

А.М. Либман

д.э.н., PhD (Economics), профессор, Берлинский Свободный университет (Берлин)

ВОЗМОЖНЫ ЛИ ОТКРЫТИЯ В ОБЩЕСТВЕННЫХ НАУКАХ?

Аннотация. История естественных наук содержит множество примеров непреднамеренных открытий — выводов, противоречащих господствующей на тот момент теории или полученных случайно; а также «открытий на кончике пера» — нетривиальных (и иногда противоречащих здравому смыслу) выводов теоретиков, впоследствии нашедших экспериментальное подтверждение. В настоящей статье рассматривается возможность подобных открытий в общественных науках на примере количественных исследований. Показывается существование своего рода «замкнутого круга»: количественные исследования в принципе способны стать источником непреднамеренных открытий, однако многие полученные результаты являются сомнительными. Теория позволяет отсеивать подобные результаты, однако активное использование «теоретического фильтра» ведет к другой проблеме — научное сообщество вместо тестирования теорий начинает заниматься интерпретацией фактов в соответствии с выводами теорий.

Ключевые слова: *непредвиденные открытия, открытия на кончике пера, количественные методы, революция достоверности, проверка устойчивости.*

JEL: B40, C51, C93.

DOI: 10.52342/2587-7666VTE_2021_4_24_40.

Введение

Одним из центральных элементов представлений о сущности науки является идея «научного открытия» — выявления принципиально новых фактов и закономерностей функционирования предмета изучения той или иной научной дисциплины. При этом и в массовом восприятии, и во внутренней мифологии науки огромную роль играют принципиально новые, «революционные» открытия и прорывы. Можно говорить о двух типичных историях такого «научного прорыва». Во-первых, речь идет о *непреднамеренном* открытии, когда исследователи в своей работе (возможно, случайно) сталкиваются с чем-то неожиданным, заранее непредсказуемым явлением природы или феноменом. Примером можно считать открытие А. Флемингом пенициллина или А. Беккерелем радиоактивности¹. Во-вторых, история науки содержит множество примеров противоположной динамики — «открытий на кончике пера», когда теоретическое предсказание получает (порой десятилетия спустя) подтверждение в исследованиях реального мира. «На кончике пера» были открыты, например, планета Нептун (в отношении которой Д. Араго, собственно говоря, и использовал эту формулировку) или бозон Хиггса. Но еще более впечатляющим примером являются открытие антиматерии или выводы теории относительности или

¹ Можно спорить, в какой степени такие открытия возможны в современной науке [Foletti, Fais, 2019].

квантовой механики — в этом случае чисто формальная теория позволила сформулировать утверждения, в каком-то смысле противоречащие здравому смыслу, основанному на повседневном опыте, и, несмотря на это, в итоге подтвержденные экспериментами.

Историю общественных наук гораздо сложнее представить как историю открытий, подобных естественно-научным [Соколов, 2015] (хотя в дисциплинах, в меньшей степени склонных к внутренней рефлексии, таких как экономическая наука, — эти нарративы безусловно присутствуют). Речь идет не об отсутствии в общественных науках «революций» и новых «поворотов» — наоборот, такие явления происходят с завидным постоянством. Однако найти примеры революций, возникающих именно под воздействием новых фактов, ведущих к пересмотру существующих ранее аргументов и теорий, в общественных науках достаточно сложно. Революции чаще возникают под воздействием внутренней логики дисциплины (например, пересмотра самого понятия «научности», заимствований из других наук, интроспекции — обращения к собственному опыту исследователя или изменений норм и ценностей общества). Порой факты заставляют учёных пересмотреть свои подходы², но это происходит далеко не во всех областях общественных наук (и к тому же часто речь идет не об открытии новых фактов, а об изменениях общественной реальности [Кошовец, 2019]). Например, расцвет экспериментальной и поведенческой экономики связан не столько с тем, что исследователям удалось выявить «ошибочность» моделей рационального выбора, сколько с желанием расширить инструментарий исследований — то, что рациональный выбор представляет собой очень серьёзное упрощение человеческого поведения, было ясно с самого начала.

Конечно, в какой-то степени описанная ситуация связана с самой природой социальной реальности — сложностью человеческого поведения (которое часто не вписывается в рамки тех или иных теорий) и изменчивостью общественных институтов и практик. Однако, несмотря на это, все равно встает вопрос о *принципиальной* способности общественных наук «производить» непреднамеренные открытия и нетривиальные «открытия на кончике пера». Настоящая статья посвящена обсуждению проблем, с которыми в этой связи сталкиваются общественные науки. Конкретно, в центре внимания находятся количественные исследования в общественных науках — выявление закономерностей общественного развития с помощью статистических методов на основе достаточно больших наборов данных. Количественные исследования, как известно, доминируют в экономике³, однако и в политологии и социологии они играют важную роль.

В основе статьи лежит важное допущение, которое необходимо сформулировать с самого начала, — представление об общественных науках как об *эмпирических*, в которых важнейшим критерием истины является соответствие научных утверждений наблюдаемой реальности, и в то же время *основанных на теории*. Это представление о сути общественных наук разделяется далеко не всеми исследователями. Антропология, например, скептически относится к попыткам «теоретизирования» и «генерализации», сосредоточившись на сборе эмпирического материала (исторически подобный подход существовал и в экономике, например в немецкой исторической школе). За пределами общественных наук история также представляет собой дисциплину, в которой вполне возможны «непреднамеренные открытия» (скажем, новые архивные документы), но которая отказывается от чрезмерного увлечения абстрактной теорией. Существуют и направления общественных наук, скептически относящиеся к эмпирике и предпочитающие ей уже упомянутую интроспекцию или чисто логическую проверку теоретических утверждений. Тем не менее я начинаю анализ с признания необходимости как теории, так и эмпирических исследований.

² Так, историю макроэкономики XX в. в какой-то степени можно считать ответом на вызовы новых кризисов.

³ Это отчасти связано с тем, что квантификация основных экономических величин (например, цен) более «естественна» и «самоочевидна», чем основных величин других общественных дисциплин [Schumpeter, 1933].

«Непреднамеренные открытия» в количественных исследованиях

Начнём наше обсуждение с феномена «непреднамеренных открытий». В принципе никаких проблем здесь в количественных исследованиях, казалось бы, возникнуть не должно. Подобно тому, как физик может в процессе эксперимента столкнуться с принципиально новым эффектом, экономист или социолог может в процессе анализа данных обнаружить корреляцию, которую существующая теория игнорировала, или установить, что связь переменных оказывается прямо противоположной той, которую предсказывала теория. Собственно говоря, в этом и заключается кажущееся преимущество количественных методов с их (по сравнению с качественными исследованиями) однозначными выводами — статистически значимая связь переменных, скажем, в регрессионном или корреляционном анализе является недвусмысленным свидетельством существования эмпирического факта — возможно, противоречащего существующей теории. Немало исследователей на практике занимаются именно подобным «поиском корреляций» в эмпирических данных.

Однако, несмотря на распространённость этой практики (так называемого «data mining»), в общественных науках она воспринимается крайне негативно [Spanos, 2000]. Для этого есть причины. Количественные исследования в общественных науках уже давно сталкиваются с рядом серьёзных проблем, в какой-то степени противоречащих главному аргументу в их пользу — независимости результатов от личности исследователя, его предпочтений и убеждений. В принципе использование количественной методологии позволяет любому исследователю, работающему с теми же самыми данными, получить одни и те же выводы. Однако это представление не подтверждается практикой. Проведённые эксперименты (когда несколько команд учёных независимо друг от друга изучали одну и ту же тему с использованием одних и тех же наборов данных) показывают, что их выводы порой принципиально отличаются друг от друга [Young, 2009; Silberzahn, Uhlmann, Martin, Anselmi, Aust, Awtrey et al., 2018] (см. также дискуссию в [Auspurg, Brüderl, 2021]⁴). Это связано с несколькими факторами, один из которых особенно важен для интересующей нас темы — речь идет о так называемом «p-hacking». Для того, чтобы повысить шансы на публикацию, исследователи «подгоняют» свои методы (скажем, набор контрольных переменных в регрессии) таким образом, чтобы получить статистически значимые результаты, т.е. «обнаружить эффект», который и может стать предметом статьи. Исследования подтверждают значимость проблемы для большого числа дисциплин [McCloskey, 1985; De Long, Lang, 1992; Ioannidis, 2005; Fanelli, 2010; Fanelli, 2012; Lenz, Sahn, 2021].

В результате вполне возможным становится следующий сценарий — исследователь случайно «сталкивается» с любопытной корреляцией, а затем выбирает дальнейшие шаги своего исследования так, чтобы «подтвердить» полученный вывод. Последнее, кстати, не всегда является следствием злонамеренной воли, сознательно «искажающей» результаты — целый ряд интуитивно кажущихся верными исследовательских практик с неизбежностью приведет к тем же самым результатам [Gelman, Loken, 2014; Loken, Gelman, 2017]; к тому же учёному психологически сложно отказаться от интересного и многообещающего вывода. Проблема усиливается вследствие существования так называемого «publication bias» — журналы не склонны публиковать статьи, в которых сформулированные авторами гипотезы не были подтверждены, т.е. статистически значимый эффект обнаружить не удалось. «Publication bias» также наблюдается во многих отраслях общественных наук [Rosenthal, 1979; Gerber, Malhotra, 2008; Franco, Malhotra, Simonovits, 2014; Auspurg, Hinz, Schneck, 2014; Berning, Weiß, 2016; Andrews, Kasy, 2019]. В результате вполне возможна следующая дина-

⁴ Любопытно, что впервые об этой проблеме количественных исследований практически при возникновении эконометрики говорил Дж. М. Кейнс [Keynes, 1940].

мика: учёные в поисках «интересных» наблюдений обнаруживают те или иные корреляции переменных в эмпирических данных; эти корреляции становятся предметом научных статей (причем в статьях приводятся только элементы анализа, подтверждающие существование корреляции); а многочисленные исследования, в которых обнаружить эффект не удалось, так и не доходят до публикации (а иногда и не покидают письменного стола исследователя). В результате при чтении научных статей создается искажённое впечатление о состоянии эмпирических исследований в той или иной области в целом.

Пример — литература, посвященная влиянию исторических факторов на современное политическое, социальное и экономическое развитие (*historical legacies*). В последние десятилетия в экономических и политологических журналах были опубликованы многочисленные статьи, показывающие невероятно высокий уровень устойчивости социально-экономических явлений во времени. Наследие европейских континентальных империй, распавшихся после Первой мировой войны, колониальных империй (и доколониальных государств), средневековых городов и гильдий, распространения крепостного права или даже неолитических земледельческих практик оказывают влияние на современные общество и политику. Вопрос состоит, однако же, в том, сколько исследований существует, в которых такую устойчивость выявить не удалось, но про которые нам ничего не известно, поскольку они никогда не были опубликованы [Libman, 2018]?

При этом устойчивый тренд последних десятилетий — постоянно растущая доступность данных для исследователей. Государственные органы, крупные исследовательские проекты, международные организации, частные компании и НКО постоянно производят всё увеличивающийся объем данных, в которых учёные могут искать (и находить) корреляции, на самом деле являющиеся лишь артефактом статистического исследования [Smith, Cordes, 2020]. Прекрасным примером является то обстоятельство, что эмпирические исследования порой действительно обнаруживают корреляцию между человеческим поведением (скажем, преступностью) и знаком Зодиака, под которым родился человек, — речь идет, конечно, о статистической иллюзии [von Eye, Losel, Mayzer, 2003]. Склонность научного сообщества «вознаграждать» исследователей за публикацию новых и оригинальных результатов с неизбежностью ведет к накоплению таких статистических иллюзий [Grunow, Schneider, Wagner, Wagner, 2018]. Проблема усиливается по мере введения в научный оборот по-настоящему гигантских наборов данных — так называемых *big data*, — открывающих особенно большие возможности для поиска значимых корреляций. Исследования показывают, что в достаточно больших базах данных учёные с неизбежностью обнаружат корреляции переменных, даже если все эти переменные были сгенерированы случайным образом [Caluda, Long, 2017].

Иначе говоря, разрешив *data mining*, общественные науки выдали бы индульгенцию на введение в научный оборот огромного числа новых «фактов», являющихся на самом деле результатом использования сомнительных методов. Центральный элемент, используемой в большинстве общественных наук количественной методологии — тесты значимости (*significance tests*), — почти неизбежно гарантирует какую-то долю «ложно-положительных» результатов, когда учёный фиксирует взаимосвязь переменных, которой на самом деле не существует. В сочетании с *p-hacking* такие результаты вполне могут получить признание в дисциплине. Соответственно, чтобы этого избежать, практика *data mining* должна быть так или иначе ограничена. Примером того, что происходит в отсутствие этих ограничений, является так называемый кризис достоверности (*credibility crisis*) в психологии — многие необычные и оригинальные утверждения, основанные на психологических экспериментах и опубликованные в научных журналах, никому, кроме авторов статей, воспроизвести не удастся [Shrout, Rodgers, 2018]. На сегодняшний день в целом ряде областей психологии читателю научной статьи следует помнить, что ее результаты с вероятностью более 50% невозпроизводимы [Open Science Collaboration, 2015].

Общественные науки предлагают два варианта «корректировки» процесса накопления эмпирических результатов количественных методов: а) фокус на причинно-следственных связях; б) проверку устойчивости результатов (*robustness*) или их чувствительности к допущениям модели (*sensitivity*). В дальнейшем мы обсудим каждый из этих подходов⁵.

«Революция достоверности»

Внимание к причинно-следственным связям находится в центре так называемой «революции достоверности» (*credibility revolution*), произошедшей в экономике в начале 2000-х гг. и постепенно распространившейся на другие общественные науки [Angrist, Pischke, 2010; Samii, 2016; Baldassarri, Abascal, 2017]. «Революция достоверности» вводит дополнительные требования к эмпирическому анализу: от него требуется не только продемонстрировать существование корреляции переменных, но и показать, что изменения одной из переменных — причина изменений значений другой переменной. «Причинность» при этом понимается в категориях *контрфактического анализа*: если при прочих равных (*ceteris paribus*) изменение значений одной из переменных сопровождается изменением значений другой переменной, то эта первая переменная является причиной второй. В реальности допущение *ceteris paribus* практически никогда не соблюдается: всегда существуют некоторые ненаблюдаемые переменные, влияющие и на потенциальную причину, и на потенциальное следствие [Gelman, 2011]⁶. «Революция достоверности» требует от исследователя использовать только такие методы, в которых влияние этих ненаблюдаемых переменных можно устранить (или оно незначительно). Идеалом для учёного становится эксперимент [Dunning, 2012]. От него ожидается или обеспечение *случайного* распределения значений интересующей переменной-причины в выборке («настоящий» эксперимент, или *randomized controlled trial, RCT*), или, как минимум, нахождение такого эмпирического кейса, в котором распределение значений переменной причины оказалось бы практически случайным (или вызванным внешними факторами, никак не связанными с исследуемым явлением), — так называемый «естественный эксперимент» [Leatherdale, 2019]. В отсутствие полноценных экспериментов, исследователь должен исключить влияние ненаблюдаемых переменных статистическими методами — или с использованием метода инструментальных переменных (который на практике часто сочетается с экспериментами) [Angrist, Krueger, 2001; Sovey, Green, 2011], или, как минимум, за счёт адекватного подбора контрольных переменных в регрессии, в той степени, в которой это позволяют данные [Morgan, Winship, 2015].

Требования «революции достоверности» в теории позволяют исследователю избежать ловушки «случайных корреляций». Научный интерес представляют исключительно результаты, соответствующие строгим критериям эксперимента. Увы, и здесь ситуация не столь благополучна. Во-первых, эксперименты все же открывают определенные возможности для манипуляции и отбора наиболее привлекательных результатов: яркий пример тут — уже упомянутая психология. Неслучайно всё большее распространение получает требование «перерегистрации экспериментов». От учёного ожидается, что он заранее, еще до проведения эксперимента, публично объявит его дизайн и основные гипотезы; тогда

⁵ Исторически существовала еще одна альтернатива — отбор «оптимальных» моделей на основе формальных статистических критериев [Du Plessis, 2006]. Однако на сегодняшний день в общественных науках роль этих подходов является маргинальной — в отличие, кстати, от статистики, в которой существуют и совершенствуются методы поиска оптимальных моделей.

⁶ Существуют сомнения, соответствует ли в принципе концепция контрафактической причинности сложной природе социальных явлений, где все переменные одновременно влияют друг на друга [Karafilidis, 2013].

опубликовать случайно найденную корреляцию, скрыв часть «неудачных» результатов, будет гораздо сложнее. Статьи, использующие инструментальные переменные, очень часто страдают от проблемы *p-hacking* [Brodeur, Cook, Heyes, 2020]. Более того, активное использование методов причинно-следственного анализа в большом числе независимых друг от друга проектов может привести к тому, что эксперименты начнут влиять друг на друга⁷, а выводы статей, использующих инструментальные переменные, станут противоречить допущениям, необходимым для применения этого метода в других статьях [Gallen, Raymond, 2021; Mellon, 2021].

Во-вторых, набор тем, которые могут быть исследованы с использованием методов «революции достоверности», ограничен. Во многих ситуациях «настоящий» эксперимент (RCT) провести невозможно, а «естественные эксперименты» или хорошие инструментальные переменные отсутствуют. В результате учёные вместо поиска интересных тем для анализа сосредотачивают свое внимание на поиске возможностей для проведения эксперимента (или уже существующих естественных экспериментов). Обычно речь идет о достаточно специфических ситуациях, на основе которых крайне сложно сформулировать обобщённые выводы [Rodrik, 2008]⁸. К тому же допущения, на которых основаны эксперименты, также нередко подвергаются критике — речь идет как о принципе экспериментов в целом, так и о его воплощении в конкретных исследованиях общественных наук [Rosenzweig, Wolpin, 2000; Sims, 2010; Sekhon, Titiunik, 2012; Deaton, Cartwright, 2018; Krauss, 2018; Ravallion, 2020] (см. также [Лубман, 2018]). Возникает такое же «лоскутное одеяло» различных, порой не связанных между собой, результатов, как и в науке до «революции достоверности». Просто место многочисленных корреляций, выявленных с помощью *data mining*, занимают многочисленные эксперименты, выводы которых вполне могут противоречить друг другу и в конечном счёте не позволяют сформулировать общую картину изучаемого явления [Капелюшников, 2018].

Проверка устойчивости

Альтернатива ориентации на поиск причинно-следственных взаимосвязей в общественных науках хорошо известна и появилась задолго до «революции достоверности» [Leamer, 2010]: речь идет о тестах устойчивости результатов, выявляющих, как зависят оценки моделей от их параметров. Вычислительные мощности, доступные сегодня практически любому исследователю, с легкостью позволяют оценить тысячи и даже миллионы альтернативных моделей и таким образом выявить, насколько стабильными являются результаты [Sala-i-Martin, 1997; Athey, Imbens, 2017; Young, Holsteen, 2017; Munoz, Young, 2018; Simonsohn, Simmons, Nelson, 2019]. Такой подход также позволяет исключить множество случайным образом найденных корреляций — для науки важны лишь те корреляции, которые остаются стабильными после многочисленных модификаций модели. Проверка

⁷ Например, в социологии для изучения дискриминации активно используются так называемые «correspondence studies». В этом случае социологи создают базу фиктивных резюме, различающихся только одной характеристикой (фамилией соискателя, выдающей его принадлежность к этническому большинству, или полом соискателя) и рассылают их компаниям в ответ на объявления о найме персонала или владельцам жилья в ответ на объявления о сдаче в аренду квартир. Если работодатели реже приглашают на собеседование женщин, чем мужчин (несмотря на идентичность всех остальных характеристик резюме), это можно считать свидетельством дискриминации. Проблема состоит в том, что если множество социологов независимо друг от друга проводит такие исследования, а их результаты обсуждаются в СМИ, раньше или позже работодатели начнут приспосабливаться к этому, исходя из возможности получения ими резюме, направленного в рамках эксперимента, что сделает исследование бессмысленным.

⁸ В этом отношении эксперименты сливаются со своей, казалось бы, полной противоположностью — сугубо качественными методами антропологии [Shadish, 1995].

устойчивости (robustness check) в том или ином виде является сегодня обязательным элементом практически всех эмпирических работ; все большее распространение получают и формальные методы, позволяющие учесть все без исключения аспекты чувствительности результатов к допущениям модели.

Тем не менее и этот подход имеет ряд недостатков. Во-первых, единого стандарта проверки устойчивости в общественных науках (и даже в отдельных дисциплинах) не существует, а это означает, что исследователи могут «отбирать» инструменты проверки устойчивости по своему усмотрению, что, в свою очередь, открывает многочисленные возможности для *p-hacking*. Во-вторых, во многих случаях проверка устойчивости сводится к оценке многочисленных моделей, каждая из которых обладает тем или иным недостатком (а порой не соответствует базовым допущениям для используемого метода оценки). Непонятно, почему мы должны доверять множеству по существу неверных моделей, вместо того чтобы использовать «лучшую из возможных» моделей (как предполагает «революция достоверности», которая исходит из существования своего рода «золотого стандарта» исследований). Наконец, в-третьих, не совсем понятно, какие модели следует считать «неустойчивыми». Например, допустим, что результаты исследования воспроизводимы лишь с определённым набором контрольных переменных. Однако, возможно, именно этот набор контрольных переменных и следует включать в анализ для корректной оценки регрессий — а многочисленные регрессии, в которых результат воспроизвести не удалось, страдают от проблемы влияния ненаблюдаемых переменных. Как следует поступать в этом случае?

Идея проверки устойчивости связана с активно обсуждающейся сегодня темой — *репликацией* научных результатов [King, 1995; Hamermesh, 2007; Freese, Peterson, 2017; Duvendak, Palmer-Jones, Reed, 2017]⁹. В естественных науках многие «случайные открытия» (как и открытия в целом) также оказываются неверными или основанными на результатах сомнительных экспериментов. Однако против таких результатов работают социальные практики науки — полученные результаты подвергаются проверке другими учёными, раньше или позже выявляющими допущенные ошибки (невоспроизводимость результатов). Конечно, этот механизм отбора работает далеко не всегда, однако он хотя бы отчасти решает проблему сомнительных выводов отдельных статей. Иначе говоря, существует своего рода внутренний механизм корректировки науки, а, следовательно, публикация неверных результатов не является проблемой (раньше или позже они будут опровергнуты другими исследователями). Сама по себе репликация, кстати сказать, также может вести к «непреднамеренным открытиям».

В общественных науках до недавнего времени репликация результатов других исследований была скорее исключением, чем правилом. Сегодня ситуация несколько улучшилась, но все равно число учёных, пытающихся воспроизвести результаты уже существующих работ, остается крайне ограниченным. Это связано с существующей системой стимулов для исследователей — экономические, социологические или политологические журналы неохотно публикуют статьи-репликации (что в принципе можно понять: для редакторов интереснее опубликовать статьи, содержащие принципиально новые выводы, которые будут более активно цитироваться). Однако в какой-то степени недостаточное внимание к репликации связано с самой сущностью социальной реальности. Многие эффекты, полученные для той или иной выборки (скажем, в конкретной стране или социальной группе или для определенного периода времени), не воспроизводятся для других выбо-

⁹ Термин «репликация» обычно используется в двух значениях — проверка опубликованных результатов с использованием тех же методов и данных, что и в исходном исследовании, и «научная репликация» — проверка чувствительности результатов к изменениям в методологии и воспроизводство результатов для других выборок (стран, регионов, фирм и т.д.), чем использовавшиеся в первоначальном исследовании. Нас интересует именно вторая форма репликации, хотя на практике уже первая форма свидетельствует о существовании серьёзных проблем в общественных науках.

рок не потому, что исходные результаты были неверными, а потому, что условия в других выборках принципиально отличаются в силу другого социального контекста. В этом случае репликация ничего не говорит о качестве исходного исследования, а лишь показывает, что его выводы можно использовать лишь в конкретном социальном контексте.

Итак, и проверка устойчивости, и «революция достоверности» не позволяют полностью решить проблемы, возникающие при использовании data mining. Неизбежностью становится применение своего рода «фильтра теории» при интерпретации эмпирических данных. Иначе говоря, от учёного следует требовать не только проработанного дизайна эмпирического исследования (позволяющего исключить влияние ненаблюдаемых переменных и проверить устойчивость результатов к допущениям модели), но и демонстрации соответствия работы существующим теоретическим аргументам. Собственно говоря, именно этот подход и считается сегодня оптимальным в общественных науках — любое количественное исследование должно начинаться с детального обсуждения лежащей в его основе теории, на основе которой формулируются гипотезы¹⁰. Как минимум, можно считать результаты, явно противоречащие господствующей теории, «менее достоверными», чем результаты, эту теорию подтверждающие (идея различной степени достоверности результатов в зависимости от исходных ожиданий исследователя лежит в основе байесовской эконометрики, см., например, [Олейник, 2021]). Такой подход обладает многочисленными преимуществами, однако делает «непреднамеренные открытия» весьма маловероятными.

«Открытия на кончике пера» в количественных исследованиях

Возможны ли в общественных науках «открытия на кончике пера», связанные с использованием количественных методов? В центре внимания последующего обсуждения будут находиться именно *нетривиальные* следствия, возможно, противоречащие здравому смыслу, которые могут быть «выведены» из формальных моделей или аргументов теории и потенциально получить подтверждение в эмпирических исследованиях. Опять же, ситуация в общественных науках отличается от ситуации в науках естественных.

Во-первых, многие теоретические модели общественных наук *в принципе* не предназначены для тестирования — скорее, они представляют собой упрощенные описания общественных процессов, облегчающие их понимание. А. Рубинштейн сравнивает теоретико-игровые модели с «баснями», ценность которых не ограничивается их способностью генерировать гипотезы для эмпирического исследования [Rubinstein, 2006]¹¹ (аналогичные выводы для политологии приводятся в [Humpfreys, 2017]), а Х. Улиг считает естественным, что модели продолжают использоваться, даже если они не соответствуют наблюдаемым фактам [Uhlig, 2012]. Допущения моделей по определению упрощают реальность [Либман, 2008]. Соответственно, нет ничего странного в том, что они непригодны для открытий «на кончике пера». Во-вторых, в значительном числе случаев теории общественных наук, находящие подтверждение в эмпирике, в принципе, могут быть выведены и из элементарного «здравого смысла». Это, конечно, не снижает значимость теоретического исследования — формальная теоретико-игровая модель или использование точной научной терминологии позволяют выявить логические нестыковки в аргументации, — но всё же ведет к принципиально иной ситуации, чем та, с которой мы сталкиваемся в естественных

¹⁰ Другой вопрос, что на самом деле идеал эмпирического исследования, основанного на теории, часто нарушается: исследователи или уделяют теории гораздо меньше внимания, чем дизайну эмпирического исследования (сводя теорию к набору ссылок на существующую литературу), или даже готовят теоретическую главу исследования уже после того, как ими были получены эмпирические результаты.

¹¹ Рубинштейн А. Дилеммы экономиста-теоретика // Вопросы экономики, 2008. №11. С. 62–80. (В русском переводе. — Прим. ред.).

науках. В-третьих, в некоторых случаях корректнее говорить не о теории, *описывающей* реальность, а о теории, *меняющей* реальность — например, если теория становится основой поведения акторов или государственной политики.

Тем не менее есть и теории, не сводимые к уточнённой формулировке утверждений, представляющихся очевидными и с точки зрения здравого смысла. Во многих случаях господствующие в обществе представления о тех или иных явлениях противоречат выводам теории, но при этом справедлива именно последняя [Krugman, 1998] (другой вопрос, противоречит ли теория в этой ситуации *здравому смыслу* — или, скорее, нормативным убеждениям, сложившимся в обществе в тот или иной момент времени). Есть и примеры однозначно нетривиальных выводов теории, которые используются на практике, например при дизайне аукционов, избирательных систем или налогообложения [Roth, 2002]. В какой степени при формулировке государственной политики следует полагаться на выводы теории, а в какой на неформализованное знание практиков, является спорным [Colander, 1992; Mankiw, 2006; Su, Kolander, 2021]. Для нас важнее то, что такой дизайн институтов, основанный на теории, отчасти может считаться «открытием на кончике пера»: нетривиальные выводы теории позволяют описать человеческое поведение лучше, чем здравый смысл.

Однако такие нетривиальные открытия все же нетипичны для общественных наук. Гораздо чаще последние сосредоточены на совершенно другой задаче — *нестандартной интерпретации* человеческого поведения на основе той или иной теории. Так, одной из важнейших задач экономического империализма было объяснение разнообразных общественных феноменов (семьи, религии, политики) с использованием экономических моделей (и, шире, теории рационального выбора). Стандартом для моделей, скажем, теории общественного выбора (public choice) или школы «экономика и права» (law and economics) было описание тех или иных политических институтов или правовых учреждений «языком» экономики¹² (современная экономика отказалась от такого подхода — если в прошлом экономисты считали теорию рационального выбора конституирующим признаком своей дисциплины, сегодня им стали эмпирические методы «революции достоверности»). Аналогично для критической социологии основная задача — «выявление» скрытых форм неравенства и иерархий, существующих в обществе¹³. Подобно тому, как для теоретика общественного выбора любая иерархия представляет собой специфическую форму обмена, для критического социолога любой обмен основан на скрытой иерархии. Теория в этой ситуации пытается объяснить, чем «на самом деле» обусловлены поведение людей и развитие общественных институтов, постулируя, что общепринятые представления о поведении людей (в том числе их собственные представления о своей мотивации) неверны, и лишь «вооруженный» теоретическим аппаратом «глаз» может выявить подлинные механизмы функционирования общества.

Такой подход связан с несколькими проблемами. Во-первых, он делает практически невозможной фальсификацию базовых положений теории: если *задача* науки сводится к интерпретации всего в обществе как эксплуатации и угнетения или как рационального поведения, практически невозможно представить себе постановку вопроса, в которой те или иные общественные явления *не были бы* продуктом эксплуатации или рационального выбора. Иначе говоря, основные допущения теории невозможно протестировать эмпирически. В каком-то смысле учёные к этому и не стремятся — для них важнее дать соответствующую их теории интерпретацию наблюдаемых ими явлений, т.е. объяснить, каким образом эмпирические факты соответствуют теории, — а не отделить факты, соот-

¹² Такой подход принято называть конструированием «аналитических нарративов» [Bates, Greif, Levi, Rosenthal, Weingast, 1999].

¹³ Следует оговориться, что критическая теория почти никогда не используется в сочетании с количественными методами эмпирического исследования.

ветствующие теории, от фактов, теории противоречащих. Во-вторых, в предельном случае подобная всеобъемлющая теория имеет все шансы превратиться в идеологию. Именно такие обвинения регулярно звучат в адрес «неолиберальных экономистов» или «гендерных исследований»¹⁴. В этом случае сомневаться в теории становится опасно для карьеры исследователя, а значит, теория становится безальтернативной.

Интерпретация человеческого поведения на основе той или иной исследовательской парадигмы сама по себе, вне всякого сомнения, является интересным интеллектуальным вызовом, и с этой точки зрения заслуживает внимания. Ставя под вопрос утверждения, кажущиеся очевидными, и предлагая нетривиальный взгляд на социальные институты и практики (а также часто рекомендации для государственной политики), наука создает потенциал для новых нетривиальных исследовательских гипотез. Однако с точки зрения дисциплины, для которой эмпирика (а не соответствие базовым допущениям теории) является критерием истины, интереснее была бы теория, формулирующая *тестируемые нетривиальные гипотезы*. А в этом отношении общественные науки далеко не так успешны.

Заключение

Итак, общественные науки оказываются в своего рода замкнутом круге. С одной стороны, количественные исследования, хотя и позволяют учёным в принципе генерировать «непредвиденные результаты», в то же время не в состоянии «отсеивать» результаты, которым действительно следует уделять внимание, от статистических артефактов и случайных корреляций. Более того, стимулы, с которыми сталкиваются учёные (жёсткая конкуренция в науке и оценка учёных на основе публикационной активности), ведут к распространению исследовательских практик, способствующих публикации все новых и новых статей, содержащих сомнительные результаты. В результате «революции достоверности» на место многочисленных и противоречащих друг другу корреляций приходят многочисленные и противоречащие друг другу результаты экспериментов (в которых к тому же исследуется лишь очень узкий спектр тем, которые в принципе могут быть изучены с использованием экспериментов и других методов причинно-следственного анализа).

Опора на теорию позволяет эффективно отсеивать «сомнительные» результаты; в итоге в науке возникает непротиворечивая (и потому привлекательная) картина мира. Однако при этом возникает другая проблема: чрезмерное внимание к теории в конечном счёте превращает науку из проекта по изучению реальности в проект по её интерпретации в категориях теоретической парадигмы. А следовательно, выйти за рамки этой парадигмы становится практически невозможно; в предельном случае многие выводы дисциплины становятся своего рода интеллектуальной игрой (или даже упражнением по использованию соответствующей той или иной идеологии риторики), а не результатом эмпирического исследования.

Существует ли возможность выйти из этого замкнутого круга? В отличие от количественных исследований, качественные исследования предполагают возможность формулировки теории в процессе эмпирического исследования «снизу вверх» (*bottom up theory building*) [Eisenhardt, Graebner, 2007; Gehman, Glaser, Eisenhardt, Gioia, Langly, Corley, 2018]. Впоследствии такая теория может стать предметом количественного исследования [Lieberman, 2005]. Однако качественные исследования также страдают многочисленными проблемами. Их выводы очень зависят от того, каким образом исследователь «организует»

¹⁴ Хорошим примером можно считать дискуссию о так называемых *grievance studies* — удавшуюся провокацию группы учёных, в 2017–2018 гг. опубликовавших в ряде журналов критической теории и гендерных исследований набор заведомо бессмысленных статей, использующих принятую в этих сообществах риторику [Pluckrose, Lindsaz, Boghossian, 2021].

доступные ему факты, на что он обращает внимание (а что игнорирует), и каким образом он интерпретирует доступные ему данные («фильтр» его восприятия может очень серьёзно исказить результаты, особенно если предметом исследования являются общества другой культуры). Иначе говоря, сам по себе факт «открытия» нового феномена или механизма при использовании качественных методов во многом зависит от интерпретации данных исследователем. Учёные, использующие качественные методы не в меньшей степени, чем их коллеги, применяющие методы статистики и эконометрики, склонны к публикации сомнительных результатов, если это позволяет им добиться успеха в науке. К тому же выводы качественных исследований, основанные на изучении конкретного кейса, могут отражать особенности этого частного случая, а не общие закономерности [Bendassolli, 2013].

Ещё один способ выхода из замкнутого круга связан с пересмотром самой сути «теории» и «эмпирики». Так, количественные исследования могут отказаться от претензии на «тестирование» теории и вместо этого сосредоточиться на *описании* данных [Gerring, 2012]. Действительно, в статистике (в отличие от экономики, политологии или социологии) огромную роль играют методы, позволяющие изучать данные и выявлять содержащуюся в них информацию без увязки с какими бы то ни было теоретическими гипотезами; этот же принцип находится в центре внимания бурно развивающейся дисциплины «data science». Однако будет ли подобный подход достаточным? Важным преимуществом теории является то, что она не только формулирует гипотезы относительно взаимосвязи переменных, но и даёт объяснения, *почему* эта связь носит тот или иной характер (т.е. описывает конкретный *механизм*, ведущий к тому или иному социальному явлению). Можем ли мы отказаться от этого аспекта общественных наук? В конце концов, одной из главных причин упадка *area studies* — междисциплинарных исследовательских сообществ, посвящённых отдельным регионам мира, — была именно слабость их теоретического аппарата, в итоге оказавшаяся более важной, чем глубокие эмпирические знания¹⁵.

В любом случае, приходится признать, что в настоящей момент привлекательной альтернативы, позволившей бы общественным наукам создавать теории, формулирующие нетривиальные выводы относительно функционирования общества, пригодные для эмпирического тестирования, и в то же время открытой для «непредвиденных открытий», не существует. Сказанное не следует воспринимать как свидетельство слабости или кризиса эмпирики и теории в общественных науках. Скорее, речь идет об определённых границах, за которые количественная политология, экономика и социология в настоящее время не могут выйти. И это требуется учитывать при интерпретации выводов данных дисциплин.

ЛИТЕРАТУРА

- Капелюшников Р.И. (2018). О современном состоянии экономической науки: полусоциологические наблюдения // Куда движется современная экономическая наука? М.: ИЭ РАН. С. 8–33.
- Кошовец А.Б. (2019). «Горизонтальный прогресс» экономической науки: между конструируемой реальностью и технотехнологией. М.: ИЭ РАН.
- Либман А.М. (2008). Теоретические и эмпирические исследования в современной экономике: проблемы коммуникации // Вопросы экономики. № 6. С. 4–19.
- Либман А.М. (2018). Эмпирические исследования в экономике: «Революция достоверности»? // Куда движется современная экономическая наука? М.: ИЭ РАН. С. 34–52.
- Олейник А.Н. (2021). О возможных методологических основаниях прагматической экономической теории: Т. Байес и Дж. Дьюи // Вопросы теоретической экономики. №3. С. 34–51.
- Соколов М.М. (2015). Социология как чудо. Процесс *sense-building* в одной академической дисциплине // Социология власти. Т. 27. № 3. С. 13–57.

¹⁵ Еще один пример — так называемая «эконофизика», попытка исследования экономики методами стохастической физики, также во многом абстрагировавшаяся от теоретизирования, но в итоге не сумевшая превратиться в реальную альтернативу экономической науке [Ball, 2006].

- Andrews I., Kasy M. (2019). Identification of and correction for publication bias // *American Economic Review*. Vol. 109. No. 8. Pp. 2766–2794.
- Angrist J.D., Krueger A.B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments // *Journal of Economic Perspectives*. Vol. 15. No. 4. Pp. 69–85.
- Angrist J.D., Pischke J.S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics // *Journal of Economic Perspectives*. Vol. 24. No. 2. Pp. 3–30.
- Athey S., Imbens G.W. (2017). The state of applied econometrics: Causality and policy evaluation // *Journal of Economic Perspectives*. Vol. 31. No. 2. Pp. 3–32.
- Auspurg K., Brüderl J. (2021). Has the credibility of the social sciences been credibly destroyed? Reanalyzing the “many analysts, one data set” project // *Socius*. Vol. 7. Pp. 1–14.
- Auspurg K., Hinz T., Schneck A. (2014). Ausmaß und Risikofaktoren des Publication Bias in der deutschen Soziologie // *Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Vol. 66. Issue 4. Pp. 549–573.
- Baldassarri D., Abascal M. (2017). Field experiments across the social sciences // *Annual Review of Sociology*. Vol. 43. Pp. 41–73.
- Ball P. (2006). Econophysics: Culture crash // *Nature*. Vol. 441. Issue 7094. Pp. 686–689.
- Bates R.H., Greif A., Levi M., Rosenthal J.L., Weingast B.R. (1999). *Analytic narratives*. Princeton: Princeton University Press.
- Bendassolli P.F. (2013). Theory building in qualitative research: Reconsidering the problem of induction // *Forum Qualitative Sozialforschung*. Vol. 14. No. 1. Pp. 1–20.
- Berning C.C., Weiß B. (2016). Publication bias in the German social sciences: an application of the caliper test to three top-tier German social science journals // *Quality & Quantity*. Vol. 50. No. 2. Pp. 901–917.
- Brodeur A., Cook N., Heyes A. (2020). Methods matter: P-hacking and publication bias in causal analysis in economics // *American Economic Review*. Vol. 110. No. 11. Pp. 3634–3660.
- Calude C.S., Longo G. (2017). The deluge of spurious correlations in big data // *Foundations of Science*. Vol. 22. No. 3. Pp. 595–612.
- Colander D. (1992). The lost art of economics // *Journal of Economic Perspectives*. Vol. 6. No. 3. Pp. 191–198.
- De Long J.B., Lang K. (1992). Are all economic hypotheses false? // *Journal of Political Economy*. Vol. 100. No. 6. Pp. 1257–1272.
- Deaton A., Cartwright N. (2018). Understanding and misunderstanding randomized controlled trials // *Social Science & Medicine*. Vol. 210. Pp. 2–21.
- Du Plessis S. (2006). The miracle of the Septuagint and the promise of data mining in economics. Economic Working Paper 15/06: Stellenbosch University.
- Dunning T. (2012). *Natural experiments in the social sciences: a design-based approach*. Cambridge: Cambridge University Press.
- Duwendack M., Palmer-Jones R., Reed W.R. (2017). What is meant by “replication” and why does it encounter resistance in economics? // *American Economic Review*. Vol. 107. No. 5. Pp. 46–51.
- Eisenhardt K.M., Graebner, M.E. (2007). Theory building from cases: Opportunities and challenges // *Academy of Management Journal*. Vol. 50. No. 1. Pp. 25–32.
- Fanelli D. (2010). Do pressures to publish increase scientists’ bias? An empirical support from US States Data // *PloS One*. Vol. 5. Issue 4. Article e10271.
- Fanelli D. (2012). Negative results are disappearing from most disciplines and countries // *Scientometrics*. Vol. 90. No. 3. Pp. 891–904.
- Foletti A., Fais S. (2019). Unexpected discoveries should be reconsidered in science—a look to the past? // *International Journal of Molecular Science*. Vol. 20. No. 16. Pp. 1–10.
- Franco A., Malhotra N., Simonovits G. (2014). Publication bias in the social sciences: Unlocking the file drawer // *Science*. Vol. 345. Issue 6203. Pp. 1502–1505.
- Freese J., Peterson D. (2017). Replication in social science // *Annual Review of Sociology*. Vol. 43. Pp. 147–165.
- Gallen T., Raymond B. (2021). Broken instruments. — SSRN Working Paper.
- Gelman J., Glaser V.L., Eisenhardt K.M., Gioia D., Langley A., Corley K.G. (2018). Finding theory–method fit: A comparison of three qualitative approaches to theory building // *Journal of Management Inquiry*. Vol. 27. No. 3. Pp. 284–300.
- Gelman A. (2011). Causality and statistical learning // *American Journal of Sociology*. Vol. 117. No. 3. Pp. 955–966.
- Gelman A., Loken E. (2014). The statistical crisis in science: data-dependent analysis — a “garden of forking paths” — explains why many statistically significant comparisons don’t hold up. // *American Scientist*. Vol. 102. No. 6. Pp. 460–466.
- Gerber A.S., Malhotra N. (2008). Publication bias in empirical sociological research: Do arbitrary significance levels distort published results? // *Sociological Methods & Research*. Vol. 37. No. 1. Pp. 3–30.
- Gerring J. (2012). Mere description // *British Journal of Political Science*. Vol. 42. No. 4. Pp. 721–746.
- Grunow M., Schneider H., Wagner G., Wagner J. (2018). Editorial // *International Journal for Re-Views in Empirical Economics*. Vol. 2. Pp. 1–6.
- Hamermesh, D.S. (2007). Replication in economics // *Canadian Journal of Economics*. Vol. 40. No. 3. Pp. 715–733.

- Humpfreys M.* (2017). *Political games: Mathematical insights on fighting, voting, lying and other affairs of state*. New York: Norton.
- Ioannidis J.P.* (2005). Why most published research findings are false. // *PLoS Medicine*. Vol. 2. Issue 8: e124.
- Karafilidis A.* (2013). Erklärungen in rekursiven Verhältnissen // *Zeitschrift für Theoretische Soziologie*. Vol. 2. No. 2. Pp. 218–239.
- Keynes J.M.* (1940). On a method of statistical business-cycle research. A comment // *Economic Journal*. Vol. 50. Issue 197. Pp. 154–156.
- King G.* (1995). Replication, replication // *PS: Political Science & Politics*. Vol. 28. Issue 3. Pp. 444–452.
- Krauss A.* (2018). Why all randomised controlled trials produce biased results // *Annals of Medicine*. Vol. 50. No. 4. Pp. 312–322.
- Krugman P.* (1998). Two cheers for formalism // *Economic Journal*. Vol. 108. Issue. 451. Pp. 1829–1836.
- Leamer E.E.* (2010). Tantalus on the road to asymptopia // *Journal of Economic Perspectives*. Vol. 24. No. 2. Pp. 31–46.
- Leatherdale S.T.* (2019). Natural experiment methodology for research: a review of how different methods can support real-world research // *International Journal of Social Research Methodology*. Vol. 22. No. 1. Pp. 19–35.
- Lenz G.S., Sahn A.* (2021). Achieving statistical significance with control variables and without transparency // *Political Analysis*. Vol. 29. No. 3. Pp. 356–369.
- Libman A.* (2018). Der lange Schatten der Vergangenheit: Historisches Erbe und quantitative Sozialwissenschaften // *Osteuropa*. Vol. 68. No. 6. Pp. 49–66.
- Lieberman E.S.* (2005). Nested analysis as a mixed-method strategy for comparative research // *American Political Science Review*. Vol. 99. No. 3. Pp. 435–452.
- Loken E., Gelman A.* (2017). Measurement error and the replication crisis // *Science*. Vol. 355. Issue 6325. Pp. 584–585.
- Mankiw N.G.* (2006). The macroeconomist as scientist and engineer // *Journal of Economic Perspectives*. Vol. 20. No. 4. Pp. 29–46.
- McCloskey D.N.* (1985). The loss function has been mislaid: The rhetoric of significance tests // *American Economic Review*. Vol. 75. No. 2. Pp. 201–205.
- Mellon J.* (2021). Rain, rain, go away: 176 potential exclusion-restriction violations for studies using weather as an instrumental variable. — SSRN Working Paper.
- Morgan S.L., Winship C.* (2015). *Counterfactuals and causal inference*. Cambridge: Cambridge University Press.
- Muñoz J., Young C.* (2018). We ran 9 billion regressions: Eliminating false positives through computational model robustness // *Sociological Methodology*. Vol. 48. No. 1. Pp. 1–33.
- Open Science Collaboration* (2015). Estimating the reproducibility of psychological science // *Science*. Vol. 349. Issue 6251.
- Pluckrose H., Lindsay J., Boghossian P.* (2021). Understanding the “grievance studies affair” papers and why they should be reinstated: A response to Geoff Cole // *Sociological Methods & Research*. Forthcoming.
- Ravallion M.* (2020). Should the randomistas (continue to) rule? National Bureau of Economic Research Working Paper No. 27554.
- Rodrik D.* (2008). The new development economics: We shall experiment, but how shall we learn? HKS Working Paper No. RWP08-055.
- Rosenthal R.* (1979). The file drawer problem and tolerance for null results // *Psychological Bulletin*. Vol. 86. No. 3. Pp. 638–641.
- Rosenzweig M.R., Wolpin K.I.* (2000). Natural “natural experiments” in economics // *Journal of Economic Literature*. Vol. 38. No. 4. Pp. 827–874.
- Roth A.E.* (2002). The economist as engineer: Game theory, experimentation, and computation as tools for design economics // *Econometrica*. Vol. 70. No. 4. Pp. 1341–1378.
- Rubinstein A.* (2006). Dilemmas of an economic theorist // *Econometrica*. Vol. 74. No. 4. Pp. 865–883.
- Sala-i-Martin X.* (1997). I just ran two million regressions // *American Economic Review*. Vol. 87. No. 2. Pp. 178–183.
- Samii C.* (2016). Causal empiricism in quantitative research // *Journal of Politics*. Vol. 78. Issue 3. Pp. 941–955.
- Schumpeter J.* (1933). The common sense of econometrics // *Econometrica*. Vol. 1. No. 1. Pp. 5–12.
- Sekhon J.S., Titiunik R.* (2012). When natural experiments are neither natural nor experiments // *American Political Science Review*. Vol. 106. No. 1. Pp. 35–57.
- Shadish W.R.* (1995). The logic of generalization: Five principles common to experiments and ethnographies // *American Journal of Community Psychology*. Vol. 23. No. 3. Pp. 419–428.
- Shrout P.E., Rodgers J.L.* (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis // *Annual Review of Psychology*. Vol. 69. Pp. 487–510.
- Silberzahn R., Uhlmann E.L., Martin D.P., Anselmi P., Aust F., Awtry E., et al.* (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results // *Advances in Methods and Practices in Psychological Science*. Vol. 1. No. 3. Pp. 337–356.
- Simonsohn U., Simmons J.P., Nelson L.D.* (2019). Specification curve: Descriptive and inferential statistics on all reasonable specifications. SSRN Working Paper.
- Sims C.A.* (2010). But economics is not an experimental science // *Journal of Economic Perspectives*. Vol. 24. No. 2. Pp. 59–68.

- Smith G., Cordes J. (2020). The phantom pattern problem: The mirage of big data. Oxford: Oxford University Press.
- Sovey A.J., Green D.P. (2011). Instrumental variables estimation in political science: A readers' guide // *American Journal of Political Science*. Vol. 55. Issue 1. Pp. 188–200.
- Spanos A. (2000). Revisiting data mining: 'hunting' with or without a license // *Journal of Economic Methodology*. Vol. 7. No. 2. Pp. 231–264.
- Su H.C., Colander D. (2021). The economist as scientist, engineer, or plumber? // *Journal of the History of Economic Thought*. Vol. 43. No. 2. Pp. 297–312.
- Uhlig H. (2012). Economics and reality // *Journal of Macroeconomics*. Vol. 34. No. 1. Pp. 29–41.
- von Eye A., Losel F., Mayzer R. (2003). Is it all written in the stars? A methodological commentary on Sachs' astrology monograph and re-analyses of his data on crime statistics // *Psychological Test and Assessment Modeling*. Vol. 45. No. 1. Pp. 78–91.
- Young C. (2009). Model uncertainty in sociological research: An application to religion and economic growth // *American Sociological Review*. Vol. 74. Issue 3. Pp. 380–397.
- Young C., Holsteen K. (2017). Model uncertainty and robustness: A computational framework for multimodel analysis // *Sociological Methods & Research*. Vol. 46. No. 1. Pp. 3–40.

Либман Александр Михайлович

Alexander.libman@fu-berlin.de

Aleksander Libman

PhD (Economics), Professor of Russian and East European Politics, Freie Universität Berlin

Alexander.libman@fu-berlin.de

ARE DISCOVERIES POSSIBLE IN SOCIAL SCIENCES?

Abstract. The history of sciences is full of examples of unintended discoveries — research findings contradicting the dominant theory or obtained by chance — as well as of discoveries made “with the point one’s pen” — non-trivial conclusions of theory (sometimes contradicting the common sense), which however, find experimental confirmation. This paper asks whether these types of discoveries are possible in social sciences and focuses on quantitative research. It shows the existence of a sort of vicious circle: quantitative research is potentially able to produce unintended discoveries but many findings are questionable. It is possible to separate these findings for the more convincing ones using theory but the active use of this “theoretical filter” results in another problem — instead of testing theories, scientific community turns to interpreting facts in line with theories.

Keywords: *unintended discoveries; discoveries made with the point of one’s pen; quantitative methods; credibility revolution; sensitivity analysis.*

JEL: B40; C51; C93.

REFERENCES

- Andrews I., Kasy M. (2019). Identification of and correction for publication bias // *American Economic Review*. Vol. 109. No. 8. Pp. 2766–2794.
- Angrist J.D., Krueger A.B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments // *Journal of Economic Perspectives*. Vol. 15. No. 4. Pp. 69–85.
- Angrist J.D., Pischke J.S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics // *Journal of Economic Perspectives*. Vol. 24. No. 2. Pp. 3–30.
- Athey S., Imbens G.W. (2017). The state of applied econometrics: Causality and policy evaluation // *Journal of Economic Perspectives*. Vol. 31. No. 2. Pp. 3–32.
- Auspurg K., Brüderl J. (2021). Has the credibility of the social sciences been credibly destroyed? Reanalyzing the “many analysts, one data set” project // *Socius*. Vol. 7. Pp. 1–14.
- Auspurg K., Hinz T., Schneck A. (2014). Ausmaß und Risikofaktoren des Publication Bias in der deutschen Soziologie. // *Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Vol. 66. No. 4. Pp. 549–573.
- Baldassarri D., Abascal M. (2017). Field experiments across the social sciences // *Annual Review of Sociology*. Vol. 43. Pp. 41–73.
- Ball P. (2006). Econophysics: Culture crash // *Nature*. Vol. 441. Issue 7094. Pp. 686–689.
- Bates R.H., Greif A., Levi M., Rosenthal J.L., Weingast B.R. (1999). *Analytic narratives*. Princeton: Princeton University Press.
- Bendassolli P.F. (2013). Theory building in qualitative research: Reconsidering the problem of induction // *Forum Qualitative Sozialforschung*. Vol. 14. No. 1. Pp. 1–20.

- Berning C.C., Weiß B. (2016). Publication bias in the German social sciences: an application of the caliper test to three top-tier German social science journals // *Quality & Quantity*. Vol. 50. 2. No. Pp. 901–917.
- Brodeur A., Cook N., Heyes A. (2020). Methods matter: P-hacking and publication bias in causal analysis in economics // *American Economic Review*. Vol. 110. No. 11. Pp. 3634–3660.
- Calude C.S., Longo G. (2017). The deluge of spurious correlations in big data // *Foundations of Science*. Vol. 22. No. 3. Pp. 595–612.
- Colander D. (1992). The lost art of economics // *Journal of Economic Perspectives*. Vol. 6. No. 3. Pp. 191–198.
- De Long J.B., Lang K. (1992). Are all economic hypotheses false? // *Journal of Political Economy*. Vol. 100. No. 6. Pp. 1257–1272.
- Deaton A., Cartwright N. (2018). Understanding and misunderstanding randomized controlled trials // *Social Science & Medicine*. Vol. 210. Pp. 2–21.
- Du Plessis S. (2006). *The miracle of the Septuagint and the promise of data mining in economics*. Economic Working Paper 15/06: Stellenbosch University.
- Dunning T. (2012). *Natural experiments in the social sciences: a design-based approach*. Cambridge: Cambridge University Press.
- Duvendack M., Palmer-Jones R., Reed W.R. (2017). What is meant by “replication” and why does it encounter resistance in economics? // *American Economic Review*. Vol. 107. No. 5. Pp. 46–51.
- Eisenhardt K.M., Graebner, M.E. (2007). Theory building from cases: Opportunities and challenges // *Academy of Management Journal*. Vol. 50. No. 1. Pp. 25–32.
- Fanelli D. (2010). Do pressures to publish increase scientists’ bias? An empirical support from US States Data // *PloS One*. Vol. 5. Issue 4. Article e10271.
- Fanelli D. (2012). Negative results are disappearing from most disciplines and countries // *Scientometrics*. Vol. 90. No. 3. Pp. 891–904.
- Foletti A., Fais S. (2019). Unexpected discoveries should be reconsidered in science — a look to the past? // *International Journal of Molecular Science*. Vol. 20. 16. No. Pp. 1–10.
- Franco A., Malhotra N., Simonovits G. (2014). Publication bias in the social sciences: Unlocking the file drawer // *Science*. Vol. 345. Issue 6203. Pp. 1502–1505.
- Freese J., Peterson D. (2017). Replication in social science // *Annual Review of Sociology*. Vol. 43. Pp. 147–165.
- Gallen T., Raymond B. (2021). *Broken instruments*. SSRN Working Paper.
- Gehman J., Glaser V.L., Eisenhardt K.M., Gioia D., Langley A., Corley K.G. (2018). Finding theory–method fit: A comparison of three qualitative approaches to theory building // *Journal of Management Inquiry*. Vol. 27. No. 3. Pp. 284–300.
- Gelman A. (2011). Causality and statistical learning // *American Journal of Sociology*. Vol. 117. No. 3. Pp. 955–966.
- Gelman A., Loken E. (2014). The statistical crisis in science: data-dependent analysis — a “garden of forking paths” — explains why many statistically significant comparisons don’t hold up // *American Scientist*. Vol. 102. No. 6. Pp. 460–466.
- Gerber A.S., Malhotra N. (2008). Publication bias in empirical sociological research: Do arbitrary significance levels distort published results? // *Sociological Methods & Research*. Vol. 37. No.1. Pp. 3–30.
- Gerring J. (2012). Mere description // *British Journal of Political Science*. Vol. 42. No. 4. Pp. 721–746.
- Grunow M., Schneider H., Wagner G., Wagner J. (2018). Editorial // *International Journal for Re-Views in Empirical Economics*. Vol. 2. Pp. 1–6.
- Hamermesh D.S. (2007). Replication in economics // *Canadian Journal of Economics*. Vol. 40. No. 3. Pp. 715–733.
- Humpfreys M. (2017). *Political games: Mathematical insights on fighting, voting, lying and other affairs of state*. New York: Norton.
- Ioannidis J.P. (2005). Why most published research findings are false // *PLoS Medicine*. Vol. 2. Issue 8: e124.
- Kapelyushnikov R.I. (2018). *O sovremennom sostoyanii ekonomicheskoi nauki: Polusotsiologicheskie nablyudeniya* [On the current state of economics: Half-sociological observations]. Moscow: Institute of Economics of the Russian Academy of Sciences. Pp. 8–33. [In Russ.].
- Karafillidis A. (2013). Erklärungen in rekursiven Verhältnissen // *Zeitschrift für Theoretische Soziologie*. Vol. 2. Issue 2. Pp. 218–239.
- Keynes J.M. (1940). On a method of statistical business-cycle research. A comment // *Economic Journal*. Vol. 50. Issue 197. Pp. 154–156.
- King G. (1995). Replication, replication // *PS: Political Science & Politics*. Vol. 28. No. 3. Pp. 444–452.
- Koshovets A.B. (2019). “Gorizontally progress” ekonomicheskoi nauki: Mezhdru konstruiruemoi real’nostyu i tekhnologii [Horizontal progress of economics: Between a constructed reality and a technoscience]. Moscow: Institute of Economics of the Russian Academy of Sciences. [In Russ.].
- Krauss A. (2018). Why all randomised controlled trials produce biased results // *Annals of Medicine*. Vol. 50. No. 4. Pp. 312–322.
- Krugman P. (1998). Two cheers for formalism // *Economic Journal*. Vol. 108. Issue. 451. Pp. 1829–1836.
- Leamer E.E. (2010). Tantalus on the road to asymptopia // *Journal of Economic Perspectives*. Vol. 24. No. 2. Pp. 31–46.

- Leatherdale S.T. (2019). Natural experiment methodology for research: a review of how different methods can support real-world research // *International Journal of Social Research Methodology*. Vol. 22. No.1. Pp. 19–35.
- Lenz G.S., Sahn A. (2021). Achieving statistical significance with control variables and without transparency // *Political Analysis*. Vol. 29. No. 3. Pp. 356–369.
- Libman A. (2018). Der lange Schatten der Vergangenheit: Historisches Erbe und quantitative Sozialwissenschaften // *Osteuropa*. Vol. 68. No. Pp. 49–66.
- Libman A.M. (2018). *Empiricheskie issledovaniya v ekonomike: Revolyutsiya dostovernosti?* [Empirical research in economics: A credibility revolution]. Moscow: Institute of Economics of the Russian Academy of Sciences. Pp. 34–52. [In Russ.].
- Libman A.M. (2008). Teoreticheskie i empiricheskie issledovaniya v sovremennoi ekonomike: Problemy kommunikatsii [Theoretical and empirical research in modern economics: Problems of Communication] // *Voprosy ekonomiki*. No. 6. Pp. 4–19. [In Russ.].
- Lieberman E.S. (2005). Nested analysis as a mixed-method strategy for comparative research // *American Political Science Review*. Vol. 99. No. 3. Pp. 435–452.
- Loken E., Gelman A. (2017). Measurement error and the replication crisis // *Science*. Vol. 355. Issue 6325. Pp. 584–585.
- Mankiw N.G. (2006). The macroeconomist as scientist and engineer // *Journal of Economic Perspectives*. Vol. 20. No. 4. Pp. 29–46.
- McCloskey D.N. (1985). The loss function has been mislaid: The rhetoric of significance tests // *American Economic Review*. Vol. 75. No.2. Pp. 201–205.
- Mellon J. (2021). *Rain, rain, go away: 176 potential exclusion-restriction violations for studies using weather as an instrumental variable*. SSRN Working Paper.
- Morgan S.L., Winship C. (2015). *Counterfactuals and causal inference*. Cambridge: Cambridge University Press.
- Muñoz J., Young C. (2018). We ran 9 billion regressions: Eliminating false positives through computational model robustness // *Sociological Methodology*. Vol. 48. No. 1. Pp. 1–33.
- Oleinik A.N. (2021). O vozmozhnykh metodologicheskikh osnovaniyakh pragmaticheskoi ekonomicheskoi teorii: T. Bayes and J. Dewey [On possible methodological foundations of a pragmatic economic theory: T. Bayes and J. Dewey] // *Voprosy teoreticheskoi ekonomiki*. No. 3. Pp. 34–51. [In Russ.].
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science // *Science*. Vol. 349. Issue 6251.
- Pluckrose H., Lindsay J., Boghossian P. (2021). Understanding the “grievance studies affair” papers and why they should be reinstated: A response to Geoff Cole // *Sociological Methods & Research*. Forthcoming.
- Ravallion M. (2020). *Should the randomistas (continue to) rule?* National Bureau of Economic Research Working Paper No. 27554.
- Rodrik D. (2008). *The new development economics: We shall experiment, but how shall we learn?* HKS Working Paper No. RWP08-055.
- Rosenthal R. (1979). The file drawer problem and tolerance for null results // *Psychological Bulletin*. Vol. 86. No. 3. Pp. 638–641.
- Rosenzweig M.R., Wolpin K.I. (2000). Natural “natural experiments” in economics // *Journal of Economic Literature*. Vol. 38. No. 4. Pp. 827–874.
- Roth A.E. (2002). The economist as engineer: Game theory, experimentation, and computation as tools for design economics // *Econometrica*. Vol. 70. No. 4. Pp. 1341–1378.
- Rubinstein A. (2006). Dilemmas of an economic theorist // *Econometrica*. Vol. 74. No. 4. Pp. 865–883.
- Sala-i-Martin X. (1997). I just ran two million regressions // *American Economic Review*. Vol. 87. No. 2. Pp. 178–183.
- Samii C. (2016). Causal empiricism in quantitative research // *Journal of Politics*. Vol. 78. No. 3. Pp. 941–955.
- Schumpeter J. (1933). The common sense of econometrics // *Econometrica*. Vol. 1. No. 1. Pp. 5–12.
- Sekhon J.S., Titiunik R. (2012). When natural experiments are neither natural nor experiments // *American Political Science Review*. Vol. 106. No. 1. Pp. 35–57.
- Shadish W.R. (1995). The logic of generalization: Five principles common to experiments and ethnographies // *American Journal of Community Psychology*. Vol. 23. No. 3. Pp. 419–428.
- Shrout P.E., Rodgers J.L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis // *Annual Review of Psychology*. Vol. 69. Pp. 487–510.
- Silberzahn R., Uhlmann E.L., Martin D.P., Anselmi P., Aust F., Awtry E., et al. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results // *Advances in Methods and Practices in Psychological Science*. Vol. 1. No. 3. Pp. 337–356.
- Simonsohn U., Simmons J.P., Nelson L.D. (2019). *Specification curve: Descriptive and inferential statistics on all reasonable specifications*. SSRN Working Paper.
- Sims C.A. (2010). But economics is not an experimental science // *Journal of Economic Perspectives*. Vol. 24. No. 2. Pp. 59–68.
- Smith G., Cordes J. (2020). *The phantom pattern problem: The mirage of big data*. Oxford: Oxford University Press.

- Sokolov M.M. (2015). Sotsiologiya kak chudo: Protsess sence-building v odnoi akademicheskoi distsipline. [Sociology as a miracle: Process of sence-building in an academic discipline] // *Sotsiologiya vlasti*. Vol. 27. No. 3. Pp. 13–57. [In Russ.].
- Sovey A.J., Green D.P. (2011). Instrumental variables estimation in political science: A readers' guide // *American Journal of Political Science*. Vol. 55. No.1. Pp. 188–200.
- Spanos A. (2000). Revisiting data mining: 'hunting' with or without a license // *Journal of Economic Methodology*. Vol. 7. No. 2. Pp. 231–264.
- Su H.C., Colander D. (2021). The economist as scientist, engineer, or plumber? // *Journal of the History of Economic Thought*. Vol. 43. No. 2. Pp. 297–312.
- Uhlig H. (2012). Economics and reality // *Journal of Macroeconomics*. Vol. 34. No. 1. Pp. 29–41.
- von Eye A., Losel F., Mayzer R. (2003). Is it all written in the stars? A methodological commentary on Sachs' astrology monograph and re-analyses of his data on crime statistics // *Psychological Test and Assessment Modeling*. Vol. 45. No. 1. Pp. 78–91.
- Young C. (2009). Model uncertainty in sociological research: An application to religion and economic growth // *American Sociological Review*. Vol. 74. No. 3. Pp. 380–397.
- Young C., Holsteen K. (2017). Model uncertainty and robustness: A computational framework for multimodel analysis // *Sociological Methods & Research*. Vol. 46. No. 1. Pp. 3–40.