



А.В. Подлазов

Реконструкция фальсифицированных результатов выборов с помощью интегрального метода Шпилькина

Рекомендуемая форма библиографической ссылки

Подлазов А.В. Реконструкция фальсифицированных результатов выборов с помощью интегрального метода Шпилькина // Проектирование будущего. Проблемы цифровой реальности: труды 4-й Международной конференции (4-5 февраля 2021 г., Москва). — М.: ИПМ им. М.В.Келдыша, 2021. — С. 193-208. — <https://keldysh.ru/future/2021/18.pdf>
<https://doi.org/10.20948/future-2021-18>

Размещено также [видео выступления](#)

Реконструкция фальсифицированных результатов выборов с помощью интегрального метода Шпилькина

А.В. Подлазов

Институт прикладной математики им. М.В. Келдыша РАН

Аннотация. Предложен простой и формальный метод реконструкции фальсифицированных результатов выборов, опирающийся на гипотезу о взаимной независимости общей явки и результата власти. Данный метод является развитием классического метода С.А. Шпилькина, однако не требует дополнительных предположений и не имеет свободных параметров, а кроме того, единообразно реконструирует все ключевые величины и использует более простую и ясную математическую процедуру. Сопоставление официальных и истинных результатов выборов федерального уровня позволило провести политологический анализ развития института электоральных фальсификаций в России в XX в. и его влияния на политическую систему.

Ключевые слова: выборы, фальсификации, электоральная статистика, большие данные, реконструкция, обратные задачи, регрессионный анализ

Reconstruction of falsified election results using Shpilkin integral method

A.V. Podlazov

RAS Keldysh Institute of Applied Mathematics

Abstract. I propose a simple and formal method to reconstruct falsified election results, based on the hypothesis of the mutual independence of the overall turnout and the result of power. This method is a development of the classical method of S.A. Shpilkin, however, it does not require additional assumptions and has no free parameters. In addition, it uniformly reconstructs all key values and uses a simpler and clearer mathematical procedure. Comparison of official and true results of federal elections allows to perform a political science analysis of the development of the institution of electoral frauds in Russia in the 20th century and its impact on the political system.

Keywords: elections, frauds, electoral statistics, big data, reconstruction, inverse problems, regression analysis

1. Введение

Сочетание развитых информационных технологий с архаичными институтами в современной России предоставляет для научного анализа такие объекты и явления, которые в других странах либо скрыты от него, либо не возникают вовсе. В частности, масштабные фальсификации на выборах, происходящие в условиях доступности их официальных результатов, детализированных до отдельных участков, способствовали бурному развитию электоральной статистики, ставшей «русской наукой» XXI в.

Фальсификации успешно и уверенно выявляются как объективными методами [1–11], так и субъективными (наблюдение на избирательных участках, просмотр видеозаписей голосования, сопоставление предварительных и итоговых протоколов и т.п.). Поэтому восприятие фальсификаций как неотъемлемого элемента отечественной политической системы давно уже стало отправной точкой любых релевантных исследований в области российских политологии и политической социологии. И первый же вопрос, который встает в этой связи, связан с определением масштабов, форм и динамики фальсификаций.

Этот вопрос далее рассматривается на материале 10 всероссийских голосований, состоявшихся в XXI в.: президентских выборов 2000, 2004, 2008, 2012 и 2018 гг., парламентских выборов 2003, 2007, 2011 и 2016 гг., а также имевшего непостижимый правовой статус плебисцита 2020 г. Для указанного временного диапазона характерны полная представленность электоральных данных в свободном доступе (за небольшими изъятиями в 2000 г.), а также вполне сформировавшаяся политическая система с одной партией власти и де-факто одним кандидатом власти, с которыми тесно был ассоциирован и плебисцит, формально не являвшийся выборами.

2. Базовая гипотеза

Реконструкция результатов выборов, как и любая некорректная обратная задача, опирается на предположение о том, какими свойствами должно обладать ее решение. Гипотеза, сформулированная С.А. Шпилькиным, постулирует независимость *результата* участника выборов (доли голосов, набранных партией или кандидатом) от *явки избирателей* (доли зарегистрированных избирателей, пришедших на участки) [12–16].

Эту гипотезу невозможно доказать в общем случае, поскольку для некоторых специальных электоральных конфигураций она заведомо не выполняется. В самом деле, представим себе страну, в которой имеются лишь две партии – урбанисты и аграрии. Первая, естественно, пользуется большей поддержкой у городских жителей, а вторая – у сельских. И если более дисциплинировано ходят на выборы, например, сельские, то очевидно, что, чем выше на некоторой территории их доля, тем одновременно выше

6. Математические модели цифрового мира

там будут явка и результат аграриев. Таким образом, возникает корреляция результата с явкой.

Однако насколько приложима к российской ситуации описанная умозрительная конфигурация? В нашей стране население характеризуется высоким разнообразием не в одном, а во многих аспектах – национальном, конфессиональном, социальном, имущественном, семейном, возрастном, образовательном, профессиональном и т.д. Каждый из них может влиять одновременно на политические предпочтения и активность. В результате, если и возникают корреляции, то они оказываются направлены в случайные стороны и в значительной мере компенсируют друг друга. Поэтому разумно предположить, что их равнодействующая может оказаться невелика, особенно если учесть широкую распространенность электоральных фальсификаций, характерную для России [2–6].

Дополнительным доводом в пользу опоры на гипотезу Шпилькина при решении задачи реконструкции может служить ее огрубленная постановка, в рамках которой интерес представляет результат не каждого участника выборов, а лишь власти. Ее представителей от других участников отличает опора на административный ресурс, который, будучи унифицирующим фактором, дополнительно нивелирует местную специфику, способную породить корреляции между явкой и результатом.

Разумеется, приведенные соображения не избавляют исследователя от необходимости проверять справедливость базовой гипотезы, но априори ее можно считать, по крайней мере, правдоподобной.

3. Дифференциальный метод реконструкции

Данный метод был предложен самим С.А. Шпилькиным и основывается на анализе графиков дифференциального распределения по общей явке числа голосов избирателей, поданных в пользу участников выборов. Чтобы найти это распределение, диапазон явок разбивается на интервалы фиксированной ширины и для участков с явкой из каждого интервала подсчитывается число голосов, поданных за рассматриваемого участника.

В предположении справедливости базовой гипотезы графики, получающиеся для разных участников, должны иметь одинаковую форму и совмещаться масштабированием, компенсирующим различную поддержку участников. Если в пользу партии/кандидата власти вбрасывались голоса, то в области высоких явок графики совместить не удастся, пример чего приведен на рис. 1 слева. Как можно видеть, имеется обусловленный фальсификациями зазор между графиком распределения голосов за власть и отмасштабированным графиком распределения голосов против нее.

Если бы вброс голосов был единственным способом фальсификаций, такой подход немедленно позволил вычислить их объем и реконструировать истинные результаты выборов. Однако в пользу партии/кандидата

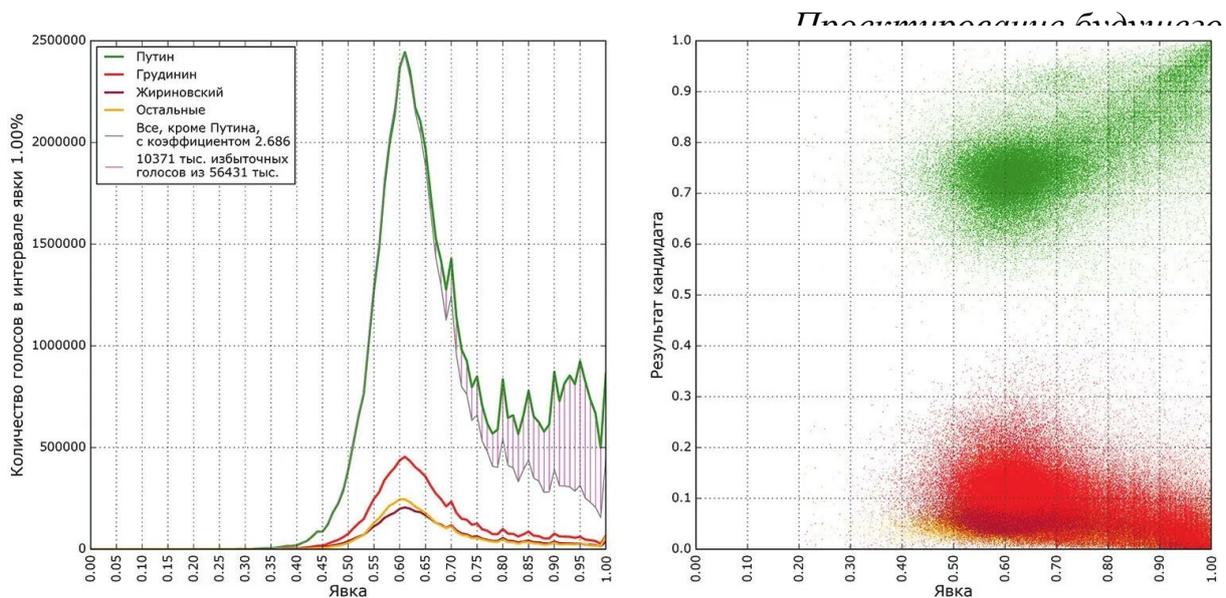


Рис. 1. Пример работы дифференциального метода Шпилькина для президентских выборов 2018 г. [15]

власти голоса не только вбрасываются, но и перебрасываются от других партий/кандидатов, что радикально усложняет реконструкцию.

Для расчетов масштабов переброса голосов С.А. Шпилькин предложил использовать диаграмму рассеяния в координатах «явка – результат» [12,13,15,16], пример которой показан на рис. 1 справа. Основной кластер точек на этой диаграмме соответствует избирательным участкам с честными результатами, а тянущийся вправо хвост – с фальсифицированными. Если отделить основной кластер и найти его центр, можно получить независимую оценку истинных явки и результатов участников. Сочетание этих оценок результатами совмещения графиков позволяет не только рассчитать объемы фальсификаций разных типов, но и проверить реконструкцию на внутреннюю согласованность [15].

4. Недостатки дифференциального метода

Исторически данный метод был ориентирован на наглядную демонстрацию наличия фальсификаций, в силу чего процедура вычисления их размеров оказывается невнятной и переусложненной.

Первой и очевидной проблемой становится наличие свободного параметра – ширины интервалов, используемых при построении распределения голосов по явке. Если взять слишком узкие интервалы, графики будут сильно зашумлены, что дополнительно осложнит их совмещение при масштабировании, осуществление которого (если делать это алгоритмически, а не визуально) и для гладких графиков нетривиально. Если же взять слишком широкие интервалы, то графики смажутся, в результате чего не только сильно перекроются диапазоны истинных и фальсифицированных явок, но и просто может не хватить количества точек на графиках для проверки их совмещения при масштабировании. Необходимо найти золотую

6. Математические модели цифрового мира

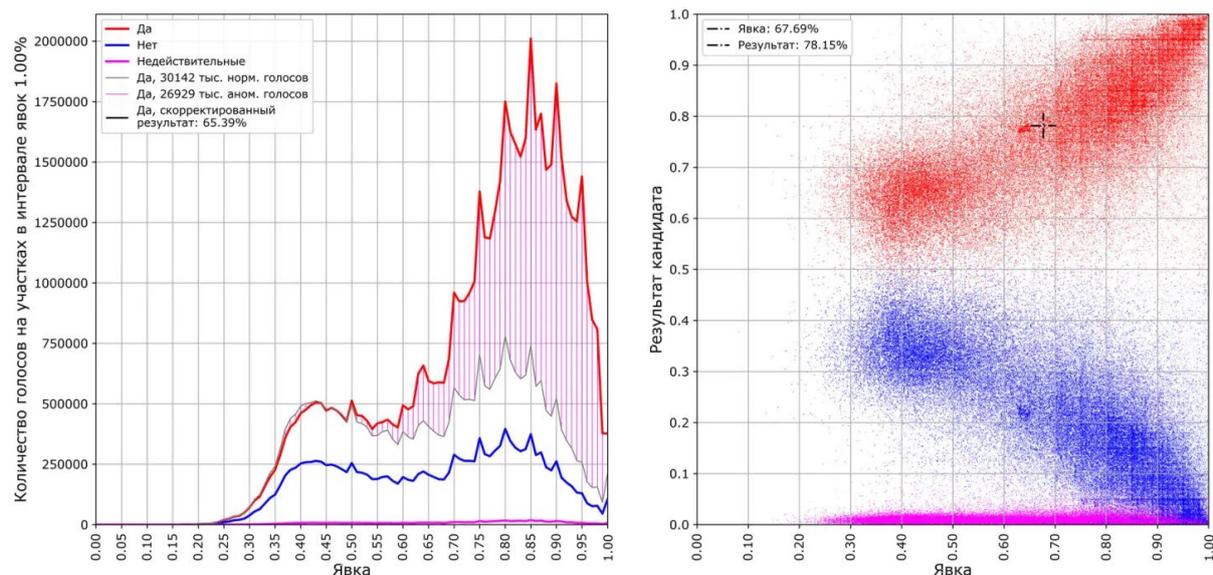


Рис. 2. Нарушение симметрии распределения голосов по явке на примере плебисцита 2020 г. [16]

середину между обозначенными крайностями. Однако вместо этого волевым образом выбираются 1%-интервалы, чтобы попутно продемонстрировать т.н. «пилу Чурова» [7,9–12,15,16] – гигантскую концентрацию голосов на полукруглых процентах явки (см. левые части рис. 1 и 2), связанную с выдумыванием результатов. Однако остается неясным, насколько такой выбор хорош для проведения расчетов.

Вторым обстоятельством, бросающимся в глаза, является то, что коэффициент масштабирования графиков, позволяющий непосредственно реконструировать результат власти, т.е. предмет первостепенного интереса, напрямую для этой цели не используется вовсе. Конечно, совмещение графиков, тем более зашумленных, неизбежно выполняется со значительной погрешностью, однако нет оснований полагать, что принципиально точнее может быть расчет, проводимый в несколько шагов, каждый из которых привносит свои ошибки.

Третья слабость метода связана с использованием для вычислений диаграммы рассеяния «явка – результат». При этом предполагается, что центр основного кластера диаграммы совпадает с истинными значениями данных величин, т.е. неявно постулируется некоторая симметричность их распределений, которая, вообще говоря, ниоткуда не следует. Более того, даже если она имеет место для честных результатов, то в случае сильных фальсификаций она может быть нарушена, пример чего демонстрирует рис. 2. Кажущийся центр основного кластера (он же – максимум нефальсифицированных частей графиков) соответствует явке примерно в 43% [16]. Однако истинная явка, которая будет реконструирована далее, на самом деле составила около 50%, что невозможно увидеть из данного рисунка. Дело в том, что, чем выше поддержка власти на участке, тем вероятнее

там фальсификации. И если их общий уровень очень высок, то вправо убегают даже участки, изначально находившиеся вблизи центра основного кластера, что приводит к его кажущемуся смещению влево.

Наконец, четвертая странность метода заключается в том, что, вроде бы, из анализа диаграммы рассеяния мы уже знаем истинные явку и результаты. Зачем же тогда нужно было рассматривать графики распределений, которые сами по себе не позволяют получить количественных результатов? По-видимому, дело в том, что совмещение графиков очень наглядно, тогда как выделение основного кластера диаграммы и поиск его центра требуют использования специальных алгоритмов, которые трудно объяснить широкой аудитории.

5. Интегральный метод реконструкции

Описанных недостатков лишен предлагаемый новый метод, основанный на использовании не дифференциальных, а интегральных распределений.

Аналогом первого этапа анализа становится рассмотрение для участков с общей явкой, не превышающей данной, зависимости числа голосов, поданных за власть, от числа голосов, поданных против нее. Иначе говоря, участки сортируются по явке, и по мере ее возрастания суммируются с накоплением голоса, поданные за и против партии/кандидата власти. Зависимость между получаемыми суммами является прямой пропорциональностью, пока фальсификаций нет, и начинает отклоняться от нее при их появлении, как схематично показано на рис. 3.

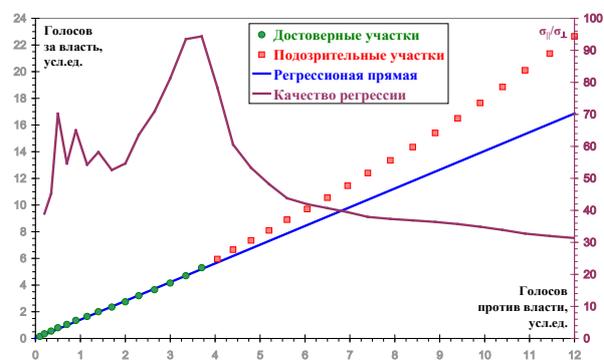


Рис. 3. Модельный пример

Для реконструкции результата власти, необходимо выделить ту часть зависимости, которая наилучшим образом приближается прямой, проходящей через начало координат, и определить ее угловой коэффициент.

Традиционный регрессионный анализ предполагает, что значения объясняющей переменной известны без погрешностей, что позволяет вычислить параметры регрессионной зависимости, минимизируя среднее квадратичное отклонение экспериментальных точек от нее по величине объясняемой переменной. Такая асимметрия переменных проявляется, в частности, в том, что если поменять местами абсциссу и ординату, то угловой коэффициент линейной регрессии отнюдь не заменится на обратный, как это было бы для функциональной зависимости.

В нашем случае обе переменные известны с погрешностями, но при этом равноправны и соизмеримы, что вынуждает и позволяет использовать

6. Математические модели цифрового мира

для определения связи между ними регрессию Деминга. Она минимизирует среднеквадратичное отклонение точек от прямой в перпендикулярном ей направлении, а не вдоль ординаты. Формулы этой регрессионной модели в общем виде довольно громоздки, но заметно упрощаются, если рассматривается прямая вида $y = k \cdot x$ с $k > 0$, т.е. проходящая через начало координат в первую координатную четверть. В этом случае коэффициент пропорциональности $k = p + (1+p^2)^{1/2}$, где $p = \langle y^2 - x^2 \rangle / 2\langle xy \rangle$, а качество регрессии, измеряемое отношением разбросов точек вдоль σ_{\parallel} и поперёк σ_{\perp} прямой, задается формулой $(\sigma_{\parallel}/\sigma_{\perp})^2 = \langle (k \cdot y + x)^2 \rangle / \langle (k \cdot x - y)^2 \rangle$.

Там, где достигается максимум отношения $\sigma_{\parallel}/\sigma_{\perp}$, и проходит граница между достоверными и подозрительными участками (см. рис. 3). Классификация участков как подозрительных означает отнюдь не то, что результаты на всех них фальсифицированы, а лишь то, что среди них, скорее всего, уже есть участки, затронутые фальсификациями, которые, однако, еще не коснулись достоверных участков.

Оптимизации подобного рода могут быть весьма чувствительны к объему рассматриваемой выборки. Чтобы не устанавливать искусственно минимальное число достоверных участков, предъявим к регрессии дополнительное требование. Подозрительная часть зависимости голосов «за» от голосов «против» должна целиком лежать над регрессионной прямой. А если какие-то из точек опускаются ниже нее, это значит, что из-за флуктуаций на скудной выборке достоверных участков случайно была получена прямая со слишком высоким значением k . В этом случае найденный максимум отношения считается недействительным и поиск границы продолжается дальше по величине явки – до тех пор, пока прямая не окажется ниже всех точек, не используемых для ее построения.

6. Реконструкция результатов участников выборов

На рис. 4 слева представлены графики для рассматриваемых выборов федерального уровня, результаты которых наверняка фальсифицированы,

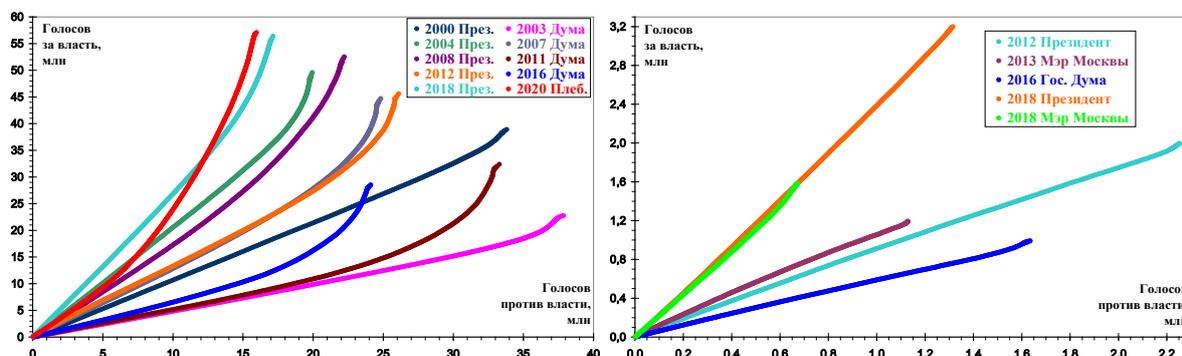


Рис. 4. Интегральные графики голосов за и против кандидата/партии власти при растущей общей явке для России и Москвы

а справа – для выборов, проходивших в Москве в 2012-18 гг., когда еще сказывался эффект протестов, вызванных массовыми фальсификациями на выборах 2011 г. и выборы в столице проводились более-менее честно. Как можно видеть, в первом случае графики сильно загибаются вверх, а во втором идут почти прямо. Здесь несущественным оказывается даже то, что разнообразие жителей Москвы существенно меньше разнообразия жителей России. Паразитные корреляции результата власти с явкой, которые могли бы нарушить линейный вид графиков, практически отсутствуют.

В отличие от дифференциального метода, здесь угловой коэффициент наилучшей регрессионной прямой k определяется с высокой точностью. Поэтому сразу же можно реконструировать истинный результат власти как отношение $k/(1+k)$. Он в сравнении с официальным результатом приведен на рис. 5, дающем представление о масштабах фальсификаций на различных выборах.

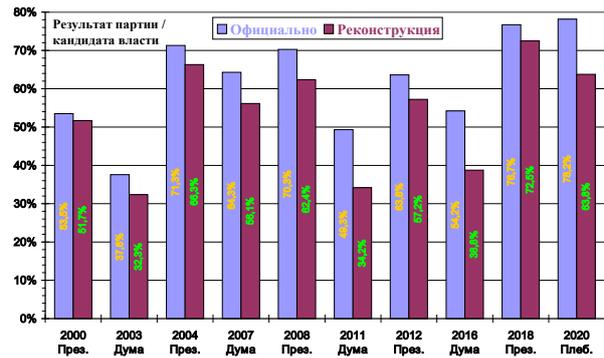


Рис. 5. Результат власти

7. Реконструкция явок

Как и в случае дифференциального метода, одного этапа обработки данных оказывается недостаточно для полной реконструкции. Нужен второй этап, чтобы отделить вброс голосов от их переброса. Но в отличие от дифференциального метода, для интегрального второй этап выполняется на той же технической основе, что и первый, однако с обращением базовой гипотезы, постулирующей теперь независимость уже общей явки от результата власти, а не результата от явки.

Для участков с результатом партии/кандидата власти, не превышающим данного, рассматривается зависимость числа избирателей, пришед-

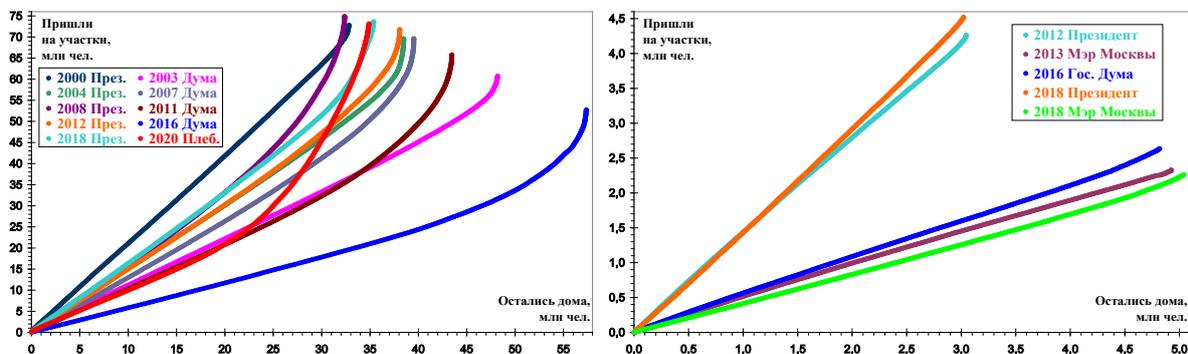


Рис. 6. Интегральные графики участия и абсентеизма избирателей при растущем результате кандидата/партии власти для России и Москвы

6. Математические модели цифрового мира

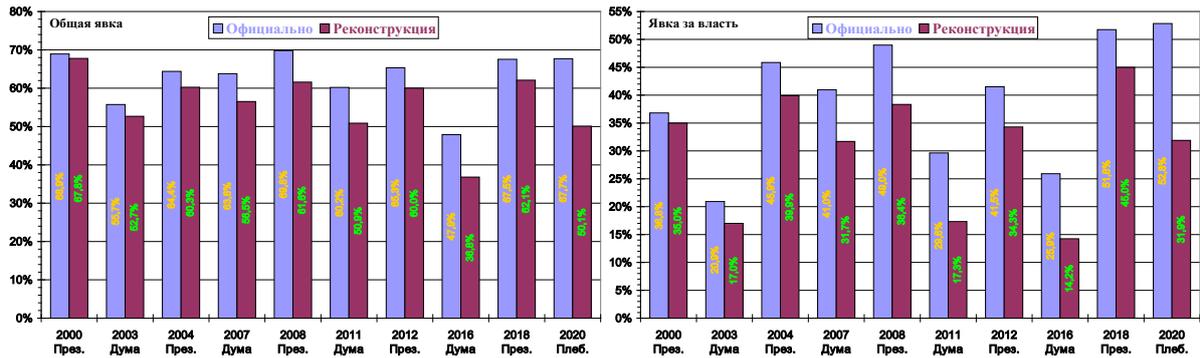


Рис. 7. Явки – общая и в поддержку кандидата/партии власти

ших на участки, от числа избирателей, оставшихся дома. Иначе говоря, участки сортируются по результату власти, и по мере его возрастания суммируются с накоплением количества избирателей, принявших участие в выборах и уклонившихся от него. Как и в прошлый раз, зависимость между получаемыми суммами является прямой пропорциональностью, пока фальсификаций нет, и начинает отклоняться от нее при их появлении. Соответствующие графики, полностью аналогичные рис. 4, даны на рис. 6.

Дальнейший анализ также полностью повторяет примененный ранее. На основе углового коэффициента k линейных частей графиков реконструируется истинная общая явка как $k/(1+k)$, из которой легко находится и истинная явка за власть. Их официальные и реконструированные значения приведены на рис. 7, вновь позволяющем убедиться в заметном завышении и общей, и провластной активности избирателей.

Рис. 5 и 7, позволяющие сравнивать фальсифицированные и истинные результаты выборов, не очень удобны с точки зрения изучения влияния фальсификаций на их исход. Более наглядна разность и отношение рассматриваемых величин, представленные на рис. 8.

На левом рисунке интереснее всего ряд данных по результату власти, максимизация которого – основная цель фальсификаций. С их помощью в избирательных циклах 2003/04 и 2007/08 гг. он увеличивался примерно на

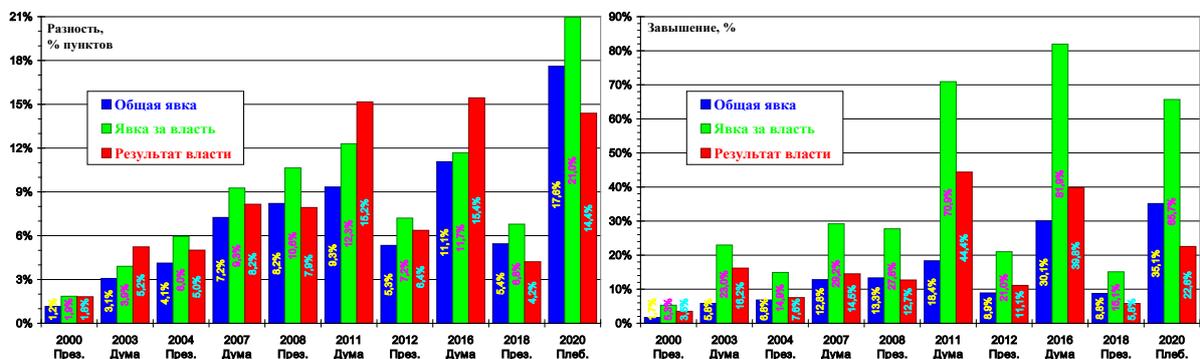


Рис. 8. Превышение официальными результатами реконструированных и завышение вторых относительно первых

одну и ту же, сравнительно небольшую, величину. Однако в 2010-е гг. ситуация с превышением официальным результатом власти его истинных значений изменилась. Оно даже немного уменьшилось на менее конкурентных, но характеризующихся высокой явкой, президентских выборах, зато существенно выросло на более конкурентных, но зато с меньшей явкой, парламентских.

В этом смысле плебисцит 2020 г. оказывается более похож на парламентские, а не на президентские выборы, хотя в силу своей политической роли (единственный победитель и прямая ассоциация с инкубентом) он должен был бы более походить именно на вторые, а не на первые. Плебисцит так же был отмечен рекордной разностью официальной и реконструированной явок за власть, что позволило перетащить эту величину через психологически важную отметку в 50% (см. рис. 7).

В случае правого рисунка основное внимание уместно сосредоточить на относительном изменении явок, прямо пропорциональных активности избирателей – общей и провластной. Роль фальсификаций в случае парламентских выборов и плебисцита здесь оказывается еще более очевидной. При этом особенно наглядны становятся как тот гигантский скачок влияния фальсификаций на состав депутатского корпуса, который произошел на рубеже нулевых/десятих годов, так и парламентский размах фальсификаций на плебисците.

8. Политологический анализ

Наличие двух временных рядов – фальсифицированных и истинных результатов выборов – дает возможность соотнести изменения, происшедшие в легальной части российской политической системы, с развитием института электоральных фальсификаций. Как можно видеть из рис. 9, имеется общий тренд к их нарастанию, из которого, однако, выпадают выборы середины 2010-х гг. Вероятно, в это время распространение фальсификаций уперлось в ограничения, налагаемые традиционной системой голосования. И лишь их преодоление на плебисците 2020 г., связанное с вне-

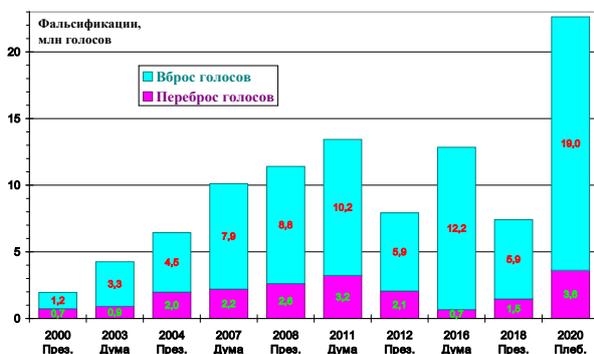


Рис. 9. Объемы фальсификаций

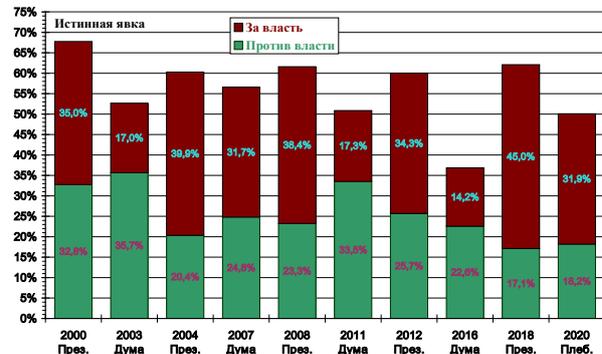


Рис. 10. Активность избирателей

дрением многодневного и электронного голосования, а также с ограниче-

6. Математические модели цифрового мира

нием наблюдения и агитации против власти, позволило вернуться на прежнюю траекторию роста. Тем не менее, насыщение возможностей по организации фальсификаций не объясняет некоторого сокращения их объемов в 2012-18 гг. Нельзя его отнести и исключительно на счет протестов зимы 2011-12 гг., т.к. они привели к отказу от масштабных фальсификаций только в Москве, где их объем в 2008 и 2011 гг. не превышал 1 млн голосов (рассчитано с помощью интегрального метода).

Отмеченная странность становится более понятной, если учесть, что для дряхлеющей российской политической системы со временем становятся всё более характерны не фальсификации выживания, решающие задачу захвата или удержания власти, а фальсификации устрашения, демонстрирующие убедительность победы. В случае президентских и парламентских выборов, различающихся как по общей, так и по провластной активности избирателей (см. рис. 10), эта задача решается разными средствами.

На всех президентских выборах XXI в. победа кандидата власти априори не вызывала сомнений. Тем не менее, в 2000 г., несмотря на принятые экстраординарные меры (начало второй чеченской войны и досрочное проведение выборов), дело едва не дошло до второго тура. Поэтому на протяжении 2000-х гг. параллельно с объективным повышением популярности власти, связанным с улучшением внешнеэкономической конъюнктуры, происходило наращивание фальсификаций. И если их объем достиг локального максимума лишь в 2011 г. (см. рис. 9), то распространенность была наибольшей именно на выборах 2008 г. [2–6,8,10,11], когда пост главы государства был передан местоблюстителю.

Дальнейшая демонстрация подавляющего превосходства кандидата власти над конкурентами уже не могла быть обеспечена посредством одних лишь фальсификаций и потому осуществлялась, в первую очередь, неэлекторальными средствами. Выборы 2012 г. представляли собой просто возврат власти от местоблюстителя и, видимо, поэтому им позволено было стать относительно конкурентными, в результате чего явка за кандидата власти оказалась наименьшей среди всех рассматриваемых президентских выборов. Надо полагать, из этого были сделаны определенные выводы, так что в 2018 г. конкурентность выборов была резко сокращена вплоть до недопуска к участию в них основного оппозиционного кандидата. Это позволило создать в обществе ощущение полной безальтернативности инкумбента и добиться максимальной явки за него, которую с помощью сравнительно небольших фальсификаций впервые удалось перетащить через психологически важный рубеж в 50% (см. рис. 7 справа). Заметим, что для повторения этого достижения на плебисците 2020 г. потребовались многократно большие фальсификации.

Более сложной оказывается ситуация с выборами парламента, способ формирования нижней палаты которого на рассматриваемом временном интервале менялся дважды. В 2007 и 2011 гг. она целиком избиралась по

пропорциональной системе с 7%-порогом. А до и после того – по смешанной: половина по пропорциональной системе с 5%-порогом, а половина по мажоритарной системе относительного большинства (в один тур).

Зачем нужен был такой зигзаг избирательного законодательства? Дело в том, что мажоритарная система при умеренных фальсификациях дает определенные преимущества оппозиционным партиям, которые за счет разнообразия могут в округах проводить в парламент своих представителей. Но при высоких фальсификациях решающее преимущество получает уже партия власти, так что даже самый активный и харизматичный оппозиционер практически не имеет шансов обойти ее ставленников [1]. У нас в стране эту проблему сейчас пытаются преодолеть с помощью т.н. умного голосования (неформальных праймериз для оппозиционных кандидатов), а в демократических странах она решается либо за счет системы абсолютного большинства (со вторым туром выборов), либо за счет устоявшейся двухпартийной системы (тогда первый тур – сразу и второй).

Отмеченную особенность мажоритарной системы наглядно демонстрирует история отечественных парламентских выборов. В 1999 г., когда две ветви партии власти еще боролись в т.ч. и друг с другом, они в сумме получили всего 139 мест в парламенте, из которых на одномандатные округа пришлось лишь 38 (27%). В 2003 г. за счет пока еще умеренных фальсификаций и неформального ограничения независимого финансирования оппозиции уже объединенная партия власти берет 223 места, из которых 103 (46%) – в одномандатных округах. А рекордные вбросы 2016 г. позволили ей поднять эти показатели аж до 343 и 203 (59%).

Вполне естественно, что партия власти, добившаяся победы на выборах 2003 г. и собравшая конституционное большинство посредством включения в свою фракцию многочисленных перебежчиков и квазинезависимых кандидатов, решила закрепить победу. Для этого был не только осуществлен переход от смешанной системы к пропорциональной, но одновременно поднят порог прохождения партий в парламент и запрещены предвыборные блоки, что позволило на несколько легислатур зафиксировать сложившуюся систему из четырех парламентских партий. На выборах 2007 г. партия власти уже в одиночку сумела взять конституционное большинство (315 мест). Тем не менее, дальнейшее падение ее поддержки более не могло быть скомпенсировано приближающимися к насыщению фальсификациями и в 2011 г. она откатывается от конституционного большинства до обычного (238 мест). Зато отработанная к этому моменту техника массовых фальсификаций и перенос предвыборной кампании с осени на лето позволили власти в дальнейшем вернуться к смешанной системе, установив на выборах 2016 г. новый рекорд представительства.

Выборы 2016 г. экстремальны и еще по одному показателю. На них доля переброшенных голосов в общей структуре фальсификаций упала почти до 5%. Даже для плебисцита эта величина, в целом испытывающая

6. Математические модели цифрового мира

тенденцию к снижению, составляла 16%, а для более ранних выборов лежала в диапазоне 20÷37%. Иначе говоря, преимущество партии власти, обеспеченное сочетанием развитого института фальсификаций и выстроенной под него избирательной системы, оказалось столь велико, что такая сложная процедура как воровство голосов у других партий отчасти утратила актуальность и применилась лишь эпизодически.

9. Выводы

Показано, что базовая гипотеза о взаимной независимости пары электоральных характеристик позволяет успешно реконструировать их истинные значения, если известно, в какую сторону те смещаются вследствие фальсификаций.

Реконструкция связана с проведением границы между достоверными и подозрительными избирательными участками в их ряду, отсортированном по величине одной из электоральных характеристик. Данная операция основывается на выделении линейной части зависимости между формирующими вторую характеристику количествами избирателей, суммируемыми при продвижении по отсортированному ряду участков. Так, если сортировка выполнена по общей явке, то это будут голоса за и против партии/кандидата власти, а если по ее результату, то – объемы участия и абсентеизма. Угловой коэффициент линейной регрессии Деминга без свободного члена, имеющей наибольшее соотношение разбросов вдоль и поперек прямой, реконструирует соотношение между рассматриваемыми величинами, а с ним – и истинное значение второй характеристики. Дополнительным условием устанавливается нахождение всей части зависимости, соответствующей подозрительным участкам, по одну сторону от прямой, рассчитанной на основе достоверных участков.

Проверить справедливость базовой гипотезы в российских условиях позволяют выборы, подвергавшиеся фальсификациям в минимальной степени. Таковы московские выборы середины 2010-х гг. и отчасти федеральные выборы начала 2000-х гг. Для них рассматриваемые зависимости оказываются практически линейными, как и предсказывает гипотеза.

Сопоставление изменений, происходивших как в предвыборной тактике власти, так и в избирательном законодательстве, показывает их тесную связь с развитием института электоральных фальсификаций в нашей стране. Для достижения желаемых результатов на выборах власть ведет активный и целенаправленный поиск решений как в легальном, так и в нелегальном поле, не делая принципиальных различий между первым и вторым. Иначе говоря, фальсификации выборов в России уже давно следует рассматривать не как порчу избирательной системы, а как ее неотъемлемую часть, т.е. фальсификации не обязательно умысел, но, возможно, и результат самоорганизации.

10. Приложение

Последние конкурентные выборы Президента России состоялись в 1996 г., когда информационные технологии были еще недостаточно развиты для сбора и обнародования результатов по каждому избирательному участку. Данные детализированы лишь до уровня территориальных избирательных комиссий, а их всего 2,7 тыс. Этого недостаточно для надежного применения дифференциального метода реконструкции, но интегральный остается применим, хотя тяготеет к некоторому завышению объемов фальсификаций при одновременном занижении числа переброшенных голосов.

Дискуссии о чистоте тех выборов не утихают и четверть века спустя. Их ключевая особенность – даже не реальная борьба действующего Президента (Б.Н. Ельцин) с сильнейшим из оппозиционных политиков (Г.А. Зюганов), а указания на фальсификации в пользу как первого, так и второго. Возможность противонаправленных фальсификаций осложняет использование метода, предполагающего лишь одного их бенефициара.

Если по очереди рассматривать каждого из упомянутых кандидатов как кандидата власти, то по аналогии с рис. 4 получаем рис. 11а и b для I и II туров выборов. Как можно видеть, для инкумбента графики загибаются не вверх, а вниз, т.е. если фальсификации в его пользу и имели место, то они потерялись на фоне фальсификаций в пользу его основного оппонента. Обработка такого перевернутого графика предполагает, что регрессионная прямая проходит уже не ниже, а выше всех точек, не используемых для ее построения. В остальном процедура остается прежней.

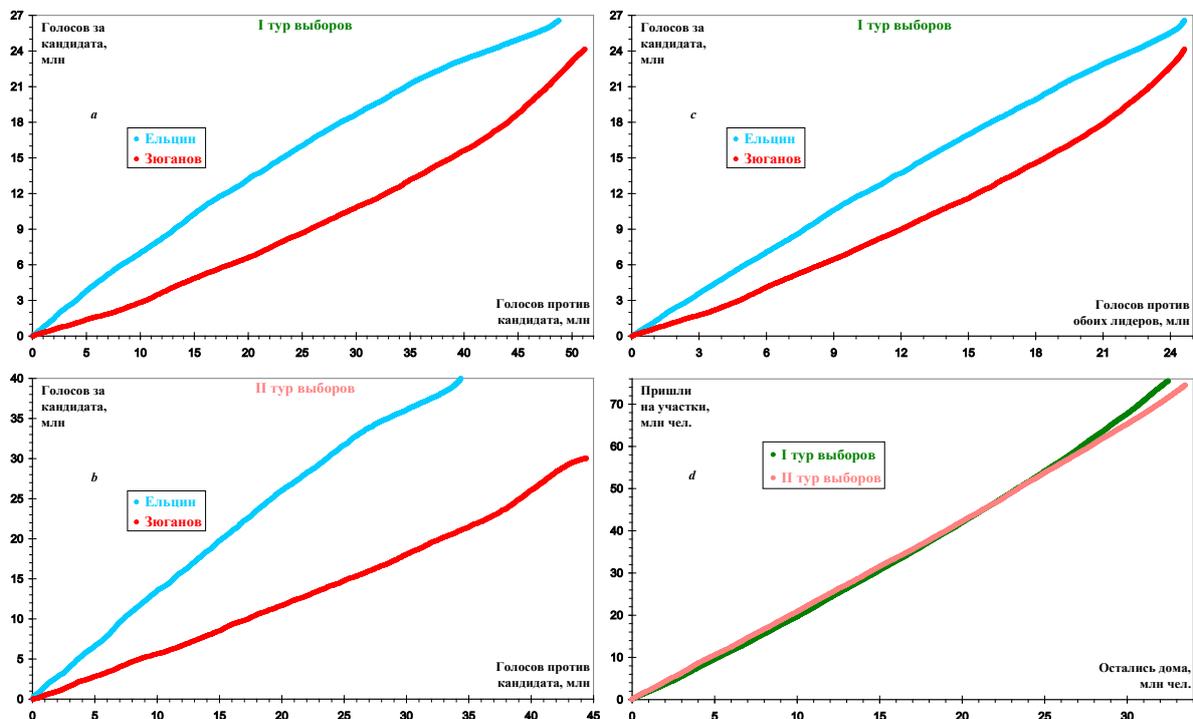


Рис. 11. Интегральные графики для президентских выборов 1996 г.

6. Математические модели цифрового мира

Итог расчетов представлен в таблице. Из-за фальсификаций разрыв между результатами инкумбента и претендента в I туре сократился на 17,8% пункта, а во II – на 6,4% пункта. Более заметно изменение отношения результатов – в 1,83 и в 1,15 раза соответственно.

Результаты выборов 1996 г.

Кандидат	Б.Н. Ельцин		Г.А. Зюганов	
	I	II	I	II
Тур выборов	I	II	I	II
Официально	35,3%	53,8%	32,1%	40,3%
Реконструкция	42,8%	56,8%	21,2%	36,9%
Проверка	42,4%	—	21,4%	—

Во II туре фальсификации уменьшились по сравнению с I. Это кажется нелогичным, т.к., скорее, следовало ожидать ожесточения борьбы, в т.ч. и незаконными средствами. Однако действия жуликов определяются не идейными соображениями, а выгодой. Выгоднее же подтасовывать результаты в пользу предполагаемого победителя. Поэтому, когда вопреки неблагоприятной предвыборной социологии, лидером по итогам I тура оказался инкумбент, фальсификаторы переметнулись на его сторону. И их лояльность изменилась столь сильно, что в 6 субъектах федерации (Дагестан, Карачаево-Черкесия, Сев. Осетия, Ингушетия, Татарстан и Кабардино-Балкария) официальная явка за оппонента власти во II туре стала даже меньше, чем в I, что невероятно.

Встает вопрос, корректно ли использовать в качестве опорной переменной голоса против одного из кандидата, включающие голоса и за второго, которые тоже могут быть фальшивыми. В случае I тура выборов, в котором участвовало множество политиков, это можно проверить, для чего по оси абсцисс откладываются голоса, поданные против лидеров, как сделано на рис. 11с. При этом, если k_i – наклон регрессионной прямой для i -го участника выборов, то его результат равен $k_i/(\sum_j k_j + 1)$. Так реконструированные результаты I тура представлены в последней строке таблицы. Они лишь несколько десятых % пункта отличаются от значений, полученных более общим способом, что позволяет считать его применимым и ко II туру, в котором уже нет других участников.

Реконструкция явки и связанных с ней величин, оказывается наиболее сложным этапом анализа. При построении графиков на рис. 6 данные сортировались просто по возрастанию результата власти. Здесь же, чтобы учесть возможные подтасовки в пользу любого из кандидатов, сортировка выполняется по сумме их результатов, отнормированных на свои истинные значения по стране в целом. Полученные графики, представленные на рис. 11d, линейризуются от I тура ко II. Это позволяет предположить, что произошла не столько компенсация фальсификациями в пользу инкумбента фальсификаций в пользу его оппонента, сколько их общее сокращение.

Объем фальсификаций составил 9,1 млн голосов в I туре (больше на президентских выборах было лишь в 2008 г.) и лишь 3,0 млн (это немного, но всё же больше, чем в 2000 г.) – во II. При этом уменьшился как вброс голосов (с 4,2 млн до 1,4 млн), так и их переброс (с 4,9 млн до 1,7 млн).

Работа выполнена при поддержке РФФИ (проект 20-511-00003).

Литература

1. Подлазов А.В. Выборы депутатов Государственной Думы VII созыва: Выявление фальсификации результатов и их реконструкция // [Социологические исследования. 2018, 405\(1\), 59-72.](#)
2. Подлазов А.В. Формальные методы выявления масштабных электро-ральных фальсификаций на материале федеральных выборов 1999-2018 гг. // [Препринт ИПМ им. М.В. Келдыша РАН. 2019, №2.](#)
3. Подлазов А.В. Исследование статистических методов выявления выдуманных результатов выборов: Часть 1. Круглые числа // [Препринт ИПМ им. М.В. Келдыша РАН. 2019, №147.](#)
4. Подлазов А.В. Исследование статистических методов выявления выдуманных результатов выборов. Часть 2. Стустки явки и результатов // [Препринт ИПМ им. М.В. Келдыша РАН. 2020, №20.](#)
5. Подлазов А.В. Формальные методы выявления масштабных электро-ральных фальсификаций // [Проектирование будущего. Проблемы цифровой реальности. 2019. Вып.2, 120-137.](#)
6. Подлазов А.В. Формальное выявление выдуманных результатов выборов // [Проектирование будущего. Проблемы цифровой реальности. 2020. Вып.3, 176-190.](#)
7. Kobak D., Shpilkin S., Pshenichnikov M.S. Statistical anomalies in 2011–2012 Russian elections revealed by 2D correlation analysis. [arXiv:1205.0741](#)
8. Kobak D., Shpilkin S., Pshenichnikov M.S. Integer percentages as electoral falsification fingerprints // [Ann. Appl. Stat. 2016, 10\(1\), 54-73.](#)
9. Kobak D., Shpilkin S., Pshenichnikov M.S. Statistical fingerprints of electoral fraud? // [Significance. 2016, 13\(4\), 20-23.](#)
10. Kobak D., Shpilkin S., Pshenichnikov M.S. Putin’s peaks: Russian election data revisited // [Significance. 2018, 8\(3\), 8-9.](#)
11. Kobak D., Shpilkin S., Pshenichnikov M.S. Suspect peaks in Russia’s “referendum” results // [Significance. 2020, 17\(5\), 8-9.](#)
12. Шпилькин С. Статистическое исследование результатов российских выборов 2007–2009 гг. // [Троицкий вариант – наука. 2009, 21\(40\), 2-4.](#)
13. Шпилькин С. Математика выборов – 2011 // [Троицкий вариант – наука. 2011, 25\(94\), 2-4.](#)
14. Шпилькин С. Двугорбая Россия // [Троицкий вариант – наука. 2016, 20\(214\), 1-3.](#)
15. Шпилькин С. Выборы 2018 года: Фактор X и «пила Чурова» // [Троицкий вариант – наука. 2018, 8\(252\), 8-10.](#)
16. Шпилькин С. Поправки на 27 миллионов // [Троицкий вариант – наука. 2020, 14\(308\), 4-5.](#)