



ПОЛИТИКА

И.А.Скулкин

МОДЕЛИРОВАНИЕ ЭЛЕКТОРАЛЬНЫХ ФАЛЬСИФИКАЦИЙ

Ключевые слова: выборы, электоральные фальсификации, состязательный авторитаризм

Постановка проблемы

Победа инкубентов с большим отрывом от ближайших соперников является нормой для состязательных авторитарных режимов. Как правило, данный феномен объясняют клиентелистской агитацией¹, покупкой голосов избирателей², контролируемым голосованием³. При этом принято считать, что электоральные фальсификации не играют ключевой роли в поддержании подобных режимов и используются лишь в качестве временного или дополнительного инструмента при истощении ресурсов на проведение политики патронажа⁴. Вместе с тем вклад фальсификаций в устойчивость состязательных авторитарных режимов до сих пор не измерен количественно. Фальсификации выступают в качестве некоей латентной переменной, которая подразумевается по умолчанию, хотя никто не знает ее реального значения.

Наиболее конвенциональный метод внесения поправок на электоральные нарушения — включение в анализ различного рода индексов демократичности, прежде всего «Freedom in the World» или «Polity IV». Однако указанные индексы строятся на экспертных оценках, что неизбежно накладывает на них отпечаток субъективизма, и измеряют не фальсификации как таковые, но лишь степень искажения электорального пространства в целом. Иными словами, эти индексы не позволяют ответить на вопрос, являются ли высокие доли голосов, обычно получаемые инкубентами при состязательных авторитарных режимах, следствием уничтожения жизнеспособной оппозиции, запугивания и подкупа избирателей, или же они представляют собой результат банальной фальсификации.

Для оценки масштабов фальсификаций используются также опросы общественного мнения. Но надежность содержащейся в них информации о предпочтениях избирателей вызывает большие сомнения. Сфальсифицировать опросные данные даже легче, чем официальные итоги выборов, а значит, опирающийся на них исследователь не может быть уверен в адекватности своих выводов.

Едва ли способны помочь в выявлении масштабов фальсификаций и отчеты наблюдателей. Фиксируя отдельные факты нарушений в ходе избирательного процесса, такие отчеты ничего не говорят о том, какой эффект оказало каждое нарушение на результаты выборов. Еще более сложная ситуация возникает в тех случаях, когда наблюдатели

¹ Wantchekon 2003.

² Magaloni 2006: ch. 4.

³ Kitschelt, Wilkinson 2007.

⁴ Magaloni 2006: 21–23; Greene 2007: 34.

просто не допускаются на избирательные участки или не присутствуют при подсчете бюллетеней. Поэтому сами по себе отчеты наблюдателей не позволяют установить, сколько голосов было на самом деле получено инкумбентами.

Еще менее эффективны в этом плане косвенные индикаторы отношения к авторитарным лидерам, в частности количество антиправительственных выступлений. Во-первых, данный показатель не учитывает превентивную силу режима. Поскольку по мере повышения репрессивности режима растут и издержки от участия в протестных акциях, очевидно, что в режимах с наибольшим репрессивным потенциалом подобные акции будут относительно редким явлением. Во-вторых, ввиду имманентно присущего ей перекоса в сторону протестно настроенных групп рассматриваемая переменная неизбежно будет преувеличивать степень отторжения авторитарных инкумбентов обществом.

Немало вопросов у исследователей состязательных авторитарных режимов вызывает и такая переменная, как явка на выборы. Согласно теории рационального выбора⁵, вероятность участия в голосовании зависит от весомости голоса, то есть от уровня неопределенности результатов, а значит — от веры в справедливость выборов⁶ и остроты конкуренции. И действительно, в развитых демократических странах при сокращении разрыва между кандидатами явка увеличивается⁷. Тем не менее при состязательных авторитарных режимах, где огромное преимущество инкумбента не вызывает сомнений, нередко фиксируется высокая явка на выборы. Голосование по принципу «наименьшего зла»⁸ не в состоянии объяснить этот феномен, поскольку перспектива победы оппозиционных сил выглядит по меньшей мере призрачной. Едва ли может служить ему объяснением и желание поддержать институты демократии или право выбора. Если в условиях состязательного авторитаризма чувство гражданского долга и мотивирует кого-то прийти на избирательные участки⁹, то только идейных сторонников оппозиции, готовых нести издержки участия, несмотря на заведомую электоральную неэффективность своих действий. Не очень убедительными выглядят и объяснения, связывающие высокую явку на выборы при состязательных авторитарных режимах с давлением на избирателей или их подкупом. Что касается подкупа, то, когда речь идет о большинстве избирателей, подобная стратегия становится чересчур затратной. Давление на избирателей тоже имеет свои ограничения, ибо широкое его применение подрывает веру в легитимность выборов. И здесь мы снова сталкиваемся с вопросом о возможных фальсификациях.

Так какую же роль играют электоральные фальсификации в поддержании авторитарных режимов? Для того чтобы ответить на этот вопрос, прежде всего необходимы надежные методы диагностики подобного рода фальсификаций, позволяющие оценить их объем. В настоящей статье я анализирую эффективность существующих на данный момент методов такой диагностики и попытаюсь обосновать альтернативный подход к решению этой задачи.

⁵ Downs 1957.

⁶ Birch 2010.

⁷ Franklin 2004.

⁸ Ferejohn, Fiorina 1974.

⁹ Riker, Ordeshook 1968; Blais 2000.

**Методы
диагностики
электоральных
фальсификаций**

¹⁰ Бузин 2008.

¹¹ См., напр.
Lukinova, Maygov,
Ordeshook 2011.

¹² Myagkov,
Ordeshook, Shakin
2008

¹³ Levin et al. 2009.

¹⁴ См., напр. Beber,
Scacco 2008;
Mebane 2008;
Мебейн, Калинин
2009.

Существует класс методов, при которых в качестве предсказанных значений истинных результатов выборов используются предположительно несфальсифицированные результаты голосования на определенной территории, на отдельных участках или на сопоставимых предшествующих выборах. В частности, речь идет о сравнении данных, полученных на обычных участках и на участках, оснащенных комплексами обработки избирательных бюллетеней, или о сравнении статистики открепительных удостоверений и выноса избирательных бюллетеней¹⁰. При анализе российских выборов широко практикуется также разбиение выборки на области и республики¹¹ (считается, что для республик характерен более высокий уровень фальсификаций). Однако такой метод позволяет оценить лишь небольшую часть фальсификаций и не может быть серьезным аналитическим инструментом.

Более сложной разновидностью методов этого класса является сравнение результатов двух следующих друг за другом выборов одного уровня. Отталкиваясь от предположения о том, что если при изменении предпочтений избирателей доля голосов, переходящих к кандидату или партии от любого из их конкурентов, не может быть меньше 0 и больше 100%, то в случае фальсификаций это правило нарушается, М.Мягков, П.Ордушук и Д.Сахин разработали так называемый «экологический метод», который был использован ими для анализа выборов 2003—2005 гг. в Московскую городскую думу и Городской совет Новосибирска¹². Сходная методика использовалась также при анализе референдумов 2007—2009 гг. в Венесуэле¹³, однако во всех трех случаях были выявлены лишь очень небольшие объемы подозрительных голосов. Главный недостаток указанного метода состоит в том, что в его рамках предшествующие выборы рассматриваются как эталонные и не подверженные фальсификациям. Кроме того, измерение перетока голосов от выборов к выборам возможно лишь при неизменности территориальных границ избирательных округов.

Для диагностики электоральных фальсификаций применяется также метод, основанный на законе Бенфорда¹⁴. Согласно этому закону, предсказанные вероятности имеют форму инверсной (перевернутой) логарифмической функции и для первой стоящей по порядку цифры ($j = 1, 2, 3 \dots 9$) составляют 0,301; 0,176; 0,125; 0,097; 0,079; 0,067; 0,058; 0,051; 0,046; а для второй ($j = 0, 1, 2 \dots 9$) — 0,120; 0,114; 0,109; 0,104; 0,100; 0,097; 0,093; 0,090; 0,088; 0,085. Оценка соответствия наблюдаемых и предсказанных вероятностей появления чисел может быть осуществлена с помощью непараметрического теста Хи-квадрат (если речь идет об одном числе) или дисперсионного анализа (когда речь идет о ряде чисел). Можно также вычислять соответствие между средним числом и предсказанной вероятностью. Для первого порядкового числа она в данном случае равна 3,441; для второго — 4,187. Предполагается, что если цифры в электоральных данных реальны, они должны соответствовать предсказанным вероятностям.

В последнее время, однако, ученые стали приходить к выводу, что отклонения от распределения Бенфорда не могут трактоваться как однозначное свидетельство фальсификаций¹⁵. Во-первых, если число избирателей на участках колеблется вокруг одной цифры (а на практике это наиболее типичная ситуация), распределение первой порядковой цифры будет стремиться к распределению долей голосов, полученных кандидатом, которое обычно соответствует нормальному (куполообразному) распределению. Во-вторых, каким бы ни было происхождение данных, распределения их в соответствии с законом Бенфорда можно ожидать только в том случае, если начальным числом измерения является единица, а каждое новое наблюдение имеет линейно возрастающую направленность. Поскольку же в отношении количества голосов, получаемых на выборах, не действует ни одно, ни другое условие, нет никаких теоретических оснований полагать, что распределение цифр будет принимать инверсную логарифмическую форму.

¹⁵ *Mebane 2010; Deckert, Myagkov, Ordeshook 2011.*

Итак, закон Бенфорда для первой и второй порядковых цифр не подходит для диагностики фальсификаций. Чуть более оправданным представляется применение его к последней цифре. Так, У.Мебейн и К.Калинин¹⁶ обнаружили, что в данных о голосовании на российских общенациональных выборах 2003—2008 гг. намного чаще остальных встречаются цифры ноль и пять. Такой перекоп действительно может указывать на фальсификации. И все же применимость закона Бенфорда к последней порядковой цифре вызывает серьезные сомнения; гораздо более правдоподобным с теоретической точки зрения здесь выглядит случайное распределение, близкое к равновероятностному. В любом случае подобного рода отклонения от предсказанных вероятностей, будь то распределение Бенфорда или случайное распределение, не позволяют определить точный объем фальсификаций.

¹⁶ *Мебейн, Калинин 2009.*

Наиболее продуктивным из существующих на сегодняшний день мне кажется метод диагностики электоральных фальсификаций, разработанный А.Собяниным и В.Суховольским¹⁷. Авторы предположили, что в случае честных выборов явка и доля голосов, получаемых любым из кандидатов, являются полностью независимыми друг от друга переменными, а значит, если построить график зависимости абсолютной доли голосов, рассчитанной как отношение количества полученных голосов к количеству *зарегистрированных* избирателей (V/E), от явки (T), то в линейной функции $V/E = \alpha + \beta T$ значение α должно быть равно нулю, а значение β — официальной доле голосов, полученных кандидатом по участкам в целом. В дальнейшем этот метод неоднократно использовался, в том числе и при анализе российских выборов¹⁸, и был подвергнут некоторым уточнениям, позволившим нивелировать высказывавшиеся в его адрес критические замечания. Рассмотрим основные из них.

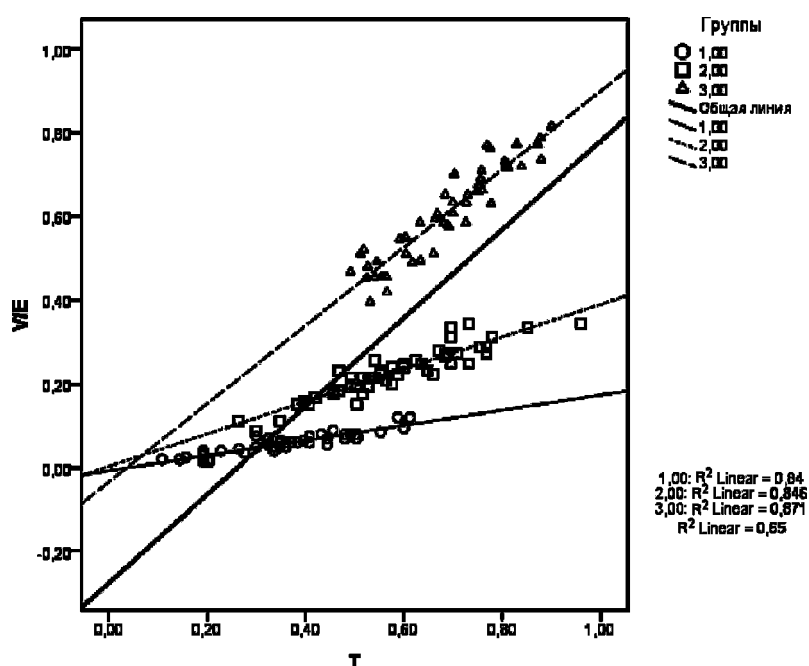
¹⁷ *Собянин, Суховольский 1995.*

¹⁸ *Myagkov, Ordeshook, Shakin 2008; Ordeshook, Myagkov 2008; Шпилькин 2009; Lukinova, Maygov, Ordeshook 2011.*

1. Важным условием корректности применения рассматриваемого метода является внутренняя однородность данных. Другими словами, внутри общего массива не должно быть групп, обладающих собствен-

ными значениями α и β . Указанный тезис проиллюстрирован на *рис. 1*, на котором симитирована ситуация, возникающая при наличии трех групп не подвергшихся фальсификации случаев, отличающихся по уровню явки и доле полученных условным кандидатом голосов. Распределение в целом создает иллюзию фальсификации: точка пересечения общей линии с осью V/E находится ниже нуля, хотя в каждой из групп значение β близко к нулю.

Рисунок 1 Явка и доля голосов по трем группам



¹⁹ Myagkov, Ordeshook, Shakin 2008: 32–42.

Включение в анализ внутренне гетерогенных данных было главным основанием для критики работы Собянина и Суховольского¹⁹. Но проблему гетерогенности можно преодолеть, если разбить агрегированные данные и проводить анализ по каждой территориальной избирательной комиссии (ТИК) в отдельности, или же посредством построения многоуровневых моделей, позволяющих оценкам α и β варьировать по ТИК.

2. Предположение о том, что явка и голосование за кандидатов являются полностью независимыми друг от друга переменными, может не соответствовать действительности, если имеется некая вмешивающаяся переменная, которая *одновременно* влияет на явку и на голосование за одного из кандидатов. Необходимо, однако, учитывать, что применительно к общенациональным выборам многие переменные, влияющие

²⁰ Jackman, Miller 1995 lais 2000.

²¹ Franklin 2004.

²² Fornos, Power, Garand 2004; Kuenzi, Lambright 2007; Dettrey, Schwindt-Bayer 2009.

²³ Wolfinger, Rosenstone 1980; Blais 2000.

²⁴ Moraski 2002; Reisinger, Moraski 2008.

²⁵ Данные любезно предоставлены Единым архивом социологических данных (<http://www.sofist.socpol.ru>).

на явку, оказываются константами. Прежде всего речь идет о таких институциональных переменных, как размер округа, избирательная система и тип выборов (парламентские или президентские)²⁰. (Проблемы, связанные с более мелкими различиями институциональной среды, легко решаются при анализе на уровне ТИК.) Сходным образом обстоит дело и с разрывом между кандидатами, уменьшение которого, как уже говорилось, способствует увеличению явки²¹: очевидно, что, оценивая весомость своего голоса, избиратели будут исходить из шансов кандидатов на победу по стране в целом, а не на конкретном избирательном участке. Что касается социально-экономических переменных (уровень урбанизации и экономическое развитие), то четко обозначенной связи между ними и явкой не обнаружено, поскольку тип связи меняется под воздействием других переменных и в различных национальных контекстах²².

Наиболее важными переменными, которые имеют положительную связь с явкой на выборы²³ и могут варьировать от участка к участку на уровне ТИК, являются возраст и образование. Вместе с тем, по заключению исследователей, доли пенсионеров и лиц с высшим образованием способны объяснить лишь незначительную часть вариации явки по регионам России²⁴. Более точное представление о влиянии этих переменных на явку дает модель логистической регрессии, построенная на основании данных всероссийского опроса, проведенного ВЦИОМ в июне 1996 г.²⁵ (см. табл. 1).

Таблица 1 Объяснение явки на президентских выборах 1996 (1 тур)

	<i>B</i>	<i>S.E.</i>	<i>Wald</i>	<i>Sig.</i>	<i>Exp(B)</i>
Пенсионер (1=да)	-0,05728	0,318201	0,032405	0,8571	0,944329
Образование (3=ниже среднего)	-0,65047	0,189133	11,8281	0,0006	0,521803
Возраст	0,042089	0,008963	22,04929	0,0000	1,042988
Пол (2=женский)	-0,0528	0,218581	0,058359	0,8091	0,948566
Экономическое положение (5=хуже)	-0,25107	0,092772	7,324097	0,0068	0,777969
Y-пересечение	2,320484	0,526735	19,40766	0,0000	10,1806
N	743				
-2 Log likelihood	592,8844				
Cox & Snell R ²	0,05755				
Nagelkerke R ²	0,099972				

Из таблицы видно, что связь между рассматриваемыми переменными носит ожидаемый характер. Однако более интересна низкая

объяснительная способность модели. Это соответствует результатам других исследований, согласно которым при переходе от минимального к максимальному значению этих переменных явка возрастает на 5—15%. Таким образом, одновременное влияние вмешивающихся переменных на явку и голосование за кандидатов можно считать несущественным. Даже если какая-то из переменных будет полностью коррелировать с голосованием за какого-либо кандидата, это окажет очень слабое воздействие на итоговые показатели.

²⁶ *Myagkov, Ordeshook, Shakin 2008; Ordeshook, Myagkov 2008; Lukinova, Maygov, Ordeshook 2011.*

3. Принято считать, что при вбросе бюллетеней распределение явки из нормального превращается в бимодальное²⁶. Распределение явки действительно будет бимодальным, если вброс произошел на части избирательных участков. Если же бюллетени вбрасывались на всех или почти всех участках, то независимо от того, был ли вброс «директивным» (примерно один и тот же процент на всех участках) или случайным, распределение явки по-прежнему останется нормальным. Исключения здесь опять же связаны с наличием особых групп в агрегированных данных.

²⁷ *Шпилькин 2009.*

Так, С.Шпилькин²⁷, основываясь на распределении явки и суммы голосов в ходе электорального цикла 2007—2008 гг., пришел к выводу, что разница между реальной и предсказанной явкой на думских и президентских выборах составила 13 и 13,77% соответственно, а разница в голосах, поданных за «Единую Россию» и за Д.Медведева, — 8,42 и 7,32%. Однако в проведенном им анализе нарушено требование внутренней однородности данных, что неизбежно должно было привести к появлению ложной взаимосвязи, которая может быть оценена как следствие фальсификации. Кроме того, в своем исследовании Шпилькин исходит из постулата, согласно которому фальсификации непременно вызывают отклонение от нормального распределения. Но что если в отдельных регионах был сфальсифицирован весь массив данных? В этом случае его оценка будет указывать лишь на часть фальсификаций.

4. Метод Собянина и Суховольского иногда критикуют за то, что с его помощью сложно отличить прямую фальсификацию результатов выборов от покупки голосов или контролируемого голосования. И действительно, если покупаются голоса только тех, кто не хотел идти на выборы, статистическая разница между покупкой голосов и вбросом бюллетеней может быть незаметна. Между тем соблазну продать свой голос могут поддаваться и те избиратели, которые изначально собирались идти на выборы и голосовать за того кандидата, за поддержку которого им были предложены материальные бонусы. Поэтому, чтобы данные выглядели как результат вброса, голосование должно быть не просто контролируемым, но и принудительным (например, голосование трудовым коллективом под угрозой увольнения). Но, как уже говорилось, такая практика подрывает легитимность выборов; к тому же широкое ее применение не осталось бы незамеченным наблюдателями.

**Оценивая
масштабы
фальсификаций:
альтернативный
подход**

²⁸ Данные взяты с сайта Независимого института выборов (<http://www.vibory.ru/elects/Total-P.zip>).

²⁹ Development Core Team 2005.

В этой части статьи представлен метод, позволяющий оценить объем фальсификаций на основании значений Y -пересечения (α) и β -коэффициентов линейных регрессионных моделей зависимости доли голосов от явки. Чтобы избежать влияния внутренней гетерогенности данных, в анализ на этом этапе будут включаться данные участковых комиссий, агрегированные до уровня ТИК. Сначала я симитирую сфальсифицированные данные, сверяя результаты с оригинальными данными президентских выборов 2008 г.²⁸, а затем создам серию линейных моделей, чтобы выяснить, как изменяются α и β в зависимости от объема фальсификаций. После этого я продемонстрирую работу предложенного метода на материале некоторых ТИК. Данные симитированы с помощью статистического пакета R²⁹ с использованием пакета Hmisc.

Исходя из сказанного выше, несфальсифицированные явка и доля голосов, рассчитанная от числа зарегистрированных избирателей, должны быть нормально распределены и иметь некую связь друг с другом, α должно равняться нулю, а β соответствовать доле голосов, полученных кандидатом по участкам в целом. В рассматриваемом случае симитировано 500 наблюдений; среднее значение для явки = 0,5; стандартное отклонение = 0,05. Переменная доли голосов (y) создается путем добавления к переменной явки (x) нормально распределенного вектора со средним значением, равным нулю, и стандартным отклонением 0,04. Взаимосвязь переменных отражена на *рис. 2*.

Полученные результаты отвечают сформулированным выше требованиям. Среднее значение $\alpha = 0,0067$; $\beta = 0,49$. Для сравнения на *рис. 3* представлена зависимость $T/E \sim T$ по одной из немногих ТИК, где, по оценке экспертов, в ходе президентских выборов 2008 г. отсутствовали фальсификации (за исключением пяти участков). Из рисунка следует, что зависимость, наблюдаемая при использовании реальных данных, в целом соответствует зависимости на симитированных данных с разницей лишь в количестве случаев.

На *рис. 4* в качестве примера комбинации вброса бюллетеней и перебрасывания голосов от одного кандидата к другому приведено распределение $V/E \sim T$ по Василеостровской ТИК Санкт-Петербурга. Из рисунка видно, что коэффициент α линии Зюганова выше среднего, то есть линия имеет отрицательный наклон. Угол наклона линии Медведева составляет около 45%, то есть β -коэффициент примерно равен единице.

Таким образом, в случае вброса бюллетеней за одного кандидата α приобретает отрицательное значение, а β -коэффициент стремится к единице. Подобное искривление объясняется тем, что функция вброса бюллетеней представляет собой простую зависимость $Y = X$, где один голос за кандидата равен одному вброшенному бюллетеню. Отрицательное значение Y -пересечения α является следствием возрастания β -коэффициента и преломления линии в месте начала фальсификации. Если же имеет место не вброс бюллетеней (или не только вброс),

Рисунок 2 Имитация зависимости доли голосов от явки
(явка и доля голосов, поданных за кандидата, – 50%)

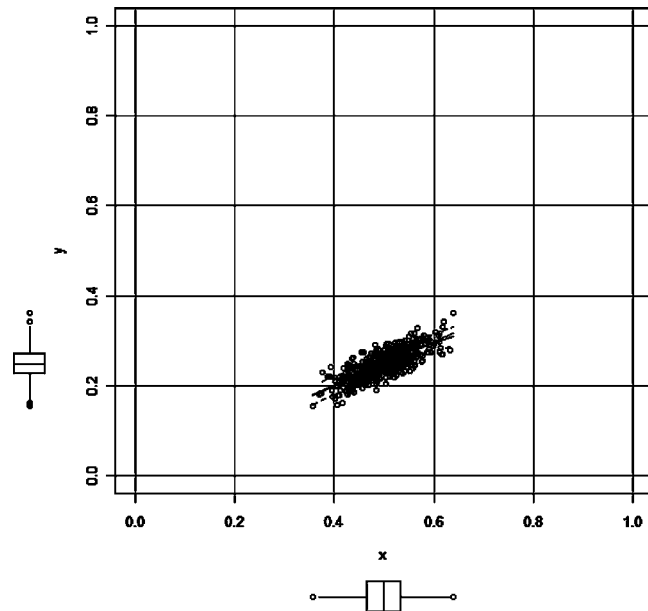


Рисунок 3 Зависимость доли голосов, поданных за кандидатов, от явки
(Владимирская область, Ковровская ТИК)

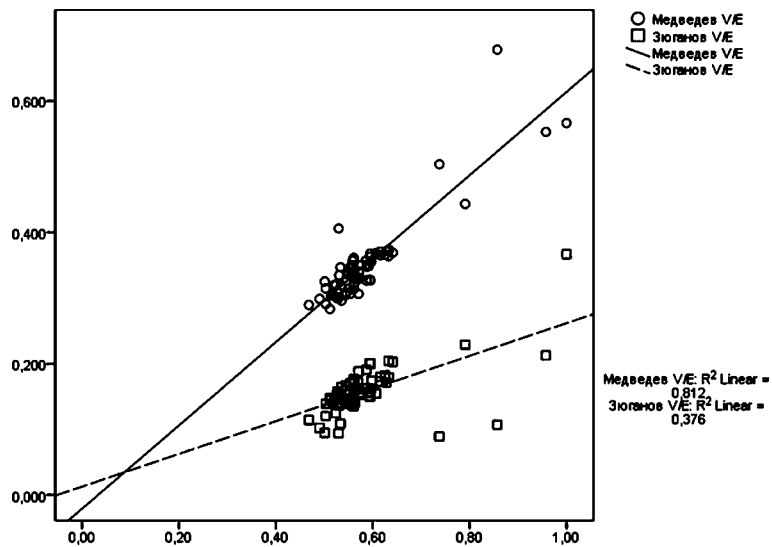
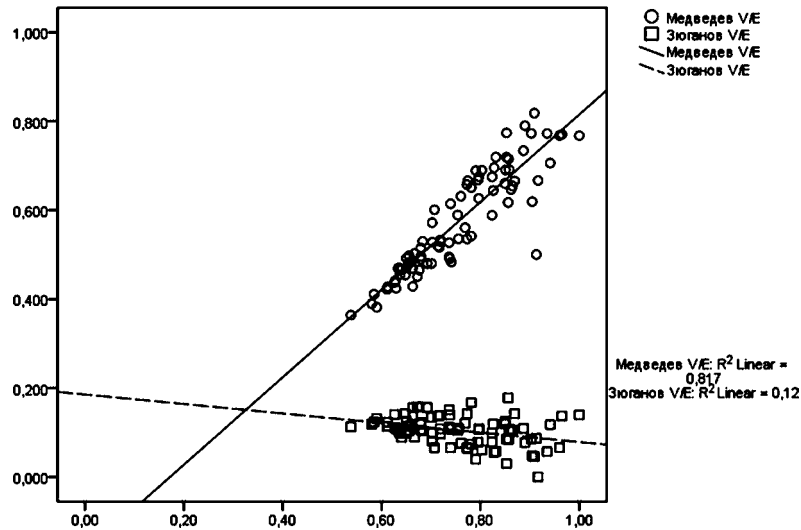


Рисунок 4 Зависимость доли голосов, поданных за кандидатов, от явки (Санкт-Петербург, Василеостровская ТИК)



но кража голосов, поданных за других кандидатов, то β может превышать единицу, а α опускаться существенно ниже нуля. В свою очередь, применительно к остальным кандидатам при вбросе β будет стремиться к нулю, а α совпадать со средним значением V/E; в ситуации же кражи голосов β будет принимать отрицательное значение, а α превышать среднее значение V/E.

На рис. 5 представлена симитированная зависимость между явкой и долей голосов, поданных за кандидата, при отсутствии фальсификаций. На рис. 6 к этим данным случайным (рандомным) способом добавлено 15% сфальсифицированных голосов. Из рисунков видно, что высказанные выше предположения относительно изменения коэффициентов α и β подтверждаются на симитированных данных. В целях экономии места графики для кандидатов от оппозиции были опущены, но тенденция очевидна: по мере вброса бюллетеней точки будут равномерно распределяться вдоль горизонтальной оси, α будет стремиться к среднему значению V/E, а β — к нулю. Нерассмотренным здесь остался и объем вброса бюллетеней, однако аналогичные графики с имитацией различных объемов фальсификаций показывают, что с увеличением количества вброшенных бюллетеней обозначенные тенденции будут усиливаться.

Рисунок 5 Имитация зависимости доли голосов от явки при отсутствии фальсификаций (явка – 40%, доля голосов, поданных за кандидата, – 50%)

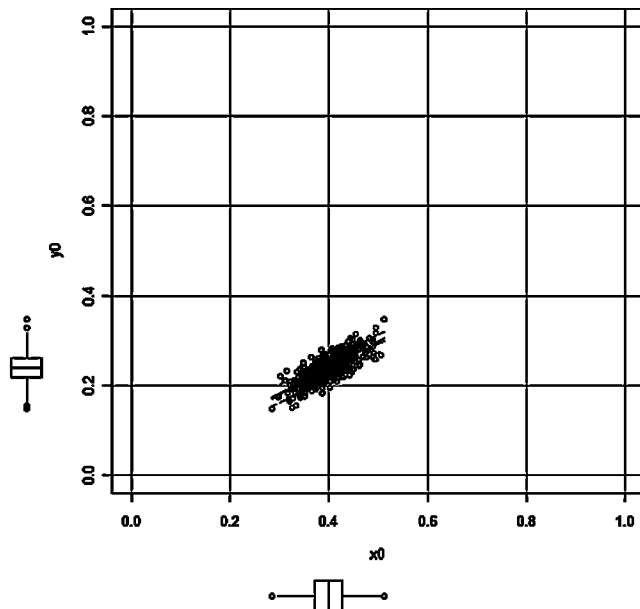
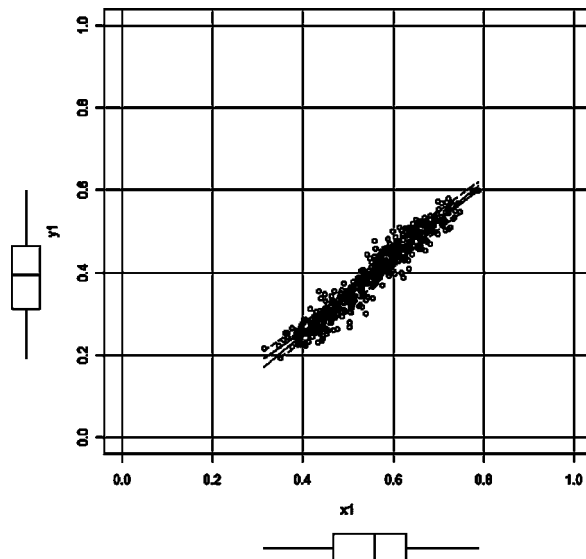


Рисунок 6 Имитация зависимости доли голосов от явки при вбросе в 15% от числа зарегистрированных избирателей (явка – 55%, доля голосов, поданных за кандидата, – 71%)



Итак, имеются три переменные, которые предположительно линейно влияют на изменение коэффициентов уравнения, — уровень первоначальной явки, доля реально полученных голосов и объем фальсификаций. Все эти переменные являются латентными, и судить о них можно только по значениям регрессионных коэффициентов. Вместе с тем эти переменные связаны между собой, и изменение одной из них приводит к изменению двух остальных. Для того чтобы установить, каким образом наблюдаемые α и β связаны с ненаблюдаемым объемом фальсификаций, я создаю взвешенную выборку из различных сочетаний явки, доли голосов и объема фальсификаций. На основе полученных сочетаний строятся линейные регрессионные уравнения, выявляющие значения α и β , а затем, после получения всех необходимых переменных, итоговая регрессионная модель с объемом фальсификаций в качестве зависимой переменной и α и β в качестве предикторов. В иллюстративных целях коэффициенты первых 19 (из 55) линейных регрессионных моделей с разными изначальными значениями явки, доли голосов и объема фальсификаций приведены в *табл. 2*.

Поскольку объем фальсификаций может объясняться не только α и β по отдельности, но и их сочетанием, на основе полученных данных создаются четыре дополнительные переменные: α opp / средняя V/E opp, α inc / средняя V/E inc, интерактивный член $\alpha \times \beta$ и отношение Y-пересечения при явке в 0% к Y-пересечению при явке в 100%: $\alpha \text{ inc} / (1 - (\alpha \text{ inc} + \beta \text{ inc}))$. В качестве зависимых переменных тестировались доля голосов, полученных инкубентом после фальсификации V/T inc, первоначальная доля голосов инкубента V/E inc, а также V/T inc, первоначальная явка T/E и доля вброшенных голосов от числа зарегистрированных избирателей. Лучше всего независимые переменные объясняют именно долю вброса, поэтому далее приводятся результаты анализа только для этой переменной (см. *табл. 3*).

Из таблицы видно, что лучше других объясняет изменение объема фальсификаций такая переменная, как соотношение Y-пересечений $\alpha \text{ inc} / (1 - (\alpha \text{ inc} + \beta \text{ inc}))$. Вместе с тем значение VIF (variance inflating factor) показывает, что объясняющие переменные, за исключением $\alpha \text{ inc} / \text{средняя V/E inc}$, сильно коррелируют друг с другом, поэтому одновременное включение в анализ хотя бы двух из них непременно вызовет проблему мультиколлинеарности. В связи с этим вместо мультивариативной модели далее будет тестироваться лучшая модель с двумя переменными (см. *табл. 4*).

Уменьшение числа переменных не сильно сказалось на объяснительной способности модели. В ней по-прежнему объясняется львиная доля вариации объема вброшенных бюллетеней. Модель имеет удобные для работы коэффициенты α и β , причем ввиду отсутствия статистической значимости коэффициент α может быть округлен до нуля. Таким образом, вброс / E = $-0,297 \times \alpha \text{ inc} / (1 - (\alpha \text{ inc} + \beta \text{ inc}))$.

Как показывает дескриптивная статистика остатков модели, взятых по модулю, при использовании рассматриваемого метода макси-

Таблица 2 Выборка коэффициентов линейных моделей с разными изначальными значениями явки, доли голосов и объема фальсификаций

$\frac{сф-й}{V/T}$ $\frac{inc}{inc}$	α_{inc}	β_{inc}	T/E	$\frac{V}{T}$ inc	$\frac{V}{E}$ inc	фальс- й/ E	$\frac{сф-й}{V/T}$ opp	α_{opp}	β_{opp}	$\frac{V}{T}$ opp	$\frac{V}{E}$ opp
0,50	-0,01	0,49	0,5	0,5	0,25	0,00	0,50	0,01	0,49	0,5	0,25
0,40	-0,08	0,64	0,3	0,3	0,09	0,05	0,60	0,08	0,37	0,7	0,21
0,48	-0,14	0,84	0,3	0,3	0,09	0,10	0,53	0,15	0,15	0,7	0,21
0,54	-0,18	0,92	0,3	0,3	0,09	0,15	0,46	0,18	0,07	0,7	0,21
0,57	-0,06	0,74	0,3	0,5	0,15	0,05	0,43	0,06	0,26	0,5	0,15
0,63	-0,10	0,88	0,3	0,5	0,15	0,10	0,38	0,11	0,11	0,5	0,15
0,67	-0,12	0,94	0,3	0,5	0,15	0,15	0,33	0,13	0,05	0,5	0,15
0,74	-0,07	0,85	0,3	0,7	0,21	0,05	0,26	0,04	0,16	0,3	0,09
0,78	-0,06	0,92	0,3	0,7	0,21	0,10	0,23	0,06	0,06	0,3	0,09
0,80	-0,07	0,96	0,3	0,7	0,21	0,15	0,20	0,08	0,03	0,3	0,09
0,52	-0,13	0,79	0,4	0,4	0,16	0,10	0,48	0,14	0,19	0,6	0,24
0,60	-0,20	0,94	0,4	0,4	0,16	0,20	0,40	0,19	0,08	0,6	0,24
0,65	-0,21	0,96	0,4	0,4	0,16	0,30	0,35	0,21	0,04	0,6	0,24
0,60	-0,11	0,82	0,4	0,5	0,20	0,10	0,40	0,12	0,16	0,5	0,20
0,67	-0,17	0,95	0,4	0,5	0,20	0,20	0,33	0,16	0,07	0,5	0,20
0,71	-0,18	0,97	0,4	0,5	0,20	0,30	0,29	0,18	0,03	0,5	0,20
0,68	-0,09	0,86	0,4	0,6	0,24	0,10	0,32	0,10	0,13	0,4	0,16
0,73	-0,14	0,96	0,4	0,6	0,24	0,20	0,27	0,13	0,05	0,4	0,16
0,77	-0,14	0,98	0,4	0,6	0,24	0,30	0,23	0,14	0,03	0,4	0,16

Примечание: V — количество голосов; E — количество зарегистрированных избирателей; T — явка; inc — инкубент; opp — оппозиция; сф-й V/T — доля голосов, полученных в результате фальсификаций, рассчитанная от явки; фальс-й/E — доля вброса по отношению к числу зарегистрированных избирателей; V/T opp — сумма голосов всех оппозиционных кандидатов.

мальная ошибка составляет 6,1% вброса / E, а средняя — 2,6% (см. табл. 5).

Рассмотрим действие предложенного метода на материале президентских выборов 2008 г. На рис. 7 представлены результаты голосования по Воскресенской ТИК Московской области. Почти горизонтальная линия голосов за Зюганова указывает на серьезные нарушения. Используя параметры полученной ранее модели, рассчитаем реальную долю голосов, поданных за Медведева ($\alpha = -0,2$; $\beta = 0,943$). Вброс/ $E = -0,297 \times (-0,2 : (1 - (-0,2 + 0,943))) = 0,231$. «За Медведева» / $T = (63689 - 115056 \times 0,23) : (87597 - 115056 \times 0,23) = 0,61$, то есть 61% вместо официальных 72%.

Таблица 3 Зависимость выброшенных бюллетеней от коэффициентов линейных моделей (линейная регрессия)

	Нестандартизованные коэффициенты		<i>t</i>	Значимость	Диагностика коллинеарности	
	β	Ст. ошибка			Tolerance	VIF
(Constant)	-0,263	0,522	-0,504	0,617		
α inc	2,134	0,842	2,5344	0,015	0,0021	478,9
β inc	0,25	0,524	0,4778	0,635	0,0014	714,6
α opp	1,73	0,631	2,7393	0,009	0,004	248,9
β opp	0,362	0,526	0,6886	0,495	0,0015	656,8
α opp / средняя V/E opp	-0,059	0,086	-0,694	0,491	0,0178	56,12
α inc / средняя V/E inc	-0,125	0,041	-3,039	0,004	0,1267	7,89
$\alpha \times \beta$	-1,039	0,453	-2,292	0,027	0,0078	129
α inc / (1 - (α inc + β inc))	-0,312	0,083	-3,757	5,00E-04	0,0159	62,99

$R^2 = 0,951$

Примечание: Зависимая переменная — доля выброшенных бюллетеней от числа зарегистрированных избирателей

Таблица 4 Зависимость выброшенных бюллетеней от α - и β -коэффициентов линейных моделей (линейная регрессия)

	β	Ст. ошибка	<i>t</i>	Значимость
Y-пересечение	0,005	0,009	0,519	0,6061
α inc / (1 - (α inc + β inc))	-0,297	0,016	-18,26	2E-24

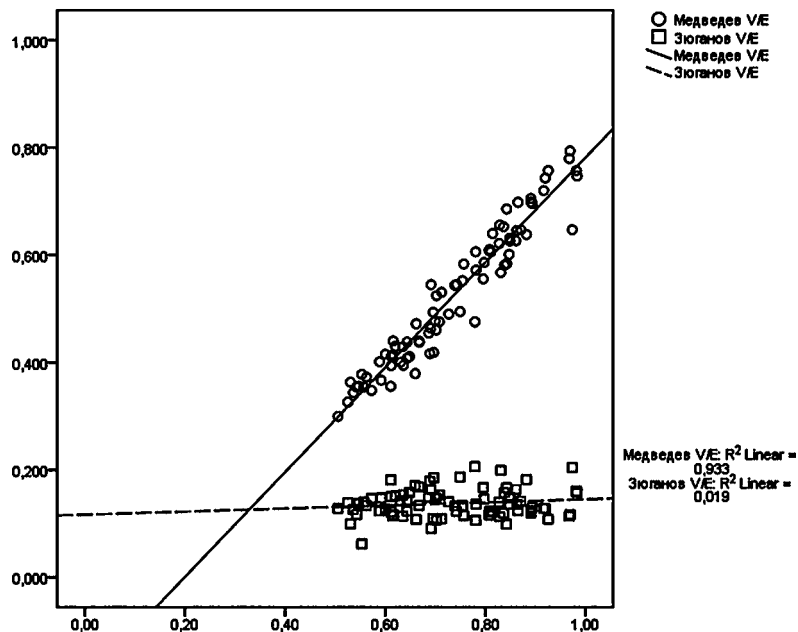
$R^2 = 0,863$

Примечание: Зависимая переменная — доля выброшенных бюллетеней от числа зарегистрированных избирателей. Остатки модели распределены нормально, гетероскедастичность и нелинейность отсутствуют.

Таблица 5 Deskриптивная статистика нестандартизованных остатков модели из таблицы 4

	<i>N</i>	Минимум	Максимум	Средняя	Ст. отклонение
Нестандартизованные остатки	55	0,001	0,0612	0,026	0,0171

Рисунок 7 Зависимость доли голосов, поданных за кандидатов, от явки (Московская область, Воскресенская ТИК)



Заключение

В отличие от большинства существующих на сегодняшний день методов диагностики электоральных фальсификаций, предлагаемый в настоящей статье метод, основанный на значениях Y-пересечений и β -коэффициентов линейных регрессионных моделей зависимости доли голосов, полученных кандидатами, от явки, позволяет относительно точно определять объем сфальсифицированных данных, а значит, и реальный уровень поддержки кандидатов, и его использование дает возможность не только продолжить исследования в области электоральных фальсификаций, но и ответить на более широкий теоретический вопрос: является ли громадный отрыв авторитарных инкубентов от ближайших соперников следствием общественной поддержки соответствующих режимов или же речь идет о банальных фальсификациях, искажающих волеизъявление избирателей?

Очевидно, что за пределами представленного выше исследования остался целый ряд важных вопросов. Применим ли предложенный метод к случаям не только вброса бюллетеней, но и кражи голосов? Позволяет ли он давать несмещенные оценки при отклонении явки от нормального распределения, корректно ли применение OLS-регрессии

или здесь требуются какие-то другие инструменты? И, наконец, какова общая схема электоральных фальсификаций в России: имеют ли они одноуровневый характер, то есть осуществляются на уровне избирательных участков, или (возможно) уже искаженные на участках данные подвергаются дальнейшему искажению на уровне региональных избиркомов или ЦИК? Поиск ответов на эти вопросы — дело дальнейших изысканий.

Библиография

- Бужин А.** 2008. *Курс аномальной электоральной статистики (рекорды и награды московских выборов)* (<http://www.votas.ru/anomstat.zip>).
- Собянин А.А., Суховольский В.Г.** 1995. *Демократия, ограниченная фальсификациями: Выборы и референдумы в России в 1991—1993 гг.* — М.
- Шпилькин С.** 2009. Статистическое исследование результатов российских выборов 2007—2009 гг. // *Троицкий вариант — Наука*. № 21.
- Мебейн У., Калинин К.** 2009. Электоральные фальсификации в России: комплексная диагностика выборов 2003—2004, 2007—2008 гг. // *Российское электоральное обозрение*. № 2.
- Beber В., Scacco A.** 2008. *What the Numbers Say: A Digit-Based Test for Election Fraud Using New Data from Nigeria*. Paper prepared for the Annual Meeting of the American Political Science Association. Boston, MA. August 28—31 (https://files.nyu.edu/bb89/public/files/Beber_Scacco_ElectionFraud.pdf).
- Birch S.** 2010. Perceptions of Electoral Fairness and Voter Turnout // *Comparative Political Studies*. Vol. 43. № 12.
- Blais A.** 2000. *To Vote or Not to Vote? The Merits and Limits of Rational Choice Theory*. — Pittsburgh (PA).
- Deckert J., Myagkov M., Ordeshook P.** 2011. Benford's Law and the Detection of Election Fraud // *Political Analysis*. Vol. 19. № 3.
- Dettrey B.J., Schwindt-Bayer L.A.** 2009. Voter Turnout in Presidential Democracies // *Comparative Political Studies*. Vol. 42. № 10.
- Ferejohn J., Fiorina M.** 1974. The Paradox of Not Voting: A Decision Theoretic Analysis // *The American Political Science Review*. Vol. 68. № 2.
- Development Core Team.** 2005. *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. — Vienna (<http://www.R-project.org>).
- Downs A.** 1957. *An Economic Theory of Democracy*. — N.Y.
- Fornos C.A., Power T.J., Garand J.C.** 2004. Explaining Voter Turnout in Latin America, 1980 to 2000 // *Comparative Political Studies*. Vol. 37. № 8.
- Franklin M.** 2004. *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies since 1945*. — Cambridge.
- Greene K.** 2007. *Why Dominant Parties Lose: Mexico's Democratization in Comparative Perspective*. — Cambridge.

- Jackman** R.W., Miller R.A. 1995. Voter Turnout in the Industrial Democracies during the 1980s // *Comparative Political Studies*. Vol. 27. № 4.
- Kitschelt** H., Wilkinson S. 2007. *Patrons, Clients, and Policies: Patterns of Democratic Accountability and Political Competition*. — Cambridge.
- Kuenzi** M., Lambright G. 2007. Voter Turnout in Africa's Multiparty Regimes // *Comparative Political Studies*. Vol. 40. № 6.
- Levin** I. et al. 2009. *Detecting Voter Fraud in an Electronic Voting Context: an Analysis of the Unlimited Reelection Vote in Venezuela* (http://www.usenix.org/event/evtwote09/tech/full_papers/levin.pdf).
- Lukinova** E., Maygov M., Ordeshook P. 2011. Metastasized Fraud in Russia's 2008 Presidential Election // *Europe-Asia Studies*. Vol. 63. № 4.
- Magaloni** B. 2006. *Voting for Autocracy: Hegemonic Party Survival and Its Demise in Mexico*. — Cambridge.
- Mebane** W. 2008. Election Forensics: The Second Digit Benford's Law Test and Recent American Presidential Elections // Alvarez R.M., Hall T.E., Hyde S.D. (eds.) *Election Fraud*. — Washington.
- Mebane** W. 2010. *Election Fraud or Strategic Voting? Can Second-Digit Tests Tell the Difference?* Paper prepared for presentation at the Summer Meeting of the Political Methodology Society, University of Iowa, July 22–24 (<http://polmeth.wustl.edu/media/Paper/pm10mebane.pdf>).
- Moraski** B. 2002. *Interpreting Voter Turnout in Russia: A Temporal and Cross-Regional Analysis*. Paper prepared for the annual meeting of the Midwest Political Science Association, Chicago, Illinois, April 25–28.
- Myagkov** M., Ordeshook P., Shakin D. 2008. *The Forensics of Election Fraud: Russia and Ukraine*. — Cambridge.
- Ordeshook** P., Myagkov M. 2008. *Russian Elections: An Oxymoron of Democracy*. CalTech/MIT Voting Technology Project Working Paper № 63.
- Reisinger** W.M., Moraski B.J. 2008. *The Relationship between Turnout and Competition Levels in Russia* (http://ir.uiowa.edu/cgi/viewcontent.cgi?article=1100&context=polisci_pubs).
- Riker** W., Ordeshook P. 1968. A Theory of the Calculus of Voting // *The American Political Science Review*. Vol. 62. № 1.
- Wantchekon** L. 2003. Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin // *World Politics*. Vol. 55. № 3.
- Wolfinger** R.E., Rosenstone S.J. 1980. *Who Votes?* — New Haven (CT).