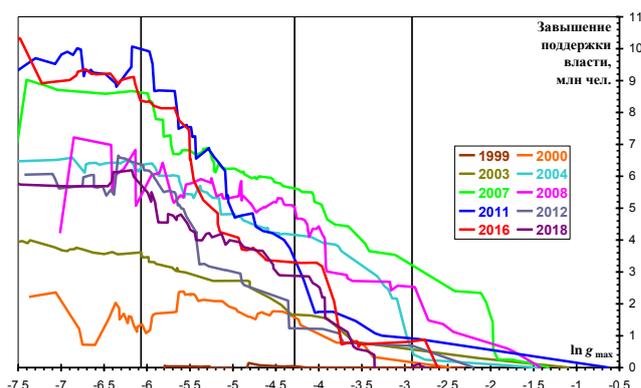


Рис. 5. Фальсификации поддержки власти, выявляемые жадностью

Избыточное число голосов, якобы поданных за кандидата/партию власти, рассчитывается распространением ее результата, полученного на субъектах с $g < g_{max}$, на всю страну.

Вертикальные линии соответствуют порогам g_1, g_2 и g_3 , разделяющим уровни жадности для незначительных, умеренных, масштабных и запредельных фальсификаций.

При движении справа налево количество ненастоящих голосов неуклонно возрастает. Начиная с 2000 г. стационара удастся достичь лишь в области незначительных флуктуаций. При этом он оказывается в целом выше для парламентских выборов по сравнению с президентскими.

низок, а представительность выборки еще достаточно высока. Если полагать, что оптимальный момент достигается лишь при прохождении порога g_1 , то количество фальшивых голосов, полученных властью, в целом оказывается значительно больше, чем при исключении только субъектов с выдуманными результатами (см. рис. 2). Оценки совпадают лишь для 2008 г., когда, по всей видимости, фальсификации были столь наглыми, что любой метод анализа позволяет элиминировать все затронутые ими субъекты.

ЦЕЛЕВЫЕ УСТАНОВКИ

Равновероятность последних цифр

Если в субъекте федерации или какой-то его части выдумывание результатов не пускают на самотек, а координируют из единого центра, то могут получаться не разные круглые числа, а одинаковые (и возможно) некруглые, что тоже нарушает равновероятность их последних цифр. Однако неизвестно, частота каких цифр и в какую сторону при этом отклоняется от $p = 1/10$. Строгий анализ в данном случае требовал бы работы с полиномиальным распределением, что представляет значительную вычислительную сложность. Но поскольку указанный способ фальсификации сравнительно редок, для него целесообразно прибегнуть к упрощенной процедуре анализа.

При большом числе испытаний биномиальное распределение приближается к нормальному, что позволяет использовать критерий согласия Пирсона для проверки гипотезы о естественном происхождении неодинаковой частоты разных последних цифр. Пусть в конце рассматриваемой характеристики n_i раз оказывается цифра i , тогда нормализованное отклонение ее частоты $\delta_i = (n_i - pn) / \sqrt{pn}$, где $n = \sum_{i=0}^9 n_i$. При условии нормального распределения величин n_i статистика $S = \sum_{i=0}^9 \delta_i^2$ подчиняется распределению χ^2 с 9 степенями свободы, что позволяет вычислить уровень значимости гипотезы $\tilde{\alpha} = I_{S/2}(9/2)$, где $I_z(x) = \Gamma_z(x) / \Gamma(x)$ – регуляризованная неполная верхняя гамма-функция. Бумажные таблицы распределения χ^2 тут бесполезны, т.к. не простираются на область столь больших S , при которых неверность гипотезы очевидна и требует не столько подтверждения, сколько количественного измерения.

Именно этот тест ранее использовался в работах [2,20] для проверки равновероятности последних цифр электоральных характеристик. Однако возникающие при фальсификациях сверхмалые уровни значимости свидетельствуют о пребывании статистики на самом хвосте распределения, где его вид, к сожалению, нельзя считать гауссовым ни при каких разумных объемах выборок. В результате вычисляемые в этом разделе значения $\tilde{\alpha}$ оказываются занижены (их десятичные показатели $p\tilde{\alpha}$ – завышены), причем тем сильнее, чем масштабнее фальсификации. Для среднего размера субъекта федерации в $n \sim 10^3$ избира-

тельных участков на количественную точность показателя \tilde{r}_α можно полагаться, лишь пока он не превосходит нескольких единиц, тогда как при бóльших значениях \tilde{r}_α его уже следует рассматривать лишь как качественную меру фальсификаций. Иначе говоря, метод вполне позволяет выявить субъекты федерации, где были фальсификации рассматриваемого типа, но при этом драматизирует картину. Поэтому нас будут интересовать не все выявляемые с помощью показателя \tilde{r}_α подозрительные случаи, а только те из них, которые не позволил выявить показатель r_α . Таких случаев не очень много, что позволяет рассмотреть их индивидуально.

Весьма показательна ситуация, сложившаяся в 2016 г. в Воронежской обл. с явкой, для которой $\tilde{r}_\alpha = 4,3$ (при $r_\alpha = 1,5$). Здесь фальсификаторы подошли к выдумыванию результатов творчески, иногда немного приподнимая явку над целыми процентами. В итоге избыточно частыми последними цифрами ее промилле оказались «0», «1», «2», «3» и «5» (цифра «4» ожидаемо проигрывает по психологической притягательности цифре «5»). Такой способ фальсификации нетривиален и поэтому весьма редок. Однако его наличие наглядно демонстрирует, что нельзя ограничиваться контролем частоты круглых чисел, даже если речь идет о результатах, выдуманных без опоры на какие-либо целевые значения.

Вполне естественно, что целевые проценты обычно устанавливаются не для явки, а для результата власти. Так, в 2008 г. в Ингушетии он тяготеет к полукруглому числу 82,5%, отклонение от которого не превышает 0,1% пункта на участках, охватывающих 15% избирателей, а 0,5% пункта – на участках, охватывающих уже почти 44% избирателей. Аналогичным образом, в 2016 г. в Карачаево-Черкесии результат власти тяготеет к значению 77,5%, отклонение от которого не превышает 0,1% и 0,5% пункта при охвате избирателей соответственно в 19% и 28%.

Целевые проценты не всегда бывают круглыми или полукруглыми. В 2016 г. в Саратовской обл. зафиксированы показатели $\tilde{r}_\alpha = 9,9$ для явки и $\tilde{r}_\alpha = 5,8$ для результата власти при концентрации этих характеристик на 64,3% и 62,2% соответственно. Саратовское голосование стало именем нарицательным [10]. Его изюминка – не только в рекордной (см. далее) концентрации результатов на целевых значениях, но и в нетривиальном способе их выбора, за которым прослеживается определенная работа мысли. Весьма вероятно, что фальсификаторы изначально ориентировались на явку за власть (долю голосов, отданную за партию власти, рассчитываемую от числа зарегистрированных избирателей) ровно в 40% [12]. Поскольку частная явка за партию равна (с точностью до унесенных избирателями бюллетеней) произведению ее результата и общей явки, получить желаемое проще всего, положив оба множителя равными $\sqrt{40\%} \approx 63,246\%$. Однако идеальное совпадение этих характеристик выглядело бы чересчур подозрительно. Поэтому целевой результат партии власти установили меньшим ровно на 1% пункт (при расчетах с округлением до 3 значащих цифр получилось 62,2%), а целевую явку – большей ровно на 1% пункт (эта величина, видимо, уже вычислялась как частное, поэтому сотые доли процента

округлились вверх и получилось 64,3%). Практически все осуществленные по данной схеме фальсификации были локализованы в городе Саратове и в окружающем его Саратовском районе области (причем львиная доля фальсификаций пришлась на Заводской, Ленинский, Октябрьский и Фрунзенский районы города, расположенные в его западной части). Столь высокая пространственная неоднородность не могла возникнуть без централизованного руководства, особенно если учесть подгонку сразу двух электоральных характеристик.

Наконец, невероятные значения $\tilde{r}\alpha$ многократно фиксируются для Чечни, как показано в табл. 6. Они преимущественно соответствуют некруглым процентам и обусловлены исключительной узостью распределения чеченских участков как по явке, так и по результату власти в 2007-16 гг., способной нарушить равномерность распределения последних цифр. Абсолютный рекорд был установлен здесь на президентских выборах 2012 г., когда разность между верхним и нижним квартилями распределения результата кандидата власти составляла менее 0,2% пункта (т.е. на участках, охватывающих половину избирателей, результат отклонялся от медианного значения не сильнее единицы третьей значащей цифры). Совокупно с круглым числом участников выборов в 2003-04 гг. (см. табл. 1) Чечня показывает невероятные значения $r\alpha$ и/или $\tilde{r}\alpha$ на всех проводившихся в ней федеральных выборах, кроме 2008 и 2018 гг., которые безупречно проходят тест на частоту последних цифр. Однако уникальная способность данного субъекта федерации произвольно включать и выключать невероятные значения показателей для разных электоральных характеристик не оставляет сомнений в том, что и в эти годы результаты там были фальсифицированы, в чем мы убедимся далее.

Таблица. 6. Величина $\tilde{r}\alpha$ для Чечни

Характеристика	2003	2004	2007	2008	2011	2012	2016	2018
Общая явка	3,0	2,8	7,3	0,4	54,5	137,4	0,6	0,4
Результат власти	1,5	2,9	14,7	0,8	49,5	82,2	18,1	0,2

В завершение заметим, что для числа избирателей, принявших участие в выборах, доминирование каких-либо последних цифр кроме «0» не наблюдается, как того и следовало ожидать. Здесь подозрительное значение показателя было зарегистрировано лишь однажды на выборах 2018 г. в Тверской обл. ($\tilde{r}\alpha = 3,4$). Однако оно, по всей видимости, является результатом стечения обстоятельств, поскольку более никаких возможных последствий массовых фальсификаций обнаружить не удается.

Сгустки явки и результатов

Ситуация, когда значения некоторой электоральной характеристики для ряда участков скапливаются в тесной близости некоторого специального значения, называется *сгустком*. Традиционно поиск сгустков основывается на обнаружении пиков эмпирической плотности вероятности рассматриваемой характеристики [7,11,12,18].

Недостатком такого подхода оказывается высокая чувствительность результатов к используемой ширине кармана гистограммы, что оставляет излишнее пространство для толкований и препятствует формализации алгоритма. Поэтому предпочтительно использование не плотности вероятности, а функции распределения. Для ее вычисления участки субъекта сортируются по величине рассматриваемой характеристики x и для каждого из них рассчитывается доля избирателей y , зарегистрированных на участках с меньшими ее значениями. При этом участки с в точности одинаковыми x объединяются в один виртуальный участок суммарного размера, чтобы предотвратить возникновение ложных сгустков на значениях, выражающихся дробями с малыми знаменателями.

Алгоритм состоит из двух стадий. На первой осуществляется предобработка данных с целью выделения потенциальных *границ сгустков*. Для этого в скользящем окне шириной в $2m+1$ точку строится непрерывное кусочно-линейное приближение зависимости $y(x)$, состоящее из двух частей равной ширины w . Они сопрягаются по абсциссе i -й точки, описывая соответственно точ-

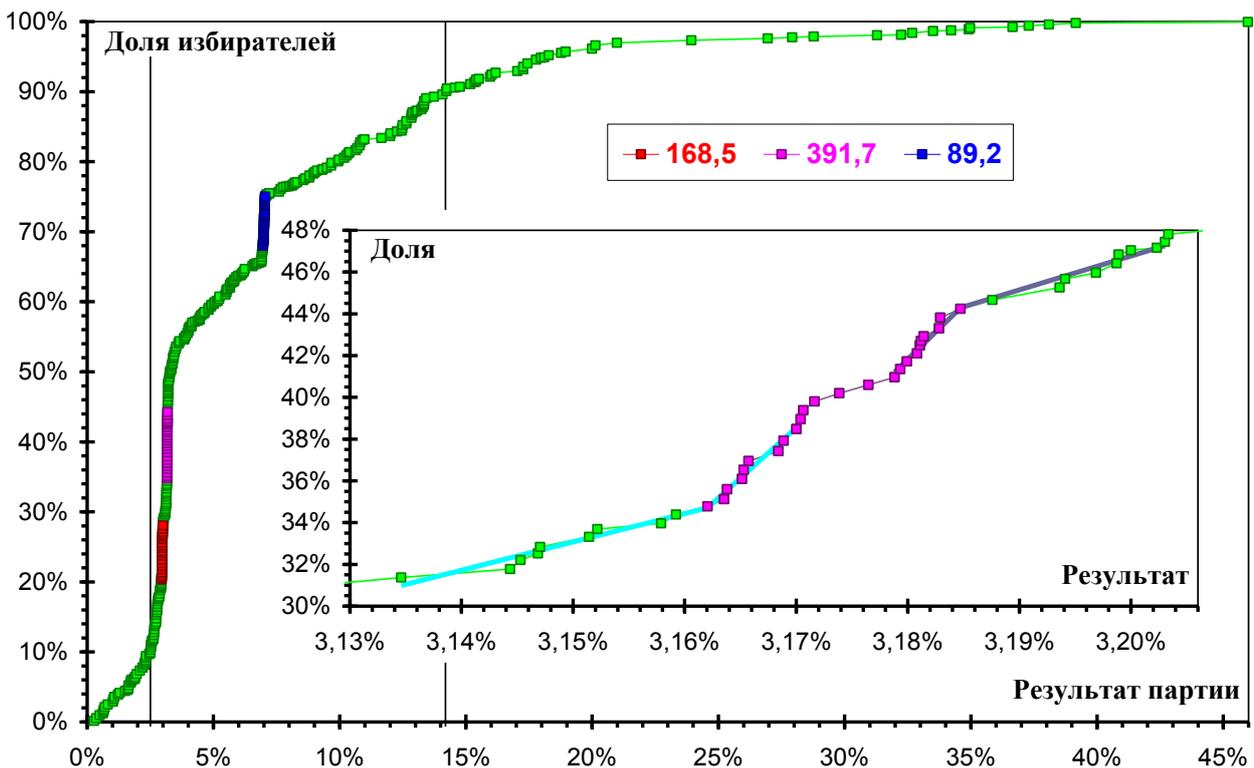


Рис. 6. Пример поиска сгустков для распределения избирателей по результату оппозиционной партии на выборах 2011 г. в Северной Осетии

Тонкие вертикальные линии обозначают положение децилей. В легенде указаны угловые коэффициенты K для трех обнаруженных сгустков.

На врезке увеличен сильнейший из сгустков и показаны кусочно-линейные приближения, с помощью которых определялись его границы. Заметим, что, как это часто бывает, данный сгусток лежит внутри заметно более крупного, но менее сильного сгустка, охватывающего все показанные на врезке точки. Для него K оказывается несколько ниже, а отношения k_+ и k_- – значительно выше.

ки с номерами $i-m \leq j \leq i$ и $i \leq j \leq i+m$. Если отношение угловых коэффициентов участков $k_+/k_- \geq t$, точка i может быть левой границей сгустка (график стал заметно круче), а если $k_-/k_+ \geq t$ – правой (график стал заметно положе).

На второй стадии анализируются все отрезки точек, способные оказаться сгустками, простирающиеся от некоторой левой границы до некоторой правой. Причем эти отрезки должны охватывать не менее p_{\min} и не более p_{\max} избирателей субъекта и содержать не менее n_{\min} и не более n_{\max} точек. Нижние ограничения имеют целью подавление случайностей, возможных при рассмотрении слишком малой части субъекта, а верхние – предназначены для ускорения работы алгоритма. При его реализации были выбраны следующие значения параметров: $m = 9$, $t = 3$, $p_{\min} = 2\%$, $p_{\max} = 25\%$, $n_{\min} = 20$ и $n_{\max} = 1\,000$, показавшие хорошие результаты. Впрочем, они слабо зависят от конкретных значений параметров.

Для каждого анализируемого отрезка вычисляется линейное приближение зависимости $y(x)$. При этом для каждого положения левой границы выбирается такое положение правой, при котором приближение имеет наибольший коэффициент детерминации. А из всех положений левой границы берется то, при котором угловой коэффициент линейного участка k , оценивающий плотность распределения, получается наибольшим. Рис. 6 иллюстрирует эту процедуру.

Последним шагом является вычисление *силы сгустка* $s = k \cdot w_{80}$, где $w_{80} = q_{90} - q_{10}$ – разность верхней и нижней децилей распределения величины x , при вычислении которых участки учитываются с весом, равным числу избирателей. Домножение на ширину распределения позволяет отстроиться от различий в диапазоне изменения рассматриваемой электоральной характеристики. А использование для измерения ширины децильной разности (а не среднеквадратичного отклонения) защищает от искажений, вносимых разного рода специализированными участками.

В случае генеральной совокупности нормального распределения сила сгустка, соответствующего максимуму его плотности, $s_0 = 2 \cdot \operatorname{erf}^{-1}(0,8) / \sqrt{\pi} \approx 1,022\dots$ И хотя реальные распределения электоральных характеристик могут быть негауссовыми, а для конечных выборок неизбежны флуктуации, меняющие силу сгустка, ее примерно единичное значение принимается как естественное. Остается лишь определить масштаб отклонений от него, свидетельствующий о фальсификациях.

Анализ сгустков проводился совместно для явки (кроме зарубежных территорий, где ее величина зачастую фиктивна), а также для результатов партий/кандидатов, показавшихся первый, второй и третий результаты (в общефедеральном зачете). Ключевым преимуществом данного метода выявления фальсификаций является именно возможность проверять официальные результаты оппозиционных партий/кандидатов (ранее эти результаты пришлось обойти вниманием, поскольку они слишком малы, чтобы диапазон их изменения позволял с уверенностью говорить о равновероятности последних цифр). Здесь мы ограничиваемся лишь тройкой лидеров, хотя, скорее всего, можно рассматривать и несколько следующих мест без риска столкнуться с эффектами, обусловленными дискретностью голосов.

Таблица. 7. Максимальная сила сгустков

Субъект	1999	2000	2003	2004	2007	2008	2011	2012	2016	2018
Саратовская обл.	1,1	1,9	3,1	1,5	1,4	1,3	5,6	1,7	50,7	1,4
Северная Осетия	—	2,6	2,0	2,1	1,5	2,4	45,9	15,3	43,4	2,5
Карачаево-Черкесия	—	—	1,7	1,5	2,0	6,9	13,9	3,0	36,9	20,0
Тюменская обл.	0,3	1,9	2,7	2,4	1,8	2,5	2,6	2,0	27,7	1,6
Дагестан	—	2,9	3,6	4,6	7,9	25,4	7,7	14,7	19,1	18,2
Ставропольский кр.	1,5	1,8	1,3	1,2	1,3	1,3	1,1	1,7	1,3	21,9
Татарстан	1,5	1,7	3,8	2,3	3,1	20,0	6,7	4,3	8,5	2,8
Коми	2,2	1,4	1,3	1,1	1,4	16,8	2,6	2,3	2,0	1,3
Ингушетия	—	—	—	—	3,1	8,7	7,6	11,9	10,2	12,5
Кабардино-Балкария	—	7,4	2,2	10,0	4,1	8,5	7,0	6,0	7,7	2,9
Мордовия	2,5	1,7	3,1	8,5	6,8	4,1	8,5	3,9	2,0	2,9
Башкортостан	3,3	1,9	2,0	6,0	2,9	7,7	2,1	2,4	4,4	1,1
Кемеровская обл.	1,4	1,5	3,1	1,3	1,3	4,4	3,3	2,0	1,6	7,5
Зарубежная территория	—	—	1,4	4,7	5,9	2,4	2,0	6,8	2,7	2,8
Калмыкия	—	—	2,1	1,2	1,3	1,6	1,7	0,4	5,5	2,3
Чечня	—	—	2,0	2,1	1,0	2,4	1,4	1,7	3,3	5,1

В табл. 7 представлены данные о максимальной силе сгустков для субъектов, где она хотя бы для одной характеристики превысила 5. Пропуски в таблице возникают не только в силу отсутствия данных, но и в силу отсутствия сгустков, удовлетворяющих сформулированным выше критериям. Выделение жирным шрифтом и малиновым цветом выполнено по порогам в 3 и 11, происхождение которых поясняется далее. Заливка кирпичным цветом соответствует случаям, которые с точки зрения анализа частоты последних цифр не были даже подозрительными ($ra, r\tilde{a} < 3$), желтым – были лишь подозрительными, но еще не исключительными ($ra, r\tilde{a} < 4$), а зеленым – уже были исключительными, но еще были не невероятными ($ra, r\tilde{a} < 5$).

Наиболее впечатляющей из фальсификаций, выявленных этим методом, стал кейс парламентских выборов 2011 г. в Северной Осетии, где результат бронзового призера («Справедливая Россия») на 24 участках, охватывающих 9,76% избирателей республики, укладывается в диапазон от 3,162% до 3,185% (при децильной ширине распределения в 11,7%). В результате возник сгусток с $s = 45,9$, использованный как иллюстративный пример на рис. 6 и уступающий по силе лишь рассмотренному выше знаменитому саратовскому голосованию, у которого $s = 50,7$ и 46,6 для явки и результата партии власти соответственно.

Для определения порогов, позволяющих интерпретировать силу сгустков, воспользуемся процедурой, аналогичной той, которая применялась в случае жадности голосования. К сожалению, здесь в отличие от предыдущего случая нет наводящих соображений, поэтому приходится опираться лишь на анализ графиков, представленный на рис. 7. При его построении, в отличие от рис. 3, критерием оптимизации был общий коэффициент детерминации данных.

Поскольку нетривиальные значения показателей, описывающих степенные участки зависимости $s(r)$, не имеют какого-либо теоретического объясне-

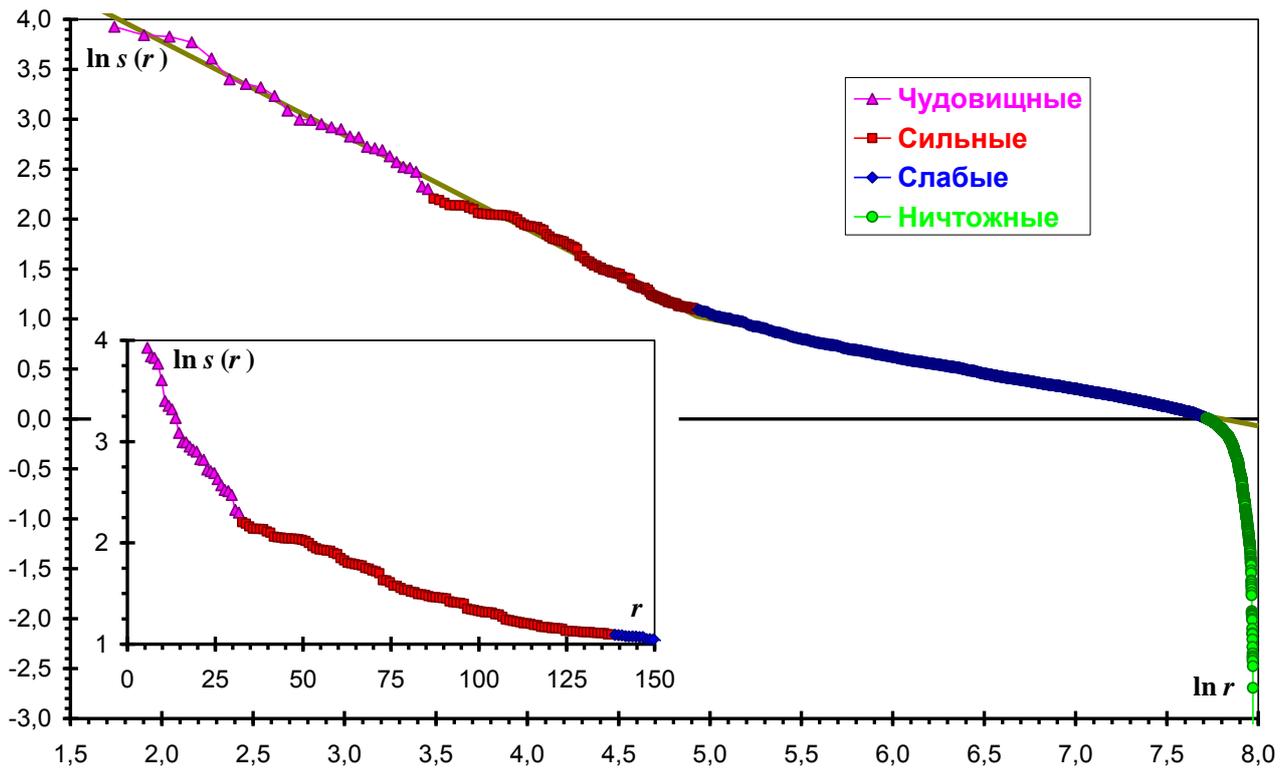


Рис. 7. Ранжировка выборов и субъектов по силе сгустков

В двойном логарифмическом масштабе зависимость распадается на два степенных участка, представленных прямыми с угловыми коэффициентами $-0,94$ и $-0,36$, и область быстрого спада. Наилучшее приближение достигается при $r_0 = 5,70$ и отнесении ко второму участку сгустков силой $0,94 < s < 2,98 = s_1$ (слабые).

На врезке в полулогарифмическом масштабе показано выделение из множества сильных сгустков подмножества чудовищных с $s > 9,52 = s_3$.

ния, деление сгустков на сильные и слабые оказывается в определенной мере условным. Да и выделение области чудовищных сгустков выполнено в значительной мере «на глаз». Поэтому полученные значения порогов следует трактовать как ориентировочные, в силу чего при раскраске табл. 7 они были округлены до ближайших целых.

Сложности с качественной интерпретацией силы сгустков, по всей видимости, обусловлены высокой зашумленностью соответствующих данных. Это же подтверждает и предпринятая на рис. 8 попытка оценить количество фальшивых голосов, приобретенных кандидатом/партией власти за счет массовых фальсификаций данного типа. В отличие от рис. 5, основанного на данных о жадном голосовании, последовательное исключение из рассмотрения субъектов с сильнейшими сгустками не позволяет достичь стационарных значений для интегрального размера фальсификаций, да и их максимальные оценки оказываются в целом ниже. Данное обстоятельство служит напоминанием о том, что в отличие от жадности голосования наличие сгустков (так же, как и нарушение равновероятности последних цифр) не является универсальным признаком фальсификаций.

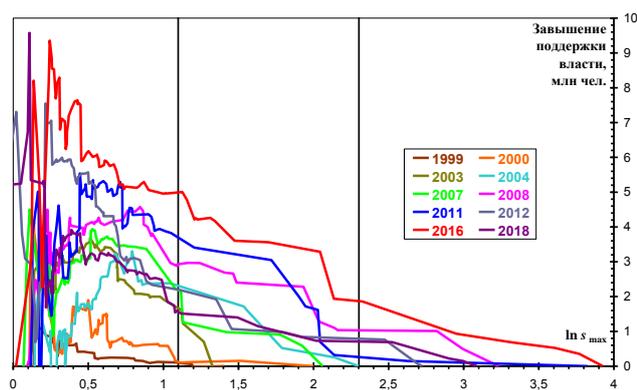


Рис. 8. Фальсификации поддержки власти, выявляемые сгустками

Рисунок аналогичен рис. 5. Избыточное число голосов, якобы поданных за власть, рассчитывается распространением ее результата, полученного на субъектах с $s < s_{\max}$, на всю страну.

Вертикальные линии соответствуют порогам, разделяющим слабые, сильные и чудовищные пики. При движении справа налево количество ненастоящих голосов неуклонно возрастает даже в области слабых пиков.

Завершая тему сгустков, заметим, что существует альтернативный подход к определению той степени уверенности, с которой они позволяют говорить о наличии фальсификаций [11,12,13,18]. Однако применяется он не к отдельным субъектам, а к стране в целом. Для нее выборка участков уже достаточно обширна, чтобы имело смысл анализировать дифференциальное распределение явки и результата власти. Вычисление его эмпирической плотности с узкими карманами диаграммы позволяет увидеть сгустки, а с широкими – тот естественный фон, над которым они поднимаются. На основе последнего с помощью метода Монте-Карло можно оценить вероятность возникновения сгустков (особенно хорошо это получается для сгустков на больших целочисленных процентах, кратных 5).

НЕДЕЙСТВИТЕЛЬНЫЕ БЮЛЛЕТЕНИ

Совершенно особое место среди электоральных характеристик занимает доля недействительных бюллетеней, которая при фальсификациях может только уменьшаться.

Известен случай, когда некий оппозиционный политик добился пересчета голосов на участке, где лично он голосовал, на том основании, что за его партию там якобы не было подано ни одного голоса. Но вряд ли какой-то избиратель, испортивший бюллетень, сумел бы опротестовать то, что на его участке все бюллетени сочтены действительными.

Кроме того, несложно представить, что результаты выборов иногда фальсифицируются и не в пользу кандидата/партии власти. Чаще всего это происходит из-за технических ошибок при фальсификациях или тихого протеста против них, выдаваемого за подобные ошибки (при этом большие цифры, предназначенные власти, приписывают кому-то еще [9,11]). Впрочем, нельзя исключать и умышленных фальсификаций в пользу оппозиционных кандидатов/партий там, где их представители имеют достаточный вес в региональных органах власти. Как бы то ни было, чего точно не бывает, так это фальсификаций в пользу «партии недействительных бюллетеней».

В свете сказанного для доли недееспособных бюллетеней имеет смысл говорить об аномально низких значениях, обнаружение которых может быть положено в основу еще одного теста на наличие массовых фальсификаций. Поскольку выборы разных лет различаются по уровню протестных настроений, доля недееспособных бюллетеней для каждого выборов нормируется на ее медианное значение, рассчитанное для всех субъектов федерации, учитываемых с весом, равным числу избирателей.

Здесь мы ограничим анализ выборами 2007-18 гг., поскольку до 2006 г. в избирательных бюллетенях присутствовала графа «против всех», наличие которой делало бессмысленной умышленную порчу бюллетеней. В результате недееспособных бюллетеней оказывалось не так много, как ныне.

Рис. 9 показывает, как выбираются пороги для сверхнизкой (фальсификации несомненны) и низкой (фальсификации вероятны) относительной долей недееспособных бюллетеней d . Чтобы убедиться, что уменьшение ее величины и впрямь может служить индикатором фальсификаций, на графиках отмечены все ранее выявленные недостоверные выборы. Недостоверность определялась суммированием штрафных баллов. За подозрительное значение $\max\{\rho, \tilde{\rho}\}$ субъекту начислялся 1 балл, за исключительное – 2, за невероятное – 3. Кроме того, i баллов добавлялись за $g > g_i$. Наконец, для унификации процедуры был введен промежуточный порог силы сгустков $s_2 = 5$ (де факто он уже использовался при построении табл. 7) и добавлено i баллов за $s > s_i$. Если сумма баллов оказывается больше или равна 3 (т.е. достаточно один раз провалить один тест, но полностью), выборы в субъекте считаются фальсифицированными. Как можно видеть из рис. 9, относительная доля недееспособных бюллетеней хорошо коррелирует с достоверностью результатов.

По аналогии с жадностью голосования и силой сгустков относительная доля недееспособных бюллетеней была использована для оценки интегрального размера фальсификации. Результат последовательного исключения субъектов с малыми значениями d , показанный на рис. 10, следует признать неоднозначным. С одной стороны, налицо ограниченная применимость метода, обусловленная изменениями в выборном законодательстве. Ранжировка регионов по величине данной характеристики имеет смысл только для 4 последних выборов. Но, с другой стороны, именно при ее использовании для завышения результата власти из-за массовых фальсификаций удается получить наибольшие оценки. В случае парламентских выборов получают значения примерно в 10 и 13 млн голосов для 2011 и 2016 гг. соответственно, отстающие от оценок общего завышения результата партии власти, полученных с помощью полной реконструкции результатов [2], лишь на 4÷5 млн голосов. А в случае президентских выборов графики вовсе не выходят на стационар, что может означать практически полное отсутствие субъектов, где бы испорченные бюллетени не засчитывались в поддержку кандидата власти и не разбавлялись ими.

В табл. 8 приведены данные о субъектах, в которых хотя бы раз регистрировалась сверхнизкая относительная доля недееспособных бюллетеней

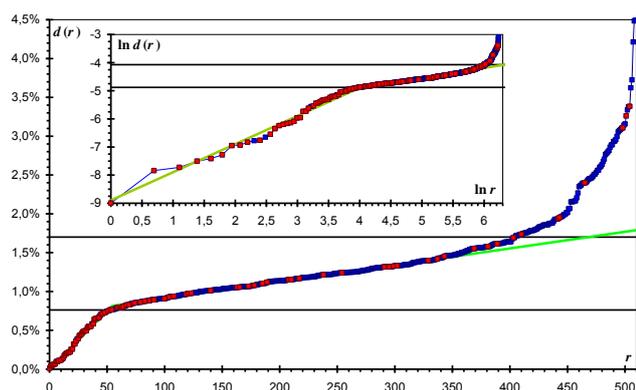


Рис. 9. Ранжировка выборов по возрастанию относительной доли недействительных бюллетеней

Данные хорошо приближаются непрерывной кусочно-линейной зависимостью, левый участок которой проходит через начало координат. При этом первые 63 значения следует считать сверхнизкими, а следующие 322 – низкими. Положения порогов $d_1 = 0,681$ и $d_2 = 1,169$ обозначены горизонтальными линиями.

Красным цветом отмечены выборы, которые по совокупности предыдущих тестов следует считать фальсифицированными. Как можно видеть, таковы за немногими исключениями все выборы с $d < d_1$, значительная часть выборов с $d_1 < d < d_2$ и очень немногие с $d_2 < d$.

(соответствующие значения выделены жирным шрифтом). Розовым залиты ячейки, соответствующие выборам, для которых массовые фальсификации ранее не были обнаружены ни одним тестом, а голубым – те, которые прошли тест на равновероятность последних цифр, но продемонстрировали масштабную жадность или сильные сгустки. Особо отметим, что относительная доля недействительных бюллетеней сверхнизка во всех цветных ячейках для Ингушетии, Кабардино-Балкарии, Карачаево-Черкесии, Мордовии и Чечни. Таким образом, в этих субъектах (равно как и в уже упомянутых ранее Башкортостане, Дагестане, Татарстане и Кемеровской обл.) за последние, по крайней мере, 12 лет ни одни выборы не обошлись без массовых фальсификаций.

Бросается в глаза резкий переход от почти белой верхней части табл. 8 к почти полностью закрашенной нижней. Вверху перечислены субъекты, уже многократно уличенные в фальсификациях другими методами, так что нехватка недействительных бюллетеней лишь подтверждает разумное предположение, что выборы здесь фальсифицируются тотально. А внизу расположились субъекты, где фальсификации тоже, скорее всего, имеют место, но обычно не являются массовыми, из-за чего их обнаружение носит вероятностный характер.

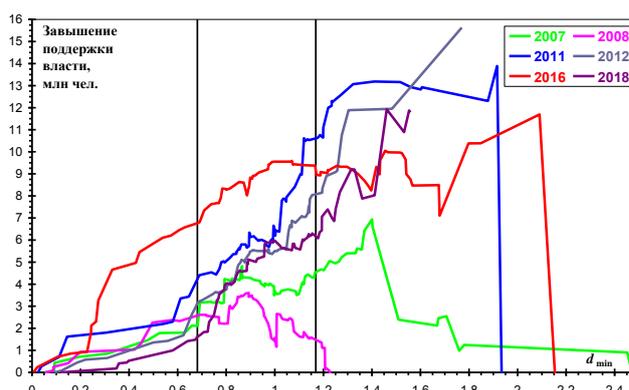


Рис. 10. Фальсификации поддержки власти, выявляемые долей недействительных бюллетеней

Избыточное число голосов власти рассчитывается распространением ее результата, полученного при $d > d_{\min}$, на всю страну. В отличие от рис. 5 и 8 здесь, чем правее точка, тем меньшее число субъектов она учитывает.

Вертикальные линии соответствуют порогам d_1 и d_2 . В области сверхнизких значений d все графики уверенно возрастают, тогда как в области низких это можно сказать только про выборы 2010-х гг. А в 2007-08 гг., когда порча бюллетеней как форма протестного голосования еще не получила широкого распространения, графики сравнительно рано переходят от роста к падению и даже проваливаются под ноль.

Таблица 8. Относительная доля недействительных бюллетеней

Субъект	2007	2008	2011	2012	2016	2018
Кабардино-Балкария	0,09	0,09	0,03	0,09	0,01	0,12
Чечня	0,09	0,09	0,04	0,12	0,02	0,35
Ингушетия	0,05	0,05	0,54	0,22	0,16	0,36
Карачаево-Черкесия	0,19	0,20	0,14	0,19	0,10	0,39
Дагестан	0,31	0,15	0,11	0,31	0,26	0,40
Кемеровская обл.	0,86	2,50	0,99	0,84	0,23	0,58
Мордовия	0,48	0,42	0,31	0,56	0,24	0,67
Северная Осетия	2,52	1,48	0,79	0,83	0,24	0,68
Татарстан	0,79	0,81	0,58	0,62	0,27	0,72
Тюменская обл.	1,10	0,83	0,76	0,71	0,33	0,77
Саратовская обл.	1,67	0,84	0,89	0,80	0,43	0,77
Башкортостан	0,68	0,45	0,65	0,66	0,44	0,73
Ямало-Ненецкий авт.окр.	1,00	0,50	0,69	0,68	0,61	0,70
Тыва	1,02	0,89	0,61	0,50	0,54	0,64
Воронежская обл.	1,01	0,96	0,81	0,86	0,57	0,71
Курганская обл.	0,70	0,58	0,84	0,77	0,77	0,81
Ростовская обл.	0,63	0,61	0,85	0,88	0,78	0,79
Брянская обл.	0,87	0,95	0,89	0,85	0,63	0,83
Костромская обл.	0,66	0,72	0,79	0,76	0,85	0,86
Рязанская обл.	0,99	1,12	1,00	0,90	0,68	0,87

Выводы

Для выявления массовых фальсификаций разработана батарея тестов. Они различаются как по степени обоснованности и надежности, так и по существу анализируемых признаков и уровню избирательности, вследствие чего сочетание тестов дает разностороннее представление о формах и масштабах фальсификаций.

Преобладание психологически привлекательных чисел – круглого количества избирателей, принявших участие в выборах, а также промилле явки избирателей и результата власти – является самым надежным признаком массовых фальсификаций, практически не оставляющим пространства для каких-либо интерпретаций. Лишь немногим хуже проверка для указанных электро-ральных характеристик одинаковости частоты последних цифр. Если значимость гипотезы о естественной природе различий в частоте чрезвычайно мала (фальсификации особо масштабны), этот тест дополнительно ее занижает. Однако сами фальсификации он выявляет столь же надежно, как и предыдущий, не будучи при этом ограниченным лишь круглыми числами.

Анализ сгустков явки и результата оставляет чуть больше свободы для трактовок. Алгоритм поиска сгустков имеет настроечные параметры, а шкала силы сгустков, обладая выделенным началом, всё же требует калибровки. И хотя есть все основания полагать, что данная калибровка должна быть единой для всех выборов, масштабы фальсификаций интерпретируются уже лишь на каче-

ственном уровне – без определения вероятности возникновения сгустков данной силы в результате стечения обстоятельств.

Еще сложнее оказывается трактовка жадности голосования, т.к. нельзя наверняка утверждать, что ее небольшие значения возникают именно в результате умышленных фальсификаций, а не процедурных нарушений. Кроме того, избыток бюллетеней, получаемых участковыми избирательными комиссиями, может варьироваться от выборов к выборам, что должно сказываться на общем уровне жадности. Поэтому выбор порогов, на основе которых он интерпретируется, опирается на чрезвычайно сложную методику анализа.

Наконец, относительная доля недействительных бюллетеней является в значительной мере оценочной характеристикой, поскольку даже на одних и тех же выборах регионы могут различаться по уровню протестных настроений и в силу естественных причин. И если сверхнизкая доля недействительных бюллетеней, наблюдаемая в 12,4% случаев, с высокой степенью уверенности позволяет говорить о массовых фальсификациях, то низкая, наблюдаемая уже в 63,1% случаев, оказывается не более чем поводом для подозрений. Кроме того, склонность избирателей выражать протест порчей бюллетеня, возникшая лишь после отмены графы «против всех», эволюционирует от выборов к выборам, что может требовать дополнительного анализа.

Сравнить эффективность тестов по выявлению массовых фальсификаций в наиболее склонных к ним субъектах позволяет табл. 9. В ней для каждого теста жирным шрифтом приведено число выборов, на которых соответствующая мера оказалась за последним порогом (невероятная частота круглых чисел или неравномерность последних цифр, запредельная распространенность жадного голосования, чудовищные по силе сгустки, сверхнизкая относительная доля недействительных бюллетеней). А обычным шрифтом приведено число выборов, когда она пересекла хотя бы первый порог (подозрительная частота круглых чисел или неравномерность последних цифр, умеренная распространенность жадного голосования, слабые сгустки, низкая относительная доля недействительных бюллетеней), но еще не достигла последнего.

При рассмотрении таблицы следует помнить, что из 10 выборов в Чечне они проводились лишь 8 раз, в Усть-Ордынском Бурятском авт.окр. – 5, в Крыму – 2, а данные по недействительным бюллетеням учитываются лишь для 6 последних выборов.

Если нас интересует не индивидуальное выявление субъектов, где имели место массовые фальсификации, а оценка их интегрального размера, то эффективность методов меняется на противоположную. Здесь лучше всего себя показывают подходы, основанные на величинах, получаемых прямым суммированием по участкам субъекта, – доле недействительных бюллетеней и доле участков с жадным голосованием. Исключение субъектов, показавших неправдоподобные значения этих величин, с распространением результатов оставшихся субъектов на всю страну дает наибольшие оценки числа голосов, приписанных кандидату/партии власти. А при учете субъектов с весами, равными значимости

статистических гипотез, или при использовании для сортировки субъектов максимальной силы сгустков значительная часть общего объема массовых фальсификаций остается незамеченной.

Таблица. 9. Сравнение тестов

Субъект	Тест	Последние цифры		Жадность голосования		Сила сгустков		Недейств. бюллетени	
Дагестан		9	0	8	2	4	4	6	0
Ингушетия		1	2	2	6	3	3	6	0
Кабардино-Балкария		4	2	1	5	1	6	6	0
Карачаево-Черкесия		2	3	3	6	3	2	6	0
Мордовия		5	0	3	6	0	6	6	0
Чечня		6	0	2	3	0	2	6	0
Татарстан		8	1	2	8	1	5	3	3
Башкортостан		7	1	3	7	0	4	4	2
Тыва		0	1	3	6	0	0	4	2
Крым		1	0	0	2	0	0	0	2
Кемеровская обл.		4	3	0	10	0	4	2	3
Коми		0	0	4	3	1	0	0	3
Саратовская обл.		4	1	0	10	1	2	1	4
Ростовская обл.		1	2	0	5	0	0	2	4
Северная Осетия		3	2	2	5	3	0	2	2
Ямало-Ненецкий авт.окр.		1	0	1	8	0	1	2	4
Краснодарский кр.		2	3	0	7	0	0	0	5
Марий Эл		0	0	2	5	0	0	0	6
Москва		2	0	0	1	0	0	0	2
Ставропольский кр.		1	0	0	2	1	0	0	4
Усть-Ордынский Бур. авт.окр.		0	0	1	3	0	0	0	1
Брянская обл.		0	0	0	2	0	0	1	5
Воронежская обл.		0	1	0	8	0	0	1	5
Костромская обл.		0	0	0	1	0	2	1	5
Курганская обл.		0	0	0	4	0	0	1	5
Рязанская обл.		0	0	0	3	0	0	1	5
Тюменская обл.		0	3	0	9	1	0	1	5
Липецкая обл.		0	1	1	5	0	1	0	4
Нижегородская обл.		0	0	1	3	0	1	0	6
Томская обл.		0	0	1	4	0	0	0	2

Литература

1. Подлазов А.В. Московская политическая социология по результатам выборов 2012-18 гг. // Препринты ИПМ им. М.В.Келдыша, 2018. №244. 16 с. DOI:10.20948/prepr-2018-244
2. Подлазов А.В. Выборы депутатов Государственной Думы VII созыва: Выявление фальсификации результатов и их реконструкция // Социологические исследования. 2018, №1, с.59-72. <http://socis.isras.ru/article/7016>, DOI:10.7868/S0132162518010075
3. Шпилькин С. Статистическое исследование результатов российских выборов 2007–2009 гг. // Троицкий вариант – наука. 27.10.2009, №21(40), с.2-4. <http://trv-science.ru/2009/10/27/statisticheskoe-issledovanie-rezultatov-rossijskix-vyborov-2007-2009-gg/>
4. Шпилькин С. Математика выборов – 2011 // Троицкий вариант – наука. 20.12.2011, №25(94), с.2-4. <http://trv-science.ru/2011/12/20/matematika-vyborov-2011/>
5. Шпилькин С. Двугорбая Россия // Троицкий вариант – наука. 04.10.2016, №20(214), с.1-3. <http://trv-science.ru/2016/10/04/dvugorbaya-rossiya/>
6. Шпилькин С. Выборы 2018 года: Фактор X и «пила Чурова» // Троицкий вариант – наука. 24.04.2018, №8(252), с.8–10. <http://trv-science.ru/2018/04/24/vybory-2018-faktor-x-i-pila-churova/>
7. Kobak D., Shpilkin S., Pshenichnikov M.S. Statistical anomalies in 2011–2012 Russian elections revealed by 2D correlation analysis. [arXiv:1205.0741](https://arxiv.org/abs/1205.0741)
8. Овчинников Б. Сто восемьдесят честных городов // Троицкий вариант – наука. 28.02.2012, №4(98), с.1-3. <https://trv-science.ru/2012/02/28/sto-vosemdesyat-chestnykh-gorodov/>
9. Пшеничников М. Пятнадцать курьезных фактов о выборах // Троицкий вариант – наука. 31.01.2012, №2(96), с.4. <http://trv-science.ru/2012/01/31/pyatnadcat-kureznykh-faktov-o-vyborakh/>
10. Штерн Б. Саратовский феномен (занимательная статистика) // Троицкий вариант – наука. 04.10.2016, №5(214), с.4. <https://trv-science.ru/2016/10/04/saratovskij-fenomen-zanimatelnaya-statistika/>
11. Kobak D., Shpilkin S., Pshenichnikov M.S. Statistical fingerprints of electoral fraud? // Significance. 2016. V.13, N4, p.20-23. DOI:10.1111/j.1740-9713.2016.00936.x
12. Kobak D., Shpilkin S., Pshenichnikov M.S. Putin's peaks: Russian election data revisited // Significance. 2018. V.8, N3, p.8-9. [arXiv:1804.09495](https://arxiv.org/abs/1804.09495), DOI:10.1111/j.1740-9713.2018.01141.x
13. Куприянов А. Гаусс против Чурова: Промежуточный итог // Троицкий вариант – наука. 08.05.2018, №9(253), с.4-5. <https://trv-science.ru/2018/05/08/gauss-protiv-churova-promezhutochnyj-itog/>
14. Подлазов А.В. Опыт изучения московской электоральной статистики (по итогам выборов) // Социологические исследования. 2014, №6(362), с.77-88. <http://socis.isras.ru/article/5695>
15. Шалаев Н.Е. Распределение явки: Норма и аномалии // Социодинамика. 2016, №7, с.49-66. DOI:10.7256/2409-7144.2016.7.19136
16. Кобак Д. Математика выборов – 2011: Дополнение // Троицкий вариант – наука. 20.12.2011, №25(94), с.5. <https://trv-science.ru/2011/12/20/matematika-vyborov-2011-dopolnenie/>
17. Enikolopov R., Korovkin V., Petrova M., Sonin K., Zakharov A. Field experiment estimate of electoral fraud in Russian parliamentary elections // Proc. Natl. Acad. Sci. USA. 2013. V.110, N2, p.448-452. DOI:10.1073/pnas.1206770110
18. Kobak D., Shpilkin S., Pshenichnikov M.S. Integer percentages as electoral falsification fingerprints // Ann. Appl. Stat. 2016. V.10, N1, p.54-73. [arXiv:1410.6059](https://arxiv.org/abs/1410.6059), DOI:10.1214/16-AOAS904
19. Подлазов А.В., Полунин Ю.А., Юданов А.Ю. Статистический анализ выручки крупных и средних российских компаний // Препринты ИПМ им. М.В. Келдыша. 2015. №105. – 20 с. http://keldysh.ru/papers/2015/2015_prep105.pdf
20. Столяров В. Про арифметику и немножко про выборы // Троицкий вариант – наука. 13.03.2012, №5(99), с.10-11. <http://trv-science.ru/2012/03/13/pro-arifmetiku-i-nemnozhko-pro-vybory/>