

*И. Н. Кандричина*

Белорусский национальный технический университет, Минск

*I. Kandrichina*

Belarusian National Technical University, Minsk

УДК 316.46(324)

## ПРОБЛЕМА ТОЧНОСТИ ЭЛЕКТОРАЛЬНОГО ПРОГНОЗА

## THE ISSUE OF THE ELECTORAL FORECAST ACCURACY

*В статье рассматривается роль опросов общественного мнения в прогнозировании результатов выборов. Акцент сделан на проблему достижения точности электорального прогноза и анализ достоинств и недостатков трех основных технологий прогнозирования итогов выборов: электорального опроса, статистического моделирования и «рынков предсказаний».*

*Ключевые слова: электоральный прогноз; электоральный опрос; статистическая прогнозная модель; рыночный прогноз.*

*The article examines the role of public opinion polls in forecasting election results. The focus is on the problem of achieving the accuracy of the electoral forecast and the analysis of the advantages and disadvantages of the three main technologies for forecasting election results: electoral polling, statistical modeling and «prediction markets».*

*Keywords: electoral forecast; electoral polling; statistical forecast model; market forecast.*

Роль опросов общественного мнения в ходе избирательных кампаний и их влияние на результаты выборов – предмет многочисленных исследований и публикаций, однако данные вопросы до сих пор остаются достаточно актуальными и дискуссионными. Электоральные опросы используются при прогнозировании итогов выборов, формировании стратегии избирательной кампании кандидатов и интерпретации поведения электората. Вместе с тем в настоящее время наблюдается тенденция снижения значимости опросов общественного мнения в электоральной борьбе и использование альтернативных технологий прогнозирования итогов выборов.

Первые электоральные прогнозы строились на основе суждений осведомленных наблюдателей, политических инсайдеров и лидеров [1]. Дж. Гэллап в 1936 г. впервые провел квалифицированный электоральный опрос, однако ненаучные соломенные опросы появились в 1824 г., когда неофициальные подсчеты возможных итогов выборов проводились в тавернах, милицейских офисах и на публичных собраниях [2].

В отличие от большинства тем опросов, электоральные опросы имеют критерий истинности – результаты выборов. Репутация исследовательских фирм в немалой степени зависит от точности данных оценок. Так, «Литерари Дайджеста» был одним из уважаемых журналов, успешно предсказывавшим результаты выборов с 1916 по 1932 г., но фиаско 1936 г. повлекло за собой потерю подписчиков и прекращение работы издания. Дж. Зогби

прославился благодаря точному прогнозу результатов президентских выборов 1996 г., а в 2010 г. он был назван «худшим социологом в мире» из-за ошибочных предсказаний [3].

Прогнозирование результатов выборов – это сложный процесс и рискованный бизнес, поэтому важно понять, почему одни опросы дают правильный результат, а другие – ложный. Сама задача оценки точности вызывает вопросы о том, как лучше ее измерить. Следует ли называть прогноз точным, если он правильно предсказывает победителя, долю голосов победителя или преимущества в процентных пунктах? В 1996 г. «CBS News» правильно предсказал победу Б. Клинтона на президентских выборах, но ошибся с долей полученных голосов. Дж. Гэллап ошибся всего на 2 процентных пункта в предсказании итогов выборов 2000 г., но он предрек неверного победителя всенародного голосования. В 2004 г. «Fox News» верно предугадал долю голосов Дж. Керри в 48 процентов, но неверно предположил его победу. Соответственно, выводы о точности могут варьироваться в зависимости от конкретного используемого критерия.

После провала электорального прогноза 1948 г. группа социологов во главе с Ф. Мостеллером выделила 8 критериев для оценки точности [4]. Сегодня используются «третья мера Мостеллера» – средняя абсолютная ошибка по всем основным кандидатам между прогнозом и фактическими результатами и «пятая мера Мостеллера» – абсолютное значение разницы между долей электората, разделяющей двух ведущих кандидатов в опросе, и разницей долей их электората в фактическом голосовании. Также используют прогностическую меру точности Э. А. Мартин, основанную на натуральном логарифме отношения шансов исхода в опросе и исхода на выборах. Однако продолжается поиск новых показателей точности.

На качество прогнозов могут влиять случайные и систематические ошибки выборки, организационные, методологические и технологические ошибки, а также ошибки покрытия, измерения, обработки и корректировки. Случайная ошибка выборки может привести к отклонению результатов от истинных (выборочная средняя выйдет за пределы заданного интервала), потому что в исследовании участвовала только выборочная совокупность респондентов, а не все представители генеральной общности. Подобные ошибки уменьшаются при увеличении размера выборки. Поэтому при публикации результатов выборочных опросов сообщаются доверительная вероятность (чаще всего 95 %) и доверительный интервал. Систематическая ошибка зависит от факторов, детерминирующих ход проведения исследования и смещение его результатов в определенную сторону. Например, занижение доли респондентов с высоким доходом в силу трудности доступа к ним и отказов от участия в исследовании. Такую ошибку можно минимизировать введением квот или перевзвешиванием данных, но ее размер сложно оценить.

При планировании и организации опросов следует учесть множество аспектов: выбор методов сбора и обработки информации, построение выбо-

рочной модели, соответствие логической программы обработки информации выдвинутым гипотезам, подготовка анкетеров/интервьюеров и операторов ввода первичной социологической информации, соблюдение правил формальной логики при операционализации основных понятий и формулировке вопросов, выбор сроков полевых исследований, времени проведения опроса и т. д. И каждое из этих методологических решений потенциально может повлиять на результаты. Эмпирически доказано, что количество дней, в течение которых проводится электоральный опрос, и их тип (выходные или будние дни) связаны с прогностической точностью и объективностью получаемых ответов [5]; телефонные опросы, исключающие звонки на сотовые телефоны домохозяйств, искажают результат за счет уменьшения доли молодежи, а порядок вопросов влияет на ответы респондентов и, как следствие, приводит к различным оценкам поддержки тех или иных кандидатов. Существует широкий спектр манипулятивных технологий и методологических решений, непосредственно влияющих на качество получаемых данных, но для электоральных опросов, направленных на прогнозирование итогов выборов, особую роль играет определение вероятных и неопределенных избирателей.

Предвыборные прогнозы могут быть неточными и в силу того, что они основаны на собранных и обобщенных мнениях респондентов, которые могут подтвердить свой выбор, поменять суждение или не явиться на избирательные участки в день выборов. Кроме того, точный прогноз выборов основывается на предпосылке о том, что респонденты дают честные ответы на вопросы о явке и намерении проголосовать. Однако здесь появляется эффект социальной желательности. Опрашиваемые могут давать ответы, которые, по их мнению, являются предпочтительными для окружающих, тем самым подтверждая намерения явки как желаемого и ожидаемого от них поведенческого акта. Правдивость электорального прогноза также зависит от склонности людей менять свое мнение о кандидатах, о явке на выборы и намерении голосовать в период с момента опроса до дня выборов. Данные панельных исследований показали, что более 40 % респондентов меняют свое намерение голосовать хотя бы один раз в течение избирательной кампании [6]. Однако по-прежнему ведутся споры об источнике этой динамики на индивидуальном уровне. Чтобы свести к минимуму подобные ошибки и нивелировать их источники, сегодня электоральные опросы все чаще проводятся в ночь перед выборами, и именно этот окончательный опрос используется в оценках точности прогнозов.

Ряд исследований доказывают, что конкретные события предвыборной кампании детерминируют изменения мнений и оценок респондентов. А. Гельман и Г. Кинг утверждают, что увеличение количества опросов позволяет достичь более истинных показателей. Анализируя динамику электоральных опросов, К. Влезейн и Р. С. Эриксон [7] объясняли до 50 % вариабельности данных опроса просто ошибкой выборки. Они также обнаружили, что неординарные и непредсказуемые события избирательной

кампании порождают реальные подвижки в общественном мнении, однако данный эффект в начале электоральной борьбы быстро рассеивался, а в конце – приводил к устойчивым изменениям мнений.

Кандидаты могут пользоваться поддержкой различных групп избирателей, например в 2008 г. за Б. Обаму проголосовала необычно большая доля меньшинств и молодежи. Следовательно, одним из наиболее важных методологических решений является выбор вероятных избирателей, ведь прогнозы могут сильно варьироваться в зависимости от метода определения их численности. Каждая исследовательская фирма имеет собственные методы и модели определения вероятных избирателей, как правило, базирующихся на показателях регистрации избирателей, истории голосований и современных исследованиях политического участия и электорального поведения.

Задача электорального опроса состоит в том, чтобы предсказать, как избиратели будут отдавать свои голоса, поэтому наиболее стандартный вопрос звучит так: «Если бы выборы проводились сегодня, за кого бы вы проголосовали?». Здесь кроется еще одна методологическая трудность – учет мнений неопределившихся респондентов. Но сокращение доли неопределившихся избирателей с помощью различных механизмов распределения не обязательно повышает точность оценок [8]. Таким образом, хотя утверждается, что неопределившиеся респонденты вносят свой вклад в ошибку прогноза, до сих пор нет единого мнения о том, как этого избежать.

Признавая, что результаты индивидуальных опросов подвержены случайной ошибке выборки и прочим искажениям, стало популярным агрегировать данные различных опросов. Во многом облегчила данный процесс широкая доступность онлайн-агрегаторов опросов, например Pollster.com, FiveThirtyEight.com и RealClearPolitics.com. Агрегирование опросов помогает снизить волатильность прогнозов. Так, например, СМИ склонны уделять большое внимание каждому изменению, даже незначительному, электоральных настроений. В то же время большинство опросов, проводимых по заказу СМИ, охватывают небольшие выборочные совокупности и, как следствие, не могут выявить относительно небольшие ежедневные или еженедельные изменения в настроениях избирателей. Объединение результатов опросов в подобных случаях повышает точность данных – ведь больший размер выборки имеет меньшую погрешность. Существует множество методов агрегирования опросов, каждый из которых имеет свои плюсы и минусы. Некоторые агрегации просто берут среднее значение всех доступных результатов опросов. Другие – базируются на различных аналитических моделях, учитывающих специфические особенности организации, проведения и обобщения данных каждого из подлежащих агрегированию опросов. Например, FiveThirtyEight.com оценивает каждый опрос с помощью трех критериев: списка опрашиваемых, размера выборки и сроков проведения опроса. Третий вид агрегаций использует алгоритмы, включающие различные источники информации.

Несмотря на растущую популярность агрегирования результатов опросов, вопросы о наиболее эффективных методах объединения данных, учитывающих особенности организации и проведения эмпирических исследований, а также размеры и ошибки выборки, остаются дискуссионными.

В целом электоральные опросы, их методология, методы и технологии проведения хорошо зарекомендовали себя, однако нельзя отрицать и наличие неудач. Уменьшение влияния электоральных опросов и снижение уровня доверия к их результатам при увеличении их числа кажется парадоксальным. С другой стороны, это вполне объяснимо: ведь увеличение числа опросов приводит к большой вариативности данных и используемых методов и технологий социологического исследования. Отсутствие методологической прозрачности способствует возникновению скептицизма в отношении результатов исследований и отрасли в целом, побуждает к поиску альтернативных методов и технологий. Поведение электората пытаются трактовать на основе экспериментальных разработок и альтернативных методов изучения установок и мотивов поведенческих актов. Выбор стратегии избирательных кампаний основывается на имеющихся информационных ресурсах и базах данных избирателей. При прогнозировании итогов выборов все чаще используются статистическое моделирование и рыночные прогнозы.

С 1970-х гг. ученые разрабатывали макроэкономические статистические модели, используя агрегированные данные для прогнозирования результатов выборов, а с 1990-х гг. – различные конкурирующие модели. Большинство статистических моделей включают показатели деятельности правительства и экономики, хотя ведутся споры о том, какой конкретно экономический показатель следует использовать: рост ВВП, рост занятости, уровень инфляции или уровень жизни населения, субъективные оценки своего благосостояния. Например, согласно модели Р. Фейра рост реального дохода граждан в расчете на душу населения на 1 % позволяет получить дополнительные 1,6 % голосов избирателей, а уровень безработицы выше 6,7 % отрицательно сказывается на рейтинге кандидата. Также дискуссионным остается вопрос о включении в статистическую модель таких переменных, как валовое накопление, взгляды на военный бюджет, милитаризацию, ядерную энергетику. Однако все модели основаны на базовой теории общественного выбора, классических теориях электорального поведения и теории голосования, согласно которым медианный избиратель склонен переизбирать действующих должностных лиц в стабильных социально-экономических условиях, а в периоды экономического спада, социального кризиса или социальной трансформации чаще имеет место электоральный успех «новых» лиц и стратегическое голосование.

Существует мнение, что электоральные прогнозы, построенные на основе статистических моделей, превосходят в точности другие виды прогнозов итогов выборов. М. Льюис-Бек утверждает, что в долгосрочной перспективе статистическое моделирование покажет себя как более эффективный и точный метод прогнозирования по сравнению с другими

методами, базирующимися, например, на данных опросов общественного мнения и на фьючерсах кандидатов, не имеющими фундаментальной теоретической основы, а просто предоставляющими точечные оценки по зависимой переменной [9].

Сегодня используется множество статистических моделей, но постулировать о 100-процентной точности подобного прогноза нельзя. Например, модель Р. Фейра предоставляла верные прогнозы итогов нескольких выборов, а в 1992 г. дала худший результат. А в 2002 г. Р. Фейр представил новую математическую модель, позволяющую оценить шансы кандидата.

Одним из недостатков статистических прогнозных моделей является предсказание итогов народного голосования, а не результата избирательной кампании. Учитывая двухпартийную систему США, всенародное голосование обычно попадает в довольно узкий диапазон значений, ведь в течение довольно продолжительного времени доля голосов победителя незначительно отклоняется от 50-процентного значения (Дж. Байден – 51,3 %, Д. Трамп – 46,1 %, Б. Обама – 51,1 и 52,9 %, Дж. Буш – 50,7 и 47,9 %).

Еще один недостаток статистического моделирования заключается в широком диапазоне альтернатив, включая победу противоположного кандидата, и в необходимости большого массива агрегированных данных, который редко удается собрать. Более того, согласно Дж. П. Грину [10], тенденция выбора модели прогнозирования на основе прошлых выборов приводит к недооценке определенных факторов, влияющих на электоральный выбор. Л. Ваврек утверждает, что неточность прогнозов, базирующихся на экономических моделях, кроется в отсутствии учета специфики содержания избирательной кампании, программ и выступлений кандидатов, их взглядов на социально-экономическое развитие, геополитическую ситуацию и прочие общественно значимые вопросы и проблемы [11].

Другой альтернативой прогнозированию выборов являются «рынки предсказаний», такие как электронный рынок Айовы. Здесь люди покупают и продают фьючерсы кандидатов, основываясь на том, кто, по их мнению, победит на выборах. Например, в 2008 г. фьючерсы Б. Обамы продавались за 0,60 доллара, что указывало на то, что Б. Обама получит 60 % голосов на выборах. В 2020 г. фьючерсы Д. Трампа на бирже FTX продавались за 0,618 доллара, соответственно, 61,8 % пользователей FTX ожидали его переизбрания.

Преимущества рыночных прогнозов достаточно широко обсуждаются. Дж. Берг [12] утверждал, что предсказания, основанные на котировках фьючерсов кандидатов, точнее, чем базирующиеся на результатах электоральных социологических исследований. Р. С. Эриксон и К. Влезьен [7] доказали, что данное преимущество исчезает при агрегировании данных опросов. Д. Ротшильд [13] обнаружил, что при условии учета предубеждения, предрассудков и стереотипов массового сознания по точности они превосходят прогнозы, сделанные на основе агрегированных опросных данных, но по мере приближения дня выборов точность обоих увеличивается в разной мере.

Хотя на основе «рынков предсказаний» и статистических моделей порой можно получить более точные прогнозы итогов выборов, необходимо отметить, что электоральные опросы и рейтинги, составленные на их основе, играют косвенную роль при их формировании. Ведь многие статистические модели включают в себя уровень доверия кандидату и/или их партиям, их рейтинги и оценки их политических программ. А инвесторы электронных рынков и бирж принимают во внимание результаты электоральных опросов, делая ставки и приобретая фьючерсы кандидатов. Кроме того, даже идеально точный электоральный прогноз мало способствует пониманию поведения избирателей, специфики выбора кандидата или динамики мнений электората. В то время как представляется важным не только верно предсказать результаты выборов, но и объяснить, почему тот или иной кандидат одержал победу. А с этой задачей статистические прогнозные модели и рыночные прогнозы не справляются, в отличие от электоральных опросов.

#### **Список использованных источников**

1. *Kernell, S.* Life before Polls: Ohio Politicians Predict the Presidential Vote / S. Kernell // *Political Science and Politics*. – 2000. – Vol. 33. – P. 569–574.
2. *Smith, T. W.* The First Straw? A Study of the Origins of Election Polls / T. W. Smith // *Public Opinion Quarterly*. – 1990. – Vol. 54(1). – P. 21–36.
3. *Silver, N.* The Worst Pollster in the World Strikes Again / N. Silver // *Public Opinion Quarterly*. – 2009. – Vol. 73(5). – P. 845–865.
4. *Mosteller, F.* The Pre-Election Polls of 1948: Report to the Committee on Analysis of Pre-Election Polls and Forecasts / F. Mosteller, H. Hyman, P. McCarthy. – NY: Social Science Research Council, 1949.
5. *Lau, R.* An Analysis of the Accuracy of “Trial Heat” Polls in the 1992 Presidential Election” / R. Lau // *Public Opinion Quarterly*. – 1994. – Vol. 58. – P. 2–20.
6. *Hillygus, D.* The Persuadable Voter / D. Hillygus, T. Shields. – Princeton, NJ: Princeton University Press, 2008.
7. *Erikson, R. S.* Are Political Markets Really Superior to Polls as Election Predictions? / R. S. Erikson, C. Wlezien // *Public Opinion Quarterly*. – 2008. – Vol. 72. – P. 190–215.
8. *Hoek, J.* Factors Affecting Poll Accuracy: An Analysis of Undecided Respondents / J. Hoek, P. Gendall // *Marketing Bulletin*. – 1997. – Vol. 8. – P. 1–14.
9. *Lewis-Beck, M.* Election Forecasting: Principles and Practice / M. Lewis-Beck // *British Journal of Political Science*. – 2005. – Vol. 7(2). – P. 145–184.
10. *Greene, J. P.* Forewarned Before Forecast: President Election Forecasting Models / J. P. Greene // *Political Science and Politics*. – 1993. – Vol. 26. – P. 17–21.
11. *Vavreck, L.* The Message Matters: The Economy and Presidential Campaigns / L. Vavreck. – Princeton, NJ: Princeton University Press, 2009.
12. *Berg, J.* Results from a Dozen Years of Election Futures Market Research. In *Handbook of Experimental Economic Results* / J. Berg. – Amsterdam: Elsevier, 2001. – P. 742–751.
13. *Rothschild, D.* Forecasting Elections: Voter Intentions versus Expectations / D. Rothschild, J. Wolfers // *Public Opinion Quarterly*. – 2010. – Vol. 73(5). – P. 895–916.

(Дата подачи: 09.02.2022 г.)