

**И.М. ЛОКШИН**

**ИГРА В БИСЕР? КОНВЕНЦИОНАЛЬНЫЕ  
КОЛИЧЕСТВЕННЫЕ МЕТОДЫ В СВЕТЕ ТЕЗИСА  
ДЮЭМА-КУАЙНА**

**I. Введение**

Прогресс современной политической науки нередко описывается в терминах все возрастающей достоверности данных и все большей объективности выводов [Алмонд, 1999, с. 69–112]. И количественные, и качественные методы имеют своей целью приближение к как можно более строгим научным стандартам. Однако остаются многочисленные проблемы, связанные с разнообразными методологическими, эпистемологическими и даже онтологическими вопросами [см., например: King, Keohane Verba, 1994; Brady, Collier, 2004; Gerring, 2012; Hay, 2006]. В настоящей статье предпринята попытка высветить некоторые проблемы, возникающие в связи с конвенциональными количественными методами (такими, как регрессионный анализ или тестирование нулевых гипотез), с точки зрения тезиса Дюэма-Куайна – важного достижения философии науки XX в. Речь не идет о том, что приложение тезиса Дюэма-Куайна к названным методам приводит к их принципиально новому восприятию или оценке; однако я полагаю, что рассмотрение конвенциональных количественных методов с позиций тезиса Дюэма-Куайна все-таки может оказаться полезным по трем причинам: 1) тезис Дюэма-Куайна может пролить новый свет на эпистемологические свойства некоторых методов конвенционального ко-

личественного анализа; 2) обращение к тезису Дюэма-Куайна позволяет «укоренить» представления о недостатках ряда конвенциональных количественных методов в философии науки, благодаря чему эти представления могут оказаться более точными и ясными; 3) привязывание методологической дискуссии к эпистемологическим, т.е. более фундаментальным, проблемам указывает на природу вызовов, стоящих перед исследователями в политической науке, а значит, и на возможные направления поиска новых методов или усовершенствования уже имеющихся.

Необходимы по меньшей мере три оговорки.

Во-первых, я не имею в виду, что статистические и эконометрические методы сводятся к их простейшим разновидностям, – однако разные типы регрессионного анализа и тестирования нулевой гипотезы занимают очень важное место в количественном крыле современной политологии. Но несмотря на серьезный прогресс в разработке более «продвинутых» методов, названные процедуры до сих пор во многом остаются «мейнстримом»; кроме того, более сложные методы основываются на простых процедурах и тем самым часто наследуют ограничения последних.

Во-вторых, многие проблемы, связанные с этими сравнительно простыми процедурами, хорошо известны. Однако несмотря на длительные поиски соответствующей литературы, мне не встретилось книг или статей, рассматривающих последствия тезиса Дюэма-Куайна для политологии<sup>1</sup>. Единственным возможным исключением является глава Фреда Черноффа из Оксфордского справочника по философии общественных наук, где он рассматривает некоторые проблемы, вытекающие из тезиса Дюэма-Куайна, но большей частью его внимание поглощено не этим тезисом *per se*, а ролью конвенциональных договоренностей в методологической плоскости для прогресса общественных наук [Chernoff, 2012, p. 229–258]. Поэтому даже названное исследование – единственное, которое, казалось бы, непосредственно касается нашей темы, – не вполне соответствует ей. Я полагаю, что такое невнимание к тезису Дюэма-Куайна незаслуженно.

---

<sup>1</sup> Для другой обществоведческой дисциплины – экономики – ситуация не такова: имеется несколько статей на эту тему, например: [Cross, 1982; Sawyer K., Beed C., Sankey H., 1997].

В-третьих, хотя я фокусируюсь на количественных методах, я не предполагаю, что качественный подход совершенно свободен от аналогичных проблем. Но репутация количественных методов как более строгих и соответствующих научным стандартам придает особую важность демонстрации именно их ограничений.

Статья выстроена следующим образом: раздел I посвящен тезису Дюэма-Куайна и некоторым из его следствий, показывающим релевантность тезиса для методологии политической науки; в разделе II освещаются отдельные методологические проблемы конвенциональных количественных методов с целью: а) привлечь внимание к их недостаткам на методологическом уровне, б) выявить связь этих недостатков с эпистемологическими сюжетами, затронутыми тезисом Дюэма-Куайна; в разделе III тезисы двух предыдущих частей статьи обобщаются и формулируется общий вывод.

## I. Тезис Дюэма-Куайна

### *I. 1. Формулировки тезиса*

История тезиса Дюэма-Куайна начинается с публикации книги французского физика Пьера Дюэма «Физическая теория: ее цель и строение»<sup>1</sup> [Duhem, 1906] в 1906 г. В ней, помимо прочего, содержится важная идея о том, что позднее было названо в англоязычной литературе «холистской недодетерминированностью» [Stanford, 2014] (holist underdetermination) физических теорий: «...Физик не может экспериментально проверить изолированную гипотезу, но только весь комплекс гипотез; когда результаты эксперимента расходятся с теоретическими предположениями, это означает, что по крайней мере одна из гипотез, составляющих этот комплекс, неверна и должна быть изменена; но эксперимент не указывает на то, какие именно из гипотез должны быть скорректированы» [Duhem, 1906, p. 307]. Содержание тезиса Дюэма состоит,

---

<sup>1</sup> Заметим, что название можно перевести и как «Физическая теория: ее предмет и строение» из-за многозначности употребленного в заглавии слова «objet».

таким образом, в том, что мы никогда не можем отвергнуть какую-то конкретную гипотезу, но только связанную совокупность гипотез и допущений.

Сегодня признается, что выводы Дюэма значимы не только для физики, но для любой науки, тестирующей гипотезы.

Еще до первой публикации книги Дюэма на английской языке в 1954 г. выдающийся философ науки и логик Уиллард Куайн обратил внимание почти на ту же проблему, что и Дюэм, но ход его мысли был совершенно иным: Куайн отталкивался от критики кантовского разделения суждений на аналитические и синтетические. Однако нас более интересуют следствия этой критики. В статье «Две догмы эмпиризма» [Quine, 1951, p. 20–43] Куайн (мы излагаем этот аспект его статьи в упрощенном виде) почти повторяет мысль Дюэма о том, что при противоречии между наблюдениями и нашими знаниями или убеждениями мы можем, чтобы последние вновь стали соответствовать опыту, пересмотреть то, какие из наших предположений являются истинными и какие – ложными, и при этом мы должны работать не с отдельными допущениями и гипотезами, а с их комплексом; как и Дюэм, Куайн полагает, что опыт ничего не диктует нам по поводу того, как должен быть произведен этот пересмотр. Но далее Куайн делает еще один шаг.

Поскольку опыт молчит в указанном отношении и поскольку вся тотальность науки есть произведенная людьми условность («a man-made fabric») [Quine, 1951, p. 39], соответствие между теорией и наблюдениями может быть восстановлено произвольно: через обращение к другим каузальным связям, к иным значимым переменным, даже к иным законам логики – и может оказаться, что таких способов чрезвычайно много.

Следовательно, одни и те же наблюдения можно объяснить более чем одним способом, при помощи более чем одной теории, хотя эти теории могут быть логически несовместимы. Заметим, что очень похожую мысль высказывал еще Джон Стюарт Милль в «Системе логики» [Mill, 1974, p. 500], но важное отличие идеи Куайна от идеи Милля заключается в том, что Куайн фактически выдвигает предположение о невозможности применения какого-либо рационального, также «научного» по своему существу, критерия

при выборе между конкурирующими гипотезами<sup>1</sup> [Quine, 1951, p. 42–43].

Из сказанного ясно, что тезис Дюэма-Куайна распадается, по сути, на два: один из них касается невозможности тестирования изолированной гипотезы («холистская недодетерминированность»), другой (в интерпретации Куайна) говорит о невозможности различения эквивалентных в смысле своих эмпирических следствий (или, как мы будем говорить для краткости, эмпирически эквивалентных), но логически несовместимых теорий согласно рациональному, но не чисто прагматическому критерию (используя неуклюжий перевод с английского, назовем это «контрастивной недодетерминированностью» – «contrastive underdetermination»<sup>2</sup>).

Другое важное следствие тезиса Куайна состоит в том, что граница между фактом и теорией всегда – в большей или меньшей степени – размыта. Факты оказываются теоретически нагруженными, и полностью отделить факты от их интерпретаций в рамках какой-либо теории невозможно. Тем самым тезис Куайна подрывает основания позитивистского проекта науки, основывающегося именно на однозначном разграничении фактов и теорий. Ниже я постараюсь показать, что проблема отсутствия четкой границы между фактами и их интерпретациями чрезвычайно актуальна для современной политической науки.

## *I. 2. Тезис Дюэма-Куайна, структурная недодетерминированность и политическая наука*

Мы еще будем обращаться к оригинальному тезису Куайна, но нам понадобится также и его «слабая» версия, акцентирующая внимание на том, что мы назовем «структурной недодетерминированностью» теорий.

---

<sup>1</sup> Куайн отмечает в той же статье, что это не подрывает его веры в науку и что он придерживается в отношении нее прагматических (в смысле философии прагматизма) установок [ibid, p. 42–43].

<sup>2</sup> Термин «contrastive underdetermination» используется потому, что он описывает невозможность предпочесть одну теорию своим альтернативам [Stanford, op. cit.].

Структурная недодетерминированность описывает ситуации, когда два или более факторов или их комбинаций претендуют на объяснение явления и вокруг этих факторов (или их комбинаций) формулируются теории, но, в отличие от случая контрастивной недодетерминированности, такие теории логически *не* противоречат друг другу. При этом у исследователя нет «научных» по своему существу способов распределения объяснительных весов (см. ниже ссылку 5) между объясняющими факторами. Структурная недодетерминированность названа так потому, что она связана с непроясненностью каузальной структуры явления<sup>1</sup>.

Контрастивная недодетерминированность есть частный случай структурной недодетерминированности. Для прояснения этого тезиса воспользуемся упрощенным примером, когда на объяснение некоего явления претендуют только две теории. В случае, если относительный вес некоего комплекса объясняющих факторов (или некой каузальной цепи, или причинного механизма) равен нулю, а вес другого – единице, но мы не можем выяснить, какому комплексу какой вес соответствует, мы имеем дело с контрастивной недодетерминированностью; если же относительный вес обоих комплексов объясняющих факторов больше 0, но меньше 1, и мы не можем выяснить конкретные значения этих весов, то мы имеем дело со структурной недодетерминированностью. Но поскольку крайние значения (0 и 1) тоже являются частью отрезка, на котором располагаются объяснительные веса теорий, контрастивную недодетерминированность можно считать частным случаем структурной.

---

<sup>1</sup> Под каузальной структурой явления здесь и далее я подразумеваю «значимость» всех факторов (по отдельности), полный учет которых позволил бы совершенно точно предсказать все параметры данного явления, интересующие исследователя. Однако само понятие «значимости» тоже нуждается в пояснении. В качестве интуитивного представления о том, что может скрываться за этой формулировкой, можно предложить следующее: «значимость» фактора А в каузальной структуре явления складывается из трех компонентов: 1) число каузальных цепочек, вызывающих явление В, в которых участвует фактор А (обозначим это число как N); 2) «правдоподобие» каждой каузальной цепочки, в которой задействован фактор А, или относительная частота случаев, когда явление В вызывается именно данной цепочкой  $i_A()$ ; 3) величина причинного эффекта, ассоциирующегося с каузальной цепочкой  $i_A()$ .

На первый взгляд, структурная недодетерминированность менее «опасна» для самого дела науки, чем недодетерминированность контрастивная, так как предполагает, что каждое объяснение исследуемого явления может быть хотя бы отчасти верно, но в то же время она гораздо более распространена в общественных науках в целом и в политологии в частности.

При достаточном усилии со стороны исследователя можно сформулировать теории, постулирующие каузальные связи между потенциально бесконечным числом факторов и явлений. Это оказывается возможным как по эпистемологическим причинам, так и благодаря специфике объектов, с которыми имеет дело политическая наука.

С эпистемологической точки зрения это оказывается возможным благодаря как минимум двум обстоятельствам: 1) возможности выделения множества каузальных цепочек, ведущих к объясняемому явлению; 2) возможности вновь и вновь переинтерпретировать «факты», т.е. эмпирические наблюдения, имеющиеся в отношении тех или иных явлений; иными словами, факты оказываются теоретически нагруженными, а не совершенно отделенными от теории (как то постулировал позитивизм), что и предполагает тезис Дюзма-Куайна.

Другой источник возможности постулировать огромное число причинно-следственных связей в рамках политологии состоит в том, что в обществе почти все хотя бы статистически связано со всем: при изучении какой-то важной переменной, отражающей масштабный социальный или политический процесс, найдется колоссальное количество признаков, хотя бы слабо коррелирующих с этой переменной и претендующих на каузальную роль. Но почти для любой статистической связи можно благодаря указанным выше эпистемологическим основаниям придумать каузальную интерпретацию.

Все это представляет плодородную почву для процветания структурной недодетерминированности и / или эмпирически эквивалентных теорий. Хрестоматийные примеры на этот счет – десяток факторов, традиционно рассматриваемых как причины демократизации [Мельвиль, Стукал, 2011, с. 165–169], или множество объяснений того, почему демократии не воюют друг с другом [Rosato, 2003, p. 585–602], или несколько причинных механизмов,

предположительно стоящих за «ресурсным проклятием» [Ross, 2001, p. 325–361]. Но даже явления не столь крупного масштаба часто бывают связанными с несколькими признаками, которые можно записать в число причин данного явления. При достаточных фантазии и времени можно разработать теории, правдоподобно связывающие почти все со всем.

## **II. Конвенциональные практики количественного анализа и тезис Дюэма-Куайна**

Ряд практик, которыми широко пользуются при проведении количественных исследований, усугубляет серьезность проблем, очерченных в предыдущем разделе. Ниже я кратко остановлюсь на распространенных практиках количественного анализа и таких их особенностях, как: 1) повсеместное употребление тестов статистической значимости; 2) недооценка последствий того, что исследуемые наблюдения могут быть каузально неоднородны; 3) стремление работать с процессами и явлениями самого общего характера.

### *II. 1. Злоупотребление тестами статистической значимости*

В экономике и особенно в психологии – науках, которые тоже очень широко, как и политология, используют количественные методы анализа, – уже десятилетия ведется острая полемика по поводу ценности тестов значимости и тестирования нулевой гипотезы. В экономике эту практику критикует в основном Д. Макклоски [McCloskey, 1985, p. 201–205], при этом, хотя далеко не все согласны с некоторыми пунктами этой критики, иные ее положения не подвергаются сомнению. В психологии же сложилась мощнейшая традиция критики тестов значимости: особенно примечательны в связи с этим имена Якоба Коэна и Пола Миля [Cohen, 1994, p. 997–1003; Cohen, 1990, p. 1304–1312; Meehl, 1967, p. 103–115; Meehl, 1990, p. 108–141]. Но в политической науке критика подобного рода начала звучать сравнительно поздно, а примеры ее – весьма



редки<sup>1</sup>. Как будет видно из дальнейшего, данный вопрос заслуживает большего внимания.

*А. Тестировать нулевые гипотезы в общественных науках почти всегда бессмысленно, потому что между переменными почти никогда не бывает **полного** отсутствия связи.*

Данное утверждение – переложение ранее сформулированного тезиса о том, что в обществе хотя бы статистически почти все связано со всем – в большей или меньшей степени<sup>2</sup>. Но тестирование нулевой гипотезы подразумевает проверку предположения о том, что связи *нет вообще*, что она в самом деле *нулевая*, но это заведомо нереалистично. Следовательно, в тестировании нулевой гипотезы нет самодостаточного смысла.

Будет или нет отвергнута нулевая гипотеза, зависит от четырех факторов: уровня значимости, статистической мощности, размера эффекта (effect size) и размера выборки [Cohen, 1992, p. 156]<sup>3</sup>. Учитывая, что размер эффекта для переменных, которыми интересуется политология, почти никогда не бывает равным *точно нулю*, при достаточно большом размере выборки нулевая гипотеза будет отвергнута на любом уровне значимости<sup>4</sup>.

Но данная проблема усугубляется эпистемологическим аспектом, который был очерчен в предыдущем разделе. Именно, практически для любой переменной, статистически связанной с объясняемым явлением, можно придумать теорию каузальной связи. Даже если тест отвергает нулевую гипотезу, исследователь может сослаться на малый размер выборки и тем самым подвергнуть сомнению полученные результаты [Zumbo, Hubleу, 1998], что вновь ведет к неопределенности в смысле того, верна или неверна тестируемая гипотеза. Поэтому тестирование нулевых гипотез, по на-

---

<sup>1</sup> Мне удалось найти лишь четыре статьи, в явной форме критикующие практику использования тестов значимости в политологии: [Gill, 1999, p. 647–654; Rainey, 2014; Ward, Greenhill, Bakke, 2010, p. 363–375; Schrodт, 2014, p. 287–300]. Хотя дело, разумеется, не в количестве статей, а в том, что они порождают слишком мало резонанса.

<sup>2</sup> Аналогичное явление в психологии отмечает П. Миль и называет это «фактором чепухи» (crud factor) [см.: Meehl, 1990, p. 123].

<sup>3</sup> О размере эффекта см.: [Cohen, 1992, p. 157].

<sup>4</sup> Прекрасный пример из реального исследования, иллюстрирующий это положение, приводит П. Миль: см.: [Meehl, 1990, p. 205].

шему предположению, само по себе очень часто лишено смысла не только в статистическом, но и в содержательном плане.

*В. Тесты значимости предоставляют информацию о статистической значимости, но не свидетельствуют напрямую о политической, или, шире, содержательной значимости.*

Хорошо известно, что, к примеру, в уравнении регрессии статистически значимыми могут оказаться коэффициенты, очень мало отличные от нуля. В то же время коэффициенты, которые исследователь может считать, полагаясь на результаты анализа, теорию и на свой опыт, достаточно большими по модулю, могут оказаться статистически незначимыми. Статистическая значимость не является надежным показателем содержательной значимости, т.е. не указывает сама по себе на то, что коэффициент при предикторе или размер эффекта достаточно велик. Тем не менее, как в силу того, что прилагательное «статистическая» часто опускается при сообщении результатов, так и, вероятно, по ряду других причин, *статистическая значимость часто интерпретируется как содержательная значимость.*

На сказанное выше можно возразить: нельзя трактовать размеры эффекта или коэффициенты статистически незначимых предикторов как устойчивые, потому что для таких предикторов сила сигнала слишком слаба по сравнению с уровнем «шума» [Hoover, Siegler, 2008, p. 15–16]. В расширенной и более репрезентативной выборке коэффициент может сильно измениться.

Но является ли статистическая значимость единственным критерием, позволяющим судить об устойчивости коэффициента при расширении выборки? Что если исследователь уверен в том, что его малая выборка включает в себя разные типы наблюдений примерно в той же пропорции, в какой они распределены в генеральной совокупности? В таком случае, по-видимому, он может быть достаточно уверен в устойчивости коэффициентов при предикторах, даже если они не являются статистически значимыми.

Но именно в политологии исследователи часто могут надеяться на репрезентативность своих выборок. Нередко исследуемые выборки являются, в сущности, генеральными совокупностями: речь идет о множестве независимых государств, о множестве посткоммунистических государств, о множестве регионов России, штатов США, множестве несостоявшихся государств, множестве

партий, принимающих участие в выборах, и т.д. И даже если предположить, что все эти «выборки – генеральные совокупности» являются частями некой «сверхгенеральной совокупности» [Western, Jackman, 1990, p. 413–414; Berk, Western, Weiss, 1999, p. 421–423] (например, выборка государств на 2014 г. принадлежит сверхгенеральной совокупности государств за период 1950–2050 гг.), исследователь может быть достаточно уверен в том, что такие выборки довольно точно репрезентируют эту сверхгенеральную совокупность (продолжая пример, можно сказать, что весь релевантный период страны делится на одни и те же группы с сохранением общего паттерна распределения значений политических переменных: есть страны Запада, образующие один кластер; страны Африки или, например, «третьего мира», образующие другой кластер, и т.д.).

Если приведенный аргумент верен, то выходит, что зачастую политологам следует гораздо внимательнее относиться к коэффициентам и размерам эффекта статистически незначимых переменных: с высокой вероятностью расширение выборки, если оно вообще возможно, не внесет в результаты статистического анализа коренных изменений.

К этому нужно добавить, что само вынесение решения о статистической значимости происходит на основании конвенциональных порогов ( $\alpha = 0,1$ ;  $\alpha = 0,05$ ;  $\alpha = 0,01$ ), которые зачастую приобретают полусвященный ореол. Они еще и потому должны играть меньшую роль в научных исследованиях, что наша задача сильно отлична от задачи инженеров или чистых экспериментаторов, для которых такого рода конвенциональные пороги особенно полезны: политологи должны не принимать решение о значимости или незначимости коэффициента при переменной или «обработки» (treatment), а уточнять степень нашей уверенности в корректности той или иной гипотезы [Rozeboom, 1960, p. 420].

Подмена понятий статистической и содержательной значимости приводит к тому, что иногда занижается важность переменных, коэффициенты при которых не признаны статистически значимыми, и завышается важность переменных со статистически значимыми коэффициентами. *Как следствие, представление о каузальной структуре изучаемого явления может сильно исказиться.* При противоречии результатов, полученных на разных выборках, при разных операционализациях переменных или при использовании

разных критических уровней усугубляется проблема структурной недодетерминированности.

Итак, акцентирование внимания на статистической значимости легко может вести к недостаточному вниманию к содержательной значимости и к оценке размеров эффекта. Не исключено, что это обстоятельство возымело последствия и в области того, как привыкли формулировать вопросы многие политологи, работающие с количественными методами. Так, много лет продолжается полемика о том, способствует ли демократия более высокому экономическому развитию; вызывают ли свободные ресурсы гражданскую войну; какие факторы повышают риск революции; опасна ли президентская форма правления для демократии – этот список можно продолжить еще многими десятками пунктов. Но с точки зрения рассмотренных выше обстоятельств такие вопросы почти бессмысленны. Нужно спрашивать не о том, является ли фактор  $X$  причиной явления  $A$ , потому что *в какой-то мере и в каких-то обстоятельствах он легко может оказаться таковым*, и, кроме него, эту же функцию, вполне вероятно, выполняют еще и многие другие факторы. В связи с этим в рамках общей логики конвенциональных количественных методов<sup>1</sup> разумнее спрашивать о том, насколько велик абсолютный и относительный причинный эффект фактора  $X$  на явление  $A$  и насколько он велик при разных конфигурациях причинных факторов – говоря более общо, насколько он велик в разных обстоятельствах и при разных условиях. Если не задавать вопросы подобного рода, то *мы сталкиваемся с проблемой структурной недодетерминированности: мы не понимаем ни того, в какой мере фактор  $X$  политически значим, ни того, как он действует в условиях разных каузальных паттернов.*

## *II. 2. Недооценка последствий каузальной неоднородности исследуемых наблюдений*

Ученые-политологи часто – явно или подспудно – стремятся обнаружить универсальные закономерности, однако многие из

---

<sup>1</sup> В заключении мы подробнее скажем о том, что понимаем под этой «общей логикой», и поставим вопрос о том, в какой мере она в принципе адекватна научным целям.

этих закономерностей напоминают морскую свинку или, если угодно, Священную Римскую империю: подобно тому как морская свинка не морская и не свинка (а Священная Римская империя, по слову Вольтера, не Священная, не Римская и даже не империя), так и универсальные закономерности в политологии иногда оказываются и не закономерностями, и не универсальными.

В самом деле, полезно различать два смысла понятия «универсальный»: с одной стороны, «универсальный» может означать «относящийся в равной степени ко всем наблюдениям»; с другой стороны, под ним фактически часто понимают «усредненный (например, в соответствии с концептом математического ожидания) по всем наблюдениям». Дэвид Бакан проводит примерно такое же по смыслу различие между «общим» (general) и «агрегированным» (aggregate) [Bakan, 1966, p. 433]. Разница между двумя этими идеями универсальности имеет далеко идущие последствия. В физике и прочих естественных науках, насколько мне известно, закономерности формулируются как универсальные в первом смысле этого слова: они применимы ко всем физическим объектам заранее определенного класса. В общественных науках и в политологии в частности, если не всегда, то, возможно, в большинстве случаев и не в последнюю очередь благодаря популярности регрессионного анализа, «универсальность» на самом деле имеет второй смысл. Установленная закономерность может быть верна в среднем, но очень плохо описывать каждое наблюдение в отдельности.

К примеру, в результате количественного анализа можно выяснить, что демократия «в среднем» способствует экономическому развитию; при использовании какого-то вида регрессионного анализа можно даже назвать вклад демократии в экономическое развитие (коэффициент при предикторе демократии). Однако возникает вопрос: что означает это открытие? Значит ли это, что демократия для всех стран выборки делает одинаковый вклад в экономическое развитие? Или мы все же допускаем, что для Швейцарии этот вклад будет не таким, как для Конго? И может ли быть, что для каких-то стран повышение уровня демократичности на один пункт согласно какому-либо рейтингу, позволившему нам операционализировать это понятие, негативно скажется на экономическом развитии? Вероятно, это предположение тоже нельзя сбрасывать со счетов. Иными словами, мы можем получить, казалось бы, ясный и

однозначный результат, но элементарные вопросы о его содержании выявляют то, что этот результат трудно распространить по крайней мере на значительную часть конкретных наблюдений. В свою очередь, это ведет к тому, что такую «универсальную» закономерность не вполне корректно использовать в прогностических целях. Эта неясность преследует ученых-политологов всякий раз, когда пространство наблюдений каузально неоднородно: именно, когда исследуемое явление в разных наблюдениях порождается в чем-то отличными друг от друга каузальными структурами<sup>1</sup>.

Таким образом, выявление «общих» или «универсальных» закономерностей приведет к результатам, успешно приложимым к конкретным наблюдениям, лишь тогда, когда каузальная структура исследуемых наблюдений достаточно однородна. Если же это требование не соблюдается, то результаты анализа оказываются весьма неустойчивыми к составу выборки, операционализации переменных и прочим аспектам исследования. В свою очередь, *такая неустойчивость приводит к путанице в каузальной структуре исследуемого явления и, следовательно, лишь усугубляет серьезность проблемы структурной недодетерминированности.*

Еще одна связь обсуждаемого в данном разделе вопроса с тезисом Дюзма-Куайна состоит в том, что при анализе каузально неоднородной выборки резко возрастает число факторов, с которыми статистически связана зависимая переменная. В свою очередь, как мы отмечали выше, это подготавливает почву для того, чтобы воспринять такие факторы как имеющие каузальную связь с интересующим исследователя явлением (во всяком случае, часто возможно придумать соответствующие теории). В свою очередь, это обычно приводит к попытке учесть все большее и большее число контрольных переменных. Однако данная стратегия чревата искажением оценок регрессионных коэффициентов – точно так же, как и отказ от учета дополнительных контрольных переменных [Clarke, 2005]. Возможное разрешение этой проблемы – сужение выборки с целью сокращения каузальной неоднородности явлений.

---

<sup>1</sup> Более подробное обсуждение этого сюжета применительно к видам регрессионного анализа [см. в: Achen, 2002, p. 446–447].

### *II. 3. Источники структурной недодетерминированности, связанные с анализом явлений самого общего характера*

Ученые-эмпирики склонны к тому, чтобы включать в свои эмпирические модели явления самого общего характера – такие, которые фигурируют в теориях, проверяемых эмпирически. Исследование причин демократизации почти всегда происходит с включением в эмпирическую модель операционализации понятия демократизации – это лишь один пример, подобный десяткам других такого же рода.

Благодаря этому эмпирическая и теоретическая работа оказываются максимально приближенными друг к другу, что, безусловно, облегчает как формулирование гипотез и теорий, так и их проверку. Но у этого удобства есть обратная сторона: *чем более общий характер носит эмпирически исследуемое явление, тем в большем количестве каузальных цепочек оно участвует и тем легче придумать правдоподобную гипотезу или теорию, эмпирически эквивалентную тестируемой.* И тогда, если теория или гипотеза не была эмпирически опровергнута, мы не можем указать «объяснительный вес» каждой из эмпирически эквивалентных теорий.

Совершенно аналогичный эффект имеет и то, что *операционализацию некой переменной можно истолковать и как операционализацию какой-то другой переменной (эта проблема имеет непосредственную связь с теоретической нагруженностью факторов, которую выявляет тезис Дюэма-Куайна).*

Итак, выше названы два источника структурной недетерминированности: первый источник – это возможность изменения каузальных цепочек, ведущих от явления А к явлению В; второй источник – это возможность изменения интерпретации операционализации явления А, ведущего к явлению В.

Для прояснения этих источников структурной недетерминированности приведем два примера из популярной ныне области исследований.

Одна из теорий, на которую мы обратим внимание, разработана Дароном Асемоглу и Джеймсом Робинсоном [Acemoglu, Robinson, 2006], другая – Рональдом Инглхартом и Кристианом Вельцелем [Welzel, Inglehart, 2008, p. 126–140].

Д. Асемоглу и Дж. Робинсон предлагают причинно-следственную цепочку, которая схематически и в упрощенном виде может быть представлена следующим образом: модернизация (включая экономическое развитие) → снижение имущественного неравенства и изменение структуры экономики, в которой повышается доля физического и человеческого капитала, рост средних доходов и среднего уровня жизни населения → ущерб, причиняемый элитам более выраженной распределительной политикой, предполагаемой демократией, снижается → элиты становятся менее склонны к сопротивлению демократизации → элиты оказываются готовыми к демократизации [Acemoglu, Robinson, 2006, p. 285–286; 319–320].

Что касается главного эмпирического основания исследования, то оно заключается в хорошо установленной связи между модернизацией и демократией; цель теории Асемоглу и Робинсона заключается в том, чтобы разработать теорию, объясняющую эту связь [Acemoglu, Robinson, 2006, p. 54–55].

Эта же задача преследуется Рональдом Инглхартом и его коллегами, в том числе Кристианом Вельцелем. Каузальная цепь, соединяющая модернизацию и демократизацию, выглядит в их теории примерно так: модернизация (включая экономическое развитие) → распространение в обществе ресурсов, а) повышающих возможности сопротивления элитам, б) повышающих субъективную ценность свободы → распространение и укрепление демократических ценностей (следствие пункта (б) → оказание давления на элиты по демократизации (следствие пунктов (а) и (б) → демократизация [Welzel, 2009, p. 74–90].

Почему сопоставление этих двух теорий порождает проблемы? На это есть две причины.

Во-первых, главное эмпирическое обоснование обеих теорий – сильная корреляция между модернизацией (экономическим развитием, уровнем грамотности и т.д.) и демократизацией; но эмпирическая проверка каждой из теорий посредством очередной оценки упомянутой связи не будет настоящей проверкой, поскольку, опираясь на идентичные переменные, мы не можем изолировать эти теории друг от друга; мы вновь сталкиваемся с проблемой структурной недодетерминированности, так как не знаем реального «объяснительного веса» этих теорий.



Во-вторых, только что указанная сложность усугубляется тем, что теория Инглхарта и Вельцеля плохо совместима с теорией Асемоглу и Робинсона; причина состоит в том, что для Инглхарта и Вельцеля переход к демократии происходит из-за давления на политические элиты «снизу», двигатель перехода – народные массы, предпринимающие успешные коллективные действия, направленные на демократизацию. В теориях, разработанных в парадигме политической экономии, напротив, демократизация случается потому, что на нее соглашается элита, отказываясь от стратегии репрессий. Возможно, это различие в теориях еще не ведет к логическому противоречию и их полной несовместимости, но оно определенно ставит проблемы на пути к их согласованию.

Второй источник структурной недодетерминированности (тот, что связан с множественностью интерпретаций операционализации абстрактных и общих явлений) удобно рассмотреть на примере теории Карлеса Боша [Voix, 2003]. Его теория предсказывает, в частности, что экономическое развитие увеличивает долю в экономике мобильного капитала, а те элиты, которые владеют мобильным капиталом, могут, в случае установления демократии и, как следствие этого, более равномерного экономического распределения, переместить свои активы за рубеж и тем самым сократить ущерб для себя от бремени, порождаемого более равномерным распределением. Вследствие этого сопротивление демократии со стороны элит, владеющих мобильным капиталом, будет слабее, чем сопротивление элит, чей капитал немобилен; значит, чем более развита экономика и чем большая доля приходится в ней на мобильный капитал, тем меньше сопротивляемость элит демократии и тем вероятнее ее установление [Voix, 2003, p. 12–13].

Однако нас интересует не теория сама по себе, а то, как операционализируются ключевые переменные. Бош использует следующие способы операционализации переменной мобильности капитала: доля в ВВП страны, которая приходится на сельское хозяйство; доля сырьевого экспорта в совокупном экспорте; среднее число лет обучения (в предположении о том, что число лет обучения формирует человеческий капитал, который мобильнее физического) [Geddes, 2007, p. 323–324].

Все три способа операционализации подвержены широким интерпретациям. Доля, которая приходится на сельское хозяйство в

ВВП страны, может рассматриваться не только как показатель мобильности капитала, но и как показатель инновационности экономики или даже как степень «традиционности» общества; при достаточной изобретательности можно предложить еще немало интерпретаций. Доля сырьевого экспорта в совокупном экспорте может рассматриваться как мера «свободных» денег у правительства, способных вести к ресурсному проклятию. Среднее число лет обучения выступает показателем образованности населения, которое, в свою очередь, может трактоваться в духе гипотезы Мартина Липсета или исследований Рональда Инглхарта о ценностных продемократических предпочтениях, свойственных более образованным гражданам. Каждая из этих интерпретаций указывает на новую теорию, посредством которой может объясняться переход к демократии. Таким образом, мы сталкиваемся с проблемой эмпирической эквивалентности: проводя анализ, подобный тому, что проделал Карлес Бош, мы не можем проверить ни одну из этих теорий в изоляции от других; следовательно, положительные результаты эмпирических тестов ни на йоту не добавляют нам уверенности в том, что какая-то из этих теорий верна<sup>1</sup>; в более широком смысле мы вновь сталкиваемся с недодетерминированностью каузальной структуры явления демократизации.

Нет сомнения, что с обоими источниками структурной недодетерминированности бороться чрезвычайно тяжело, хотя бы потому, что многие общественные процессы сильно коррелируют друг с другом, а мы ограничены в своих возможностях собирать те или иные данные для формирования все новых операционализаций переменных. Тем не менее не исключено, что в некоторых случаях полезно было бы работать не на уровне очень общих и абстрактных явлений, а на более «низком» уровне переменных, которые в каузальной цепи достаточно близки к эффекту. В этом смысле идея поиска «microfoundations» весьма резонна, но она не играет большой роли, если касается лишь придумывания новых теорий, а не попытки их прямого тестирования.

Кроме того, продуктивным может оказаться использование более узких, менее подверженных разным интерпретациям опера-

---

<sup>1</sup> На амбивалентность выводов К. Боша указывает, например, Барбара Геддес в упомянутой главе: [Geddes, op. cit.].

ционализаций. Однако сложно найти какой-либо хороший критерий или признак, которым можно руководствоваться при выборе таких операционализаций. Вероятно, в каждом отдельном случае нужно полагаться на уже имеющееся представление о каузальной сложности явления и на опыт и проницательность исследователя.

Заметим еще, что освещенный в этом разделе сюжет тесно связан с предыдущим: анализ масштабных явлений часто проводится в контексте каузальной неоднородности наблюдений. Поэтому стратегия сокращения этой неоднородности тоже может оказаться полезной.

### **Заключение**

Выше я попытался показать, как в конвенциональных методах количественного анализа воспроизводятся и усугубляются эпистемологические проблемы, связанные с тезисом Дюэма-Куайна. Особенное внимание было уделено трем распространенным в рамках количественного подхода исследовательским практикам: 1) использование тестов значимости как самодовлеющего аналитического инструмента; 2) невнимание к каузальной неоднородности явлений в больших выборках; 3) применение операционализаций переменных (представляемых как «данные», т.е. факты), подверженных большому числу интерпретаций. Эти практики значительно усугубляют проблему структурной недодетерминированности теорий: исследователь оказывается в ситуации, когда разные факторы, в том числе такие, которые плохо совместимы друг с другом, претендуют на объяснение данных, а строгий метод предпочтения одних теорий перед другими неочевиден.

На предполагаемое отсутствие надежного критерия выбора теорий нужно обратить особенно внимание. Об обманчивой эффективности тестов значимости, которые часто выступают в роли волшебной палочки, позволяя легко отвергать гипотезы, говорилось выше. Преимущества квазиэкспериментального дизайна, основанного на идее «контрольных переменных», теоретически несостоятельны [Clarke, 2005]. Большие выборки, нацеленные на выявление универсальных закономерностей, имеют оборотную сторону медали (казуальная неоднородность явлений), а потому

выводы, сделанные на их основании, нельзя по умолчанию признать более надежными, чем заключения, полученные на меньших выборках.

Все это означает, что в рамках конвенциональных количественных методов вопрос о критерии предпочтения одних теорий перед другими стоит весьма остро. Более того, на первый взгляд вовсе неочевидно и то, каким мог бы быть хотя бы прагматический критерий выбора теорий (именно прагматический характер научных теорий защищал Куайн).

В действительности пресловутая *проблема эндогенности* – частный случай *структурной недодетерминированности*: мы не можем корректно оценить, как соотносятся между собой две (в случае «классической» проблемы эндогенности, когда есть сомнения о том, в какую сторону и с какой силой направлена каузальная связь) не исключающие друг друга теории. Но если этой проблеме справедливо уделяется столь большое внимание, то ее обобщение бросает политической науке еще более серьезный вызов.

Кроме того, разработка новых теорий, «проверяемых» на давно известных и многократно использованных данных, часто несет в себе мало смысла, потому что можно придумать десятки правдоподобных, но эмпирически эквивалентных теорий, и в результате нельзя будет по-настоящему проверить ни одну из них.

Нередко в политической науке речь идет о попытке выяснить, влияет ли одна переменная на другую, но довольно редко прилагаются усилия к тому, чтобы понять, насколько сильно это влияние и в какой мере оно сильнее или слабее влияния прочих факторов, одинаково ли оно для наблюдений разных классов, необходимое оно или достаточное, может ли оно компенсировать слабость влияния других переменных и т.д. Иными словами, ученые-политологи редко стремятся составить обоснованное *представление о каузальной структуре явления в ее целостности*. Но без такого представления простой вывод о влиянии одной переменной на другую зачастую оказывается совершенно тривиальным: политика настолько сложная область, что почти все коррелирует почти со всем, и при достаточном воображении исследователь может вычленить многие десятки факторов, имеющих причинное воздействие на изучаемый процесс в тех или иных обстоятельствах. Но какие-то из этих факторов будут иметь большой объяснительный вес

или по иным причинам окажутся важными, а какие-то, напротив, будут играть лишь маргинальную роль.

Несомненно, задача составления целостного представления о каузальной структуре явления чрезвычайно сложна. В этом смысле неудивительно, что Джон Герринг, рассматривающий в одной из своих статей сюжеты, близкие к настоящему рассуждению<sup>1</sup>, призывает к снижению строгости стандартов, когда речь заходит об эмпирическом тестировании каузальных механизмов [Gerring, 2010, p. 1518]. Хотя такая идея по ряду причин может показаться сомнительной, она вскрывает сложности, заложенные в стремлении как можно более полно учесть каузальную структуру исследуемых явлений.

Однако можно поставить под вопрос даже и саму задачу о выяснении каузальной структуры явления, потому что она сформулирована в предположении, что исследователь заинтересован в выявлении причинных эффектов всех важных (в каком бы то ни было смысле этого слова) факторов изучаемого явления. Использование стратегии явного учета размера причинных эффектов – это и есть то, что ранее мы назвали «общей логикой» конвенциональных количественных методов. Но эта стратегия чревата большими трудностями.

Одна из явных альтернатив ей – оценка размера только одного причинного эффекта при изоляции всех остальных. Это и есть идея, стоящая за экспериментальным дизайном. В связи с этим можно сформулировать такую методологическую дилемму:

а) явный учет размеров причинных эффектов технически легко осуществим (если только речь не идет о попытке оценить эффекты *всех* важных факторов; это и делается в рамках конвенциональных количественных методов), но приводит к проблемам, которым посвящена настоящая статья;

б) явный учет лишь одного причинного эффекта – в котором более всего заинтересован исследователь – при изоляции всех ос-

---

<sup>1</sup> Предмет размышлений Дж. Герринга – проблемы, проистекающие из стремления эмпирически тестировать каузальные механизмы; многие выводы, к которым он приходит, имеют немало общего с положениями настоящей статьи, но Герринг рассматривает каузальные механизмы без привлечения тезиса Дьюэма-Куайна.

тальных (что и происходит в эксперименте) технически тяжело осуществить, но, по-видимому, эта стратегия свободна от многих проблем, выявляемых в связи с тезисом Дюэма-Куайна.

Разумеется, эта дилемма не нова для общественных наук вообще и для политологии в частности. Но в настоящей статье я попытался продемонстрировать, что тезис Дюэма-Куайна позволяет: а) выявить ряд серьезных недостатков первой стратегии и явно их сформулировать; б) «укоренить» эти формулировки в философии науки и таким образом придать им более точный и ясный характер.

Как бы то ни было, несомненно, что методологический спор (*Methodenstreit*) в политической науке будет продолжен, но также несомненно и то, что соответствие строгим научным стандартам и самим целям науки требует дальнейшего совершенствования альтернативных методологических подходов и более аккуратного применения конвенциональных количественных методов.

### Список литературы

- Алмонд Г.* Политическая наука: история дисциплины // Политическая наука: новые направления / Под ред. Гудин Р., Клингеманн Х.-Д. – М., 1999. – С. 69–112.
- Мельвиль А.Ю., Стукал Д.К.* Условия демократии и пределы демократизации // Полис: Политические исследования. – М., 2011. – № 3. – С. 164–183.
- Acemoglu D., Robinson J.* Economic origins of dictatorship and democracy. – Cambridge: Cambridge univ. press, 2006. – 434 p.
- Achen C.* Toward a new political methodology: microfoundations and ART // Annual review of political science. – 2002. – N 5. – P. 423–450.
- Bakan D.* The test of ignificance in psychological research // Psychological bulletin. – 1966. – Vol. 66, N 6. – P. 423–437.
- Berk R., Western B., Weiss R.* Statistical influence for apparent populations // Sociological methodology. – 1995. – Vol. 25. – P. 421–458.
- Boix C.* Democracy and redistribution. – Cambridge: Cambridge univ. press, 2003. – 284 p.
- Chernoff F.* The impact of duhemian principles on social science testing and progress. // Oxford handbook of philosophy of social science / Ed. by Harold Kincaid. – New York: Oxford univ. press, 2012. – P. 229–258.
- Clarke K.* The Phantom menace: Omitted variable bias in econometric research // Conflict management and peace science. – 2005. – N 22. – P. 341–352.
- Cohen J.* Things I have learned (So Far) // American psychologist. – 1990. – Vol. 45, N 12. – P. 1304–1312.
- Cohen J.* A power primer // Psychological bulletin. – 1992. – Vol. 112, N 1. – P. 155–159.

- Cohen J.* The Earth Is round ( $p < 0.05$ ) // *American psychologist*. – 1994. – Vol. 49, N 12. – P. 997–1003.
- Cross R.* The Duhem-Quine thesis, Lakatos and the Appraisal of theories in macroeconomics // *The economic journal*. – 1982. – Vol. 92. – P. 320–340.
- Duhem P.* *La theorie physique. Son objet et sa structure.* – Paris, 1906. – 450 p.
- Geddes B.* What causes democratization? // *The Oxford handbook of comparative politics* / Ed. by Carles Boix, Susan C. Stokes. – Oxford: Oxford univ. press. – 2007. – P. 317–339.
- Gerring J.* Causal mechanisms: Yes, But... // *Comparative political studies*. – 2010. – Vol. 43, N 11. – P. 1499–1526.
- Gerring J.* *Social science methodology: A unified framework.* – Cambridge, Cambridge univ. press, 2012. – 522 p.
- Gill J.* The insignificance of null hypothesis significance testing // *Political research quarterly*. – 1999. – Vol. 52, N 3. – P. 647–654.
- Hay C.* Political ontology. // *The Oxford handbook of contextual political analysis* / Ed. by Robert E. Goodin, Charles Tilly. – Oxford: Oxford univ. press., 2006. – P. 78–96.
- Hoover K., Siegler M.* Sound and fury: McCloskey and significance testing in economics // *Journal of economic methodology*. – 2008. – Vol. 15, N 1. – P. 1–37.
- King G., Keohane R., Verba S.* *Designing social inquiry: Scientific inference in qualitative research.* – Princeton: Princeton univ. press, 1994. – 300 p.
- McCloskey D.* The loss function has been misled: The rhetoric of significance tests // *American economic review*. – 1985. – Vol. 75, N 2. – P. 201–205.
- Meehl P.* Testing Theories in psychology and physics: A methodological paradox // *Philosophy of science*. – 1967. – Vol. 34, N 2. – P. 103–115.
- Meehl P.* Appraising and amending theories: The strategy of Lakatosian defence and two principles what warrant it // *Psychological inquiry*. – 1990. – Vol. 1, N 2. – P. 108–141.
- Meehl P.* Why summaries of research on psychological theories are often uninterpretable // *Psychological reports*. – 1990. – Vol. 66. – P. 195–244.
- Mill J.* *A System of logic ratiocinative and inductive.* – Toronto, Toronto univ. press, 1974. – 1254 p.
- Quine W.* Main trends in recent philosophy: Two dogmas of empiricism // *The philosophical review*. – 1951. – Vol. 60, N 1. – P. 20–43.
- Rainey C.* Testing hypotheses of no meaningful effect // *The society for political methodology*. – 2014. – 31 p. – Mode of access: <http://www.polmeth.wustl.edu/media/Paper/nme.pdf> (Дата посещения: 23.08.2014.)
- Rethinking social inquiry: diverse tools, shared standards* / Eds. Brady H., Collier D. – L.: Rowman and Littlefield publishers, 2004. – 428 p.
- Rosato S.* The flawed logic of democratic peace theory // *American political science review*. – 2003. – Vol. 97, N 4. – P. 585–602.
- Ross M.* Does oil hinder democracy? // *World politics*. – 2001. – Vol. 53. – P. 325–361.
- Rozeboom W.* The fallacy of the null-hypothesis significance test // *Psychological bulletin*. – 1960. – Vol. 57, N 5. – P. 416–428.

- 
- Sawyer K., Beed C., Sankey H.* Underdetermination in economics. The duhem-quine thesis // *Economic and philosophy*. – 1997. – Vol. 13, N 1. – P. 1–23.
- Schrodt P.* Seven deadly sins of contemporary quantitative political analysis // *Journal of peace research*. – 2014. – Vol. 51, N 2. – P. 287–300.
- Stanford K.* Underdetermination of scientific theory // *Stanford encyclopedia of philosophy*. – 2009. – 12 August. – Mode of access: <http://plato.stanford.edu/entries/scientific-underdetermination/> (Дата посещения: 08.03.2014.)
- Ward M., Greenhill B., Bakke K.* The perils of policy by p-value // *Journal of peace research*. – 2010. – Vol. 47, N 4. – P. 363–375.
- Welzel C.* Theories of democratization // *Democratization* / Ed. by Christian Haerpfer, Patrick Bernhagen, Ronald Inglehart, Christian Welzel. – Oxford: Oxford univ. press, 2009. – P. 74–90.
- Welzel C., Inglehart R.* The role of ordinary people in democratization // *Journal of democracy*. – 2008. – Vol. 19, N 1. – P. 126–140.
- Western B., Jackman S.* Bayesian inference for comparative research // *The American political science review*. – 1994. – Vol. 88, N 2. – P. 412–423.