
Сравнительный анализ данных электоральной статистики и выявление электоральных аномалий на материале стран Восточной Европы

Аннотация

В данной статье делается попытка эмпирически установить, существуют ли универсальные свойства, присущие данным электоральной статистики. Несмотря на то, что рядом исследователей предпринимаются попытки статистически установить, например, наличие признаков манипуляций результатами волеизъявления избирателей методами статистического анализа, политическая наука не готова предложить теоретических обоснований для характеристик, предполагающихся нормальными. Данная статья призвана отчасти восполнить пробел в знаниях о свойствах электоральной статистики, обратившись к материалу общенациональных выборов в постсоциалистическом пространстве Восточной Европы. Как показало исследование 98 случаев парламентских и президентских выборов, действительно существуют общие и стабильно воспроизводимые характеристики, которым соответствуют данные электоральной статистики большинства стран, а расхождения с нормой редки.

Ключевые слова: президентские выборы, парламентские выборы, статистический анализ, явка, электоральные аномалии, Восточная Европа.

Автор

Шалаев Никита Евгеньевич

Аспирант факультета политологии
СПбГУ (Россия, Санкт-Петербург)



Контроль за честностью и прозрачностью проведения выборов является одной из важнейших проблем в современном мире. Стандартным способом достижения этой цели является привлечение наблюдателей, в особенности международных, для контроля за ходом выборов. Но это решение, которое представляется нам естественным и надежным, не лишено своих недостатков. Подобного рода миссии требуют огромных

финансовых и человеческих ресурсов, а цели достигают не всегда. Ведь даже авторитарные режимы охотно приглашают наблюдателей, считая это вопросом престижа или жестом доброй воли, который международным сообществом будет засчитан в их пользу в любой ситуации [10], а внутри страны может быть использован в качестве инструмента давления на оппозицию [15]. По мнению ряда авторов, наблюдение за выборами может

Сравнительный анализ данных электоральной статистики и выявление электоральных аномалий

даже оказать негативное влияние на ситуацию в стране в целом, вытесняя электоральные манипуляции в другие, трудно поддающиеся наблюдению и контролю сферы [20], разлагая таким образом институты государства в целом.

Электоральная криминалистика — направление электоральных исследований, ставящее своей задачей разработку и применение методов статистического контроля результатов выборов — претендует на роль если не замены, то компаньона для традиционного института наблюдателей. При этом, как отметил в своей диссертации Дж. Декерт, очевиден разрыв между исследованиями поведения избирателей и изучением электоральных фальсификаций [4]. Политическая наука не может предложить готовых теоретических обоснований для исходных положений методов, предлагаемых для обнаружения искажения результатов голосования. Например, не существует «всеобщей теории явки», которая могла бы предсказать, как должна быть распределена эта величина в случае честных выборов. Нет консенсуса и относительно того, является ли успех кандидатов не зависящим от уровня явки — существуют исследования, как подтверждающие отсутствие связи [6; 21], так и утверждающие ее наличие [14; 13]. Наконец, нет оснований утверждать, что цифры в протоколах будут встречаться с той или иной частотой — за исключением, пожалуй, психологических исследований, которые могут пролить свет на то, как систематически ошибается человек, пытающийся сымитировать случайные числа [1].

В такой ситуации единственным выходом представляется обращение к понятию электоральных аномалий — тех самых отклонений от ожидаемой картины, поиском которых занята электоральная криминалистика, и наличие которых она считает индикаторами фальсификаций. Если нельзя достоверно определить свойства нормы, исходя из чисто теоретических предпосылок, то можно попытаться выявить наличие и соотношение нормы

и отклонений сравнительным методом на основании эмпирических данных. Если никаких универсальных черт в параметрах электоральной статистики не наблюдается, то само направление электоральной криминалистики бессмысленно. Если же общие закономерности прослеживаются, то будет полезным определить их свойства и степень распространенности. Чем более редкими и существенными будут отклонения от нормы, тем более значимым будет поиск такого рода аномалий.

В этом смысле страны Восточной Европы предоставляют уникальную возможность для проведения сравнительных электоральных исследований: большое число государств практически одновременно начали переход от социалистических режимов «народной демократии» к рыночному капитализму и демократии электоральной. При этом общий культурно-исторический контекст позволяет называть эти страны одними из наиболее сравнимых между собой, лишенными в определенной степени специфических особенностей. В то же время в этом регионе присутствуют как страны с хорошей репутацией в разрезе прозрачности выборов, так и страны, к демократичности выборов в которых выдвигались претензии (например, Армения или Украина, особенно в 2004 г.).

Первой гипотезой, подвергнутой проверке, стало предположение о том, что явка является нормально распределенной величиной. Существует несколько способов операционализировать распределение явки, в данном случае было выбрано распределение числа избирательных участков по уровню достигнутой на них явки. Особый интерес именно к явке не случаен. Из показателей электоральной статистики явку можно назвать одним из самых известных и широко применяемых. Ее могут использовать для оценки степени легитимности выборов, как индикатор уровня доверия к политическим институтам, степени поддержки режима, уровня гражданской активности.

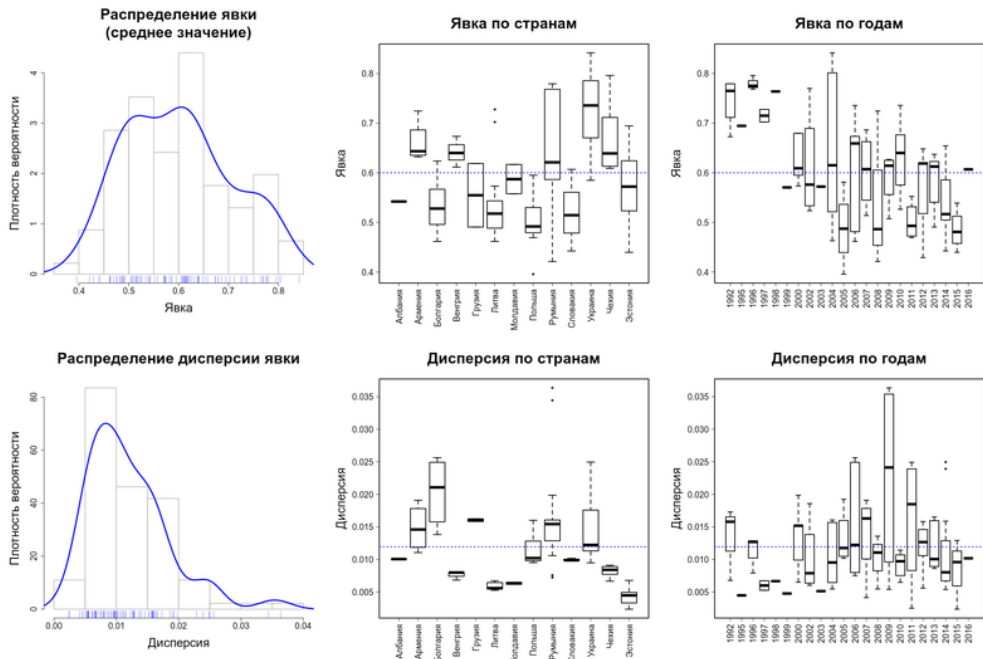


Рисунок 1. Свойства распределений УИК по уровню явки: средняя и дисперсия

Нужно отметить, что подобные попытки осложняются тем, что свойства явки как статистической величины, как правило, не представляли большого интереса для исследователей. Вообще говоря, проблемы свойств распределений политически релевантных статистических величин являются темой, обделенной вниманием в исследователей, и явка здесь не является исключением. В качестве редких примеров таких работ можно привести исследование распределения числа голосов внутрипартийных списков на выборах в Польше [9] (которое оказалось логнормальным), распределение числа голосов, полученных кандидатами на выборах в Бразилии [7; 8] (был выявлен степенной закон), исследование соблюдения закона Ципфа для распределений результатов голосования [16] и работы о распределении уровня явки и результатов кандидатов в городах Франции [3] и ряде других стран Европы [2]. В подавляющем же большинстве случаев явка интересует исследователей лишь как показатель, полученный

по стране или региону в целом, а число голосов представляет интерес только как итоговый результат участия в выборах.

В результате на сегодняшний день политическая наука не готова дать уверенный и обоснованный ответ на вопрос, какое распределение является нормой для явки (или, например, для числа голосов) на выборах — нормальное (гауссовское) или какое-то иное, или существуют какие-то факторы, по которым можно будет заранее предсказать ожидаемое распределение. Меж тем, анализ распределений (хотя и другого рода) применяется для выявления искаженных и сфальсифицированных данных, например, в области финансового аудита [17]. Это дает основания полагать, что и в области электоральных исследований существенные расхождения свойств наблюдаемых распределений с ожидаемыми могут являться признаками внесения искажений в истинные результаты волеизъявления граждан.

Для анализа были взяты данные электоральной статистики на общенациональ-

Сравнительный анализ данных электоральной статистики и выявление электоральных аномалий

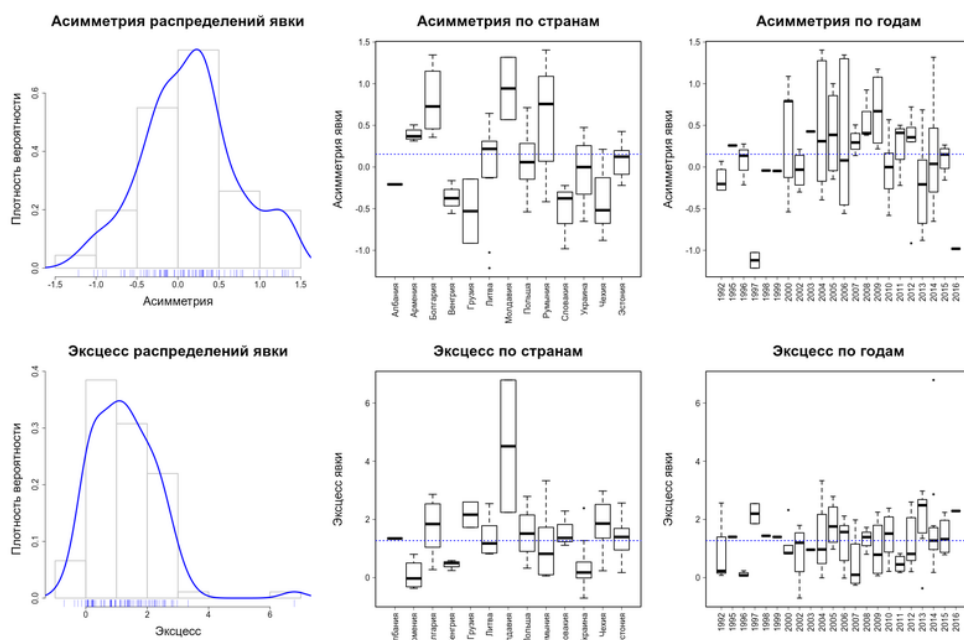


Рисунок 2. Свойства распределений УИК по уровню явки: асимметрия и эксцесс

ных выборах стран Восточной Европы, зафиксированные на уровне участков для голосования (УИК). Основных причин тому две: во-первых, это дает гораздо больше замеров и, следовательно, большую точность оценки распределения. Во-вторых, поскольку в перспективе результат подобных исследований может быть применен для мониторинга демократичности выборов, анализ нужно проводить на том уровне, на котором ожидается внесение искажений в результате волеизъявления граждан. Поскольку наиболее распространенные методы (принуждение избирателей, вбросы, фальсификация протоколов, «карусели» и т. п.) носят локальный характер и действуют в рамках отдельных участков (а не, скажем, муниципалитетов или регионов), то и данные необходимо рассматривать на этом же уровне. В результате эмпирическая база исследования охватила выборы в 13 странах. В их число входят как бывшие республики СССР (Литва, Эстония, Украина, Грузия, Армения, Молдавия), так

и страны «народной демократии» (Польша, Чехия, Словакия, Болгария, Румыния, Албания, Венгрия), входившие в Организацию Варшавского договора. Для анализа были взяты случаи общенациональных выборов — мажоритарных президентских и пропорциональных парламентских (нижняя палата для бикамеральных; в случае смешанной системы мажоритарная часть не учитывалась). В большинстве случаев данные были опубликованы в виде отдельных веб-страниц, и для их агрегации автору пришлось разрабатывать специализированные программы на языке Perl. В некоторых случаях использовались файлы данных, опубликованные центральными избирательными комиссиями стран (или аналогичными структурами).

К сожалению, несмотря на то, что публикация подробных результатов выборов в сети Интернет сейчас становится все более популярной, далеко не все страны могут похвастаться архивами, которые охватывали бы период 1990-х, а иногда и 2000-х гг. В большинстве случаев публи-

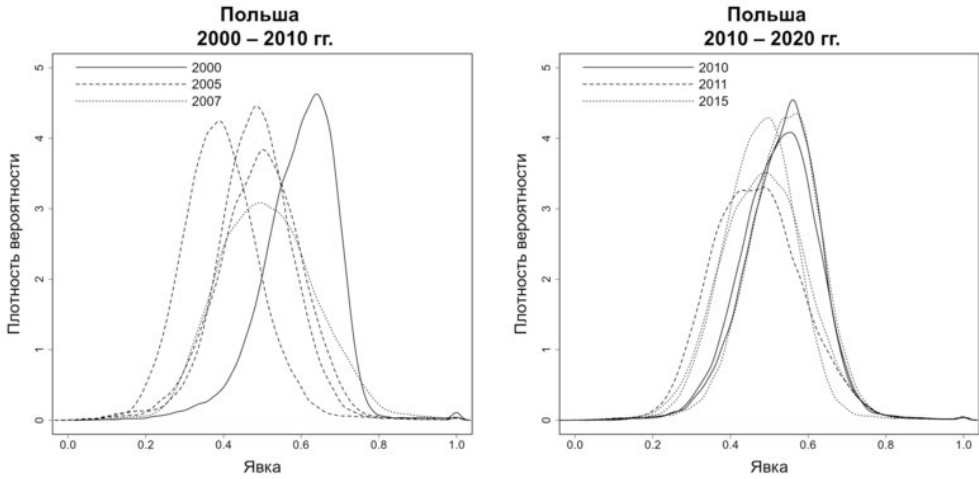


Рисунок 3. Распределение явки на уровне избирательных участков, Польша

кация начинается с середины или даже с конца 2000-х, а данные о более ранних выборах не публикуются вовсе или публикуются в крайне фрагментарном виде (например, только в виде результатов по стране в целом). Тем не менее можно подобрать достаточно представительную группу стран, архивы электоральной статистики которых уходят в прошлое на достаточную глубину. Прежде всего это Румыния и Эстония (данные доступны с начала 1990-х), Чехия и Литва (данные доступны с середины 1990-х), а также Польша (с начала 2000-х). С середины 2000-х стали публиковать данные на уровне УИК и остальные страны. В Латвии же, несмо-

тря на регулярную публикацию данных с конца 1990 гг., традиционно не указывают списочную численность избирателей на участках, что не дает вычислить явку на этом уровне, поэтому она была исключена из рассмотрения. Общие характеристики данных электоральной статистики исследуемых стран приведены в табл. 1.

Мы видим, что в части стран средний участок для голосования обслуживает небольшое количество избирателей (около полутысячи), а в других — примерно вдвое больше. Аналогично для числа участков: в некоторых странах их количество велико (десятки тысяч), в других — на порядок меньше. В рассматриваемой подборке

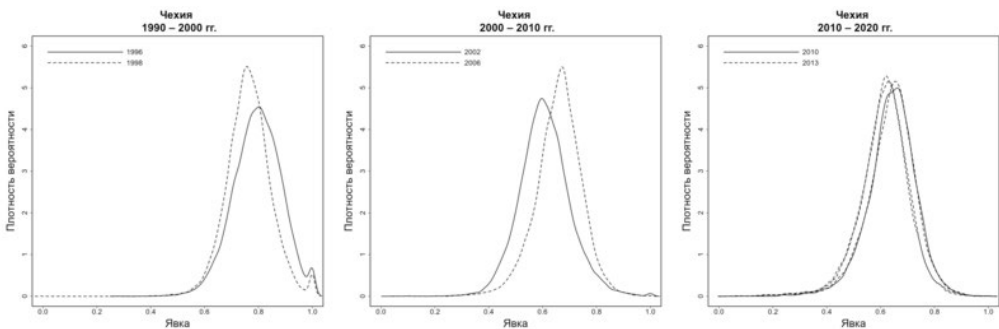


Рисунок 4. Распределение явки на уровне избирательных участков, Чехия

Сравнительный анализ данных электоральной статистики и выявление электоральных аномалий

представлены почти все варианты сочетаний: большое число крупных участков (Польша, Румыния), большое число малых участков (Болгария, Венгрия, Чехия), малое число крупных участков (Литва, Эстония). Сочетание малой величины участка

и их малого количества не реализуется, но и по европейским меркам оно соответствовало бы карликовому государству с примерно 250 тыс. избирателей, т. е. в три раза меньше Эстонии.

**Таблица 1. Средняя списочная численность избирателей
и число участков для голосования**

Страна	Выборы, число случаев	Среднее количество избирателей на участке	Среднее число участков для голосования
Албания	2	594	5133
Армения	4	1243	1951
Болгария	6	583	11633
Венгрия	3	749	10741
Грузия	2	1000	3648
Литва	9	1299	2024
Молдавия	2	1427	1972
Польша	8	1170	25711
Румыния	12	1073	16810
Словакия	2	740	5968
Украина	8	1093	32647
Чехия	7	561	14778
Эстония	7	1305	622

Также следует отметить, что в отличие от остальных стран Венгрия, Эстония и Чехия представляют собой парламентские республики без прямых выборов президента (в Чехии президент был выбран на прямом всенародном голосовании в 2013 г. впервые; прямые президентские выборы 2016 г. в Молдавии не рассматривались). Румыния и Литва — парламентско-президентские республики. Таким образом, в рассматриваемой группе стран оказалось представленным и некоторое разнообразие форм государственного правления. Наконец, часть рассматриваемых выборов по пропорциональной системе проходила с использованием открытых списков, а часть — закрытых.

Итак, соответствует ли распределение числа избирательных участков по уровню достигнутой на них явки распределению Гаусса? Для оценки формы распределений,

помимо визуального сравнения кривых ядерной оценки плотности вероятности, были использованы коэффициенты асимметрии и эксцесса (с поправкой Пирсона). От нормального распределения ожидаются нулевые значения обоих коэффициентов. Отрицательным значениям асимметрии соответствует смещение основной «массы» графика вправо, положительным — влево. Эксцесс выше нуля означает, что экстремальные значения вносят большой вклад в дисперсию распределения, ниже нуля — что основной вклад в дисперсию принадлежит большому числу незначительных отклонений от средней.

В результате сравнения оказалось, что типовое распределение действительно существует, но для большинства стран и случаев его характеристики отличаются от нормального (Рис. 1 и 2; пунктиром обозначены средние значения параметра для

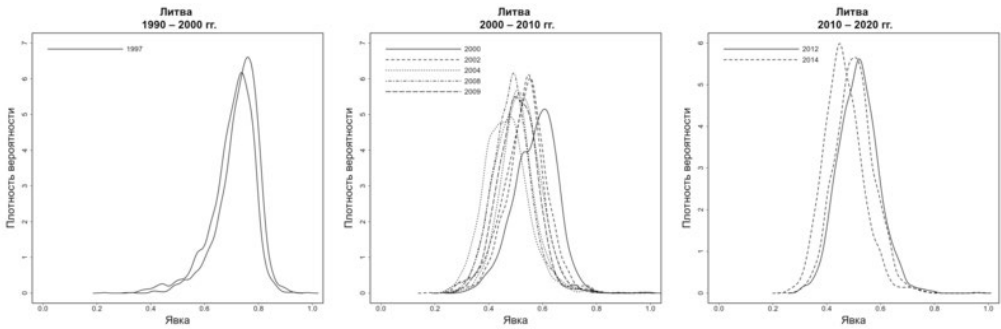


Рисунок 5. Распределение явки на уровне избирательных участков, Литва

всех случаев).

Можно уверенно говорить о том, что в общем случае распределение характеризуется значениями эксцесса выше нуля (Рис. 2). Аномальными являются как раз случаи в Армении и Украине, когда эксцесс был примерно равен нулю. С точки зрения симметрии графика нормой является или симметричность (Литва, Польша, Украина, Эстония), или небольшое смещение основной массы графика вправо (Венгрия, Грузия, Словакия, Чехия). Смещение влево — характерное для Армении, Болгарии, Молдавии и частично Румынии — является наименее распространенным вариантом.

Таким образом, мы видим, что в свойствах электоральной статистики действительно прослеживаются некоторые универсальные черты — пусть и в рамках конкретной группы стран. Однако описание распределений с помощью коэффициен-

тов не является достаточно наглядным. Далее будут рассмотрены некоторые наиболее интересные и представительные примеры с помощью более наглядных графиков.

Для оценки свойств распределения явки (на уровне избирательных участков) было также использовано сопоставление кривых ядерной оценки плотности вероятности. Ядерная оценка плотности вероятности (далее ЯОПВ) — непараметрический метод для оценки функции плотности распределения случайной величины [19]. В отличие от гистограмм, результат применения ЯОПВ представляет из себя отдельные кривые, которые легко сравнивать между собой (пример их соотношения можно найти в левых частях графиков 1 и 2). В то же время такой подход снимает необходимость обосновывать адекватность идентификации наблюда-

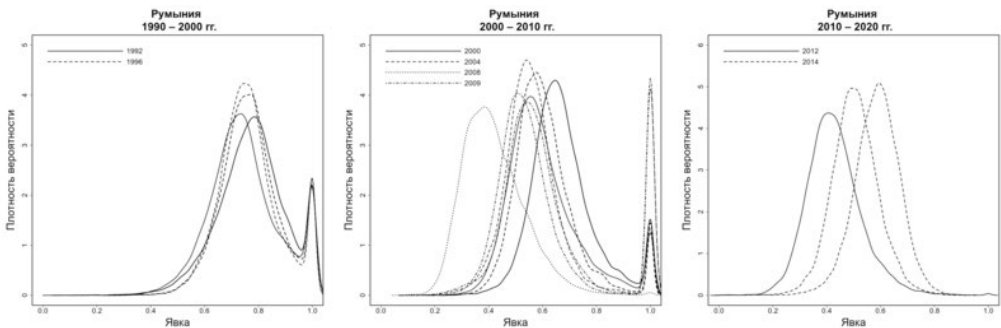


Рисунок 6. Распределение явки на уровне избирательных участков, Румыния

Сравнительный анализ данных электоральной статистики и выявление электоральных аномалий

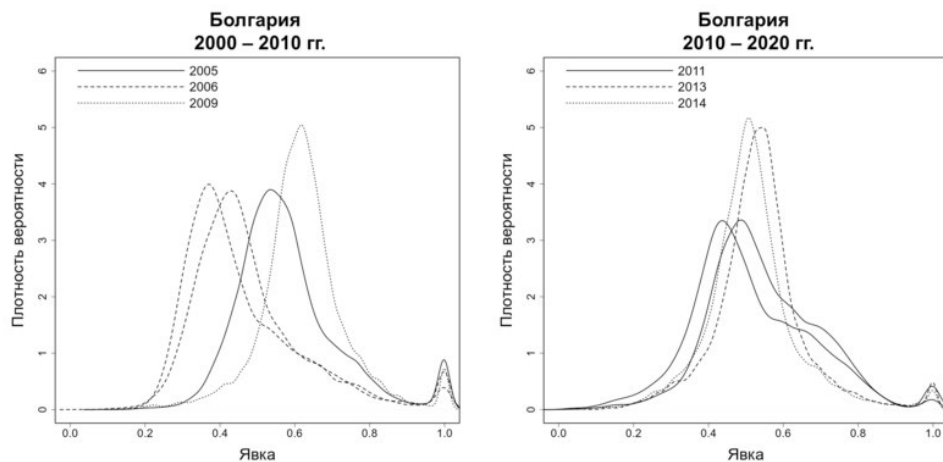


Рисунок 7. Распределение явки на уровне избирательных участков, Болгария

емого распределения, что являлось бы существенной проблемой при подборе параметрических распределений. Пример сравнения кривых ЯОПВ представлен на рис. 3 (Польша). На этом графике мы видим довольно однотипные симметричные кривые, которые мы ожидали бы увидеть от нормально распределенной величины. Однако, насколько «стандартными» являются такие распределения? Изменялись ли они с течением времени? Ответить на этот вопрос можно только в сравнительной перспективе, что и будет сделано далее.

На последующих графиках распределения участков для голосования по уровням явки в отдельных странах будут представлены по декадам; в результате для большинства стран случаи делятся на примерно равные группы. Поскольку не во всех странах доступна информация о выборах в 1990 г., то в некоторых случаях представленными оказываются только две декады (2000–2010 гг. и с 2010 по настоящее время). В рамках каждой декады кривые ЯОПВ для проведенных в эти годы выборов наложены друг на друга (президентские выборы — в случае проведения

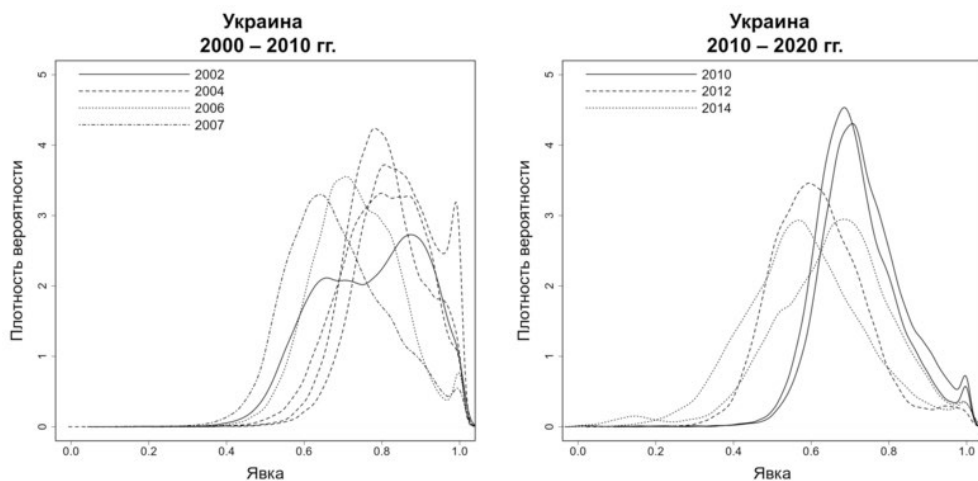


Рисунок 8. Распределение явки на уровне избирательных участков, Украина

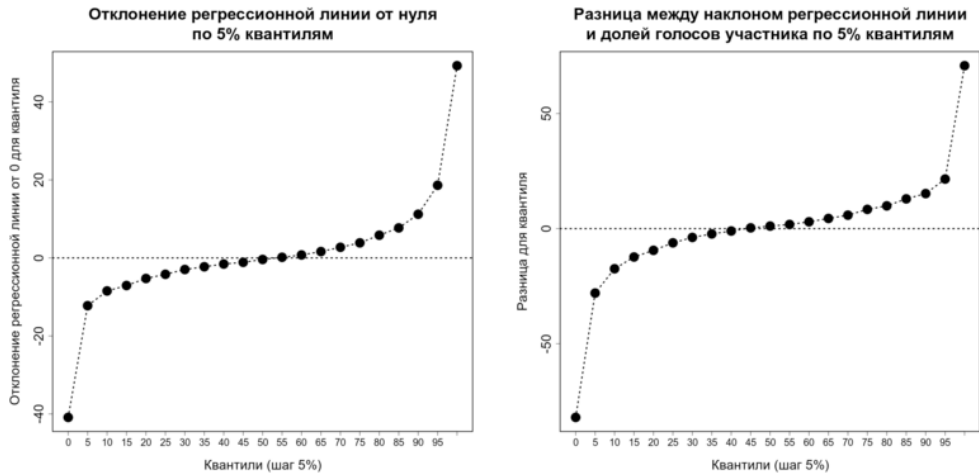


Рисунок 9. Отклонение параметров регрессионных моделей от ожидаемых значений

второго тура — дают две кривые).

Ближе всего к ожидаемым показателям оказались распределения явки в Польше (Рис. 3). После 2004 г. наблюдается некоторое снижение уровня явки (что видно по смещению пика кривых ЯОПВ влево), но затем она снова растет и к 2015 стабилизируется в районе 50 %. Форма распределения симметричная, «хвосты» распределения тонкие. Средняя асимметрия распределений равняется 0,1; коэффициент эксцесса (здесь и далее указывается с поправкой Пирсона) в среднем составляет 1,57. Таким образом, распределения похожи на нормальные в том, что являются симметричными относительно вершины, но в то же время отличаются тем, что общая дисперсия распределения в большей мере определяется редкими экстремальными значениями (а не большим количеством умеренных отклонений от средней).

Аналогичная траектория наблюдается и для явки в Чехии (Рис. 4): явка в 1990 гг. была существенно выше, чем в последующих. Несмотря на то, что в 2010-х наблюдается смещение явки в область более высоких значений, до уровня 1990 гг. она не доходит. Коэффициент асимметрии для распределений в Чехии в среднем составляет — 0,42. Коэффициент эксцес-

са оказался равен 1,94. При этом в Чехии более ранние распределения носят слабо выраженный бимодальный характер — в области высоких значений (~100 %) наблюдается вторая, хотя и небольшая по величине, вершина. Аналогичный пик можно рассмотреть и в распределениях, наблюдаемых на выборах в Польше, но там он выражен гораздо слабее. В его наличии, впрочем, нет ничего удивительного — всегда существуют участки, где высокая явка получается совершенно естественным образом (например, на кораблях или в больницах). Более странным является то обстоятельство, что после середины 2000-х такие участки в Чехии исчезли.

В Литве можно наблюдать иную тенденцию: явка, снизившись в целом по сравнению с 1990 гг., в 2000-х зафиксировалась в районе одних и тех же значений (Рис. 5). При этом хорошо видно, что форма распределений воспроизводится из года в год с большой точностью. Интересным моментом является наличие более тяжелого «хвоста» слева (а не справа) у распределений на выборах 1997 г. Средняя же асимметрия распределений равна 0,03, а эксцесс — 1,33. По этим параметрам распределения в Литве ближе к распределению Гаусса, чем рассмотренные выше.

Сравнительный анализ данных электоральной статистики и выявление электоральных аномалий

Итак, несмотря на то, что с точки зрения тенденций в изменении уровня явки, в пяти рассмотренных выше странах наблюдается различная картина, характер распределения явки остается аналогичным. То есть безотносительно уровня активности избирателей совокупный результат их поведения имеет одни и те же статистические характеристики (в т. ч. как до, так и после вступления в ЕС; как на выборах парламента, так и президента). Это наблюдение могло бы показаться тривиальным, если бы на фоне этого единообразия свойств распределения явки на избирательных участках не выделялось два исключения: Румыния и Болгария. Несмотря на то, что в случае Румынии (Рис. 6) можно обнаружить те же тенденции к постепенному снижению явки, в распределениях явки мы видим нечто, чего в других странах не наблюдалось: «толстый» правый хвост распределения. Сразу следует оговориться, что второй пик на значении явки в 100 % во многом вызван техническими причинами: в архиве электоральных данных Румынии за более ранние годы указывалась только списочная численность зарегистрированных на территории участка избирателей; для многих участков она была равна нулю (как, например, в случае избирательных участков на вокзалах). Для таких УИК явка по необходимости принималась равной 100 %. Однако большое число участков с явкой, приближающейся к 100 %, артефактом обработки данных не является.

Румыния вошла в Европейский союз в 2007 г. И на выборах 2008 г. впервые исчезает второй пик графика в правой части. Кроме того, постепенно сходил на нет и правый «хвост» распределения в целом. На выборах 2009 г. вторая вершина графика, впрочем, снова проявляется — уже в последний раз, и не сопровождаясь «толстым» правым хвостом (следовательно, в большей мере являясь артефактом представления данных). После 2009 г. распределение участков по уровню явки становится идентичным тому, что можно

было зафиксировать в остальных странах. В результате средняя асимметрия распределений в Румынии оказывается одной из максимальных (уступая только Болгарии) и равна 0,56. Средний эксцесс при этом равняется единице.

На фоне всех остальных стран (включая и Румынию) Болгария (Рис. 7) показывает совершенно уникальную картину. Некоторые распределения имеют не две, а три вершины: в районе собственно среднего значения явки (в середине графика), в области максимальных значений (100 %) и между ними (выраженную слабее) (в особенности 2011 г., первый и второй туры выборов президента). В то же время распределения, наблюдавшиеся в 2013 и 2014 гг. аналогичны распределениям в других странах. Следует отметить также, что с течением времени снижается количество участков, на которых фиксируется 100 % явка. Средняя асимметрия распределений в Болгарии максимальна и составляет 0,8.

Кроме того, появление «промежуточной вершины» не связано с типом выборов — ее можно заметить в распределениях 2011 г. (выборы президента) и 2005 г. (парламентские выборы). Можно отметить также, что форма распределения заметно меняется в течение года или двух. Если бы подобное распределение было порождено сочетанием в Болгарии двух типов электората (с «обычной» явкой и со склонностью к более высокой явке), то можно было бы ожидать устойчивого воспроизведения соответствующего паттерна — такого рода группы обычно связываются с социальными расколами (например, город — село), а их структура обладает большой инерцией и не склонна к постоянному изменению. В качестве аналогичного случая можно привести страну, которая также шла по пути обратной трансформации и демократичность выборов в которой неоднократно становилась спорным вопросом, — Украину (Рис. 8). Здесь наблюдается большое разнообразие форм распределений, которое сложно

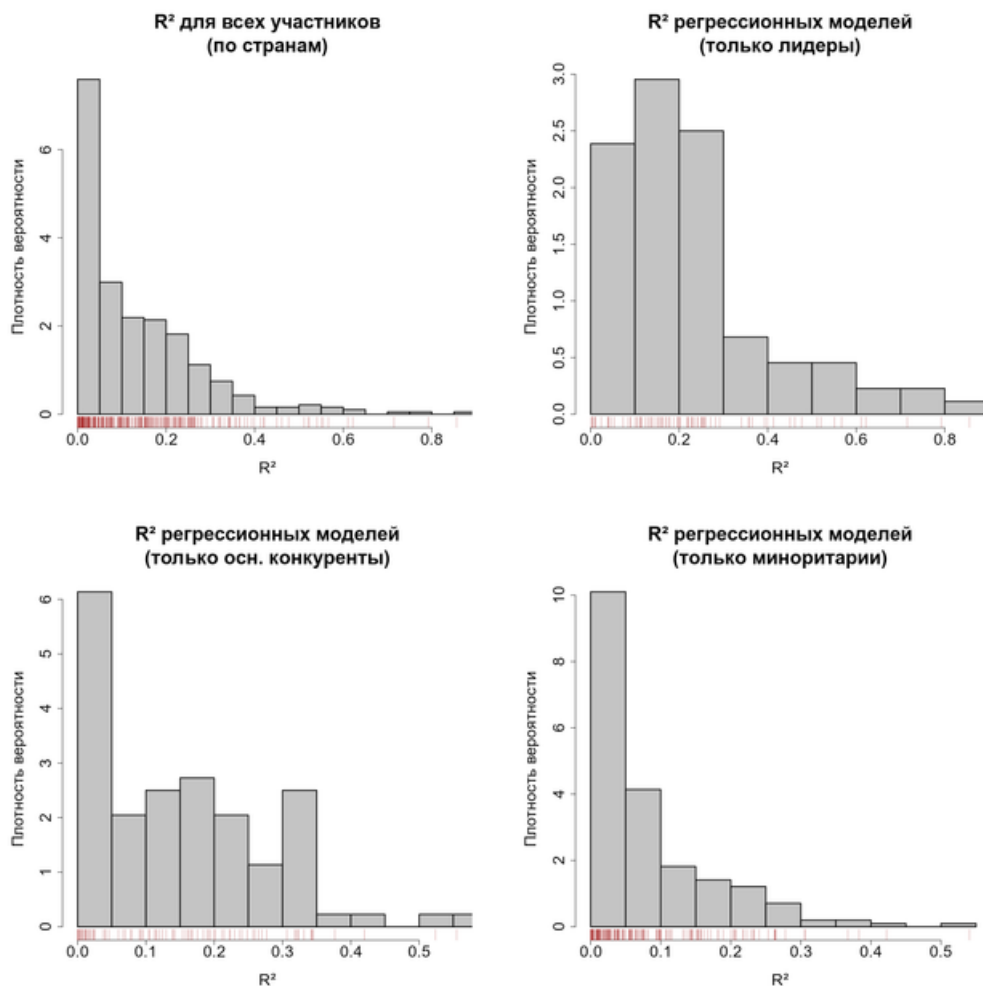


Рисунок 10. Величина R² для регрессионных моделей (в целом и по категориям участников выборов)

было бы объяснить только неоднородностью электората — в таком случае требовались бы также факторы, которые приводили бы к постоянному изменению его характеристик. Особенно яркие различия между распределениями, наблюдавшимися в 2004 г., когда в течение трех месяцев украинский электорат три раза повел себя совершенно различным образом.

Начиная с 2010 г., можно также наблюдать тенденцию распределений к тому, чтобы сблизиться с теми, что были обнаружены в других странах. И, очевидно, эти изменения нельзя объяснить событиями

2014–2015 гг., когда в электоральной статистике Украины перестали быть представлены регионы, с «уникальностью» электората которых можно было бы связать предшествовавшие аномалии.

Можно сделать предварительный вывод о том, что в подавляющем большинстве рассмотренных стран тенденции в электоральном поведении граждан схожи: на парламентских и президентских выборах распределение активности граждан по избирательным участкам одинаково, и форма распределения воспроизводится с большой точностью из года в год даже в

Сравнительный анализ данных электоральной статистики и выявление электоральных аномалий

разных странах. Может изменяться средний уровень явки (положение вершины графика), но характеристики распределения остаются в одних и тех же (довольно узких) пределах. За исключением Румынии, мы не наблюдаем перекоса в сторону больших уровней явки — но в Румынии можно также отметить и последовательную трансформацию распределений, в результате которой картина электоральной активности в этой стране сблизилась с «типовой». Был ли массивный «правый» хвост распределения наследием коммунистического режима или результатом фальсификаций, сейчас сказать затруднительно. Но в любом случае эти аномалии с течением времени сошли на нет. Это наблюдение согласуется как с демократизацией электоральных институтов, так и с нормализацией электорального поведения избирателей. Этот вопрос, безусловно, должен быть исследован полнее в дальнейшем.

Картина, наблюдаемая в Болгарии, напротив, не поддается разумному объяснению. В некоторых случаях избиратели этой страны ведут себя так же, как и в соседних, — особенно на выборах в 2013 и 2014 гг. В других почему-то проявляют совершенно иные свойства, как если бы электорат страны был разделен на две группы с существенно разной склонностью к абсентеизму. При этом переходы от одной модели к другой происходят быстро и резко и не зависят от типа выборов. В то время как объяснить такие метаморфозы изменением самих избирателей затруднительно, на подозрения в искусственном их характере наводит параллель с распределениями, наблюдавшимися на Украине. Там резкие смены видов распределений сопутствовали беспрецедентным скандалам относительно фальсификации результатов выборов. Поэтому из рассмотренных выше стран Болгария более, чем какая-либо иная страна, заслуживает пристального внимания исследователей.

Второй подвергнутой проверке гипотезой стало предположение о том, что в нормальных условиях результаты канди-

датов не зависят от уровня явки на участке. Это предположение является основой, например, метода Собянина-Суховольского, использующего линейную регрессию для анализа связи между явкой и долей голосов (относительно списочной численности избирателей), набранных участниками выборов. Здесь предполагается, что в нормальных условиях константа регрессионного уравнения (т. н. СРЛ) будет равна нулю, коэффициент наклона (т. н. ПДИ) будет соответствовать итоговой доле голосов, набранных участником выборов, а объяснительная сила регрессионной модели будет невелика (значение коэффициента детерминации R^2 будет малым). Этот метод и был применен к имеющимся данным.

На рис. 9 отражено расхождение наблюдаемых показателей регрессионных прямых с ожидаемыми для 5 % квантилей из числа рассматриваемых моделей. Можно отметить, что в целом предположения о характере регрессионных моделей соответствуют эмпирическим данным: прямые проходят близко к началу координат (разница между СРЛ и 0 обычно невелика), а коэффициент наклона в целом соответствует доле голосов, полученных участником. Конечно, идеального соответствия не наблюдается, но для подавляющего большинства 5 % квантилей расхождение крайне мало, и резко возрастает только для 5 или 10 % наиболее экстремальных случаев.

На рис. 10 представлены результаты регрессионного анализа для участников выборов (партий и кандидатов) в рассматриваемых странах. Как и следовало ожидать — объяснительная сила моделей в подавляющем большинстве случаев крайне мала. Только для единичных случаев регрессионная модель смогла объяснить более половины дисперсии. В 95 % моделей доля объясненной дисперсии составила менее 0,395. При этом наиболее «хорошо» объясняемые случаи (т. е. с аномально высокими значениями R^2) относятся к лидерам выборной гонки, и только в двух случаях — к их главным конкурентам (т. е. тем участникам, которые в итоге заняли

Тема номера: «Политические выборы в России и за рубежом»

второе место по числу голосов). Всего в одном случае R^2 превысил 0,5 для миноритарного участника, а в целом среди моделей миноритариев доминируют предельно низкие значения коэффициента детерминации.

Оставшейся гипотезой стало предположение о том, что в полученных «естественным образом» данных электоральной статистики распределение цифр в различных разрядах чисел соответствует закону Ньюкомба-Бенфорда, а в искаженных данных — нет. Кроме закона Ньюкомба-Бенфорда была проверена его модификация, предложенная Л. Перикки и Д. Торресом [18] и призванная учесть ограничение на диапазон допустимых значений (в силу ограниченного количества избирателей на участках). Также для первой цифры чисел было проверено соответствие наблюдаемых частот закону Стиглера [11]. Наконец, была проверена альтернативная методика, предложенная Б. Бебером и А. Скакко [1], заключающаяся в том, что распределение парных цифр в двух последних разрядах чисел должно отличаться от естественного (в качестве которого они полагали равновероятное) в случае, когда данные генерируются человеком, пытающимся создать видимость «случайных чисел». В этом случае полученные частоты были сопоставлены также с предсказываемыми законом Ньюкомба-Бенфорда (использовался принцип вычисления ожидаемой вероятности в совокупности чисел различной разрядности, предложенный Л. Лееманном и Д. Бохслером [12]). На этом этапе исследования оказалось возможным учесть также данные из Латвии, доведя число исследуемых стран до 14.

Все вышеперечисленные тесты потерпели неудачу. G-тест (тест максимального правдоподобия; тест χ^2 является его аппроксимацией) по первому разряду (для результатов всех участников вместе взятых) для закона Ньюкомба-Бенфорда из 98 случаев (различные туры президентских выборов учитываются отдель-

но) прошли только 2, по второму разряду — 16, по последнему разряду — 10. Для закона с поправкой Перикки-Торреса результаты составили соответственно 0,6 и 3 случая. Совпадения с равновероятным распределением наблюдались в 17 случаях, но только для последних разрядов. Наконец, частоты парных цифр на концах чисел совпали с ожидаемыми по закону Ньюкомба-Бенфорда в 11 случаях, по нему же с поправкой Перикки-Торреса в 47 случаях, и с ожидаемыми при равновероятном распределении — в 37 случаях. Нужно также отметить, что, например, все парламентские выборы и все первые туры президентских выборов в Польше тест по исходной методике Бебера-Скакко не прошли, а все вторые туры президентских выборов — прошли. Очевидно, что это является результатом разного уровня фрагментации, которая никак не учитывается такого рода тестами, а не тем, что в Польше повально фальсифицируют выборы, но только при числе участников больше двух.

Следовательно, как раз совпадение частот появления цифр в том или ином разряде с предсказываемыми законом Ньюкомба-Бенфорда (а равно и иными протестированными законами) является аномалией, а отсутствие совпадения — нормой. Этот результат согласуется с предыдущими работами [5], опровергавшими принципиальную применимость закона Бенфорда для анализа электоральной статистики. Впрочем, это оставляет возможность использования такого подхода в случае выявления дальнейшими исследованиями истинного закона распределения цифр, если таковой в принципе существует. В то же время перспективы анализа свойств явки и зависимостей между показателями электоральной статистики выглядят гораздо более оптимистично. Мы могли наблюдать, что в этих областях прослеживается общность характеристик для большинства рассмотренных случаев выборов, а расхождения с ожидаемыми значениями являются редкими и заслуженно претендующими на статус аномальных.

Сравнительный анализ данных электоральной статистики
и выявление электоральных аномалий

Литература

1. *Beber B., Scacco A.* What the Numbers Say: A Digit-Based Test for Election Fraud // *Political Analysis*. — 2012. — Vol. 20. — P. 211–234.
2. *Borghesi C.* Election Turnout Statistics in Many Countries: Similarities, Differences, and a Diffusive Field Model for Decision-Making / *Borghesi C., Raynal J. — C., Bouchaud J. — P.* // *PLoS ONE* — 7 (5). — 2012. — P. e36289.
3. *Borghesi C.* Spatial correlations in vote statistics: a diffusive field model for decision-making / *Borghesi C., Bouchaud J. — P.* // *Eur. Phys. J. B.* — 2010. — Vol. 75. — P. 395–404.
4. *Deckert J.* Patterns of Fraud: Tools for Election Forensics. University of Oregon. — 2013. — URL: <https://scholarsbank.uoregon.edu/xmlui/handle/1794/13331>.
5. *Deckert J., Myagkov M., Ordeshook P. C.* The Irrelevance of Benford's Law for Detecting Fraud in Elections // *Caltech/MIT Voting Technology Project Working Paper*. — 2010. — № 9.
6. *Ferwerda J.* Electoral consequences of declining participation: A natural experiment in Austria // *Electoral Studies*. — 2014. — Vol. 35. — P. 242–252.
7. *Filho C.* Brazilian elections: voting for a scaling democracy / *Costa Filho, R. N.; Almeida, M. P.; Moreira, J. E.; Andrade, J. S.* // *Physica A: Statistical Mechanics and its Applications*. — 2003. — Vol. 322. — P. 698–700.
8. *Fortunato S.* Scaling and Universality in Proportional Elections / *Fortunato S., Castellano C.* // *Phys. Rev. Lett.* — 99. — 2007. — 138701. — 4 p.
9. *Gradowski T. M.* Statistical Properties of the Proportional Voting Process / *Gradowski T. M., Kosinski R. A.* // *ACTA PHYSICA POLONICA A*. — 2008. — Vol. 114. — № 3. — P. 575–580.
10. *Hyde S. D.* Catch Us If You Can: Election Monitoring and International Norm Diffusion // *American Journal of Political Science*. — April 2011. — Vol. 55. — № 2. — P. 356–369.
11. *Lee J., Tam Cho W. K., Judge G. G.* Stigler's approach to recovering the distribution of first significant digits in natural data sets // *Statistics & Probability Letters*. — 2010. — Vol. 80. — Issue 2. — P. 82–88.
12. *Leemann L., Bochsler D.* A systematic approach to study electoral fraud // *Electoral Studies*. — 2014. — Vol. 35. — P. 33–47.
13. *Lijphart A.* Unequal Participation: Democracy's Unresolved Dilemma // *The American Political Science Review*. — Vol. 91. — № 1 (Mar., 1997). — P. 1–14.
14. *Linder W.* *Swiss Democracy: Possible Solutions to Conflict in Multicultural Societies* (Third Edition). — Palgrave MacMillan. — New York. — 2010.
15. *Little A. T.* Fraud and Monitoring in Non-competitive Elections // *Political Science Research and Methods*. — Vol. 3. — Issue 01. — January 2015. — P. 21–41.
16. *Lyra M. L.* Generalized Zipf's law in proportional voting processes / *Lyra, M. L., Costa, U. M. S., Costa Filho, R. N., Andrade (Jr), J. S.* // *Europhysics Letters*. — 2003. — Vol. 62. — P. 131.
17. *Nigrini M. J.* *Benford's law: applications for forensic accounting, auditing, and fraud detection*. — John Wiley & Sons, Inc. — Hoboken, New Jersey. — 2012. — 330 p.
18. *Pericchi L., Torres D.* Quick Anomaly Detection by the Newcomb-Benford Law, with Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela // *Statistical Science*. — 2011. — Vol. 26. — № 4. — P. 502–516.
19. *Silverman B.* *Density estimation for statistics and data analysis*. — Chapman and Hall. — London. — 1986. — 175 p.
20. *Simpser A., Donno D.* Can International Election Monitoring Harm Governance? // *The Journal of Politics*. — Vol. 74. — № 2. — April 2012. — P. 501–513.
21. *Tóka G., Popescu M.* Inequalities of Political Influence in New Democracies // *International Journal of Sociology*. — Vol. 37. — № 4, Causes and Consequences of Political Inequality in Cross-National Perspective (Winter, 2007/2008). — P. 67–93.