В.В. КАБЕРНИК*

ЛОВУШКИ СТАТИСТИКИ И ОПЫТ ИХ ОБХОДА[1]

Аннотация. В статье с критической точки зрения рассматриваются подходы к государственному управлению, основанному на данных, а также характерные ошибки в принятии решений, которые проистекают из ненадежных данных и их неверных интерпретаций. Приводятся исторические примеры заблуждений, допущенных министром обороны США Р. Макнамарой, рассматривается пагубная политика отбрасывания качественных оценок в пользу использования только и исключительно измеримых параметров. На примере некорректной интерпретации демографической статистики в Косово рассмотрено влияние заведомо ложной и некритически воспринимаемой информации на принятие решений в части внутренней национальной политики бывшей Югославии и дальнейшей гуманитарной интервенции под эгидой защиты албанского населения от геноцида. Рассматриваются причины этих заблуждений, приводятся примеры корректных интерпретаций. Далее автор фокусируется на прикладном аспекте: на основе опыта автора в предварительном анализе и обработке данных, используемых в реализации научных проектов, предлагаются способы обхода типичных «ловушек» статистики, ведущих к некорректным интерпретациям.

Ключевые слова: управление, основанное на данных; данные; статистика; заблуждение Макнамары; закон Гудхарта; временные ряды.

Для цитирования: Каберник В.В. Ловушки статистики и опыт их обхода // Политическая наука. – 2024. – № 2. – С. 237–261. – DOI: http://www.doi.org/10.31249/poln/2024.02.11

---

* **Каберник Виталий Владимирович**, ведущий эксперт Управления научных и инновационных проектов, Московский государственный институт международных отношений (Университет) Министерства иностранных дел Российской Федерации (Москва, Россия), e-mail: vic@inno.mgimo.ru

[1] Статья подготовлена в рамках консорциума МГИМО МИД России и НИУ ВШЭ из средств гранта на реализацию программы стратегического академического лидерства «Приоритет-2030».

© Каберник В.В, 2024 DOI: 10.31249/poln/2024.02.11

Введение

Не будет большим преувеличением утверждать, что управление современным миром осуществляется с интенсивным использованием данных (или на основе данных). Принятие решений в самых разных областях бизнеса и политики стало невозможным без подкрепления этих решений аналитическими материалами, оперирующими данными. С использованием собираемой статистической информации составляются влиятельные, регулярно издаваемые обзоры политики, экономические сводки и множество научных публикаций, которые, в свою очередь, формируют представления о мире, в том числе у элит. Сегодня кажется невозможным принятие решений, основанных на интуитивном понимании обстановки. Любое решение требует научного или хотя бы наукообразного обоснования.

Этот подход в публикациях последних лет получил наименование «управление, основанное на данных», или «политика, управляемая данными» (data-driven policy), и общепринятой стала точка зрения, всячески поддерживающая подобные практики. Среди его преимуществ указываются [van Veenstra, Kotterink, 2017] возможности по предсказанию проблем с использованием данных, поступающих в реальном времени (в идеальном рассмотрении процесса), а также повышенная прозрачность в принятии политических решений в непрерывном цикле мониторинга эффектов проводимой политики. Идеализированный цикл принятия решений в этой модели представлен на иллюстрации ниже.

Однако одновременно с ростом количества входных данных и все более глубоким проникновением бизнес-практик в решение задач государственного и политического управления мы парадоксальным образом наблюдаем снижение (или отсутствие повышения) качества принимаемых решений. Причины этого многообразны, но в качестве основной можно выделить недостаточное понимание методов работы с входными данными, их анализа, а также выбор слабо релевантных индикаторов эффективности.

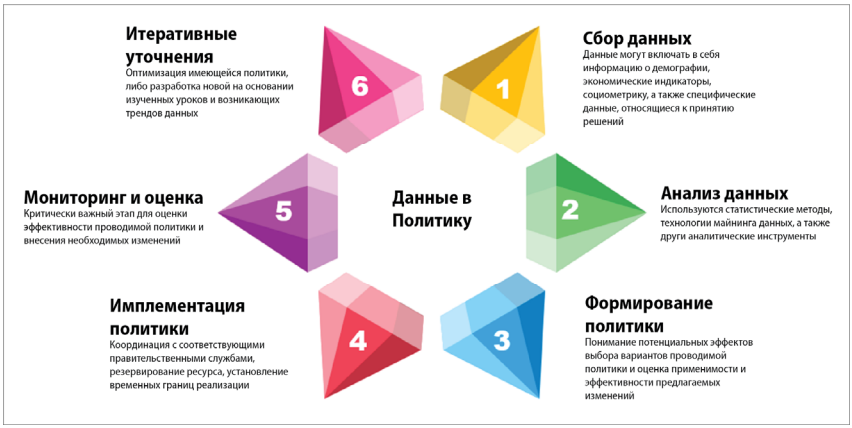

Рис. 1
Цикл принятия решений и формирования политики в идеализированной модели

Публикации последних лет, исследующие подходы к управлению, базирующемуся на данных, тем не менее возлагают большие надежды на этот метод несмотря на то что примеры успеха проводимых нововведений найти намного сложнее, чем отчеты о внедрении инноваций. Так, например, внедренная администрацией президента Дж. Буша-мл. в 2003 г. система оценки эффективности федеральных программ PART (Performance Assessment Rating Tool) рекламировалась как инновация в государственном управлении, позволяющая оптимизировать процессы распределения бюджета. Практически такой же подход к оптимизации бюджета Министерства обороны США использовал Р. Макнамара, возглавлявший ведомство в 1961–1968 гг. Однако в той же системе оценок программы восстановления, проводимые Федеральным агентством по управлению в чрезвычайных ситуациях (*Federal Emergency Management Agency,* FEMA), были отмечены как «адекватные» сразу после выявленных серьезных проблем в реагировании на разрушения, вызванные ураганом «Катрина» [Esty, Rushing, 2007]. Это яркая иллюстрация описываемого ниже «принципа Гудхарта» (или «закона Гудхарта»).

При всех описываемых проблемах авторы публикаций, посвященных политике (или управлению), основанной на данных, со-

храняют оптимизм при условии, что «будут разрушены некоторые барьеры». В качестве примеров успеха приводятся в основном частные решения из области муниципального управления и системы автоматизации отдельных служб, например, внедрение геоинформационных систем в задачах управления службой скорой помощи. Самое примечательное, однако, в том, что проблемы управления, основанного на данных, отнюдь не являются инновацией XXI века. Ниже мы рассмотрим один из наиболее ярких примеров.

Заблуждения Макнамары, «заблуждение Макнамары» и «закон Гудхарта»

Популярность практики принятия решений, основанных на анализе статистических данных, может быть связана с именем Роберта Макнамары, который, занимая пост министра обороны США, стоял у истоков разнообразных и отнюдь не бесспорных реформ американской армии и военной стратегии, многие из которых основывались на принципах системного подхода в управлении. Впрочем, это была далеко не первая попытка Макнамары интегрировать использование статистики в принятие решений на самых различных уровнях.

Вся карьера Макнамары так или иначе была связана с обработкой значительных массивов данных, и его знакомство с инновационными на тот момент практиками состоялось еще в 1939 г. в бухгалтерской компании Price Waterhouse в Сан-Франциско. Уже годом позже после возвращения в Гарвардский университет он встал у истоков программы по обучению офицеров Армии США аналитическим подходам, используемым в бизнес-процессах. В 1943 г. он получил пост в Офисе статистического надзора ВВС США и был напрямую ответственен за анализ применения и повышение эффективности стратегических бомбардировщиков США на тихоокеанском театре военных действий. Именно при нем была закреплена практика оценки эффективности налетов через процент разрушенной площади цели (как и выбора площадных целей – японских городов) (см., например, отчет о результатах ударов по четырем японским городам, включая Токио, 10 марта 1945 г.[1]).

[1] Tactical Mission Report, Mission No. 40 Flown 10 March 1945, Headquaters XXI Bomber Command APO 234 // National Security Archive. – Mode of access:

В 1946 г. совместно с еще несколькими отставными офицерами из Офиса статистического надзора Макнамара был приглашен в переживавшую тяжелые времена автомобильную компанию «Форд» для проведения всеобъемлющих реформ, использующих принципы «научного менеджмента». Там при его активном участии впервые было закреплено использование компьютерной обработки табличной информации в интересах анализа значительных массивов статистики [Byrn, 1993]. В 1960 г. Макнамара непродолжительное время возглавлял компанию «Форд», выведя на рынок – вопреки мнению владельца компании в силу нетипичности конструкции – одну из самых успешных на тот момент моделей автомобиля, после чего был приглашен президентом Дж. Кеннеди возглавить Министерство обороны США с правом полностью сформировать управленческую команду и схему управления Пентагоном. Полученный в бизнесе опыт Макнамара стал переносить на задачи государственного управления в этом очень большом и сложном ведомстве.

Важнейшим нововведением Макнамары на посту министра обороны США стало создание института системного анализа, проводившего декомпозицию индикаторов, описывающих текущее состояние вооруженных сил, на основании чего должны были приниматься бюджетные решения. Результатом работы на этом направлении стало создание Системы планирования, программирования и бюджетирования (Planning, Programming and Budgeting System – PPBS), которая была призвана оптимизировать процессы военного строительства в рамках пятилетних планов. Следствием проводимой политики оптимизации стала консолидация оборонных инициатив (для сокращения расходов), в частности создание единой истребительно-бомбардировочной платформы TFX, позже воплотившейся в многоцелевом самолете F-111, показавшем себя крайне посредственным истребителем-перехватчиком, который к тому же не мог эксплуатироваться с авианосцев, как было заявлено при его разработке. Впрочем, был и успешный результат таких усилий в виде появления изначально предназначенного для ВМФ истребителя-перехватчика F-4 на вооружении ВВС (хотя и в роли истребителя-бомбардировщика).

---

https://nsarchive.gwu.edu/document/28509-document-8-headquarters-xxi-bomber-command-tactical-mission-report-mission-no-40 (accessed: 25.01.2024).

Закрепленный в зарубежной литературе термин «заблуждение Макнамары» (McNamara fallacy) типично ассоциируется с принятием решений в ходе войны во Вьетнаме. Основным проявлением этого заблуждения применительно к оценке эффективности ведения боевых действий было «арифметическое» оценивание конфликта, не учитывающее более сложных и комплексных качественных и количественных составляющих. Предполагалось, что одно только увеличение потерь противника – при сомнительности оценки этих потерь в трудных условиях театра боевых действий – и одновременное сокращение собственных потерь обязательно должно в итоге привести к победе в конфликте [Tucker, 2000, p. 42]. Ключевой индикатор эффективности – "body count" (букв. подсчет тел) – в полном соответствии с принципом Гудхарта (см. ниже) привел к масштабным приписками и в течение долгого времени поддерживал совершенно неадекватную оценку хода конфликта.

Суть заблуждения Макнамары при этом глубже, и его описание в контексте войны во Вьетнаме является в большей степени следствием критики его действий на посту министра обороны США. В полном рассмотрении заблуждение состоит из трех последовательных принципов, слепо применяемых при оценке ситуации, а также оцениваемых эффектов проводимой политики – и эти принципы были впервые сформулированы Макнамарой еще в ходе реформирования компании «Форд»: (1) любые измеримые эффекты принятия решений должны быть измерены и зафиксированы; (2) плохо измеримые (в том числе качественные) индикаторы не могут быть использованы в математической модели оценки эффектов проводимой политики управления, и поэтому должны быть исключены из этой модели; (3) любые индикаторы, которые плохо поддаются измерению, предполагаются несущественными для проводимых оценок. При точном следовании данным принципам сложно измеримые индикаторы эффективности либо успешности проводимой политики не только исключаются из рассмотрения, но и могут быть объявлены ложными [Yankelovich, 1971, p. 26]. В воспоминаниях бригадного генерала Э. Лонсдейла приводится пример о том, что Макнамара при выборе индикаторов для оценки хода войны во Вьетнаме исключил из модели ряд качественных параметров, сославшись на то, что, если они не могут быть надежно измерены, они являются совершенно несущественными [Phillips, 2008]. Для оценки ситуации во Вьетнаме, опираясь на свой

опыт оценки результативности работы стратегической авиации во время Второй мировой войны, Макнамара, как выяснилось, предпочитал рассматривать ее ход преимущественно через призму статистики бомбардировок, игнорируя данные, поступающие «с земли» (в том числе искаженные из-за политики «подсчета тел»).

Заблуждение Макнамары (и в кавычках, и без них) не является чем-то уникальным, – не только директор, а потом министр Макнамара сознательным отбрасывал значительный массив плохо поддающейся квантификации информации для принятия решений. Более того, этот подход хорошо описывается (или дополняется) «принципом Гудхарта» (иное название – «закон Гудхарта»), который гласит: «Когда мера становится целью, она перестает быть хорошей мерой»[1], то есть становится объектом манипуляций, нацеленных как на улучшение собственно индикатора путем прямых фальсификаций, так и на оценку эффективности собственных достижений преимущественно через выбранный индикатор с сознательным отбрасыванием иных важных показателей (подробнее см.: [Goodhart, 1984]). Принцип Гудхарта был подтвержден в ходе анализа денежной политики и экономики США и Великобритании в 1984–1985 гг. [Muller, 2018, p. 22]. При рассмотрении «заблуждения Макнамары» в контексте Вьетнамской войны можно выделить как раз необоснованно завышенный вес количества потерь в оценке эффективности собственных действий и планомерную работу по улучшению преимущественно этого показателя (в том числе посредством завышения чужих потерь).

Тем не менее, несмотря как на осознанность заблуждения Макнамары, так и на общее согласие научного сообщества с принципом Гудхарта вплоть до сегодняшнего дня статистические данные активно используются для подкрепления политических решений с различным качеством этих данных. Зачастую качество решений страдает от неверных интерпретаций исходных данных либо от их неполноты. Прецеденты наблюдались при принятии

[1] В оригинальной несколько шутливой формулировке, использованной в докладе 1975 г., выглядит следующим образом: "…Any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes" [Goodhart, 1984, p. 96] (Любая наблюдаемая статистическая закономерность тяготеет к разрушению, как только на нее оказывается давление с целью контроля). В приведенном выше виде встречается в широко цитируемой работе К. Хоскин [Hoskin, 1996, p. 265].

решений о вторжении в Ирак (вторая кампания), бомбардировках Югославии и наиболее массово в последние годы при работе со статистикой по коронавирусной инфекции и планировании мероприятий по контролю развития пандемии – несмотря на то что заблуждение Макнамары конкретно упомянуто как негативный эффект в исследованиях адекватности клинических тестов в работе 2009 г. [Basler, 2009]. Та же проблема в полной мере проявилась и в прогнозной оценке влияния санкций против РФ после 2014 г., но эта тема требует отдельного рассмотрения, для которого пока данных недостаточно.

В ряде случаев – возможно даже в большинстве – мы сталкиваемся не со злонамеренными манипуляциями данными, а с естественными проблемами, связанными с их неполнотой либо ошибками интерпретации. В других случаях мы наблюдаем эффект фиксации на измеримых индикаторах [Muller , 2018, p. 68] и некорректное построение модели. Тем не менее лица, принимающие решения, обычно не обладают достаточными компетенциями и ресурсами для самостоятельного анализа данных (и их предварительной проверки) и часто вынуждены полагаться на чужие данные и интерпретации. Значительную долю такой информации предоставляют статистические агрегаторы[1], и нередко эти данные используются без понимания граничных условий применения и особенностей их сбора. К сожалению, эта же ошибка нередко допускается и в научных исследованиях.

Ниже мы рассмотрим несколько характерных примеров, иллюстрирующих типичные ошибки, которые допускаются при работе со статистическими показателями в научных исследованиях и при принятии решений. Примечательно, что большинство таких примеров являются результатом переплетения некорректной работы с данными и типичными когнитивными искажениями при их оценке. При этом, однако, собственно исходные данные и их интерпретации могут играть ведущую роль.

[1] Примечательно, что Роберт Макнамара после отставки с поста министра обороны США возглавил Всемирный банк, который сегодня считается одним из наиболее авторитетных агрегаторов глобальной статистики.

Демография косовских албанцев и гуманитарные бомбардировки НАТО

Широко известен афоризм Марка Твена о том, что «существует три вида лжи: ложь, наглая ложь и статистика», который также приписывается британскому политику Б. Дизраэли. Менее известно, что эта фраза сопровождает ироничные размышления о снижении с возрастом производительности писателя в словах, написанных за час. Несмотря на очевидную иронию таких рассуждений, слепое использование статистических данных и ложных экстраполяций, в том числе вырванных из корректного контекста, получило широкое распространение в принятии политических решений и пропаганде. При этом нередко заблуждение, основанное на статистике, является несознательным, но может приводить к самым катастрофическим последствиям.

В материалах расследования палаты общин парламента Великобритании от 23 мая 2000 г. содержится утверждение[1], что в 1996 г. под эгидой Сербской академии наук и искусств был опубликован отчет, описывающий прогнозируемое состояние демографии Югославии в 2015 г. В отчете подчеркивалось, что при сохраняющемся уровне фертильности на косово-албанскую женщину, превышающего 7 рождений, уровень демографического давления албанцев к 2015 г. может привести к ситуации, когда до 25% сербских мужчин призывного возраста будут иметь албанские корни. На основании этого отчета были предложены меры по изоляции региона Косово-Метохия за счет создания подковообразной по форме буферной зоны. Соответствующий план широко обсуждался в бывшей Югославии.

В 1999 г., опираясь на конфигурацию предложенной буферной зоны, министерство обороны Германии опубликовало разведывательную сводку, где описывалась операция «Подкова», в рамках которой якобы предлагалось массовое переселение косовских албанцев и последующие этнические чистки. Министр иностранных дел Германии Йошка Фишер тогда же заявлял, что соответ-

[1] Fourth Report // Select Committee on Foreign Affairs. – 23.05.2000. – Mode of access: https://publications.parliament.uk/pa/cm199900/cmselect/cmfaff/28/2802.htm (accessed: 20.01.2024). Оригинал упомянутого в материалах расследования отчета найти не удается.

ствующий план уже запущен. Ему же принадлежит высказывание о том, что этнические чистки сравнимы с преступлениями нацистов в годы Второй мировой войны[1].

И та и другая оценки, которые в итоге привели к гуманитарной интервенции НАТО в Югославию, в значительной степени основаны на наблюдении об албанских косоварах «7 рождений на женщину». Если в Югославии этот индикатор использовался для решения внутриполитических проблем и пропаганды националистической сербской повестки, то в Германии и в дальнейшем на Западе, этот же показатель использовался для оценки объемов этноцида. В обоих случаях использовались простейшие методы экстраполяции показателя, которые не могут быть корректны сами по себе, поскольку никак не учитывают, например, показатели детской смертности. Но наиболее примечательно, что и само значение «7 рождений на женщину» на 1990-е годы никак не соответствовало действительности!

В работе Дж. Фолкингэм и А. Гьонча [Falkingham, Gjonça, 2001] показано, что несмотря на высокий коэффициент фертильности албанцев в послевоенный период, к 1990 г. уровень фертильности упал до примерно типичных для Балкан трех рождений на женщину, что незначительно превышало средний показатель по Югославии в целом. Откуда же взялись показатель «более 7» и последующие заявления об этнических чистках косовских албанцев? Более того, показатели фертильности в 7,58 рождений на женщину в 1950 г. регулярно встречаются в различных публикациях и активно цитируются[2].

Ответ становится более понятным при рассмотрении периодичности переписей населения. Предвоенная перепись 1931 г. указывает численность косовских албанцев в 331 549 человек. Во время Второй мировой войны Косово было присоединено к итальянской зоне контроля Албании, что привело к выселению

[1] Auszüge aus der Fischer-Rede // Spiegel. – 13.05.1999. – Mode of access: https://www.spiegel.de/politik/deutschland/wortlaut-auszuege-aus-der-fischer-rede-a-22143.html?sara_ref=re-xx-cp-sh (accessed: 20.01.2024.). Полное выступление доступно по адресу: Joschka Fischer auf dem Kosovo-Sonderparteitag in Bielefeld 1999 (youtube.com).

[2] См., например: Salihu F. The politicization of having children // Kosovo 2.0. – 31.10.2018. – Mode of access: https://kosovotwopointzero.com/en/the-politicization-of-having-children/ (accessed: 20.12.2023).

местных сербов за пределы региона. Одновременно с этим итальянская оккупационная администрация способствовала переселению 72 000 албанцев из собственно Албании в Косово [Sullivan, 1999, p.15]. Результатом всех перемещений этнических групп стал рост албанского населения по переписи 1948 г. до 498 242 человек – более чем 50%-ный прирост албанского населения за 17 лет. Такого прироста албанского населения за отчетный период не наблюдалось ни в один из других периодов между проведением переписей. При этом почти половина этого прироста вызвана миграцией албанцев в Косово, а не увеличением численности за счет новых рождений.

Более подробное рассмотрение исторических источников позволяет установить, что цитируемый показатель фертильности относится к 1939 г., а собственно данные по фертильности албанского населения в ходе всех последующих переписей населения не оценивались. Однако этот же показатель присутствовал в ежегодниках CIA World Factbook путем переноса из года в год вплоть до середины 1990-х годов. В итоге данные от 1939 г. использовались для принятия политических решений и пропаганды вплоть до начала XXI в., не подвергаясь сомнению.

Наряду с отсутствием контроля данных на этом примере можно также говорить и о непрозрачности источников. Ежегодники CIA World Factbook – один из наиболее популярных справочников, но одновременно, вероятно, один из самых непрозрачных. Никакие данные здесь не подкрепляются источниками и приводятся абсолютно анонимно (и этому есть «оправдание» в виде статуса ведомства, под эгидой которого они издаются).

История с ложной трактовкой демографических тенденций в Косово на этом, впрочем, не заканчивается. Последняя официальная перепись до начала гуманитарной интервенции проводилась в 1991 г. и была подвергнута бойкоту местным албанским населением [Bugajski, 2002, p. 479]. Оценки численности были проведены с использованием демографической модели, и на основании этой же оценки базировались спекуляции политиков и СМИ о масштабах проводимых сербами этнических чисток.

Рис. 2
Динамика численности населения Косово в период 1950–2050 гг. (прогноз)[1]

Актуальные данные о численности албанских косоваров были получены только в 2011 г. – в ходе первой послевоенной переписи. Оценка Евростата, сделанная на основе демографической модели, была скорректирована на 400 000 человек в меньшую сторону, причем численность населения Косово планомерно падает с перспективой сокращения на 11% к 2050 г.[2] Из этого следует, что использование единственного ложного значения и последующая некорректная его экстраполяция стали одной из важнейших причин последовавшего политико-гуманитарного кризиса, боевых дей-

[1] Judah T. Kosovo's demographic destiny looks eerily familiar // Balkan Insight. – 7.11.2019. – Mode of access: https://balkaninsight.com/2019/11/07/kosovos-demographic-destiny-looks-eerily-familiar/ (accessed: 20.12.2023).

[2] Ibid.

ствий в центре Европы, вмешательства НАТО и существенных изменений во внешней политике ряда государств мира, включая Россию.

Трансформации суверенитетов, ошибки базового процента и долевые показатели

В 1991–2023 гг. мир менялся драматически, и эти изменения требуют учета при анализе статистических данных. Даже если оставить за рамками рассмотрения распад СССР, произошли значимые изменения на политической карте: несколько стадий распада Югославии, мирный распад Чехословакии, переход Гонконга и Макао под управление Китая, появление непризнанных территориальных образований (Косово, Абхазия, Южная Осетия и др.). Данные международной статистики обычно отражают эти трансформации с запозданием либо просто игнорируют их в угоду поддержания целостности наблюдений.

Характерен пример Гонконга и Макао: вплоть до сегодняшнего дня и Всемирный банк, и МВФ ведут отдельные наблюдения для этих территорий, не включая их в состав КНР. В то же время Гонконг генерирует более 2% ВВП Китая и значительную часть международных инвестиций и торговли. При работе с наблюдением «Китай» эти данные просто отбрасываются, хотя по объему эта «прибавка» превосходит ВВП ЮАР[1].

Аналогичные проблемы наблюдаются с зависимыми территориями Великобритании, Нидерландов, заморскими департаментами Франции, членами Датского содружества и др. Для этих территорий ведется независимый подсчет статистики, который не включается в статистику государства-суверена. Простое контрольное суммирование всех суверенных наблюдений по ВВП показывает отбрасывание из расчетов до 2 трлн долларов мирового ВВП в разные годы, что существенно меняет место ряда суверенных государств в «распределении» международного влияния, по крайней мере экономического.

Проблема описанной утечки данных разрешима, если их подвергнуть предварительной обработке. Например, при агрегиро-

[1] World Development Indicators // World Bank. – Mode of access: https://databank.worldbank.org/source/world-development-indicators (accessed: 1.12.2023).

вании показателей ВВП по суверену объем неучтенного мирового ВВП сокращается до примерно 0,5 трлн долларов, что сопоставимо (и на удивление похоже) на ВВП Тайваня, который не имеет отдельного наблюдения, но учтен в агрегированных подсчетах.

Метод агрегирования по суверену, на первый взгляд, решает часть проблем в части корректировки расчетов ВВП, но здесь тоже требуется проявлять внимательность. До 1997 г. упомянутый Гонконг входил в состав Великобритании, а Макао – в состав Португалии до 1999. Простой просмотр архивной статистики показывает, что в составе Великобритании в 1997 г. Гонконг генерировал около 10% ВВП. Перейдя во владение Китая, территории создавали более 16% ВВП Китая. Это та часть ВВП и экономической мощи, которая сменила принадлежность, перешла от одного суверена к другому, изменив распределение влияния как минимум в регионе. На иллюстрации ниже демонстрируется различие динамики ВВП Китая и Великобритании без коррекции и с проведенной коррекцией по суверену. В последнем случае ВВП Китая превосходит ВВП Великобритании не в 2006, а в 2005 г., а также одновременно сокращается отрыв по этому показателю в период 2001–2005 гг.

Корректным способом при анализе временных рядов для проведения агрегирования данных является не простое суммирование по суверену, но поддержание вспомогательного временного ряда суверенитетов. С такой коррекцией данные позволяют учесть, например, трансформации Югославии и Чехословакии.

Однако адекватным подходом к анализу данных для распадающихся государств является не агрегирование данных, а отбрасывание нерелевантных наблюдений. Например, на 2003 г. приведена статистика для Сербии и Черногории в составе конфедерации, но эта статистика едва ли является верной – адекватный подсчет выделенной частной экономики внутри любых союзных образований является практически невозможным. Точно так же можно поступать с наблюдениями, которые выделяют отдельные республики внутри СССР как независимые страновые наблюдения.

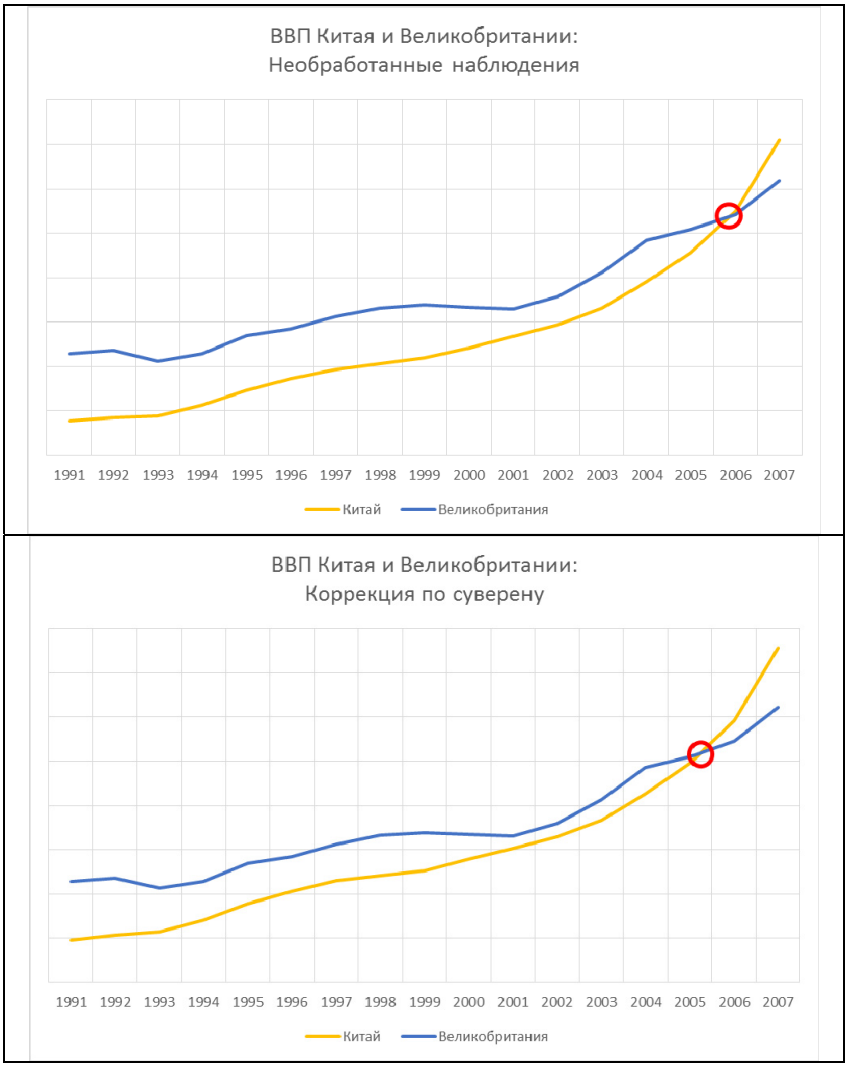

Рис. 3
Динамика ВВП Китая и Великобритании с коррекцией по суверену и на исходных нескорректированных данных

Граничные условия применения метода агрегирования по суверену очевидны также при внесении корректировок по временным границам суверенитета: этот метод показывает некорректные результаты при работе с долевыми и подушевыми показателями (ВВП на душу населения, например). Эта проблема осознана, и соответствующие корректировки могут быть внесены при расчете – об этом в следующем разделе.

Впрочем, проблема учета объемов торговли в общем случае не имеет решения. Например, торговля Гонконга или Французской Гвианы в значительной степени (свыше 70%) является внутренней для суверена, хотя учитывается как внешняя. Выделение этих объемов требует анализа диадной статистики, и здесь требуется решение исследователя о том, как ее интерпретировать.

Еще одной распространенной ошибкой интерпретации статистических показателей является ошибка базового процента при использовании долевых или подушевых показателей. Суть этой ошибки заключается в том, что проводимые расчеты вероятностей оперируют долями и процентными значениями без учета абсолютных цифр, а часто и в смещенной выборке. Сходные эффекты могут наблюдаться при интерпретации статистических наборов, которые для более убедительной демонстрации какого-либо эффекта оперируют процентными, долевыми или подушевыми показателями. К таким параметрам статистики относятся, например, ВВП на душу населения, процентные распределения подушевых показателей и иные показатели, которые в отрыве от абсолютных значений могут привести к ложным выводам.

В качестве абстрактного примера рассмотрим показатель процента пахотной земли от общей площади территории. Этот показатель – часто именно в процентном или подушевом выражении – используется для расчета уровня продовольственной безопасности. Предполагается, что при некотором пороговом уровне пахотной земли на душу населения продовольственная безопасность государства не подвергается рискам.

Но процент пахотных земель для Канады и для Кубы – это принципиально разные абсолютные значения, отличающиеся на порядок, как отличается и их абсолютная площадь, от которой рассчитывается этот процент. Согласно имеющейся статистике, порядка 30% земли Кубы являются пригодными для возделывания, в то время как в Канаде соответствующий показатель составляет

лишь около 4%. Однако в абсолютных цифрах этот показатель различается более чем на порядок: свыше 41 млн гектаров для Канады и порядка 3,6 млн гектаров для Кубы по данным на 2020 г.[1] Использование долевого показателя в подобных расчетах, как демонстрирует приведенный пример, может приводить к некорректной интерпретации.

Проблема расчета подушевых показателей дополнительно усугубляется описанной выше проблемой пересчета абсолютных значений по суверену для зависимых территорий. Например, у Великобритании их не меньше 17, и каждая из них в агрегаторах проходит независимым наблюдением. Для уже упоминавшегося случая Гонконга показатель уровня жизни через ВВП на душу населения превышает показатели материкового Китая почти в четыре раза[2], что относит Гонконг к другому классу наблюдений: экономикам с высоким доходом, что, очевидно, неверно для этой территории в составе единой экономики Китая.

Обойтись без долевых показателей нельзя, поскольку они выражают важные характеристики государств. Но надо помнить, что это всего лишь одно из представлений абсолютных показателей, и нет смысла умножать сущности путем их переноса из статистических баз. Практически всегда корректным методом работы будет самостоятельный расчет нужных долей по прозрачной методике (включая коррекцию по суверену), что также исключает возможные ошибки пересчета, которые, к сожалению, могут вкрадываться даже в самые авторитетные источники.

Доверие к данным и их ретроактивные корректировки

Перейдем теперь к вопросу о том, насколько вообще можно доверять исходным данным и проводимым экстраполяциям. Ведь проблема заполнения пропусков характерна для всех статистических агрегаторов. Пропусков данных не может не быть – статистические данные поступают неравномерно и в отдельные годы отсутствуют. В иные годы данные оцениваются крайне грубо или

[1] World Development Indicators // World Bank. – Mode of access: https://databank.worldbank.org/source/world-development-indicators (accessed: 01.12.2023).

[2] Ibid. (accessed: 1.12.2023).

приводятся по заведомо ангажированным расчетам. Все авторитетные статистические агентства используют те или иные методы заполнения пропусков. Настораживает, что эти методики далеко не всегда бывают хорошо документированы.

Различные способы интерполяции и экстраполяции характерны для заполнения пропусков по сложным расчетным параметрам. Характерным примером является расчет паритета покупательной способности (ППС), который подчинен трехлетним циклам. Но последний полный и согласованный пересчет продуктовой корзины для вычисления корректных коэффициентов был проведен в 2017 г. В 2021 г. на фоне пандемии коронавируса было принято решение не рассчитывать его в срок из-за очевидных дисбалансов национальных экономик и неспособности ряда из них собрать адекватные данные. Таким образом, на протяжении пяти лет, до финальной корректировки значений ППС осенью 2022 г., расчеты эконометрических параметров государств мира опирались на экстраполяцию значений от 2017 г., полученных для предсказанного экономического роста в то время. Очевидно, что эти данные все пять лет не отражали экономических шоков пандемии коронавируса.

В свете регулярного пересчета различных опорных коэффициентов экономической статистики архивные данные также регулярно корректируются ретроактивно. Упомянутый выше индикатор ВВП по ППС регулярно корректируется на три года назад при получении значений новых коэффициентов. Но в сентябре 2022 г. произошло значительно более масштабное изменение статистики Всемирного банка, в том числе и в архивных временных рядах: показатели ВВП по ППС были скорректированы не на три года назад, а на весь период с 2011 по 2017 г. Одновременно с этим на глубину до 2011 г. были скорректированы показатели бедности, демографической статистики, продовольственной безопасности и др.[1] Из этого буквально следует, что попытка воспроизвести любые исследования периода 2011–2017 гг., основанные на расчетных показателях развития, пользуясь обновленными статистическими показателями, обречены на провал. Любой исследователь столкнется

[1] Data Updates and Errata World Development Indicators // World Bank. – Mode of access: https://datahelpdesk.worldbank.org/knowledgebase/articles/906522-data-updates-and-errata (accessed: 11.02.2024).

попросту с несовпадением исходных данных, и получаемый результат при сохранении метода расчета окажется другим. Автор статьи столкнулся с данным эффектом, пытаясь воспроизвести расчеты исследований 2015 г. на новых панельных данных Всемирного банка, обнародованных в 2023 г.

Но не только параметры статистики подвергаются пересчетам, порой целые массивы данных изымаются из статистических агрегаторов, корректируются либо заменяются новыми, не являющимися прямыми аналогами. Так происходит с показателями, которые перестали собираться академическими институтами и другими международными агентствами, – их просто перестают публиковать. В качестве примера можно привести показатель производства энергии в нефтяном эквиваленте, который собирался Всемирным банком на основе данных Международного энергетического агентства по 2014 г. включительно[1]. Этот показатель использовался рядом исследований для оценки экономической безопасности и качества жизни, но с 2015 г. он заменен на показатели потребления взамен оценки объемов производства, что произошло по соображениям продвижения «зеленой повестки» в попытке подчеркнуть значимость возобновляемых источников.

Не менее опасна некорректная экстраполяция и интерполяция, подобная описанной в предыдущем разделе. Как ни печально, они нередко используются в совершенно непрозрачных вариантах, без должного документирования изменений. При отсутствии актуальных данных ряд агентств просто копируют данные предыдущего периода, либо используют методы усреднения. Такой подход замечен автором статьи в данных по этническому распределению проекта Ethnic Power Relations (EPR), собираемых Группой исследования международных конфликтов в Цюрихе[2], в данных Стокгольмского международного института исследований проблем мира (Stockholm International Peace Research Institute, SIPRI / СИПРИ) и ряда других. Например, для статистики EPR образца 2021 г. характерны зафиксированные на протяжении десятилетий показате-

[1] Metadata Glossary // The World Bank. – Mode of access: https://databank.worldbank.org/metadataglossary/world-development-indicators/series/EG.EGY.PROD.KT.OE (accessed: 24.01.2024).

[2] Ethnic Power Relations (EPR) Dataset Family 2021 // International Conflict Research (ICR) group at ETH Zurich. – Mode of access: In https://icr.ethz.ch/data/epr/ (accessed: 24.01.2024).

ли доминирующего этноса в Скандинавии и Прибалтике, а также продолжающийся учет доли фактически вымершего этноса айну в составе населения Японии. Эти данные, за неимением доверенных источников для обновления статистики, просто копировались из данных предыдущих отчетных периодов.

Некритичный перенос данных прошлых лет, интерполяция и экстраполяция во временных рядах, плохое управление источниками и банальные ошибки заполнения приводят к появлению существенных, заметных невооруженным взглядом выбросов данных, а также значительных девиаций нормальных трендов. Это не единственная причина появления выбросов, но она является довольно распространенной. Характерно повышена частотность выбросов в постковидной статистике. Причины этого частично описаны выше, и они связаны с пропуском цикла пересчета коэффициентов ППС, изменениями моделей потребления и др. Но намного более банальной причиной является простое отсутствие достоверных данных, запаздывание их поступления или целенаправленные манипуляции.

Возможно ли вообще корректное заполнение пропусков? Да, для ряда длинных трендов, – но только *при корректной модели построения недостающих данных.* Например, длинными трендами характеризуется демографическая статистика, хотя и для нее, как было показано выше, некорректная экстраполяция данных может приводить к появлению совершенно неадекватных оценок. Однако в общем случае заполнение пропусков методами интерполяции и (или) экстраполяции данных требует инерционного развития и отсутствия внешних шоков. Правда, сейчас отсутствие внешних шоков на протяжении продолжительных временных отрезков – это скорее редкость, чем норма.

Проблема некритичного отношения к исходным данным, к сожалению, кочует от одного исследования к другому. Очень часто авторитет источника является достаточным аргументом в пользу того, чтобы ему доверять безоговорочно. При этом не проводится даже простого первичного анализа данных.

В связи с этим можно рекомендовать хотя бы предварительные проверки данных перед их использованием. Признаки, которые сигнализируют о некорректных методах интерполяции и экстраполяции во временных рядах, включают в себя: постоянный повтор одного и того же значения на протяжении длительных вре-

менных отрезков, нарушения монотонности, не связанные с какими-либо катастрофическими изменениями (например, появление нулей после серии адекватных значений), подмену наблюдения «данные отсутствуют» явными усреднениями, апеллирующими к предыдущим наблюдениям, и т.п. Ряд таких девиаций выявляется легко, а другие (выявление усреднений и взвешивания, например), требуют более пристального анализа и обработки.

Проблема ретроактивного непрозрачного пересчета исходных также влечет за собой проблему воспроизводимости результатов исследования, которое было проведено на наборах данных, измененных впоследствии. С точки зрения дизайна исследования можно лишь дать рекомендации по фиксации наборов данных, чтобы не сталкиваться с нарушениями непрерывности рядов, которые могут быть вызваны появлением новых наборов данных в доступных статистических источниках.

Следует отметить, что девиации, хорошо выявляемые во временных рядах, на мгновенном срезе могут быть совершенно незаметны для исследователя. Поэтому анализ временных рядов в их непрерывности в целом дает хороший потенциал по повышению релевантности исследований и расчетных методик, что позволяет надеяться на повышенную точность получаемых результатов.

Таким образом, от исследователей и экспертов в современном мире требуется крайне внимательное отношение к исходным данным. Статистические агентства порой используют изощренные математические модели, которые применяются для предсказания трендов и заполнения пропусков и которые хороши лишь настолько, насколько хороша модель внешних условий, в которых работает предсказание.

Выводы и рекомендации

К фундаментальным проблемам исходных данных для любого исследования относятся проблемы доверия к статистике (включая проблемы ретроактивной замены данных) и проблемы их интерпретации.

Проблема доверия в общем случае, к сожалению, не решаема. Статистикой манипулируют, данные могут быть ангажированы, данные могут пересчитываться и напрямую фальсифицироваться.

Тем не менее есть методы, которые позволяют хотя бы устанавливать уровень доверия – ряд из них упомянут выше. Как поступать дальше с данными, которые вызывают сомнения, – вопрос к дизайну исследования, и в каждом конкретном случае он должен решаться отдельно. Вместе с тем рекомендуется проводить первичный анализ и отсев данных, либо хотя бы их маркировку.

Проблема некорректных интерпретаций часто может быть решена. Несмотря на то что статистика предлагает уже готовые наборы интерпретаций, на них просто не стоит полагаться. Любые долевые, подушевые и иные счетные показатели, если они могут быть вычислены самостоятельно, должны вычисляться самостоятельно. Нет никакого практического смысла в использовании чужих интерпретаций и умножении сущностей, если интересующий параметр легко рассчитывается.

Кроме того, в части интерпретаций рекомендовано минимизировать число источников, особенно использующих долевые показатели. В ситуации непрозрачности подсчетов разумным видится подход по самостоятельному исчислению параметров, которые могут быть получены из абсолютных значений. Долевые и подушевые показатели, если такой параметр требуется в исследовании, надежнее рассчитать самостоятельно по одному доверенному вспомогательному источнику. Разумеется, по возможности следует оперировать хорошо документированными источниками. При этом недопустимо смешивать показатели различных источников, которые могут вычислять показатели по разным методикам.

Можно суммировать рекомендации по работе с данными следующим образом.

1. Набор данных для исследования должен быть «герметичным», т.е. все данные должны быть зафиксированы к моменту начала исследования и в дальнейшем не дополняться. Допустимо полное обновление набора данных, но ни в коем случае нельзя обновлять данные частями – это может нарушить целостность набора.

2. Количество источников должно быть минимизировано. Недопустимо смешение источников данных одного и того же характера из-за их, возможно, различной размерности и методов подсчета.

3. Количество параметров также должно быть по возможности минимизировано. Любые вычисляемые параметры должны исчисляться только и исключительно внутри счетных моделей с использованием прозрачных, документированных методов и алго-

ритмов. Недопустимо использовать внешние исчисляемые параметры, поскольку всегда есть риск непрозрачной методики, коррекции и попросту случайных ошибок. Исключение допускается для параметров со сложными статистическими методиками подсчета (ППС), но, если есть возможность ими не пользоваться, лучше ими и не пользоваться.

4. При абсолютной необходимости использования исчисляемых показателей в исследовании их источником должен выступать тот же, который используется как источник для показателей абсолютных (см. п. 2).

5. Все данные должны проходить предварительную обработку и фильтрацию, по возможности анализироваться на предмет выявления нарушений монотонности и маркироваться по доверию (необязательно, но желательно). В предварительную обработку по возможности должны входить коррекции данных по государству-суверену (если релевантно для исследования) и учет временных рядов в части трансформаций суверенитета.

«Заблуждение Макнамары» в принятии решений, в том числе политических, с нами, видимо, навсегда. В эпоху интернета и компьютеров статистике при принятии решений порой доверяют больше, чем здравому смыслу, и любое решение может быть подкреплено если не сырыми данными, то выгодной их интерпретацией. В настоящей статье описано лишь несколько примеров подобных интерпретаций, но в период недавней пандемии коронавируса можно найти сотни примеров искажений, которые становились причиной принятия стратегически важных решений в области здравоохранения, социальной политики, экономики и др. В хорошо построенных исследованиях подобным заблуждениям не должно быть места.

Данная статья не претендует на полноту, но все же типичные ошибки, методы их обхода и повышения релевантности исследований описать удалось. Хочется надеяться, что они окажутся полезными для исследователей, работающих со статистическими данными, помогут повысить точность исследований и привлекут должное внимание, собственно, к данным, которые не являются просто числами, а несут в себе важные смыслы.

V.V. Kabernik*

Traps set by statistics and how to evade them[1]

Abstract. The article takes a critical view of the data-driven approaches to public management and the common pitfalls in decision-making stemming from unreliable data and misinterpretations. Historical examples of misconceptions made by US Secretary of Defense R. McNamara are given, and the disastrous policy of discarding qualitative assessments in favor of using only measurable parameters are examined. By the example of incorrect interpretation of demographic statistics in Kosovo, the influence of deliberately false and uncritically perceived information in decision-making regarding the national policy in the former Yugoslavia and further humanitarian intervention under the auspices of protecting the Albanian population from genocide are examined. The reasons for these misconceptions are studied and examples of correct interpretations are provided. The author also focuses on the means to avoid traps set by available statistics. Basing on the author's experience in preliminary analysis and processing of data used in academic projects, methods are proposed to circumvent the typical "traps" of statistics that lead to incorrect interpretations and findings.

Keywords: data-driven policy; data; statistics; McNamara fallacy; Goodhart's law; time series.

For citation: Kabernik V.V. Traps set by statistics and how to evade them. *Political science (RU)*. 2024, N 2, P. 237–261. DOI: http://www.doi.org/10.31249/poln/2024.02.11

References

Basler M. Utility of the McNamara Fallacy. *BMJ*. 2009, N 339, P. b3141. DOI: https://doi.org/10.1136/bmj.b3141

Bugajski J. *Political parties of Eastern Europe: a guide to politics in the post-Communist era*. New York: Routledge, 2002, 1120 p.

Byrne J.A. *The Whiz Kids: the founding fathers of American business - and the legacy they left us*. New York: Doubleday, 1993, 581 p.

Esty D., Rushing R. The promise of data-driven policymaking. *Issues in science and technology*. 2007, Vol. 23, N 4, P. 67–72.

Falkingham J., Gjonça A. Fertility transition in Communist Albania, 1950-90. *Population studies*. 2001, Vol. 55, N 3, P. 309–318. DOI: https://doi.org/10.1080/00324720127699

Goodhart C.A.E. *Monetary theory and practice: the UK experience*. London: Macmillan press, 1984, 282 p.

* **Kabernik Vitaly**, MGIMO University (Moscow, Russia), e-mail: vic@inno.mgimo.ru

[1] The article was prepared within the consortium of MGIMO University and HSE University and funded by the grant for the implementation of the Priority 2030 Strategic Academic Leadership Program.

Hoskin K. The ‘awful idea of accountability’: inscribing people into the measurement of objects. *In:* Munro R., Mouritsen J. (eds). *Accountability: power, ethos and the technologies of managing*. London: International Thomson Business Press, 1996, P. 265–285.

Muller J.Z. *The tyranny of metrics*. Princeton, Oxford: Princeton university press, 2018, 248 p.

Phillips R. *Why Vietnam matters: an eyewitness account of lessons not learned*. Annapolis: Naval Institute press, 2008, 398 p.

Sullivan B. The Balkans: of what is past, or passing, or to come. In: Murray W. (ed.). *The emerging strategic environment: challenges of the twenty-first century*. Westport, Conn.: Praeger, 1999, P. 1–33.

Tucker S.C. (ed.). *The encyclopedia of the Vietnam war: a political, social and military history*. Oxford: Oxford university press, 2000, 578 p.

van Veenstra A.F., Kotterink B. Data-driven policy making: the policy lab approach. In: Parycek P., Charalabidis Y., Chugunov A.V., Panagiotopoulos P., Pardo T.A., Tambouris Øystein Sæbø Efthimios (Eds.) *Electronic participation: 9th IFIP WG 8.5 International Conference, ePart 2017, St. Petersburg, Russia, September 4-7, 2017, Proceedings*. Lecture Notes in Computer Science, Vol. 10429. Springer, Cham, 2017, P. 100–110. DOI: https://doi.org/10.1007/978-3-319-64322-9_9

Yankelovich D. Interpreting the new life styles. *Sales Management: the Marketing Magazine*. 1971, Vol. 107, N 11, P. 26–27.