

# МЕТОД УСЛОВНОЙ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ В СОВРЕМЕННОЙ ПОЛИТИЧЕСКОЙ НАУКЕ<sup>1</sup>

Седашов Е.А., Чернов Д.Н.

Национальный Исследовательский Университет «Высшая Школа Экономики»,  
Москва, Россия

esedashov@hse.ru, dnchernov@hse.ru

*Аннотация.* В данной статье рассматривается один из основных подходов к эмпирическому моделированию ситуаций качественного выбора альтернатив – условная логистическая регрессия МакФадена. Для иллюстрации особенностей метода приводятся примеры его применения для решения конкретных исследовательских задач в политической науке.

*Ключевые слова:* условная логистическая регрессия, электоральный выбор, формирование коалиций, интервенции в гражданские конфликты.

## Введение

В политической науке часто возникают задачи эмпирического моделирования качественного выбора альтернатив (qualitative choice). В техническом смысле качественный выбор альтернатив означает, что есть некоторое количество ситуаций для принятия решений (decision opportunities), в каждой из которых субъект (избиратель, государство, потребитель) может выбрать одну или несколько альтернатив. При этом предполагается, что субъекту принятия решения известен некоторый набор характеристик для всех альтернатив, а аналитику также известен данный набор и, в дополнение к этому, известны конкретные характеристики самих субъектов. Классическим примером ситуации качественного выбора альтернатив может служить голосование за партию на выборах: избирателю известны идеология партии, её текущая популярность, характеристики лидера партии, и решение о поддержке той или иной партии принимается на основе данной информации. Другой известной проблемой является выбор конкретного бренда внутри гомогенной товарной категории.<sup>2</sup>

В данной статье рассматривается один из основных подходов к эмпирическому моделированию качественного выбора альтернатив, а именно условная логистическая регрессия МакФадена [1]. Основной фокус статьи – прикладное применение данной модели для решения исследовательских задач в политической науке. Будут рассмотрены задачи эмпирического моделирования выбора политической партии, формирования правительственной коалиции и вмешательства третьих сторон в гражданские конфликты. Каждой из задач будет посвящен свой подраздел.

## 1. Условная логистическая регрессия в задачах моделирования партийного выбора

Условная логистическая регрессия прочно вошла в эмпирический инструментарий исследователей электорального поведения (см. [2, 3]). Причина популярности подхода в рамках данной задачи будет более понятна после рассмотрения технических деталей. Предположим, что некоторому избирателю  $i$  предоставляется «меню» из  $j \in \{1, 2, \dots, J\}$  возможных политических партий, среди которых он должен выбрать одну конкретную партию.<sup>3</sup> Избирателю известны релевантные характеристики партий, и вероятность выбора, таким образом, может быть записана как:

$$\pi_{ik} = \frac{e^{X_{ik}\mathbf{b}}}{\sum_{j=1}^J e^{X_{ij}\mathbf{b}}}, \quad (1)$$

где  $k$  – конкретная партия из «меню»,  $i$  – индекс конкретного избирателя,  $X$  – матрица, в которой хранятся характеристики партий и избирателей, индекс  $ik$  определяет конкретную строку в матрице, соответствующую избирателю  $i$  и партии  $k$ ,  $\mathbf{b}$  – вектор-столбец, содержащий регрессионные коэффициенты. Данный подход к определению вероятности выбора партии обладает рядом привлекательных свойств. Во-первых, он позволяет определить расстояния между позициями

<sup>1</sup> Статья подготовлена при поддержке гранта Российского научного фонда, реализуемого в МГИМО МИД России (грант № № 22–18–00664–П, <https://rscf.ru/project/22-18-00664/>)

<sup>2</sup> Например, выбор бренда смартфона.

<sup>3</sup> Базовым предположением условной логистической регрессии является ИА (independence of irrelevant alternatives, независимость нерелевантных альтернатив): иерархия предпочтений избирателя не изменится, если из «меню» партий убрать некоторые из них.

избирателя и партии в разных измерениях политического пространства, а затем включить эти расстояния в эконометрическую модель в качестве предикторов.<sup>4</sup> На данный момент общепринятыми подходами к определению данных расстояний являются расчёт Евклидова расстояния между позицией избирателя и позицией партии [4, 5] или скалярного произведения позиционных векторов избирателя и партии [6]. Во-вторых, появляется возможность включения в модель других важных партийных характеристик: идентификация избирателя с конкретной партией [7] и её электоральная сила [8]. Итоговая модель позволяет взглянуть на процесс выбора партии как на процесс сравнения альтернатив друг с другом, как с точки зрения близости партийной идеологии, так и с точки зрения возможности партии получить места в парламенте. Функция правдоподобия для данной модели определяется следующим уравнением:

$$L(\mathbf{b}|X) = \prod_{i=1}^I \prod_{j=1}^J \pi_{ij}^{y_{ij}}, \quad (2)$$

где  $y_{ij} = 1$ , если избиратель  $i$  выбрал партию  $j$ , в остальных случаях переменная равна 0. Оценка параметров регрессии, задаваемых вектором  $\mathbf{b}$ , производится с помощью стандартных процедур, принятых в методе максимального правдоподобия. Задача описывается следующим образом:

$$\hat{\mathbf{b}} = \arg \max_{\mathbf{b}} L(\mathbf{b}|X) \quad (3)$$

Отметим, что в недавно опубликованных работах [13, 15] представленный подход был подвергнут обоснованной критике. Наиболее уязвимой точкой подхода считается предположение о равенстве множеств выбора (consideration sets, choice sets): каждый избиратель рассматривает все партии как возможные опции для выбора. В действительности процесс устроен не так: некоторые партии могут изначально исключаться избирателем из спектра возможных объектов поддержки, потому что находятся слишком далеко от его позиции в политическом пространстве, не претендуют на распределение мест в парламенте или обладают общей плохой репутацией. Таким образом, разные избиратели выбирают из разных, хотя и возможно пересекающихся, «меню», и эмпирическая модель должна учитывать данный факт. [13] предложили достаточно элегантное обобщение условной логистической регрессии для данной задачи, добавив дополнительный параметр в функцию вероятности выбора:

$$\pi_{ik} = \frac{\varphi_{ik} e^{X_{ik}\mathbf{b}}}{\sum_{j=1}^J \varphi_{ik} e^{X_{ij}\mathbf{b}}}, \quad (4)$$

где  $\varphi_{ik}$  – вероятность того, что партия  $k$  входит в «меню» приемлемых альтернатив для избирателя  $i$ . Для параметризации  $\varphi_{ik}$  используется стандартная логистическая функция:

$$\varphi_{ik} = \frac{1}{1 + \exp(-Z_{ik}\gamma)}, \quad (5)$$

где  $Z_{ik}$  – матрица переменных, влияющих на включение партии  $k$  в «меню» избирателя  $i$ ,  $\gamma$  – вектор-столбец параметров регрессии. В матрицу  $Z$  обычно входят электоральная сила партии (т.е. ожидаемый процент на выборах по данным предвыборных опросов), экстремальность партийной идеологии и факт идентификации избирателя с партией (strong affinity to a party). Фактически, данный подход позволяет учесть нежизнеспособность некоторых партий на выборах за счёт дополнительного штрафного параметра. Как будет видно ниже, такой подход имеет применение и в других исследовательских задачах.

## 2. Условная логистическая регрессия в задачах моделирования внешней поддержки участников гражданских конфликтов

Поддержка участников гражданских конфликтов также хорошо вписывается в область применения условной логистической регрессии, но данная модель до настоящего времени не использовалась в профильных исследованиях. В рамках общепринятого в литературе подхода обычно используется стандартная логистическая регрессия на диадных данных, где в качестве диады выступает потенциальная страна-спонсор и потенциальный объект поддержки (правительство или повстанческая группировка) в конкретный год (см. [9]). У данного подхода есть очевидные проблемы. Во-первых, в таком подходе не предусмотрено изначального отсеивания из массива государств, которые никогда не оказывают поддержки участникам гражданских конфликтов в других странах. Другими словами,

<sup>4</sup> Данные предикторы будут элементами  $X_{ik}$ .

поддержка участников зарубежного военного конфликта не входит во внешнеполитический репертуар таких государств. Во-вторых, внешняя поддержка участника конфликта – это именно стратегическое решение, а значит и эмпирический анализ должен отражать данную логику. Последующее обсуждение предполагает, что участник конфликта – повстанческая группировка, но логика поддержки правительств подчиняется тем же принципам.

В контексте условной логистической регрессии вероятность того, что страна  $i$  поддержит участника  $k$  некоторого конфликта в конкретный год  $t$  задаётся в виде:

$$\pi_{ikt} = \frac{e^{X_{ikt}\mathbf{b}}}{\sum_{j \in K_t} e^{X_{ij}\mathbf{b}}}, \quad (6)$$

где  $K$  – множество всех когда-либо существовавших повстанческих группировок,  $K_t$  – подмножество этого множества, соответствующее группировкам, активно ведущим деятельность в год  $t$ ; вектор  $\mathbf{b}$ , как и раньше, содержит коэффициенты регрессии. Матрица  $X$  включает в себя разнообразные структурные характеристики государства, в котором ведёт операции группировка  $k$ , характеристики самого конфликта и повстанческой группировки, а также геополитические факторы. К структурным характеристикам обычно относят соотношение уровней экономического развития и режимных характеристик, интенсивность конфликта [10] и сила повстанческой группировки [11] также считаются важными предикторами внешней поддержки. Наконец, в число важных геополитических предикторов поддержки входят расстояние [9] и наличие альянса со страной конфликта [9]. Условная логистическая регрессия позволяет рассматривать решение о внешней поддержке в контексте наличия разнообразных альтернатив, из которых страна-спонсор выбирает ту, которая лучше всего соответствует её геополитическим и экономическим интересам. Другим преимуществом является тот факт, что ситуации, когда страна-спонсор  $i$  не поддерживала ни одну повстанческую группировку в конкретный год  $t$ , автоматически удаляются из массива.<sup>5</sup>

Заметим, что уравнение (6) следует логике уравнения (1), но тип и структура данных в данной задаче существенно отличаются от первой. Наиболее важным является то, что список повстанческих группировок, из которых государства-спонсоры выбирают объекты поддержки, может быть (и часто бывает) разным для разных лет: одни группировки выбывают из списка в силу окончания конфликта или в силу упразднения, другие появляются в силу начала новых конфликтов. В первой задаче список альтернатив и их характеристики (партийные позиции, электоральная жизнеспособность партий и т.д.) был одинаковым для всех избирателей.

### 3. Условная логистическая регрессия в задачах моделирования формирования коалиций

Эмпирическое моделирование формирования коалиций в парламентских республиках является, пожалуй, самой неочевидной областью применения условной логистической регрессии. В данной задаче нет конкретного субъекта, принимающего решения (избиратель или государство-спонсор), вместо этого партии, присутствующие в парламенте, пытаются договориться о формировании правительства, которое будет поддержано большинством голосов депутатов.<sup>6</sup> Каким образом условная логистическая регрессия вписывается в данную картину? [12] дают детальные инструкции, касающиеся данного вопроса. Базовая идея достаточно проста: в рамках каждого момента, соответствующего переговорам о формировании коалиции, может быть сформировано  $2^N - 1$  коалиций, где  $N$  – количество партий, имеющих места в парламенте. Каждой из этих потенциальных коалиций соответствуют конкретные переменные, а именно идеологическая гомогенность, наличие контроля большинства голосов в парламенте, присутствие медианного законодателя во фракции партии-члена коалиции, наличие электорального альянса, количество коалиций идентичной структуры, сформированных в прошлые периоды. Фактически, мы получаем набор данных, в которых для каждого момента времени, соответствующего переговорами о создании коалиции, есть идентификатор реально

<sup>5</sup> Для понимания логики данной процедуры можно обратиться к модифицированной версии формулы (2). Ситуация, когда страна-спонсор  $i$  не поддержала ни одной группировки в год  $t$  соответствует кейсу  $y_{ikt} = 0$  для всех  $k$  в год  $t$ . Это, в свою очередь, означает, что данный конкретный спонсор в данный конкретный год не внёс вклада в функцию правдоподобия (умножение на единицу).

<sup>6</sup> Существует достаточно большое количество нюансов, связанных с формальными способами получения полномочий коалициями в разных странах. С целью экономии пространства мы рассматриваем упрощенный сценарий, в котором поддержка большинства – обязательное требование для любой коалиции.

сформированной коалиции. Вероятность того, что коалиция  $i$  получит власть в переговорный момент  $m$ , таким образом, может быть записана как:

$$\pi_{im} = \frac{e^{X_{im}b}}{\sum_{j \in C_m} e^{X_{jm}b}}, \quad (7)$$

где  $C_m$  – множество всех возможных коалиций, которые могут быть сформированы в переговорный момент  $m$ ,  $X$  содержит переменные, описанные выше. Снова мы можем заметить, что концептуальная структура подхода остаётся неизменной, несмотря на значительную содержательную специфику.

В контексте такого подхода к эмпирическому моделированию формирования коалиций возникает проблема «невозможных коалиций», похожая на уже обозначенную в разделе 1 проблему нежизнеспособных партий. Часто, несмотря на гипотетическую возможность, коалиция не может быть сформирована из-за институциональных барьеров (как правило, это требование поддержки сформированной коалиции парламентским большинством) или возможного нарушения взятых на себя обязательств (например, созданные в ходе выборов альянсы, которые партии стремятся сохранять и в процессе формирования правительственных коалиций). Таким образом, так же, как и в случае с моделированием партийного выбора, моделирование формирования коалиций требует учёта присутствия в массиве таких «невозможных» коалиций.<sup>7</sup> [14] представили именно такой подход, адаптировав метод из [13]. В разработанном подходе так же, как и в случае с партиями, «невозможные» коалиции получаю дополнительный штраф.

#### 4. Заключение

В данной короткой статье были рассмотрены три задачи эмпирического моделирования, решение которых может быть осуществлено с помощью условной логистической регрессии МакФадена, а также показаны содержательные преимущества данного подхода. Кроме того, мы показали некоторые важные тренды развития моделей качественного выбора альтернатив, сделав основной акцент на моделировании множеств выбора (множеств жизнеспособных альтернатив). Наш доклад показывает высокую степень универсальности подхода МакФадена, т.к. подход не только доказал свою полезность в моделировании ситуаций качественного выбора альтернатив, но и показал значительный потенциал для обобщений и адаптации к конкретным задачам. Модель условной логистической регрессии нашла своё применение в подчас совершенно неочевидных для этого областях (формирование коалиций – яркий пример), и с высокой долей уверенности продолжит находить подобные неожиданные применения для решения исследовательских задач в политической науке.

#### Литература

1. *McFadden D.* Conditional logit analysis of qualitative choice behavior // *Frontiers in Econometrics*. – New York: Academic Press, 1974. – P. 105–142.
2. *Adams J., Merrill S.* Modeling party strategies and policy representation in multiparty elections: why are strategies so extreme? // *American Journal of Political Science*. – 1999. – Vol. 43, № 3. – P. 765–791.
3. *Bargsted M.A., Kedar O.* Coalition-targeted duvergerian voting: how expectations affect voter choice under proportional representation // *American Journal of Political Science*. – 2009. – Vol. 53, № 2. – P. 307–323.
4. *Downs A.* An economic theory of democracy. – New York: Addison-Wesley Publishing, 1957.
5. *Enelow J.M., Hinich M. J.* The spatial theory of voting: an introduction. – New York: Cambridge University Press, 1984.
6. *Rabinowitz G., Macdonald S.E.* A directional theory of issue voting // *American Political Science Review*. – 1989. – Vol. 83, № 1. – P. 93–121.
7. *Green D.P., Palmquist B., Schickler E.* Partisan hearts and minds: political parties and the social identities of voters.
8. *Alvarez M.R., Nagler J.* A new approach for modeling strategic voting in multiparty elections // *British Journal of Political Science*. – Vol. 30, № 1. – P. 57–75.
9. *Koga J.* Where do third parties intervene? Third parties' domestic institutions and military interventions in civil conflicts // *International Studies Quarterly*. – 2011. – Vol. 55, № 4. – P. 1143–1166.
10. *Reagan P.* Choosing to intervene: outside interventions in internal conflicts // *Journal of Politics*. – 1998. – Vol. 60, № 3. – P. 754–779.
11. *Gent S.* Going in when it counts: military intervention and the outcome of civil conflicts // *International Studies Quarterly*. – 2008. – Vol. 52, № 4. – P. 713–735.

---

<sup>7</sup> *Может возникнуть вопрос, почему такие коалиции не могут быть просто убраны из массива для анализа. Проблема заключается в том, что данные критерии «нежизнеспособности» не являются жесткими: коалиции без поддержки большинства иногда создаются, как и коалиции, нарушающие избирательные альянсы.*

12. *Martin L., Stevenson R.* Government formation in parliamentary democracies // *American Journal of Political Science*. – 2001. – Vol. 45, № 1. – P. 33–50.
13. *Moral M., Zhirnov A.* Issue voting as a constrained choice problem // *American Journal of Political Science*. – 2018. – Vol. 62, № 2. – P. 280–295.
14. *Sedashov E., Zhirnov A., Moral M.* Obstacles to executive power: institutional and policy constraints on government formation // Working Paper.
15. *Steenbergen M., Willi T.* What consideration sets can teach us about electoral competition: a two-hurdle model // *Electoral Studies*. – 2019. – Vol. 57. – P. 263–274.