

Е.А. СЕДАШОВ*

МЕТОДЫ КАУЗАЛЬНОГО АНАЛИЗА В СОВРЕМЕННОЙ ПОЛИТИЧЕСКОЙ НАУКЕ

Аннотация. В настоящей статье описываются методы каузального анализа, наиболее популярные в современной политической науке. Мы не фокусируемся на технических деталях, стоящих за каждым методом, а описываем основные идеи с целью сделать их более доступными для широкой аудитории исследователей. Мы также описываем алгоритм исследовательского дизайна для каждого метода.

Во-первых, мы анализируем общую мотивацию каузального анализа. Мы обсуждаем, как проблема каузальности возникает в тестировании гипотез и показываем эту проблему на примере взаимосвязи демократии и экономического развития. Во-вторых, мы иллюстрируем общую проблему каузальности в рамках каузальной модели Рубина (КМР). Мы представляем все основные понятия данной модели и затем демонстрируем, как КМР формулирует общую проблему каузальности. В-третьих, мы описываем наиболее часто используемые методы каузального анализа: рандомизированные эксперименты, метод разрывной регрессии, метод difference-in-difference и инструментальные переменные. Для каждого метода мы даем читателю общее описание, а также шаги исследовательского дизайна. Мы также приводим примеры известных исследований, в которых применяется тот или иной метод, и кратко обсуждаем достоинства и недостатки каждого метода. Читатель, вооруженный этим знанием, сможет найти метод, наиболее подходящий к его исследовательской проблеме. В заключение мы обсуждаем, как идеи каузального анализа могут применяться в качественных исследованиях, в частности в кейс-стади, и делаем вывод о важности этих идей как для количественных, так и для качественных политических исследований.

* Седашов Евгений Александрович, PhD, доцент, Департамент политики и управления, Национальный исследовательский университет «Высшая школа экономики» (Москва, Россия), e-mail: esedashov@hse.ru

Ключевые слова: каузальный анализ; каузальная модель Рубина; рандомизированный эксперимент; регрессионный разрыв; инструментальные переменные; матчинг.

Для цитирования: Седашов Е.А. Методы каузального анализа в современной политической науке // Политическая наука. – 2021. – № 1. – С. 98–115. – DOI: <http://www.doi.org/10.31249/poln/2021.01.04>

Введение

Проблема каузальности (causal inference) лежит в основе наиболее известных вопросов позитивной политической науки. Взаимосвязь демократии и экономического развития выступает в качестве яркого примера. С одной стороны, в ряде известных статей постулируется положительный эффект экономического развития, оказываемый на оформление демократических институтов [Lipset, 1959; Bates, Lien, 1985; North, Weingast, 1989]. С другой стороны, демократию часто видят ключевой предпосылкой экономической развитости государства [Acemoglu, Robinson, 2012; Democracy does cause growth, 2019]. Эта проблема известна в политической науке как проблема обратной каузальности (reverse causality problem), и ее эффективное разрешение невозможно без применения специализированных статистических методов.

Целью данной статьи является обзор подходов, позволяющих находить надежные каузальные связи между политическими феноменами. Мы постарались рассмотреть наиболее популярные в современных исследованиях методы. Экспозиция каждого метода сопровождается кратким обзором статей, в которых он применяется. Мы постарались организовать материал так, чтобы область применения каждого метода была интуитивно понятна читателю. Оставшийся текст поделен на четыре раздела: в первой части описывается «каузальная модель Рубина» [Imbens, Rubin, 2015]; во второй части рассматриваются рандомизированные эксперименты; третья часть посвящена естественным экспериментам и квазиэкспериментальным подходам; четвертая часть является заключением.

Каузальная модель Рубина

Каузальная модель Рубина [Imbens, Rubin, 2015] (далее – КМР) выступает в качестве одного из ключевых подходов к определению проблемы каузальности¹. Прежде чем перейти непосредственно к рассмотрению проблемы каузальности, мы опишем несколько базовых понятий КМР. Первым ключевым понятием является *экспериментальная единица* – объект исследовательских действий. Например, в медицине экспериментальной единицей может быть пациент-испытуемый; в опросных экспериментах экспериментальной единицей является респондент. Вторым важным понятием является *вмешательство* (treatment) – исследовательское действие, производимое над экспериментальными единицами. В стандартном исследовательском дизайне, как правило, применяются два типа вмешательства: *активное вмешательство* (например, в опросном эксперименте интервьюер читает респонденту определенную информацию перед тем как задать вопросы) и *пассивное, или нулевое, вмешательство* (респонденту не предоставляется информация, которая может повлиять на его ответы). При этом группу экспериментальных единиц, к которым применяется активное вмешательство, называют *экспериментальной группой*, а экспериментальные единицы с пассивным вмешательством – *контрольной группой*. Третьим ключевым понятием является *потенциальный исход* – результат применения определенного типа вмешательства для конкретной экспериментальной единицы в определенный момент времени. Если в исследовании определено два типа вмешательства, то и потенциальных исходов для каждой экспериментальной единицы будет тоже два. Таблица 1 иллюстрирует, как три ключевых понятия КМР связаны между собой. Экспериментальной единицей в данном случае является пациент; для каждого пациента определено два типа вмешательства, препарат и плацебо, и два потенциальных исхода, соответствующие двум типам вмешательства. Как видно из таблицы, для каждого пациента наблюдаем только один потенциальный исход из двух возможных.

¹ КМР не является единственным подходом к определению проблемы каузальности. Например, в последние несколько лет все большую популярность приобретает подход «каузальных графов» [Pearl, 2009]. С нашей точки зрения, КМР лучше подходит для описания общей проблемы каузального анализа, но подход «каузальных графов» также имеет свои уникальные преимущества.

Определив основные понятия КМР, мы можем перейти к рассмотрению непосредственно проблемы каузальности.

КМР определяет *каузальный эффект* как сравнение потенциальных исходов для конкретной экспериментальной единицы [Imbens, Rubin, 2015, p. 6–7]. Исходя из этого определения и табл. 1, проблему каузальности можно сформулировать как проблему отсутствующих данных: так как в конкретный момент времени наблюдаем только один потенциальный исход, прямое вычисление каузального эффекта является невозможным. Все методы, рассматриваемые в данной статье, так или иначе направлены на решение данной проблемы.

Таблица 1

Потенциальные исходы и экспериментальные единицы

	Потенциальный исход 1: Препарат	Потенциальный исход 2: Плацебо
Пациент 1	Улучшение состояния	-
Пациент 2	-	Состояние не изменилось
Пациент 3	-	Состояние ухудшилось
Пациент 4	Состояние не изменилось	-

Рандомизированные эксперименты

Рандомизированные эксперименты являются наиболее «чистым» решением проблемы каузальности. Общий алгоритм исследовательского дизайна для рандомизированных экспериментов состоит из нескольких последовательных шагов. Во-первых, исследователь определяет генеральную совокупность. Во-вторых, на основе теоретической рамки и гипотез исследования определяются типы вмешательства. В-третьих, разрабатывается способ измерения результатов эксперимента. В-четвертых, разработанный исследовательский дизайн и основные гипотезы пререгистрируются. Конечной стадией исследования выступает анализ полученных данных с помощью стандартных статистических техник. Ниже мы проследим непосредственную реализацию каждого из этих шагов.

Определение генеральной совокупности является первым шагом исследовательского дизайна и, как правило, вытекает из

общих рамок исследования. В качестве генеральной совокупности могут выступать, например, деревенские поселения в Индии [Chat-topadhyay, Duflo, 2004], индонезийские деревни, участвующие в государственной программе грантов на развитие инфраструктуры [Olken, 2007], московские УИК на выборах в Госдуму в 2011 г. [Field experiment estimate ..., 2013], население Российской Федерации [Frye, 2019] и т.п. Генеральная совокупность определяет не только границы, на которые распространяются выводы эксперимента, но и экспериментальные единицы – ключевой компонент КМР.

После определения генеральной совокупности следует стадия определения типов вмешательства. Типы вмешательства, как правило, вытекают из основных гипотез исследования. Например, в исследовании [Olken, 2007] анализируется эффективность различных способов борьбы с коррупцией, возникающей в ходе реализации грантов на развитие инфраструктуры, и в частности эффективность прямого аудита. Для оценки эффективности прямого аудита выборка была разбита на две группы: в деревнях из экспериментальной группы было публично объявлено об ожидающихся визитах аудиторов для оценки реализации проекта, а в деревнях из контрольной группы таких объявлений не делалось. Помимо теоретических рамок и гипотез исследования, характеристики экспериментальных единиц также прямым образом влияют на выбор типов вмешательства. В качестве примера можно привести уровень образования респондентов. Предположим, что исследователь хочет выяснить, возможно ли снизить риск распространения туберкулеза путем информирования населения о болезни и способах ее профилактики. Эксперимент предполагается проводить в Эфиопии – стране с высоким уровнем заболеваемости туберкулезом и одновременно низким средним уровнем грамотности. Типы вмешательства должны учитывать данные обстоятельства. Так, распространение памфлетов с текстом о туберкулезе будет неэффективным, так как многие жители не умеют читать. Вместе с тем визуальная информация вроде картинок может хорошо восприниматься людьми с самым разным уровнем грамотности. Стоит заметить, что использование преимущественно картинок как обучающего механизма несет в себе дополнительные сложности. Во-первых, дизайн картинок необходимо сделать простым и в то же время информативным. Во-вторых, необходимо убедиться, что между испытуемыми нет систематических различий в восприятии

информации на картинках – существование подобных различий затруднит интерпретацию результатов эксперимента. Для устранения подобных различий, как правило, проводятся пилотные исследования, в рамках которых тестируются различные варианты дизайна картинок и их восприятие испытуемыми.

После определения типов вмешательства исследователем разрабатывается механизм измерения результатов эксперимента. В зависимости от типа исследования замер может производиться как сразу же после осуществления вмешательства, так и по истечении определенного времени. В опросных экспериментах, как правило, результат замеряется в ходе самого опроса. Например, в эксперименте [Frye, 2019] интервьюер просил респондентов оценить свое отношение к российскому руководству по 5-балльной шкале. В экспериментальных оценках экономических и политических программ замер результатов может значительно отстоять во времени от момента осуществления вмешательства. Так, в эксперименте [Olken, 2007] объявление о планируемой аудиторской проверке производилось в начале реализации гранта, а замер коррупции происходил по итогам реализации гранта: эксперты оценивали качество и стоимость проведенных работ, а затем эти оценки сравнивались с отчетностью, предоставленной самими деревнями.

Четвертым этапом реализации программы рандомизированного эксперимента является пререгистрация исследовательского дизайна и основных гипотез. Сама по себе пререгистрация исследовательских дизайнов не является новой идеей; в социальных науках, однако, пререгистрация только становится официально признанным мейнстримом. Основным аргументом сторонников пререгистрации является проблема «фишинга», т.е. проведения множества статистических тестов для поиска статистически значимых связей и последующей «подгонки» гипотез под эти связи. В исследованиях, прошедших пререгистрацию, осуществлять фишинг гораздо сложнее, так как основные гипотезы и исследовательские процедуры представляются на суд академического сообщества до сбора данных [Humphreys, Sanchez de la Sierra, van der Windt, 2013; Monogan, 2015; Gelman, 2013]. В качестве дополнительных аргументов в пользу пререгистрации приводятся четкое разделение исследований на индуктивные и дедуктивные и возможность избежать влияния профессиональной и карьерной мотивации на публикуемые результаты [Monogan, 2015]. В целом поль-

за пререгистрации экспериментальных исследований очевидна. Вместе с тем *что именно* необходимо пререгистрировать является предметом дебатов. «Минималистский» подход выступает за пререгистрацию основных гипотез исследования и процедур по сбору данных. На другой стороне спектра [Humphreys, Sanchez de la Sierra, van der Windt, 2013] выступают не просто за пререгистрацию, а за написание полноценной статьи с использованием симулированных данных (mock report). Такой подход в хорошем смысле связывает руки исследователю, так как заранее известны не только гипотезы и методы сбора данных, но и спецификации моделей, и аналитические методы. По сути, после сбора данных исследователю остается только проанализировать реальные данные и вставить их в таблицы.

Финальным этапом экспериментального исследования является анализ полученных данных. Наиболее простым и интуитивным способом анализа в данном случае является Average Treatment Effect, или АТЕ [см.: Imbens, Rubin, 2015, p. 83–112] – разница в средних результатах эксперимента между экспериментальной и контрольной группами¹. В качестве основного недостатка АТЕ можно выделить возможную несбалансированность характеристик двух сравниваемых групп. Для оценки несбалансированности характеристик двух сравниваемых групп по признаку X (например, возраст) можно применить следующую формулу [Imbens, Rubin, 2015, p. 311]:

$$\Delta_{ct}^{\wedge} = \frac{\bar{X}_t - \bar{X}_c}{\sqrt{(s_t^2 + s_c^2)/2}}$$

где субиндексы t и c обозначают экспериментальную и контрольную группу соответственно, s обозначает стандартное отклонение, а верхняя горизонтальная линия обозначает среднее арифме-

¹ Либо двумя произвольными группами, если типов вмешательства больше двух.

тическое¹. Если обнаружена значительная несбалансированность между сравниваемыми группами по одному или нескольким ключевым признакам, применение АТЕ может быть проблематичным, и необходимо обратиться к регрессионным методам. Для интервальных данных, как правило, применяется стандартная МНК-регрессия, в то время как для ординальных или номинальных данных применяются модели дискретного выбора (например, логит-регрессия или упорядоченная логит-регрессия).

Естественные эксперименты и квазиэкспериментальные подходы

Несмотря на свои очевидные преимущества, проведение рандомизированных экспериментов может быть невозможно в силу высоких затрат или этических ограничений. В подобных обстоятельствах исследователи вынуждены опираться на поиск естественных экспериментов и применение квазиэкспериментальных подходов. Основная цель данных методов – приближение анализа к стандарту рандомизированного эксперимента и максимально возможное удовлетворение критерия внутренней валидности². Поиск данных, подходящих для таких методов, – непростая задача, требующая от исследователя изрядной креативности и эрудиции. Мы проведем краткий обзор наиболее часто используемых подходов и, как и в случае рандомизированных экспериментов, проанализируем практику их применения.

Метод разрывной регрессии. Первым методом, заслуживающим упоминания, является метод разрывной регрессии (regression discontinuity design). Данный метод ставит целью максимизацию внутренней валидности исследования за счет возмож-

¹ Обозначение «[^]» (хэт) в данном случае обозначает степень несбалансированности и указывает на то, что величина является статистической оценкой, а не величиной, рассчитываемой для генеральной совокупности. Для более подробного описания методов анализа баланса характеристик сравниваемых групп см.: [Imbens, Rubin, 2015, p. 310–317].

² Чтобы соответствовать критерию внутренней валидности, исследовательский дизайн должен максимально изолировать эффект независимой переменной на зависимую переменную. Внешняя валидность определяет степень общности выводов исследования.

ного уменьшения внешней валидности. Основные идеи исследовательского дизайна, применяющего разрывную регрессию, легко описать в виде нескольких последовательных шагов.

Во-первых, необходима переменная, разделяющая данные на экспериментальную и контрольную группы. Довольно часто в качестве такой переменной служит граница определенной территории. Например, в исследовании [Dell, 2010] сравниваются две группы географически близких поселений: первая группа находится на бывшей территории Мита¹, а вторая группа – вне ее пределов. В работе [Ferwerda, Miller, 2014] сравнивается интенсивность сопротивления нацистам в оккупированной и вишистской Франции. В качестве механизма разделения на экспериментальную и контрольную группы здесь служит установленная нацистами географическая граница между оккупированными территориями и вишистской Францией. Граница, где происходит «разрыв», может быть и более фигуральной (пример – использование регрессионного разрыва в исследовании [Lee, 2001] для измерения эффекта инкумбента).

Во-вторых, необходимо определить так называемый интервал сравнения (bandwidth). Так как основной целью метода разрывной регрессии является максимизация внутренней валидности, идеальной является ситуация, при которой наблюдения не отличаются друг от друга ничем, кроме принадлежности к экспериментальной и контрольной группам. Очевидно, что этот идеал недостижим на практике – целью интервала сравнения как раз является максимально возможное приближение к этому идеалу. Вернемся к исследованию [Dell, 2010]. Для того чтобы четко выявить каузальный эффект Мита на потребление домохозяйств и замедленный рост у детей, необходимо сравнить поселения, идентичные во всем, кроме принадлежности к территории Мита. Логично предположить, что степень схожести поселений тем выше, чем они ближе к границе Мита. В работе [Dell, 2010] представлены результаты для трех интервалов сравнения: первый интервал включает поселения, расположенные на расстоянии < 50 км от границы Мита, второй интервал включает поселения на расстоянии < 75 км от

¹ Мита – институт принудительного труда, существовавший в колониальных Перу и Боливии, согласно которому определенная часть коренного населения была обязана работать на серебряных и ртутиевых шахтах.

границы Мита, третий интервал включает все поселения ближе 100 км от границы Мита. Наиболее валидным является первый интервал, так как поселения в нем находятся ближе друг к другу, но устойчивость результатов для второго и третьего интервалов только усиливает выводы статьи. Определение правильного размера интервала сравнения – открытый методологический вопрос [Imbens, Lemieux, 2008; Imbens, Kalyanaraman, 2012], поэтому стандартом в эмпирических исследованиях является представление результатов для нескольких интервалов сравнения.

Финальным шагом реализации исследовательского дизайна является анализ наблюдений, попавших в заданный интервал сравнения. Здесь есть несколько возможных альтернатив. Наиболее простым подходом является сравнение средних значений зависимой переменной для экспериментальной и контрольной групп. Еще одним относительно простым подходом является полиномиальная регрессия, включающая бинарный индикатор принадлежности к экспериментальной группе в качестве основной независимой переменной. Стандартное регрессионное уравнение выглядит следующим образом:

$$y_i = \alpha + \beta \text{Группа}_i + \gamma_1 \text{Расстояние}_i + \gamma_2 \text{Расстояние}_i^2 + \epsilon_i,$$

где в качестве переменной *Расстояние* выступает расстояние до точки «разрыва» (географическая или условно определенная граница), а в качестве переменной *Группа* выступает индикатор принадлежности к экспериментальной группе. В исследованиях могут использоваться и полиномы расстояния более высоких степеней (например, [Dell, 2010] использует кубический полином расстояния); в работе [Gelman, Imbens, 2018], однако, не рекомендуется использовать полиномы со степенью выше квадрата. Альтернативой полиномиальной регрессии является подход локальной линейной регрессии [Imbens, Lemieux, 2008] с использованием переменной взаимодействия:

$$y_i = \alpha + \beta_1 \text{Группа}_i + \beta_2 \text{Расстояние}_i + \beta_3 \text{Расстояние}_i \text{Группа}_i + \epsilon_i.$$

Такой подход используется в исследовании [Ferwerda, Miller, 2014]. В работе [Gelman, Imbens, 2018] также отмечают преимущества данного подхода.

Method difference-in-difference. Аналогично методу разрывной регрессии основной задачей метода difference-in-difference

является максимизация внутренней валидности исследования. Как и разрывная регрессия, difference-in-difference опирается на естественные эксперименты. Структура этих естественных экспериментов существенно отличается от используемой в разрывной регрессии.

Во-первых, для каждой экспериментальной единицы должны быть доступны наблюдения за два временных промежутка. Например, в известном исследовании [Card, Krueger, 1994] используется два периода – до повышения в Нью-Джерси минимального размера оплаты труда и после повышения. Во-вторых, и контрольная, и экспериментальная группы в начальный период времени t не получают активного вмешательства (сам период так и называется – *pre-treatment period*). В-третьих, экспериментальная группа получает вмешательство в момент времени $t+1$, а контрольная группа – нет (период $t+1$ называется *post-treatment period*). В работе [Card, Krueger, 1994] в качестве контрольной группы взяли соседний с Нью-Джерси штат Пенсильвания, в котором не производилось повышения минимальной оплаты труда. В-четвертых, в рамках выбранного естественного эксперимента должно соблюдаться предположение о параллельных трендах (*common trends assumption*) [Angrist, Pischke, 2008, p. 171–172]. Суть данного предположения состоит в том, что в отсутствие вмешательства разница между экспериментальной и контрольной группами в момент t и в момент $t+1$ должна быть одинаковой. С учетом данного предположения становится понятна логика подбора контрольной группы в исследовании [Card, Krueger, 1994]: штат Пенсильвания граничит с Нью-Джерси, поэтому можно предположить, что в отсутствие повышения минимального размера оплаты труда безработица в двух штатах развивалась бы по схожему сценарию. Насколько оправдана данная географическая логика в подборе контрольной группы? Указывается [Angrist, Pischke, 2008, p. 172], что при пристальном анализе трендов безработицы в двух штатах можно обнаружить существенную вариацию: например, безработица в Нью-Джерси повышалась между июнем и октябрём 1995 г., но в то же время снижалась в Пенсильвании. Данный пример свидетельствует в пользу необходимости количественного анализа трендов, так как логичное, на первый взгляд, предположение об их схожести в географически близких территориях может на поверку оказаться не совсем верным.

После выбора подходящего естественного эксперимента исследователь оценивает следующую регрессию:

$$y_{it} = \alpha + \beta_1 * \text{Группа}_i + \beta_2 * \text{Период}_t + \beta_3 * \text{Группа}_i \text{Период}_t + \epsilon_{it},$$

где *Группа* – бинарный индикатор экспериментальной и контрольной групп, а *Период* – бинарный индикатор временного промежутка (до или после получения вмешательства). Индекс *i* обозначает экспериментальную и контрольную группы, а индекс *t* – *pre-treatment* и *post-treatment* временные периоды.

Метод инструментальных переменных. Как уже было указано, методы разрывной регрессии и difference-in-difference используют естественные эксперименты для получения выводов о каузальных связях между переменными. В случае отсутствия подходящего естественного эксперимента исследователям приходится опираться на альтернативные статистические методы. Одним из наиболее популярных таких методов является метод инструментальных переменных.

По сути, метод инструментальных переменных в политической науке призван решить так называемую проблему обратной каузальности, о которой мы коротко упоминали в начале статьи. Действительно, нередки случаи, когда эффект, оказываемый независимой переменной на зависимую, сложно изолировать в силу возможного «обратного эффекта», т.е. влияния зависимой переменной на независимую, или «искажения опущенной переменной» (omitted variable bias), т.е. отсутствия в анализе третьей переменной, объясняющей как зависимую, так и независимую переменную. Упомянутая в начале статьи взаимосвязь демократии и экономического развития является одним из ярких примеров, где проявляется проблема «обратного эффекта». В качестве примера «искажения опущенной переменной» можно упомянуть эффект экономических условий, оказываемый на гражданские конфликты: с одной стороны, низкий уровень экономического развития и безработица снижают альтернативные издержки участия в повстанческих движениях и, соответственно, повышают вероятность гражданских конфликтов [Collier, Hoeffler, 2004]; с другой – и низкий уровень экономического развития, и гражданский конфликт могут объясняться репрессивным политическим режимом. Метод инструментальных переменных позволяет изолировать эффект независимой переменной на зависимую при соблюдении ряда условий.

Первым важным условием является существование третьей переменной-инструмента, которая является сильным предиктором независимой переменной. Например, в работе [Acemoglu, Johnson, Robinson, 2001] предлагается использовать смертность первых колонистов в качестве инструмента-предиктора институтов: в колониях, имевших комфортные климатические условия для жизни, страны-метрополии создавали «нео-Европы», а в странах, где таких условий не было, создавались эксплуататорские системы. В исследовании [Miguel, Satyanath, Sergenti, 2004] используется уровень осадков в 41 африканской стране в качестве переменной-инструмента для экономического роста: засухи приводят к уменьшению урожая и, соответственно, рецессии в аграрных экономиках.

Вторым важным условием является «критерий исключенности» (exclusion restriction), т.е. отсутствие прямого влияния переменной-инструмента на зависимую переменную – единственным каналом влияния переменной-инструмента на зависимую переменную должна быть независимая переменная. Данное условие является краеугольным камнем метода инструментальных переменных. Несмотря на существование ряда статистических тестов, имеющих целью оценку соблюдения данного условия (например, тест Саргана – Хансена [Sargan, 1958; Hansen, 1982]), наиболее важным критерием почти всегда является логическое обоснование переменной-инструмента, предложенное в исследовании. Другими словами, исследователь должен убедительно объяснить, почему переменная-инструмент не имеет прямого влияния на зависимую переменную. Уровень осадков, используемый в исследовании [Miguel, Satyanath, Sergenti, 2004] в качестве переменной-инструмента для экономического роста, является ярким примером: сложно предположить, как уровень осадков может повлиять на возникновение гражданских конфликтов, кроме как через экономический рост. Уровень смертности первых колонистов, используемый в исследовании [Acemoglu, Johnson, Robinson, 2001], также имеет под собой хорошее обоснование: смертность колонистов несколько сотен лет назад вряд ли может повлиять на экономическое развитие сегодня по каналам, отличным от институциональных.

Если оба условия соблюдаются, метод инструментальных переменных выступает в качестве надежного механизма выявления каузальных связей. Стоит, однако, заметить, что поиск хорошей

переменной-инструмента зачастую так же труден, как и поиск подходящих для исследования естественных экспериментов.

Заключение

В настоящей статье мы рассмотрели методы каузального анализа, наиболее популярные в современной политической науке. Мы продемонстрировали область применения каждого метода и схематично описали разработку соответствующих исследовательских дизайнов. Следует отметить, что принципы, на которые опираются методы каузального анализа, применимы как для количественных, так и для качественных исследований. Например, принципы метода difference-in-difference хорошо применимы для кейс-стади, в рамках которого анализируется каузальный эффект конкретного события. В качестве исследовательского дизайна могут быть взяты два схожих по заранее определенным параметрам кейса, при этом в одном из них интересующее событие произошло, а в другом нет. Сравнение подобных кейсов может дать хорошую картину каузального эффекта рассматриваемого события.

Список литературы

- Acemoglu D., Johnson S., Robinson J.A.* The colonial origins of comparative development: An empirical investigation // *American economic review*. – 2001. – Vol. 91, N 5. – P. 1369–1401. – DOI: <https://doi.org/10.1257/aer.91.5.1369>
- Acemoglu D., Robinson J.A.* Why nations fail. – New York, NY : Crown business, 2012. – 546 p.
- Angrist J.D., Pischke J.-S.* Mostly harmless econometrics: an empiricist's companion. – Princeton, NJ : Princeton university press, 2008. – 392 p.
- Bates R.H., Lien D.D.* A note on taxation, development, and representative government // *Politics & Society*. – 1985. – Vol. 14, N 1. – P. 53–70. – DOI: <https://doi.org/10.1177/003232928501400102>
- Card D., Krueger A.B.* Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania // *American economic review*. – 1994. – Vol. 84, N 4. – P. 772–793. – DOI: <https://doi.org/10.1257/aer.90.5.1397>
- Chattopadhyay R., Duflo E.* Women as policy makers: evidence from a randomized policy experiment in India // *Econometrica*. – 2004. – Vol. 72, N 5. – P. 1409–1443. – DOI: <https://doi.org/10.1111/j.1468-0262.2004.00539.x>

- Collier P., Hoeffler A. Greed and grievance in civil war // Oxford economic papers. – 2004. – Vol. 56. – P. 563–595. – DOI: <https://doi.org/10.1093/oep/gpf064>
- Democracy does cause growth / D. Acemoglu, S. Naidu, P. Restrepo, J.A. Robinson // Journal of political economy. – 2019. – Vol. 127, N 1. – P. 47–100. – DOI: <https://doi.org/10.1086/700936>
- Dell M. The persistent effects of Peru's Mining Mita // Econometrica. – 2010. – Vol. 78, N 6. – P. 1863–1903. – DOI: <https://doi.org/10.3982/ECTA8121>
- Ferwerda J., Miller N.L. Political devolution and resistance to foreign rule: A natural experiment // American political science review. – 2014. – Vol. 108, N 3. – P. 642–660. – DOI: <https://doi.org/10.1017/S0003055414000240>
- Field experiment estimate of electoral fraud in Russian Parliamentary Elections / R. Enikolopov, V. Korovkin, M. Petrova, K. Sonin, A. Zakharov // Proceedings of the National Academy of Sciences. – 2013. – Vol. 110, N 2. – P. 448–452. – DOI: <https://doi.org/10.1073/pnas.1206770110>
- Frye T. Economic sanctions and public opinion: survey experiments from Russia // Comparative political studies. – 2019. – Vol. 52, N 7. – P. 967–994. – DOI: <https://doi.org/10.1177/0010414018806530>
- Gelman A. Preregistration of studies and mock reports // Political analysis. – 2013. – Vol. 21, N 1. – P. 40–41. – DOI: <https://doi.org/10.1093/pan/mps032>
- Gelman A., Imbens G.W. Why high-order polynomials should not be used in regression discontinuity designs // Journal of business & Economic statistics. – 2018. – Vol. 37, N 3. – P. 447–456. – DOI: <https://doi.org/10.1080/07350015.2017.1366909>
- Hansen L.P. Large sample properties of generalized method of moments estimators // Econometrica. – 1982. – Vol. 50, N 4. – P. 1029–1054. – DOI: <https://doi.org/10.2307/1912775>
- Humphreys M., Sanchez de la Sierra R., Van der Windt P. Fishing, commitment, and communication: a proposal for comprehensive nonbinding research registration // Political analysis. – 2013. – Vol. 21, N 1. – P. 1–20. – DOI: <https://doi.org/10.1093/pan/mps021>
- Imbens G.W., Kalyanaraman K. Optimal bandwidth choice for the regression discontinuity estimator // Review of economic studies. – 2012. – Vol. 79, N 3. – P. 933–959. – DOI: <https://doi.org/10.1093/restud/rdr043>
- Imbens G.W., Lemieux T. Regression discontinuity designs: A guide to practice // Journal of econometrics. – 2008. – Vol. 142, N 2. – P. 615–635. – DOI: <https://doi.org/10.1016/j.jeconom.2007.05.001>
- Imbens G.W., Rubin D.B. Causal inference for statistics, social, and biomedical sciences. An introduction. – New York, NY : Cambridge university press, 2015. – 625 p. – DOI: <https://doi.org/10.1017/CBO9781139025751>
- Lee D.S. The electoral advantage to incumbency and voters' valuation of politicians' experience: A regression discontinuity analysis of elections to the US house // NBER working paper series. – 2001. – N w8441.
- Lipset S.M. Some social requisites of democracy: economic development and political legitimacy // American political science review. – 1959. – Vol. 53, N 1. – P. 69–105. – DOI: <https://doi.org/10.2307/1951731>

- Miguel E., Satyanath S., Sergenti E.* Economic shocks and civil conflict: An instrumental variables approach // *Journal of political economy*. – 2004. – Vol. 112, N 4. – P. 725–753. – DOI: <https://doi.org/10.1086/421174>
- Monogan J.* Research preregistration in political science: the case, counterarguments, and a response to critiques // *PS: Political science & Politics*. – 2015. – Vol. 48, N 3. – P. 425–429. – DOI: <https://doi.org/10.1017/S1049096515000189>
- North D.C., Weingast B.R.* Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England // *Journal of economic history*. – 1989. – Vol. 49, N 4. – P. 803–832. – DOI: <https://doi.org/10.1017/S0022050700009451>
- Olken B.* Monitoring corruption: evidence of a field experiment in Indonesia // *Journal of political economy*. – 2007. – Vol. 115, N 2. – P. 200–249. – DOI: <https://doi.org/10.1086/517935>
- Pearl J.* *Causality: models, reasoning and inference*. Second edition. – N.Y. : Cambridge university press, 2009. – 459 p. – DOI: <https://doi.org/10.1017/CBO9780511803161>
- Sargan J.D.* The estimation of economic relationships using instrumental variables // *Econometrica*. – 1958. – Vol. 26, N 3. – P. 393–415. – DOI: <https://doi.org/10.2307/1907619>

E.A. Sedashov*

Methods of causal inference in contemporary political science

Abstract. This paper serves as an exposition of the causal inference methods that are most popular in political science. Rather than focusing on technical details we present a brief summary of main ideas behind each method with the goal of making them accessible to a broad audience of researchers. We also provide a research design algorithm for each method.

First, we focus on a general motivation behind causal inference methods. We discuss how the problem of causality arises in hypothesis testing and describe the relationship between democracy and economic development as a case in point. Second, we give an exposition of a general causality problem within the framework of Rubin Causal Model (RCM). We provide all basic definitions and then demonstrate how the problem of causal inference arise within RCM. Third, we describe the most frequently used methods of causal inference such as randomized experiments, regression discontinuity design, difference-in-difference design, and instrumental variables. For each method we give a reader a general description as well as steps of a research design. We also briefly discuss advantages and disadvantages of each method. Armed with this knowledge, a reader can use it to find the method that is the most appropriate for a research problem at hand. We conclude by arguing that the ideas of causal inference are useful for both quantitative and qualitative research.

* **Sedashov Yevgeniy**, HSE University (Moscow, Russia), e-mail: esedashov@hse.ru

Keywords: causal inference; Rubin Causal Model; randomized experiment; regression discontinuity design; matching; instrumental variables.

For citation: Sedashov E.A. Methods of causal inference in contemporary political science. *Political science (RU)*. 2021, N 1, P. 98–115. DOI: <http://www.doi.org/10.31249/poln/2021.01.4>

References

- Acemoglu D., Johnson S., Robinson J.A. The colonial origins of comparative development: An empirical investigation. *American economic review*. 2001, Vol. 91, N 5, P. 1369–1401. DOI: <https://doi.org/10.1257/aer.91.5.1369>
- Acemoglu D., Naidu S., Restrepo P., et al. Democracy does cause growth. *Journal of political economy*. 2019, Vol. 127, N 1, P. 47–100. DOI: <https://doi.org/10.1086/700936>
- Acemoglu D. Robinson J.A. *Why nations fail*. New York, NY : Crown business, 2012, 546 p.
- Angrist J.D., Pischke J.-S. *Mostly harmless econometrics: an empiricist's companion*. Princeton, NJ : Princeton university press, 2008, 392 p.
- Bates R.H. Lien D.D. A note on taxation, development, and representative government. *Politics & Society*. 1985, Vol. 14, N 1, P. 53–70. DOI: <https://doi.org/10.1177/003232928501400102>
- Card D., Krueger A.B. Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *American economic review*. 1994, Vol. 84, N 4, P. 772–793. DOI: <https://doi.org/10.1257/aer.90.5.1397>
- Chattopadhyay R. Duflo E. Women as policy makers: evidence from a randomized policy experiment in India. *Econometrica*. 2004, Vol. 72, N 5, P. 1409–1443. DOI: <https://doi.org/10.1111/j.1468-0262.2004.00539.x>
- Collier P. Hoeffler A. Greed and grievance in civil war. *Oxford economic papers*. 2004, Vol. 56, P. 563–595. DOI: <https://doi.org/10.1093/oeq/gpf064>
- Dell M. The persistent effects of Peru's Mining Mita. *Econometrica*. 2010, Vol. 78, N 6, P. 1863–1903. DOI: <https://doi.org/10.3982/ECTA8121>
- Enikolopov R., Korovkin V., Petrova M., et al. Field experiment estimate of electoral fraud in Russian Parliamentary Elections. *Proceedings of the national academy of sciences*. 2013, Vol. 110, N 2, P. 448–452. DOI: <https://doi.org/10.1073/pnas.1206770110>
- Ferwerda J. Miller N.L. Political devolution and resistance to foreign rule: A natural experiment. *American political science review*. 2014, Vol. 108, N 3, P. 642–660. DOI: <https://doi.org/10.1017/S0003055414000240>
- Frye T. Economic sanctions and public opinion: survey experiments from Russia. *Comparative political studies*. 2019, Vol. 52, N 7, P. 967–994. DOI: <https://doi.org/10.1177/0010414018806530>
- Gelman A. Preregistration of studies and mock reports. *Political analysis*. 2013, Vol. 21, N 1, P. 40–41. DOI: <https://doi.org/10.1093/pan/mps032>

- Gelman A. Imbens G.W. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of business & Economic statistics*. 2018, Vol. 37, N 3, P. 447–456. DOI: <https://doi.org/10.1080/07350015.2017.1366909>
- Hansen L.P. Large sample properties of generalized method of moments estimators. *Econometrica*. 1982, Vol. 50, N 4, P. 1029–1054. DOI: <https://doi.org/10.2307/1912775>
- Humphreys M., Sanchez de la Sierra R., Van der Windt P. Fishing, commitment, and communication: a proposal for comprehensive nonbinding research registration. *Political analysis*. 2013, Vol. 21, N 1, P. 1–20. DOI: <https://doi.org/10.1093/pan/mps021>
- Imbens G.W., Kalyanaraman K. Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*. 2012, Vol. 79, N 3, P. 933–959. DOI: <https://doi.org/10.1093/restud/rdr043>
- Imbens G.W., Lemieux T. Regression discontinuity designs: A guide to practice. *Journal of econometrics*. 2008, Vol. 142, N 2, P. 615–635. DOI: <https://doi.org/10.1016/j.jeconom.2007.05.001>
- Imbens G.W. Rubin D.B. *Causal inference for statistics, social, and biomedical sciences. An introduction*. New York, NY : Cambridge university press, 2015, 625 p. DOI: <https://doi.org/10.1017/CBO9781139025751>
- Lee D.S. The electoral advantage to incumbency and voters' valuation of politicians' experience: A regression discontinuity analysis of elections to the US house. *NBER working paper series*. 2001, N 8441.
- Lipset S.M. Some social requisites of democracy: economic development and political legitimacy. *American political science review*. 1959, Vol. 53, N 1, P. 69–105. DOI: <https://doi.org/10.2307/1951731>
- Miguel E., Satyanath S., Sergenti E. Economic shocks and civil conflict: An instrumental variables approach. *Journal of political economy*. 2004, Vol. 112, N 4, P. 725–753. DOI: <https://doi.org/10.1086/421174>
- Monogan J. Research preregistration in political science: the case, counterarguments, and a response to critiques. *PS: Political science & Politics*. 2015, Vol. 48, N 3, P. 425–429. DOI: <https://doi.org/10.1017/S1049096515000189>
- North D.C., Weingast B.R. Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England. *Journal of economic history*. 1989, Vol. 49, N 4, P. 803–832. DOI: <https://doi.org/10.1017/S0022050700009451>
- Olken B. Monitoring corruption: evidence of a field experiment in Indonesia. *Journal of political economy*. 2007, Vol. 115, N 2, P. 200–249. DOI: <https://doi.org/10.1086/517935>
- Pearl J. *Causality: models, reasoning and inference. Second edition*. New York, NY : Cambridge university press, 2009, 459 p. DOI: <https://doi.org/10.1017/CBO9780511803161>
- Sargan J.D. The estimation of economic relationships using instrumental variables. *Econometrica*. 1958, Vol. 26, N 3, P. 393–415. DOI: <https://doi.org/10.2307/1907619>