

Руководство по индексу потребительских цен: концепции и методы, 2020 год

Международная организация труда / Международный валютный фонд / Организация экономического сотрудничества и развития / Европейский Союз / Организация Объединенных Наций / Всемирный банк

Глава 10. Данные сканирования

Введение

10.1 Среда, в которой работают статистические агентства, меняется. Появляются новые возможности для доступа к большим данным и их изучения, что увеличивает потенциал для возникновения новых идей и возможностей для составления индексов потребительских цен (ИПЦ). Статистический ландшафт становится все более сложным, ожидания лиц, принимающих решения, растут, и перед национальными статистическими службами (НСС) стоит задача предоставлять наилучшие статистические программы более эффективными и инновационными способами.

10.2 Запуск технологии сканирования штрих-кода в 1970-х годах и ее распространение в двадцатом веке позволили розничным продавцам получать подробную информацию об операциях в местах продажи. Данные сканирования имеют большой объем и содержат информацию об отдельных операциях, включая дату, количества и стоимость, а также подробные характеристики товара. Данные, предоставляемые НСС, обычно агрегируются по потребителям и по времени (неделя или месяц). Это богатый источник данных, который потенциально можно использовать для повышения точности ИПЦ и снижения затрат на физический сбор данных; однако это может увеличить затраты на анализ и обработку данных. Кроме того, данные сканирования предоставляют возможности для повышения качества в других аспектах, таких как публикация более подробной информации и выпуск новых информационных продуктов.

10.3 В этой главе обсуждаются возможности и проблемы, возникающие при использовании данных сканирования для составления ИПЦ, и цель состоит в том, чтобы предоставить НСС рекомендации по различным методам обработки данных сканирования. В главе изложены несколько практических соображений, касающихся получения массивов данных со сканеров, оценки и подготовки данных, а также вопросов внедрения. Представлены новые методы, разработанные для построения индексов цен по данным сканирования, так называемые многосторонние методы. Эта глава завершается обсуждением оценки новых методов и эмпирических результатов, коммуникации с пользователями и заинтересованными сторонами, а также публикации и распространения «новых» индексов цен.

10.4 Учитывая динамичный характер использования данных сканирования для составления ИПЦ, в этой главе представлен обзор практики разных стран. Работа по использованию данных сканирования продолжается, особенно в отношении разработки формул для элементарных агрегатов, пригодных для расчета индекса с использованием данных сканирования. Работа с данными сканирования была включена в программу исследований в области ИПЦ для дальнейшего обсуждения и разработки.

Практические соображения

Введение

10.5 Наличие данных сканирования дает возможность улучшить ИПЦ. Массивы данных со сканеров обычно обеспечивают полный охват товаров, продаваемых розничным продавцом во всех его торговых точках, и включают как проданное количество, так и выручку, полученную розничным продавцом за эти товары. Эта информация может: повысить точность данных о ценах, используемых для составления ИПЦ,

путем расчета стоимости единицы однородных продуктов (см. пункты 10.26–10.61); улучшить выборки товаров-представителей, на которые собирают цены, с возможностью использования полного круга товаров для категорий продуктов и торговых точек, охваченных данными сканирования, например, категорий продуктов в данных, предоставленных конкретной сетью супермаркетов; и использования информации о количествах или выручке для определения весов товаров-представителей в соответствии с их экономической значимостью. Данные со сканеров, как правило, не охватывают всю совокупность, входящую в сферу разработки ИПЦ. Например, в большинстве стран данные сканирования не охватывают услуги, арендную плату, автомобили, рестораны или кафе. Кроме того, эта информация может быть доступна только для крупных розничных сетей, но не для небольших независимых магазинов или других типов торговых точек.

10.6 Хотя массивы данных сканирования предоставляют возможности для повышения точности расчета ИПЦ, существуют также проблемы, которые необходимо решить, прежде чем НСС смогут использовать данные сканирования для составления ИПЦ.

Получение массивов данных сканирования

10.7 Данные сканирования существуют уже несколько десятилетий, и их ценность для производства официальной статистики со временем становится все более очевидной. Одной из проблем, с которыми сталкиваются НСС, является получение массивов данных сканирования. Доступны два основных варианта: НСС могут запрашивать массивы данных сканирования непосредственно у предприятий розничной торговли или у сторонних поставщиков данных. Оба варианта имеют свои преимущества и недостатки

10.8 Некоторые НСС успешно договорились о поставке данных сканирования непосредственно от предприятий розничной торговли и использовали эти данные при составлении своего ИПЦ. Непосредственное получение НСС массивов данных о розничной торговле имеет потенциальные преимущества. К ним относятся умение вести переговоры по следующим вопросам:

- Предоставление массива данных бесплатно (или с минимальными затратами)
- Количество товаров, включенных в массив данных
- Уровень агрегирования товаров для обеспечения однородной информации
- Временной охват и детализация (день, неделя или месяц)
- Согласованный график предоставления массива данных для удовлетворения требований обработки данных для ИПЦ.
- Контактное лицо в розничном бизнесе, знакомое с массивом данных, чтобы отвечать на запросы данных от НСС.

10.9 Переговоры о поставках массивов данных сканирования напрямую с предприятиями розничной торговли также сопряжены с трудностями. Основная проблема заключается в том, что двусторонние переговоры о массивах данных сканирования касаются данных, которые могут считаться конфиденциальными, поскольку они содержат информацию об обороте и количествах на уровне товаров. Еще одним моментом являются правовая и институциональная среда, регулирующая отношения между НСС и предприятиями розничной торговли. В некоторых странах может быть необходимо, чтобы закон (о статистике) устанавливал, какие данные должны быть предоставлены, тогда как в других странах достаточно устного соглашения между сторонами. Опыт стран, использующих данные сканирования, показывает, что на завершение этих переговоров уйдет не менее шести месяцев. Переговоры касаются широкого круга тем: от систем информационных технологий (ИТ) и форматов отчетности до вопросов конфиденциальности. Соглашение, достигнутое между НСС и предприятиями розничной торговли, обычно оформляется в виде меморандума о взаимопонимании (или аналогичного документа), в котором документируются роли и обязанности каждой стороны и который направлен на обеспечение непрерывной поставки данных сканирования в НСС в соответствии с согласованным графиком.

10.10 Вместо получения массивов данных сканирования непосредственно от предприятий розничной торговли можно получать данные от посредников или компаний, занимающихся исследованиями рынка. Компании, занимающиеся исследованиями рынка, владеют массивами данных сканирования, получаемыми некоторыми НСС для целей оценки и составления ИПЦ (Krsinich 2015). Такие компании не

обязаны по закону предоставлять данные сканирования; тем не менее, за определенную плату они могут предоставить более старые данные, которые позволят НСС изучить данные и лучше ознакомиться с ними. Более глубокое понимание этих данных может прояснить требования перед началом переговоров с предприятиями розничной торговли или компаниями, занимающимися исследованиями рынка. Основным преимуществом этого подхода является возможность согласовать поставку нескольких массивов данных, относящихся к разнообразному набору продуктов, с одним поставщиком данных или небольшим их числом.

10.11 Опыт НСС, использующих массивы данных сканирования для составления своего ИПЦ, показывает, что предпочтительно получать данные непосредственно от предприятий розничной торговли. Однако получение данных сканирования от компаний полезно в тех случаях, когда получение данных непосредственно от розничных предприятий невозможно или отсутствуют ресурсы для заключения двусторонних соглашений о поставке данных. Доступ к данным сканирования от компаний, занимающихся исследованиями рынка, в большинстве случаев требует ресурсов для покупки этих данных.

Оценка и подготовка данных сканирования для использования

10.12 Если НСС удастся обеспечить доступ к массивам данных сканирования, эти массивы данных затем следует превратить в информацию, которая может быть эффективно и действенно использована для составления ИПЦ. НСС необходимо преодолеть несколько проблем для достижения этих целей.

Разработка ИТ-системы

10.13 Данные со сканеров по своей природе являются *большими* данными. Размер файлов зависит от характеристик исходных данных. Например, файлы с ежедневными данными по торговым точкам будут более объемными, чем файлы с еженедельными данными, агрегированными на уровне розничной сети. Для использования этой информации НСС требуется ИТ/вычислительная система, которая может получать, хранить и обрабатывать большие массивы данных сканирования, если эта информация будет использоваться для составления ИПЦ. ИТ-система должна иметь возможность получать и обрабатывать массивы данных, которые имеют различные структуры классификации, форматы и содержание. Это связано с тем, что предприятия розничной торговли (и сторонние поставщики данных) обычно разрабатывают уникальные системы для собственных целей внутренней отчетности. Эти массивы данных могут быть использованы для минимизации нагрузки на розничных продавцов и повышения своевременности предоставления данных в НСС. Разработка ИТ-системы требует человеческих и финансовых ресурсов. Некоторые НСС задокументировали вопросы, возникающие в связи с необходимостью разработки ИТ-системы. Решение зависит от индивидуальных обстоятельств НСС. Ресурсы НСС потребуются для ИТ-системы, если НСС будет использовать данные сканирования для составления ИПЦ, независимо от того, какой будет поставщик данных.

10.14 Учитывая эти крупные инвестиции, для НСС важно получить опыт работы с тестовыми данными (например, от компаний, занимающихся исследованиями рынка) и взаимодействовать с другими НСС, уже имеющими соответствующий опыт. Система должна быть рассчитана на некоторое время, но в то же время она должна быть способна адаптироваться к меняющейся методологии и быть масштабируемой для обслуживания большого объема данных, получаемых от все большего числа поставщиков данных. Рекомендуется, чтобы, независимо от методов, применяемых для расчета ИПЦ, система оставалась легко адаптируемой к новым методам расчета по мере их появления.

Классификация данных сканирования

10.15 Массивы данных сканирования обычно бывают основаны на классификациях продуктов, которые уникальны для отдельных розничных продавцов. НСС, скорее всего, получит массивы данных, содержащие различные классификации продуктов, которые необходимо сопоставить с единой классификацией, используемой для составления ИПЦ. Классификация массивов данных сканирования может потребовать значительных ресурсов от НСС. Наибольшие инвестиции ресурсов необходимы, когда массивы данных впервые получены НСС. Однако существует потребность в постоянных ресурсах для поддержания классификации по мере того, как в массивах данных появляются новые продукты.

10.16 Классификация товаров, представленных в массивах данных со сканеров, в соответствии с

классификацией ИПЦ осуществляется НСС различными способами, во многих из которых использовались классификации отдельных розничных продавцов. Такие классификации предоставляют важную информацию и могут быть очень полезными, если они имеют тот же (или более высокий) уровень детализации, что самый низкий уровень Классификации индивидуального потребления по целям (КИПЦ). Если имеется соответствие 1:1 или n:1 (продавец: КИПЦ), то данные сканирования могут сопоставляться автоматически. В других случаях данные сканирования либо должны классифицироваться НСС, либо исключаться. Время от времени продавец также может менять свою классификацию. ИТ-система и процесс классификации должны быть гибкими, чтобы изменения в классификации розничных продавцов могли быть своевременно учтены.

10.17 НСС пытались найти решение с учетом своих обстоятельств. Некоторые страны классифицировали товары в данных сканирования по классификации ИПЦ, приобретая метаданные исследования рынка (Muller 2010). Одно европейское НСС использует наиболее подробную классификацию, предоставленную розничными продавцами, а затем проверяет правильность сопоставления и при необходимости вносит соответствующие изменения (van der Grient and de Haan 2010). Некоторые НСС по разным причинам полностью выполнили сами классификацию товаров из данных сканирования в соответствии со своей классификацией ИПЦ (Howard and others 2015).

10.18 Несколько НСС изучают использование алгоритмов машинного обучения для классификации данных сканирования (см. Van Loon and Roels 2018). Эти методы используют массив входных данных либо из предварительно промаркированных товаров-представителей (контролируемое обучение), либо из немаркированных товаров (неконтролируемое обучение), чтобы предсказать правильную таксономическую метку для каждого товара. Полученную модель затем можно использовать для классификации новых массивов данных. Методы машинного обучения особенно перспективны там, где существует несоответствие между классификациями продуктов, используемыми розничными продавцами и классификацией, используемой для расчета ИПЦ. Как и в случае со всеми методами классификации, требуется постоянное поддержание классификации, чтобы товары с новыми характеристиками, которые ранее не были идентифицированы, были классифицированы правильно.

10.19 Трудности с распределением товаров-представителей из данных сканирования по категориям классификации ИПЦ в первую очередь возникают, когда массивы данных сканирования были получены непосредственно от предприятий розничной торговли. Получение массивов данных сканирования от компаний, занимающихся исследованиями рынка, может позволить НСС договориться о поставке данных, которые уже были классифицированы в соответствии с классификацией ИПЦ. Некоторые НСС рассматривают это как особое преимущество получения данных сканирования от компаний, занимающихся исследованиями рынка.

10.20 Необходимо постоянно контролировать надежность системы классификации. Ошибки, допущенные на этом этапе, будут отражены в полученных субиндексах, которые могут быть составлены на основе ошибочно классифицированных товаров.

Обеспечение качества данных сканирования

10.21 По сравнению с традиционной регистрацией цен в торговых точках массивы данных сканирования представляют собой новый источник данных для расчета ИПЦ. Как и в случае любого изменения в источнике данных, составители статистических рядов должны провести ряд проверок, чтобы убедиться, что новый источник данных обеспечивает основу для получения статистических данных, соответствующих цели.

10.22 Проверки должны стать обычными и выполняться автоматически в каждом производственном цикле. Поскольку массивы данных сканирования являются новыми, важно, чтобы НСС приобрели некоторый опыт работы с ними, прежде чем использовать эти данные в производстве. Полученный опыт облегчит установку значений для проверок.

10.23 Эти проверки данных сканирования можно разделить на *глобальные* проверки и *детальные* проверки. *Глобальные* проверки происходят, когда данные поступают в производственный процесс и являются частью процедуры приемки. *Детальные* проверки обычно проводятся ближе к концу производственного процесса.

10.24 *Глобальные* проверки относятся к общим мерам обеспечения качества, обычно применяемым в момент получения НСС массива данных. Эти проверки обеспечивают общее соответствие массива данных массивам, полученным НСС от того же поставщика данных в предыдущие периоды. Проверки могут касаться формата массива данных, общего количества товаров в массиве данных и общей выручки торговой точки. Эти глобальные проверки должны выявить значительные ошибки в данных.

10.25 *Детальные* проверки обычно применяются на уровне разновидностей или групп товаров. Эти проверки направлены на выявление существенных изменений в количестве проданных товаров, в выручке и ценах на товары в массиве данных. Эти детальные проверки традиционно назывались микро-редактированием данных о ценах. Неожиданные изменения динамики цен, товарооборота или количества вызовут эти проверки. Обработка данных сканирования означает обработку гораздо больших массивов данных и может потребовать другого подхода, чем обработка традиционно собираемых данных.

10.26 Как *глобальные*, так и *детальные* проверки должны быть автоматизированы для формирования отчетов для анализа персоналом НСС. Эти проверки могут потребовать контакта с поставщиком данных, а также сравнения данных с другими источниками информации о ценах (например, буклеты и онлайн-цены). Окончательные составленные индексы должны быть рассмотрены и обоснованы для обеспечения правдоподобия.

Внедрение — от конфронтации до новых методов

Преимущества и проблемы использования данных сканирования

10.27 Использование информации, содержащейся в массивах данных сканирования для составления ИПЦ, может привести к значительным изменениям в практике сбора данных и методах составления индекса цен, традиционно используемых НСС. Это говорит о том, что этими изменениями необходимо тщательно управлять как в отношении их воздействия на статистическую программу, так и в отношении коммуникации с пользователями, ключевыми заинтересованными сторонами и персоналом. Сотрудникам НСС необходимо понимать, как управлять этими данными, поскольку массивы данных сканирования намного больше, чем традиционные наборы данных для расчета ИПЦ. При традиционном подходе влияние каждой цены можно было отследить, и часто оно должно было быть видимым. При работе с данными сканирования такое внимание к деталям может оказаться неосуществимым, и эти данные могут потребовать скорее подхода сверху-вниз. Большое значение имеет общение с пользователями и ключевыми заинтересованными сторонами. Для пользователей данных важно полностью понимать, как данные со сканеров используются в процессе составления индекса. Это повышает прозрачность и доверие пользователей.

10.28 Использование данных сканирования потенциально повышает точность ИПЦ несколькими способами и предоставляет значительно больше данных при меньших затратах. Массивы данных сканирования можно использовать для (1) сравнения и проверки данных о ценах; (2) замены цен, собираемых в торговых точках (включая лучшее рассмотрение данных о распродажах, рекламных акциях и скидках); (3) расширения выборки для сбора цен; (4) расширения периода сбора цен; (5) взвешивания товаров на самых низких уровнях агрегирования ИПЦ, чтобы отразить их экономическую значимость; и (6) внедрения новых улучшенных методов расчета индекса и обеспечения автоматизации процессов. Каждое из этих усовершенствований объясняется в ниже.

Использование массивов данных сканирования для валидации данных и обеспечения качества

10.29 Наличие данных сканирования дает НСС возможность проводить валидацию или обеспечивать качество данных, используемых для построения ИПЦ. Массивы данных со сканеров содержат информацию о разновидностях проданных товарах и о выручке, полученной розничным продавцом от продажи этих видов за некоторый период времени, обычно за неделю или месяц. Эта информация позволяет НСС рассчитать цену на отдельный вид товара путем деления выручки от продажи этого вида товара на проданное количество. Эта цена, называемая *стоимостью единицы товара*, представляет собой среднюю цену, с которой сталкиваются потребители за определенный период времени. Заметим, однако, что данные о выручке могут не полностью соответствовать цели и концепции национального ИПЦ, поскольку они могут включать расходы домохозяйств-нерезидентов, предприятий или даже органов

государственного управления (Fenwick 2014).

10.30 Для однородного товара стоимость единицы более точно отражает цены, уплачиваемые потребителями в течение всего периода, чем сбор цен на определенный момент времени (Balk 1998). Стоимость единицы содержит скидки и влияние этих скидок на количество проданных разновидностей товаров. Период, за который рассчитывается стоимость единицы, важен с точки зрения точности этого показателя. Diewert и другие (2016) утверждают, что цены на единицу товара, используемые для расчета ИПЦ, должны относиться к тому же периоду, что и индекс, который необходимо построить, а не к субпериоду.

10.31 Признано, что НСС могут использовать сдвинутый базисный период из-за необходимости обеспечить своевременность предоставления данных и сроки публикации. Привносимые этим действием смещение и дисперсия могут быть оценены путем сравнения индексов, составленных с использованием данных за сдвинутый базисный период, с индексами, составленными с использованием полного базисного периода (Krsinich 2015).

10.32 Ценовые аналитики могут сравнивать цены, собранные в поле, с ценами, рассчитанными на основе массивов данных сканирования. Этот анализ дает представление о любых смещениях, вносимых в ИПЦ из-за сбора данных о ценах в определенный момент времени по сравнению с данными о стоимости единицы товара. Ценовые аналитики НСС могут провести анализ выручки от продажи разновидности товара и проданных количеств, чтобы указать, где выборки ИПЦ можно улучшить.

Использование массивов данных сканирования для замены данных о ценах, собранных в поле

10.33 В большинстве стран большая часть данных о ценах, используемых для составления ИПЦ, собирается путем посещения выбранных предприятий розничной торговли. Эти визиты осуществляются регистраторами НСС, которые наблюдают за ценами на определенный момент времени, а также обсуждают с респондентом скидки, специальные предложения и товары массовой продажи. Регистраторы записывают эту информацию во время визита, часто на портативные электронные устройства. Регулярные посещения торговых точек позволяют регистраторам НСС проводить мониторинг развития рынка и наблюдать за изменением качества.

10.34 Замена собранных в торговых точках цен ценами (стоимостью единиц) из данных сканирования обычно приводит к экономии ресурсов НСС, поскольку регистраторам НСС больше не нужно посещать торговые точки, где они регистрировали цены. Потенциал экономии ресурсов для НСС зависит от сокращения численности регистраторов и увеличения ресурсов, необходимых НСС для управления данными и обработки массивов данных сканирования.

10.35 Стоимость единицы товара должна относиться к одному однородному виду товара, характеристики которого остаются неизменными во времени, поскольку изменения в составе продаваемых разновидностей и их качества не должны отражаться в виде изменения цены. Эти требования создают некоторые проблемы при замене цен, собранных в поле, информацией из массивов данных сканирования. Переговоры между НСС и поставщиком данных необходимы для обеспечения доступа к данным на соответствующем уровне агрегации (или дезагрегации) товаров, необходимом для поддержки расчета стоимости единиц для использования при составлении ИПЦ. Непосредственное предоставление характеристик продукта может облегчить классификацию товаров. Такая информация, если она доступна, может затем использоваться для выполнения явной корректировки качества.

10.36 Некоторые НСС имеют опыт производства данных о стоимости единицы товара из массивов данных сканирования. На самом детальном уровне товары в массивах данных со сканеров обычно идентифицируются по штрих-коду или соответствующему Глобальному идентификационному номеру единицы товара (GTIN) или его подвариантам, универсальному коду товара и европейскому номеру товара. Хотя стандартизированные идентификаторы, такие как GTIN, позволяют отслеживать товары в разных точках розничной торговли, они могут быть слишком подробными, различая разновидности по характеристикам, таким как упаковка, которые не имеют значения для потребителей (Dalen 2017). В этом случае отток товаров будет завышен, и возникнет потенциальная проблема повторного запуска, что может затруднить расчет ИПЦ. Например, при использовании GTIN в качестве идентификатора товара изменение цены однородной разновидности, чей GTIN изменяется в то же время, не будет измеряться. В

некоторых странах использование единицы учета запасов (SKU) вместо GTIN оказалось успешным (Howard and others 2015).

10.37 Существенной частью измерения цен является учет изменения качества и введения новых товаров. НСС осуществляет это, когда регистраторы посещают розничные торговые точки с целью измерения изменения цен на идентичные или эквивалентные товары в последовательные периоды и выявления новых товаров. По мере изменения характеристик разновидностей товаров регистраторы НСС собирают описательную информацию, которая позволяет отделить влияние изменения качества от изменения цены, так чтобы ИПЦ измерял только чистое изменение цены.

10.38 Учет изменения качества особенно сложен при использовании данных сканирования. Массивы данных со сканеров, как правило, показывают высокий уровень оттока разновидностей товаров из месяца в месяц. На рынке появляются новые модели (и версии моделей) товаров, а старые модели уходят с рынка по мере устаревания. Поэтому рассчитать цены с поправкой на качество затруднительно.

Использование массивов данных сканирования для обновления выборок для сбора цен

10.39 Сбор данных о ценах на определенный момент времени регистраторами цен из НСС, посещающими торговые точки, требует значительных ресурсов. На все товары практически невозможно собрать цены в каждом периоде, что приводит к необходимости в выборочном подходе. Например, товары-представители из выборки могут быть отобраны для включения в корзину ИПЦ регистраторами цен из НСС, которые обсуждают с респондентом, какие товары продаются наиболее широко, или изучают полочное пространство товаров и делают выводы об их относительной важности. Затем регистраторы НСС стремятся выбрать репрезентативную корзину товаров для сбора цен. Это подход к формированию целенаправленной выборки.

10.40 Традиционно использовалась невероятностная или целенаправленная выборка, поскольку не было совокупностей для выборки приобретаемых товаров, и отсутствовали подробные данные о количествах или выручке для измерения экономической значимости товаров (дополнительную информацию о выборке см. в Главе 4). Невероятностная или целенаправленная выборка могут привести к систематическим ошибкам, если выбранные товары-представители не являются репрезентативными для совокупности продуктов.

10.41 Этот традиционный подход к выборке может быть заменен более научными методами отбора благодаря наличию данных сканирования. Поскольку данные со сканеров обычно предоставляются для всех товаров, массивы данных сканирования можно использовать в качестве совокупности для обновления выборок для сбора цен. Выборка для сбора цен обычно двумерна; это комбинация выборки торговых точек и выборки товаров/разновидностей продуктов. Если охвачены все магазины розничной сети, массив данных сканирования можно использовать в качестве совокупности для отбора как для торговых точек, так и для товаров (см. также главы 4 и 5). Например, можно применить двухэтапный подход к отбору: сначала отбираются торговые точки, а затем товары в выбранных торговых точках.

10.42 Доли выручки по каждому товару (или комбинации товара/торговой точки) можно использовать для определения значимости каждого товара в группе продуктов. Затем продукты отбираются для включения в «корзину» ИПЦ на основании доли выручки либо путем отбора пропорционально выручке, либо методом отсека (de Naan and others 1999).

10.43 Однако со временем продукты в выборке могут потерять актуальность или даже прекратить свое существование. В этих ситуациях требуется замещение продукта для сохранения релевантности выборки. Тесты для проверки релевантности можно использовать для выделения в выборках тех товаров-представителей, которые стали неподходящими, а также для выделения и ранжирования подходящих товаров для замещения.

10.44 Основной принцип, лежащий в основе этих тестов для проверки релевантности, заключается в том, что товары-представители должны характеризоваться стабильной долей в выручке (то есть постоянной долей в выручке по сравнению с другими продуктами) в группе продуктов для расчета ИПЦ. Эти группы продуктов называются элементарными агрегатами. Стабильная доля в выручке важна, поскольку товары-представители могут иметь большие продажи при выводе на рынок из-за новизны или

начальных продажных цен, после чего приносить незначительную выручку и, следовательно, не быть представительными для более широкого рынка.

10.45 Чтобы смягчить эти проблемы, выручка от возможных товаров для замещения должна быть стабильной и значительной в течение определенного периода (например, от трех до шести месяцев), прежде чем их можно будет рассматривать для включения в выборку для сбора цен. Затем аналитики ИПЦ могут вручную просмотреть все товары, помеченные как подлежащие замене, и выбрать товары из списка, ранжированного в соответствии со среднемесячной долей выручки за предыдущие шесть месяцев.

10.46 Многие продукты питания и предметы домашнего обихода будут иметь разновидности одного и того же базового товара, которые имеют аналогичную, если не идентичную динамику цен. Консервированный тунец определенной марки, например, имеется со многими вкусами, и составители ИПЦ будут знать, что цены на консервы с разным вкусом одного и того же бренда будут вести себя одинаково: поступят в продажу в одно и то же время, и их цена изменится в одно и то же время. Таким образом, наличие продукта с одним вкусом в выборке будет отражать движение цен для гораздо более значительной части рынка, чем можно было бы предположить, исходя из значения выручки по этой одной разновидности консервированного тунца.

10.47 Процессом отбора, используемым для обеспечения репрезентативности выборок товаров, обычно управляют вручную, требуя от аналитиков ИПЦ выбора замещающего товара из ранжированного списка потенциальных продуктов, которые соответствуют определенным критериям приемлемости. Такой подход к формированию выборки на основе данных сканирования требует дополнительных ресурсов аналитика ИПЦ, которые в идеале компенсируются сокращением ресурсов для полевого сбора данных.

Использование данных сканирования для обновления структур индексов и применения весов

10.48 Выборки разновидностей товаров традиционно были небольшими. Когда дополнительные ресурсы аналитиков ИПЦ действительно компенсируются сокращением ресурсов для полевого сбора данных, НСС может принять решение о расширении выборки разновидностей товаров без изменения формулы индекса цен на уровне элементарного агрегата или процедуры выборки.

10.49 Однако, возможно, стоит пересмотреть структуру индекса и процедуру выборки, когда НСС получает данные сканирования непосредственно от розничных сетей. Традиционно индекс элементарного агрегата составляется на основе данных о ценах, собранных в торговых точках, принадлежащих разным торговым сетям (или самостоятельным магазинам). Когда НСС хочет использовать намного больше информации о ценах от розничной сети, чем раньше, представляется предпочтительным рассматривать комбинации элементарных агрегатов¹ и розничных сетей как отдельные страты в процессе составления индекса.

10.50 Если НСС решит использовать систему классификации, предоставленную розничным продавцом, будет необходимо расширить структуру индекса ниже уровня элементарного агрегата с отдельными элементарными индексами цен для каждой розничной сети. Это ставит несколько вопросов. Первый вопрос заключается в том, следует ли рассматривать магазины, принадлежащие рассматриваемой сети, как отдельные торговые точки. В этом случае стоимость единиц для отобранных товаров должна рассчитываться на уровне магазина. С другой стороны, когда уровни обслуживания одинаковы в магазинах, принадлежащих розничной сети, может быть полезно рассчитать стоимость единицы товара на уровне сети (Ivancic and Fox 2013). В этом случае данные со сканеров напрямую из розничной сети представляют все магазины этой сети. При использовании данных на уровне сети необходимо убедиться, что каждая розничная сеть имеет вес в итоговом индексе. Некоторые розничные сети эксплуатируют разные типы магазинов с разным ассортиментом товаров и уровнем цен. В этом случае может быть введена стратификация, разграничивающая типы магазинов, входящих в розничную сеть. У некоторых НСС нет выбора, если они получают данные сканирования на уровне сети.

¹ Элементарный агрегат относится к самому низкому уровню, для которого доступны весовые коэффициенты расходов. Поскольку данные сканирования содержат подробную и своевременную количественную информацию, элементарные индексы цен представляют собой индексы, которые агрегируются для получения индексов более высокого уровня в структуре классификации ИПЦ.

10.51 Региональная дезагрегация данных сканирования может потребоваться, если составляются региональные ИПЦ. Сети, магазины, продукты и цены, включенные в конечные индексы, должны быть репрезентативными для соответствующего региона (регионов). С другой стороны, данные сканирования можно использовать непосредственно на национальном уровне. Затем необходимо определить, могут ли индексы цен, полученные на основе данных сканирования, которые являются репрезентативными для страны в целом, также использоваться для составления региональных ИПЦ.

10.52 Следующий вопрос, который необходимо рассмотреть НСС, заключается в том, в какой степени следует улучшить существующие процедуры формирования выборки. Предположим, НСС раньше использовало выборку товаров, пропорциональную выручке из данных сканирования. Эту процедуру также можно использовать для отбора товаров-представителей из элементарных индексов цен конкретных сетей, где товары определены (и рассчитывается стоимость единицы) либо на уровне торговой точки, либо на уровне сети. Если НСС хочет значительно увеличить размеры выборки, чтобы использовать существенную часть информации о ценах, содержащейся в массивах данных сканирования, необходимо пересмотреть процедуры выборки.

10.53 Еще один вопрос заключается в том, как интегрировать элементарные индексы цен по конкретным сетям из данных сканирования с информацией о ценах из других источников. Поскольку эти элементарные индексы цен отличаются от элементарных индексов цен в традиционной структуре индексов, индексы цен по данным сканирования должны быть агрегированы до уровня — возможно, самого низкого уровня агрегирования продуктов, для которого НСС публикует индексы цен, — где они могут быть объединены с индексами цен из других источников. Если используется классификация конкретных розничных продавцов, она должна быть такой же подробной, как и самый низкий уровень агрегирования продуктов, публикуемый НСС. В противном случае данные должны быть соответствующим образом переклассифицированы. Другими словами, требуется два шага агрегирования: агрегирование элементарных индексов цен для конкретной сети до некоторой товарной категории более высокого уровня и агрегирование результирующих индексов по данным сканирования с индексами цен на этом уровне, относящимися к другим розничным сетям и независимым магазинам.

10.54 Данные о выручке дают НСС возможность чаще взвешивать индексы цен, используя более своевременные данные. Это может быть достигнуто различными способами, в зависимости от наличия у НСС данных сканирования для нескольких сетей. Предлагается, чтобы веса для объединения индексов цен по данным сканирования обновлялись ежегодно с использованием данных о выручке от продажи товаров за предыдущие 12 месяцев. Для объединения индексов по данным сканирования с индексами цен, составленными из других источников, требуются данные о расходах для последних индексов, которые может быть трудно получить или оценить.

10.55 При обследовании бюджетов домохозяйств подробные данные по товарам (или комбинациям товаров/торговых точек) недоступны. Поэтому большинство НСС применяют методы построения невзвешенного индекса цен на самых низких уровнях агрегирования в ИПЦ: цены или изменения цен товаров-представителей, отобранных из элементарного агрегата, объединяются без взвешивания товаров в явном виде в соответствии с их экономической значимостью. В большинстве случаев НСС используют формулу индекса Джевонса.

10.56 Массивы данных со сканеров содержат данные о выручке на самом подробном уровне разновидностей товаров. Эти данные можно использовать для построения выборки разновидностей товаров с вероятностями, пропорциональными выручке от их продаж, как упоминалось ранее, но это вызывает ряд вопросов. Вероятности включения в выборку служат неявными весами. То есть элементарный индекс цен будет взвешенным в неявном виде индексом, а вероятности включения в выборку должны соответствовать целевому индексу (Balk 2005). Кроме того, распределение выручки внутри категории товара, наблюдаемое в данных сканирования, часто сильно асимметрично. Следовательно, при отборе с вероятностью, пропорциональной выручке, скорее всего, будут выбраны некоторые разновидности с высокой выручкой с вероятностью, равной единице, и товары, обеспечивающие высокую выручку, в этой «самоотбирающейся» подвыборке должны быть взвешены в явном виде — без явного или неявного² взвешивания этих товаров (невзвешенный) построенный на

² Неявные веса получаются, когда выборки формируются с использованием метода отбора с вероятностью, пропорциональной размеру

основе выборки индекс Джевонса не может быть объективной оценкой взвешенного геометрического целевого индекса.

10.57 Предпочтительнее в явном виде отражать экономическую значимость товаров с помощью формулы взвешенного индекса, а не в неявном виде с помощью вероятностей включения в выборку для расчета невзвешенного индекса.

Использование данных сканирования для внедрения новых методов составления ИПЦ

10.58 Описанные выше подходы позволяют НСС продолжать использовать методы, основанные на выборке, для составления своего ИПЦ. Повышение точности ИПЦ будет достигнуто за счет того, что цены (стоимость единицы) будут лучше представлять то, что фактически заплатили потребители. Кроме того, отобранные разновидности товара отражают крупных продавцов, а веса, используемые для расчета агрегированных показателей изменения цен, основаны на более своевременной информации и могут обновляться чаще.

10.59 Основная проблема, с которой сталкиваются НСУ, внедряющие эти подходы, связана с увеличением необходимых ресурсов. Поддержание подхода, основанного на выборке, особенно при расширении выборки разновидностей товаров, требует значительного ручного вмешательства, прежде всего потому, что сменяемость видов товаров может быть быстрой. Когда одна европейская страна впервые включила в ИПЦ данные сканирования от супермаркетов, использовался индекс Лоу (Schut and others 2002.). Идея заключалась в том, чтобы воспроизвести традиционные методы и процессы на выборке из примерно 10 000 товаров (штрих-кодов) из каждой сети супермаркетов. Этот подход требовал очень больших усилий для проведения ручного выбора товаров для замещения исчезающих товаров и корректировки качества, когда это считалось необходимым.

10.60 В идеале, с учетом практических ограничений, НСС должна использовать всю имеющуюся информацию из массивов данных сканирования, а не строить выборки. Однако ручная обработка всех разновидностей товаров из массивов данных сканирования является непомерно дорогой и не может быть предпринята при необходимости соблюдать сроки производства ИПЦ. Требуется автоматизация процессов составления ИПЦ.

10.61 Кроме того, при использовании всех разновидностей товаров, а не их выборки, предпочтительна формула взвешенного индекса. Опять же, частая смена разновидностей товаров представляет собой серьезную проблему. Чтобы максимизировать количество совпадений в данных, потребуется часто проводить цепную увязку. Это может привести к значительному смещению индекса. Многосторонние методы расчета индекса цен, в которых смещение отсутствует по построению, в настоящее время считаются подходящим методом для обращения со всеми товарами и разновидностями из данных сканирования; однако продолжается работа над тем, как лучше всего применять многосторонние методы для составления элементарных агрегатов с использованием данных сканирования. Данные сканирования открывают множество возможностей для новых исследований и разработок.

Многосторонние методы построения индекса цен

Введение

10.62 Данные сканирования могут быть внедрены в расчеты ИПЦ с использованием традиционных методов, основанных на выборке. Цены, ранее наблюдаемые регистраторами, посещающими торговые точки, могут быть заменены значениями стоимости единиц товара из данных сканирования без изменения дизайна выборки или используемой формулы индекса цен. Если НСС решит использовать все имеющиеся данные, а не формировать выборки, предпочтительным подходом являются многосторонние методы построения индекса цен. Многосторонние методы изначально были разработаны для сравнения уровней цен в разных странах, но их можно адаптировать для сопоставления цен с течением времени. Эти методы особенно полезны для данных сканирования, когда сменяемость товаров высока и часто проводятся рекламные распродажи.

10.63 Наиболее важные многосторонние методы построения индекса цен описаны ниже; всестороннее обсуждение см. в главе 7 *Теории индекса потребительских цен*. После определения разновидностей дается краткий обзор традиционных двусторонних индексов цен и цепных индексов.

Определение разновидности

10.64 Прежде чем применять какой-либо метод расчета индекса, необходимо определить разновидность товара, на который будут собираться цены. Основной принцип заключается в том, чтобы сравнивать подобное с подобным и отслеживать цену одного и того же вида товара во времени. Уровень кода товара обычно представляет собой наиболее детальный уровень однородности в данных. В дополнение к этому аспекту продукта необходимо также учитывать параметры торговой точки и времени. Часто можно рассматривать один и тот же товар, продаваемый в разные моменты времени в одних и тех же или аналогичных торговых точках, для определения достаточно однородной разновидности, чтобы можно было рассчитать среднюю стоимость операции (стоимость единицы) для этой разновидности.

10.65 В некоторых случаях коды товаров, такие как номера GTIN, стабильны и долговечны. В некоторых странах есть доступ к кодам продуктов розничных продавцов, например, к внутреннему коду SKU, который уже является более агрегированным, чем код GTIN. Однако в других случаях уровень этих кодов может быть слишком подробным для расчета индекса цен. В некоторых категориях продуктов, таких как одежда и обувь, коды товаров часто появляются и исчезают, что затрудняет их сопоставление во времени, и поэтому изменения цен не измеряются должным образом. Ниже описаны различные стратегии, позволяющие справиться с изменениями кода товара.³

10.66 Один из подходов состоит в том, чтобы сгруппировать различные отдельные коды со схожими характеристиками. НСС могут выбрать более широкое или более узкое определение разновидности. Важно создавать группировки так, чтобы потребители были более или менее безразличны к различным отдельным товарам в этих группах. Расчет стоимости единиц на этом уровне не только позволяет выявить эффекты замещения между сопоставимыми товарами, но и облегчает включение новых товаров, поступающих на рынок. НСС сталкиваются с необходимостью уравнивания этих группировок. Если группировка слишком широкая, это может привести к смещению стоимости единицы товара (и высокой волатильности), поскольку отдельные товары не являются строго сопоставимыми. С другой стороны, слишком узкое определение группировки может привести к отсутствию сопоставления между выходящими и новыми или возвращающимися товарами. Решения, принятые на этом этапе, могут существенно повлиять на индексы цен, которые в конечном итоге будут получены. Это может быть особенно актуально для технологических продуктов, особенно для моделей с высокой сменяемостью (подробнее см. главу 6). Если возможно, следует проверить чувствительность определений групп к результатам.

10.67 На практике построение этих группировок может оказаться сложной задачей. НСС требуется информация о характеристиках товара, включая марку и размер, а также внутренние классификационные коды, используемые розничными продавцами. Некоторые розничные продавцы могут указывать характеристики только в определенной текстовой строке, в то время как другие могут иметь несколько различных переменных, описывающих характеристики разных товаров. Характеристики, собранные в одну переменную (текстовую строку), требуют некоторой формы анализа текста, чтобы сделать ее полезной для классификации. Не все характеристики одинаково важны и в одинаковой степени влияют на цену. Группировка отдельных товаров должна определяться наиболее важными характеристиками, определяющими цену.

10.68 Другой подход может заключаться в условном исчислении цен на новые и исчезающие товары в те периоды, когда их нет в наличии. Цены могут, например, быть условно исчислены с помощью гедонической функции. Вместо использования характеристик товаров для формирования групп, эта информация теперь используется для оценки отсутствующих цен. Однако такая стратегия подходит только для таких формул индекса, как индексы Торнквиста или Джини—Элтето—Кёвеша—Шульца (GEKS)-Торнквиста, которые реагируют на условное исчисление отсутствующих цен.

Двусторонние индексы цен и цепная увязка

10.69 Предположим сначала, что набор продаваемых разновидностей товаров фиксирован во времени (то есть статистик имеет дело со *статической совокупностью*). Этот фиксированный набор товаров обозначается S , а его размер обозначается N . Период выборки состоит из $(T + 1)$ периодов времени $t =$

³ С подобной же проблемой можно столкнуться при использовании данных веб-скрейпинга (см. Глава 5, Приложение 5.6).

$0, \dots, T$. Цены (стоимости единиц) товара $i \in S$ ($i = 1, \dots, N$) в периоды 0 и t ($t = 1, T$) обозначены P_i^0 и P_i^t ; а q_i^0 и q_i^t обозначают соответствующие проданные количества. Цель состоит в том, чтобы построить индексы цен, которые сравнивают период 0, начальный период временного ряда, с каждым периодом t .

10.70 В ситуации отсутствия информации о расходах, в главе 8 данного Руководства рекомендуется использовать *индекс цен Джевонса* - невзвешенное среднее геометрическое соотношений цен.

$$I_J^{0:t} = \prod_{i \in S} \left(\frac{P_i^t}{P_i^0} \right)^{\frac{1}{n}} = \frac{\prod_{i \in S} (P_i^t)^{\frac{1}{n}}}{\prod_{i \in S} (P_i^0)^{\frac{1}{n}}} \quad (10.1)$$

10.71 НСС традиционно формирует выборку товаров из всей совокупности S , чтобы снизить затраты на производство ИПЦ. Без доступа к данным сканирования S неизвестно, а детализированная совокупность для выборки отсутствует. В связи с этим, большинство выборок для ИПЦ были целенаправленными, что приводит к риску систематической ошибки в индексе.

10.72 Поскольку данные сканирования содержат информацию о расходах для всех товаров, построение гиперболических индексов цен возможно для всего набора S . Здесь основное внимание уделяется индексу Торнквиста, а не индексу Фишера или другим гиперболическим формулам индекса. Хотя индексы Фишера и Торнквиста дают очень похожие результаты, индекс Торнквиста позволяет использовать более простые выражения. Индекс цен Торнквиста определяется так

$$I_T^{0:t} = \prod_{i \in S} \left(\frac{P_i^t}{P_i^0} \right)^{\frac{s_i^0 + s_i^t}{2}} \quad (10.2)$$

где $s_i^0 = p_i^0 q_i^0 / \sum_{i \in S} p_i^0 q_i^0$ и $s_i^t = p_i^t q_i^t / \sum_{i \in S} p_i^t q_i^t$ обозначают доли расходов в периоды 0 и t

10.73 В *динамической совокупности* имеются новые и исчезающие товары, так что не все товары могут быть сопоставлены с течением времени. Наборы товаров в периоды t ($t = 0, \dots, T$) обозначаются S^t с размером N^t . Для максимизации количества совпадений в данных цепная увязка гиперболических индексов цен для сравнимых моделей представляется полезной, например, цепная увязка индексов цен Торнквиста от периода к периоду:

$$I_T^{t-1:t} = \prod_{i \in S_M^{t-1,t}} \left(\frac{P_i^t}{P_i^{t-1}} \right)^{\frac{s_{i,M}^{t-1} + s_i^t}{2}} \quad (10.3)$$

где S_{iM}^{t-1} и S_{iM}^t – доли расходов в двух периодах для набора $S_{iM}^{t-1,t} = S^{t-1} \cap S^t$ сравнимых товаров, которые имеются в обоих периодах: периоде $(t-1)$ и периоде t .

10.74 Однако эмпирическая работа показала, что частое цепная увязка взвешенных индексов цен, включая гиперболические индексы цен, может привести к сильному «цепному смещению». В случае продаж в рамках рекламных акций по сниженным ценам покупаемые количества часто значительно увеличиваются. Но когда цены возвращаются к своему первоначальному уровню, количество приобретаемых товаров длительного хранения может не вернуться к своему «нормальному» уровню. Этот тип асимметричного поведения может вызвать цепное смещение гиперболических индексов цен, которое обычно направлено вниз. Ivancic (2007), используя данные сканирования для исследования рынка товаров, продаваемых в супермаркетах, обнаружил смещение в сторону понижения в цепных индексах цен Фишера (см. также Ivancic and others 2009, 2011). Смещение цепных гиперболических индексов цен для сравнимых моделей также было задокументировано для товаров длительного пользования. Здесь смещение, вероятно, связано с сезонными колебаниями цен и объемов. De Naan and Krsinic (2014), используя данные сканирования, обнаружили смещение в сторону понижения в цепных индексах цен Торнквиста на проданные товары бытовой электроники. Silver and Heravi (2005) представили доказательства смещения в сторону понижения цепных индексов Фишера, полученных с использованием данных сканирования, для

телевизоров.

Таблица 10.1 Численный пример цепного смещения

Период	P1	P2	Q1	Q2	Доля 1 (%)	Доля 2 (%)	Средняя доля в t и t - 1 для товара 1 (%)	Средняя доля в t и t - 1 для товара 2 (%)	Период-к периоду индекс Торнквиста	Цепной индекс Торнквиста
1	3	4	12	10	47.4	52.6				100.00
2	3	2	12	30	37.5	62.5	42.4	57.6	67.10	67.10
3	1	4	40	5	66.7	33.3	52.1	47.9	78.66	52.78
4	3	4	5	10	27.3	72.7	47.0	53.0	167.53	88.42
5	3	4	12	10	47.4	52.6	37.3	62.7	100.00	88.42
6	3	2	12	30	37.5	62.5	42.4	57.6	67.10	59.33
7	1	4	40	5	66.7	33.3	52.1	47.9	78.66	46.67
8	3	4	5	10	27.3	72.7	47.0	53.0	167.53	78.18
9	3	4	12	10	47.4	52.6	37.3	62.7	100.00	78.18

10.75 В таблице 10.1 показан численный пример смещения в сторону понижения цепного индекса цен Торнквиста. Здесь выделено два товара и девять периодов. «Обычные» цены товаров 1 и 2 составляют 3,00 и 4,00 соответственно, но цена товара 1 временно снижается в периоды 3 и 7, а цена товара 2 снижается в периоды 2 и 6. Обратите внимание, что в последний период (период 9) цены и количества точно такие же, как и в первый период. Тем не менее, цепной индекс цен Торнквиста от периода к периоду заканчивается на уровне 78,18, таким образом, показывая снижение цен почти на 22 процента. Для многосторонних индексов индекс равен единице (как и прямой индекс Торнквиста). Это смещение вниз для цепного индекса Торнквиста от периода к периоду связано с тем, что, количества не сразу возвращаются к своему «нормальному» уровню после скидки, и изменение цены от нормальной цены к сниженной цене имеет больший вес, чем последующее изменение цены от сниженной цены обратно к нормальной цене.

10.76 Один из способов избежать цепного смещения, вызванного продажами по акции для товаров, подлежащих хранению, будет состоять в том, чтобы не взвешивать товары и построить временной ряд посредством цепной увязки от периода к периоду индекс цен Джевонса для сравнимых моделей:

$$I_J^{t-1,t} = \prod_{i \in S_M^{t-1,t}} \left(\frac{p_i^t}{p_i^{t-1}} \right)^{\frac{1}{N_M^{t-1,t}}} \quad (10.4)$$

Где $N_M^{t-1,t}$ – количество совпавших товаров между периодами $t - 1$ и t .

Это не означает, что использование цепного индекса Джевонса для сравнимых моделей не вызывает проблем. Например, распродажи могут оказать понижающее давление на индекс. Чтобы смягчить эту проблему, можно использовать фильтр для удаления товара. Фильтр удаляет товар, если цена и количество проданного товара резко падают. Смещение вниз может также возникать для модных товаров, таких как одежда, которые покидают выборку по низким распродажным ценам и больше не возвращаются. Одежда требует особого рассмотрения, как указано в главе 11.

10.77 Отсутствие взвешивания также является проблемой. Расходы на товары обычно сильно асимметричны, поэтому многим товарам с низкими расходами на них будет придан тот же вес, что и нескольким товарам с высокими расходами. В грубой форме взвешивание в неявном виде можно получить, просто исключив товары с низким уровнем расходов (то есть, придав им нулевую вероятность включения в выборку), например, используя порог, основанный на средней доле расходов на товар в смежные месяцы. Этот подход, иногда называемый «динамическим подходом» (см. Eurostat 2017), был реализован в нескольких европейских странах (например, см. van der Grient and de Haan 2010, 2011). Этот метод снижает риск цепного смещения, поскольку веса используются неявно при отборе товаров, а не в явной форме при расчете индекса. Преимущество этого метода заключается в том, что он опирается на обычные методы построения двустороннего индекса, в то же время наилучшим образом используя информацию, содержащуюся в массивах данных сканирования. Поэтому метод легко объяснить пользователям. Тем не менее, это не оптимальная ситуация. Лучшим решением было бы взвешивание товаров в явном виде и построение взвешенных многосторонних индексов цен.

Многосторонние методы

10.78 Многосторонние методы позволяют получить транзитивные индексы цен. Для сравнения цен во времени это означает, что индексы не зависят от выбора базисного периода, могут быть записаны в цепной форме и, следовательно, не подвержены цепному смещению. Общим для многосторонних методов является то, что индексы цен строятся одновременно для всего периода выборки.

10.79 Можно определить два типа многосторонних методов. Первый тип начинается с сопоставления цен по сравнимым моделям между любой парой периодов времени в течение всего периода выборки, а затем привносит транзитивность в этот набор двусторонних индексов цен. Наиболее известным методом является GEKS (Eltetö and Kovács, 1964; Gini, 1931; Szulc, 1964). Другой метод основан на связующих деревьях (Hill 1999a, 1999b), где связующее дерево является мостом между странами. Для определенного связующего дерева двусторонние индексы являются цепными для построения сравнений цен между любой парой стран или, в соответствии с нашим контекстом, между периодами времени. Однако неясно, в чем заключаются теоретические и практические преимущества по сравнению с более простыми в построении индексами GEKS. Второй тип многостороннего метода достигает транзитивности другим способом, который будет объяснен ниже, и включает в себя метод Гири-Камиса (Geary, 1958; Khamis, 1972) и метод условных переменных для страны и продукта (Summers, 1973).

Метод GEKS

10.80 Индекс GEKS между периодом 0 и периодом t рассчитывается как среднее геометрическое отношений двусторонних индексов цен для сравнимых моделей $I^{0,j}$ и $I^{0,k}$, построенных по той же формуле индекса, где каждый период l принимается за базисный. При условии, что двусторонние индексы удовлетворяют критерию обратимости времени, индекс GEKS может быть записан следующим образом (de Naan and Van der Grient 2011; Ivancic and others 2011):

$$I_{\text{GEKS}}^{0:t} = \prod_{l=0}^T \left(\frac{I^{0,l}}{I^{t,l}} \right)^{\frac{1}{T+1}} = \prod_{l=0}^T \left(I^{0,l} \times I^{l,t} \right)^{\frac{1}{T+1}} \quad (10.5)$$

10.81 Критерий обратимости времени требует, чтобы при перемене базисного периода и периода сравнения местами результат был равен обратной величине исходного индекса. В своей стандартной форме метод GEKS использует двусторонние индексы Фишера, которые удовлетворяют критерию, но возможны и другие варианты, включая двусторонние индексы Торнквиста. Индексы GEKS-Tornqvist также называют индексами Caves, Christensen и Diewert.

Метод Гири-Камиса

10.82 Метод Гири-Камиса (GK) применительно к сравнениям во времени дает следующий индекс цен:

$$I_{\text{GK}}^{0:t} = \frac{\frac{\sum_{i \in S^t} p_i^t q_i^t}{\sum_{i \in S^t} p_i^0 q_i^0}}{\frac{\sum_{i \in S^0} p_i^0 q_i^0}{\sum_{i \in S^0} p_i^t q_i^t}} = \frac{\left[\sum_{i \in S^t} s_i^t \left(\frac{p_i^t}{p_i^0} \right)^{-1} \right]^{-1}}{\left[\sum_{i \in S^0} s_i^0 \left(\frac{p_i^0}{p_i^t} \right)^{-1} \right]^{-1}} \quad (10.6)$$

10.83 Числитель уравнения 10.6 представляет собой индекс цен (с использованием количеств за период t) с «базисными ценами» которые фиксируются в течение периода выборки. Индекс должен быть равен единице в начальном периоде 0, поэтому необходимо нормализовать индекс, разделив его на значение в периоде 0, которое является знаменателем уравнения 10.6. Базисные цены задаются так:

$$\hat{p}_i = \frac{\sum_{\tau \in S_i} q_i^\tau \left(\frac{p_i^\tau}{I_{\text{GK}}^{\tau, 0, \text{tau}}} \right)}{\sum_{\tau \in S_i} q_i^\tau} \quad (10.7)$$

где S_i ряд периодов времени, в которые фактически продается товар i и для которых имеются цены.

Уравнение 10.7 показывает, что \widehat{P}_t равняется взвешенному среднему арифметическому дефлятированных наблюдаемых цен, где весами являются доли в каждом периоде в общем количестве продаж товара за весь период выборки.

10.84 Поскольку индекс GK действует как дефлятор в уравнении 10.7, уравнения 10.6 и 10.7 определяют систему уравнений, которые необходимо решать одновременно. Это можно сделать итеративно (что может быть проще реализовать), но есть и другие способы решения системы (Balk 2008).

Метод условных переменных времени и продуктов (TPD)

10.85 Это подход основан на регрессионной модели. Предполагая, что за весь период выборки $0, \dots, T$ наблюдается n различных товаров (большинство из которых, как правило, не будет продаваться в каждый период времени), регрессионная модель условных переменных времени и продуктов (TPD) для объединенных данных будет иметь следующий вид:

$$\ln p_i^t = \alpha + \sum_{t=i}^T \delta^t D_i^t + \sum_{i=t}^{N-1} \gamma_i D_i + \varepsilon_i^t \quad (10.8)$$

Где D_i - это условная переменная, равная единице, если наблюдение относится к товару i , и равная нулю в противном случае, а D_i^t - это условная переменная, равная единице, если наблюдение относится к периоду t и равная 0 в противном случае; условные переменные для товара n и периода 0 исключены из спецификации модели.

10.86 Diewert (2005) предложил использовать для оценки регрессионной модели (10.8) взвешенный метод наименьших квадратов, где весами являются доли расходов на товары в каждом периоде.

Возведение в степень расчетного параметра условной переменной времени, δ^t , дает индекс TPD между

периодом 0 и t ; $I_{TPD}^{0,t} = \exp(\delta^t)$. Взвешенный метод условных переменных времени и продуктов может быть записан как система уравнений, которая аналогична геометрической системе в стиле GK, определяемой формулами (10.6) и (10.7), и таким образом модель TPD может быть решена итеративно, а также методами прямой регрессии (Rao 2005):

$$I_{TPD}^{0,t} = \frac{\prod_{i \in S^t} \left(\frac{p_i^t}{\exp(\gamma_i)} \right)^{s_i^t}}{\prod_{i \in S^0} \left(\frac{p_i^0}{\exp(\gamma_i)} \right)^{s_i^0}} \quad (10.9)$$

$$\exp(\gamma_i) = \prod_{\tau \in S_i} \left(\frac{p_i^\tau}{I_{TPD}^{0,\tau}} \right)^{\sum_{\tau \in S_i} s_i^\tau} \quad (10.10)$$

10.87 Уравнение 10.10 показывает, что возведенные в степень оценки с фиксированным эффектом товара, γ_i , или базисные цены, равны взвешенным по долям расходов средним геометрическим значениям дефлятированных цен, где индекс TPD выступает в качестве дефлятора. И метод GK, и TPD явно приводят к базисным ценам. В случае GK это означает, что индекс согласуется с национальными счетами, поскольку он поддается аддитивной декомпозиции, а TPD, будучи геометрическим индексом - нет.

10.88 Обратите внимание, что индекс TPD (10.9) можно рассматривать как нормализованный геометрический индекс Пааше с условно исчисленными ценами периода 0 на основании базисных цен (10.10). Точно так же индекс GK (10.6) можно рассматривать как нормализованный (обычный) индекс Пааше с условно исчисленными ценами периода 0 на основании базисных цен (10.7).

Недостаточная сравнимость и поправки на качество

Методы поправки на качество в неявном виде

10.89 Как и GEKS, GK и TPD являются методами для сравнимых моделей в том смысле, что товары (услуги)-представители с одним наблюдением за весь период выборки не влияют на индекс. Товары-представители вносят вклад в совокупное изменение цен только тогда, когда соотношения цен могут быть рассчитаны на основе цен, наблюдаемых в оба сравниваемых периода, если только информация о характеристиках не будет доступна для выполнения поправок на качество в явном виде. Одним из следствий метода сравнимых моделей является то, что товары-представители, введенные в самый последний период T, игнорируются.

10.90 Поправка на качество в неявном виде также может быть проиллюстрирована с помощью другой интерпретации индекса GK. Деление индекса стоимости категории продукта на отношение «количеств с поправкой на качество» определяет *индекс стоимости единицы продукта с поправкой на качество* (de Naan 2004, 2015):

$$I_{QAUV}^{0:t} = \frac{\sum_{i \in S^t} P_i^t q_i^t}{\sum_{i \in S^0} P_i^0 q_i^0} = \frac{\sum_{i \in S^t} P_i^t q_i^t}{\sum_{i \in S^t} \lambda_{i/b} q_i^t} = \frac{\left[\sum_{i \in S^t} s_i^t \left(\frac{P_i^t}{\lambda_{i/b}} \right)^{-1} \right]^{-1}}{\left[\sum_{i \in S^0} s_i^0 \left(\frac{P_i^0}{\lambda_{i/b}} \right)^{-1} \right]^{-1}} \quad (10.11)$$

10.91 Если $\lambda_{i/b}=1$ для всех i , уравнение 10.11 упрощается до обычного индекса единицы стоимости. Факторы поправки на качество $\lambda_{i/b}$ направлены на то, чтобы выразить купленные количества каждого товара i по отношению к количеству произвольного товара b . Отношения $P_i^t / \lambda_{i/b}$ и $P_i^0 / \lambda_{i/b}$ во втором выражении в (10.11) являются ценами с поправкой на качество. В случае статической совокупности (без новых и исчезающих товаров), если фактор поправки на качество товара-представителя соответствует его базисной цене, индекс стоимости единицы с поправкой на качество превращается в индекс цен Пааше. Von Auer (2014) утверждал, что многие традиционные индексы цен можно рассматривать как то, что он назвал обобщенным индексом стоимости единицы продукта.

10.92 Сравнение уравнений 10.6 и 10.11 показывает, что GK-индекс можно рассматривать как индекс стоимости единицы продукта с поправкой на качество, где факторы поправки на качество измеряются при помощи базисных цен в уравнении 10.7. Точно так же индекс TPD в 10.9 можно рассматривать как его геометрический аналог, где факторы поправки на качество измеряются с помощью базисных цен (10.10). Правильно ли отражают базисные цены в индексах GK и TPD различия в качестве, скорее всего, зависит от рыночной ситуации.

Поправка на качество в явном виде

10.93 Если позволяют данные о характеристиках товаров-представителей, предпочтительной является поправка на качество в явном виде, в частности, с помощью методов гедонической регрессии. Полезной отправной точкой является многосторонняя гедоническая модель с условной переменной времени (TDH):

$$\ln p_i^t = \alpha + \sum_{t=1}^T \delta^t D_i^t + \sum_{k=1}^K \beta_k z_{ik} + \varepsilon_i^t \quad (10.12)$$

Где z_{ik} - количество характеристики k ($k = 1, \dots, K$) для товара-представителя i .

Заметим, что, как указывали Aizcorbe и др. (2003) и Krsinich (2016), модель TPD (10.8) возникает из модели TDH (10.12) путем замены гедонических эффектов фиксированными эффектами, специфичными для отдельных товаров, $\exp(\gamma_i)$.

10.94 Полученный в результате взвешенный индекс TDH, $I_{TDH}^{0,t} = \exp(\delta^t)$, может быть выражен

аналогично индексу TPD (10.9) с оцененными гедоническими эффектами $\exp\left(\sum_{k=1}^K \tilde{\beta}_k z_{ik}\right)$ вместо расчетных фиксированных эффектов товаров $\exp(\tilde{\gamma}_i)$, действующими как базисные цены. Формулу

$\exp\left(\sum_{k=1}^K \tilde{\beta}_k (z_{ik} - z_{bk})\right)$ также можно использовать для оценки факторов поправки на качество в уравнении 10.11. Ожидается, что полученный индекс стоимости единицы с поправкой на качество в явном виде — или «гедонический вариант» индекса GK — будет очень близок к индексу TDH (de Naan and Krsinich 2018).

10.95 De Naan и др. (2016) сравнили методы TPD и TDH. Они утверждали, что модель TPD страдает от чрезмерно близкой подгонки, потому что она имеет слишком много параметров и «искажает остатки регрессии в сторону нуля». При определенных стратегиях ценообразования розничных продавцов, таких как скиминг цен (новые товары) и демпинг (старые товары), индекс TPD может сильно отличаться от индекса TDH. Точно так же GK-индекс может сильно отличаться от своего гедонического аналога — индекса стоимости единицы товара с поправкой на качество.

10.96 Если повторный запуск однородных товаров-представителей с разными штрих-кодами или кодами SKU является основной причиной низкой степени сравнимости, то определение товаров по их характеристикам, а не по штрих-коду или коду SKU, может быть выходом (Chessa 2016). Однако массивы данных сканирования, предоставляемые розничными продавцами, обычно включают довольно общие описания товаров, из которых возможно извлечь лишь несколько характеристик, таких как размер упаковки и торговая марка. В этом случае цены, рассчитанные как стоимость за единицу по всем штрих-кодам SKU, которые принадлежат к различным «группам», могут страдать от систематической ошибки стоимости за единицу.

10.97 В отличие от TPD и GK, метод GEKS не нацелен на корректировку изменения качества в неявном виде. Однако потенциальное преимущество GEKS по сравнению с TPD и GK заключается в том, что «отсутствующие цены» на несравнимые новые и исчезающие товары-представители могут быть условно исчислены. Таким образом, можно оценить в явном виде скорректированные на качество индексы GEKS, заменив двусторонние индексы цен Торнквиста для сравнимых моделей условно исчисленными двусторонними гедонистическими индексами Торнквиста, например, как это было предложено в работе Naan and Krsinich (2014) или de Naan (2019). Это означает, что нет необходимости определять товары-представители по их характеристикам; штрих-кода или кода SKU будет достаточно, чтобы различать товары. Гедонические условные исчисления для несравнимых товаров корректируют изменения качества, а также решают проблему перезапуска. De Naan (2019) предлагает использовать ту же информацию о характеристиках в гедонических регрессиях, которая будет использоваться для определения «групп» для работы с перезапусками в методах TPD и GK.

10.98 Отсутствие информации о важных характеристиках делает использование гедонической поправки на качество или явной поправки на качество в целом проблематичным, поскольку это может привести к смещению, связанному с пропуском переменных. Кроме того, как упоминалось ранее, это может привести к смещению стоимости единицы товара в «групповом подходе». Несколько НСС изучают возможность использования веб-скрейпинга для наблюдения за характеристиками качества на веб-сайтах розничных продавцов или производителей для обогащения массивов данных сканирования. Данные сканирования, полученные от компании, занимающейся исследованиями рынка, могут уже содержать подробную информацию о характеристиках товара. Одно НСС, например, производит индексы цен GEKS с поправкой на качество для товаров бытовой электроники на основе массивов данных сканирования от одной исследовательской компании, которые включают множество характеристик товара (Krsinich 2015). Все эти методы требуют больших объемов данных, поскольку требуют детальных данных о ценах, товарообороте и характеристиках, определяющих цену.

Пересмотры в многосторонних индексах

10.99 При появлении новых данных ранее оцененные многосторонние индексы изменяются при

обработке новых данных. Это проблематично, поскольку ИПЦ не подлежит пересмотру. Кроме того, с течением времени на недавнее движение цен все больше будут влиять цены и изменения цен в далеком прошлом. В контексте индекса GEKS это известно, как *потеря характеристики*. Были предложены различные методы расширения временного ряда многосторонних индексов без пересмотра опубликованных индексов (и смягчения потери характеристики). Методы можно охарактеризовать рядом вариантов: настройка окна (скользящее окно или расширяющееся окно), период увязки и индекс в периоде увязки (Chessa 2019).

10.100 Методы *скользящего окна* оценивают многосторонние индексы в скользящем окне фиксированной длины, скажем, $T + 1$, которая сдвигается вперед каждый период. Таблица 10.2 иллюстрирует скользящее временное окно из 13 периодов. Многосторонний индекс, составленный в период 13, охватывает периоды 1-13. Многосторонний индекс, составленный в период 14, охватывает периоды 2-14 и так далее. Затем результаты последнего окна увязываются с ранее рассчитанным индексом. Например, соединение связывает при изменении цен самое последнее изменение индекса от периода к периоду с индексом предыдущего периода (то есть с последним опубликованным индексом). В таблице 10.2 индекс соединения при изменении цен для периода 14 получается путем привязки к периоду 13 изменения цены между периодами 13 и 14, полученного из соответствующего многостороннего индекса.

10.101 Альтернативой соединению при изменении цен является соединение окон, которое увязывает изменение индекса за весь период с последним рассчитанным индексом T периодов назад. Соединение изменения цен было предложено Ivancic и др. (2011) для метода GEKS, а соединение окон было предложено Krsinich (2016) для метода TPD. Однако каждый метод соединения можно объединять с любым многосторонним методом. Эти методы соединения увязывают изменения индекса в одном периоде увязки. Diewert и Fox (2017) предложили соединение средних значений, взяв среднее геометрическое индексов цен, полученных при использовании каждого возможного периода увязки. Это делает результат независимым от выбора периода увязки.

10.102 В работе Chessa (2019) указано, что на самом деле существует два основных варианта методов соединения (кроме соединения изменения цен, где индекс, опубликованный в предыдущий период, является единственным вариантом для увязки). Последовательные сдвиги окна создают последовательность пересчитанных или «пересмотренных» индексов наряду с первоначально опубликованным индексом за тот же период. Как пересчитанные, так и опубликованные индексы являются кандидатами для увязки нового ряда индексов. В эмпирических исследованиях и в применении двумя НСС использовался первый вариант. Второй вариант (то есть увязка с опубликованными индексами) недавно был предложен Chessa (2019). Увязка с опубликованными индексами имеет свои преимущества. Например, годовые показатели инфляции, рассчитанные для сдвинутых окон, также будут опубликованными показателями, если период увязки соответствует периоду 12 месяцев назад. Это повышает прозрачность методов соединения. Более того, каждый годовой показатель выводится из ряда транзитивных индексов и в этом смысле свободен от цепного смещения. Это не относится к соединению последнего пересчитанного индекса, который, следовательно, может привести к некоторому смещению (см. Chessa 2019).

10.103 О выборе длины окна следует подумать. Ivancic и др. (2011) выступали за 13-месячный (или пятиквартальный) период, поскольку это самое короткое окно, в которое можно рассматривать сильно сезонные товары. Однако недавние исследования показывают, что индексы для сильно сезонных товаров могут быть значительно улучшены с окном в 25 месяцев (или девять кварталов) (Chessa 2020). Можно построить взвешенные индексы GEKS, которые учитывают надежность двусторонних индексов цен (Rao 2001b). Melser (2018) предложил взвешенный метод GEKS, в котором веса зависят от степени сравнимости товаров-представителей, например, в отношении долей расходов. Здесь выбор длины окна менее важен, поскольку двусторонним индексам с более низкой степенью сравнимости будут приданы более низкие веса.

10.104 Методом прямого расширения с годовой цепной увязкой (Chessa 2016) можно строить ряды многосторонних индексов, скажем, за 13 месяцев, начиная, например, с декабря и заканчивая декабрем следующего года, и сцеплять их в декабре каждого года для получения долгосрочного временного ряда. Длина окна оценки краткосрочных индексов увеличивается каждый месяц — индекс за январь в

краткосрочном ряду оценивается на основе данных за два месяца (что представляет собой скорее двустороннее, а не многостороннее сравнение) и так далее, пока в декабре не будут использованы данные за 13 месяцев.

10.105 Потенциальная слабость метода прямого расширения заключается в том, что индексы цен за первые пару месяцев каждого года основаны на немногочисленных данных и, как ожидается, будут волатильными. Кроме того, декабрь выступает в качестве краткосрочного базисного периода индекса, и ему придается особое значение. Если по какой-то причине декабрь является необычным месяцем, это может отрицательно сказаться на результатах. Чтобы смягчить эти проблемы, Lamborau (2017) предложил сочетать метод прямого расширения с годовой цепной увязкой с подходом скользящего окна.

Таблица 10.2 Увязка методом сращивания движения цен со скользящим окном в 13 месяцев

Период	1	2	3	4	5	...	11	12	13	14	15
Первый раунд расчета в период 13	100	100.7	100.6	101.6	102.7	...	104.3	106	103.8		
Второй раунд расчета в период 14		100	100.2	101.1	102.2	...	103.8	105.5	103.3	104.6	
Третий раунд расчета в период 15			100	101	102	...	103.5	105.3	103.2	104.4	104.1
Опубликованный индекс (соединение изменения цен)	100	100.5	100.6	101.6	102.7	...	104.3	106	103.8	105.1	104.8

Соединение начинается в периоде 14 (выделено жирным шрифтом). Публикуемые индексы за периоды 1-13 получаются на первом раунде расчета. Публикуемый индекс за 14-й период получается путем применения изменения между индексами за 13-й и 14-й периоды второго раунда расчета к опубликованному индексу за 13-й период ($103,8 \times 104,6/103,3 = 105,1$). Публикуемый индекс за 15-й период получается путем применения изменения между индексами за 14-й и 15-й периоды третьего раунда расчета к опубликованному индексу за 14-й период ($105,1 \times 104,1/104,4 = 104,8$).

Внедрение многосторонних методов

Оценка многосторонних методов

10.106 Внедрение новых источников данных и методов в любой статистический ряд требует тщательного рассмотрения статистического воздействия, а также выгод и затрат. Лишь несколько НСС внедрили многосторонние индексы цен в ИПЦ.

10.107 Предлагаемые критерии для оценки многосторонних методов должны учитывать понятие качества статистических продуктов в широком смысле. Система может включать семь аспектов качества:

- Институциональная среда — институциональная и организационная ситуация, в которой работает производитель статистических данных.
- Актуальность — насколько хорошо статистика соответствует потребностям пользователей.
- Своевременность — насколько быстро и часто публикуется статистика.
- Точность — насколько хорошо статистика измеряет желаемое понятие.
- Согласованность — насколько статистические данные соответствуют источникам соответствующей информации.
- Интерпретируемость — имеющаяся информация, обеспечивающая понимание статистики.
- Доступность — касается легкости доступа к статистике.

10.108 Следует отметить, что многосторонние методы более сложны, чем стандартные двусторонние индексы, и создают проблемы с коммуникацией для НСС. Большое значение следует придавать прозрачности для объяснения опубликованных статистических данных, а также описанию и обоснованию используемых методов. Это имеет решающее значение для обеспечения доверия к опубликованному ИПЦ. Необходимо рассмотреть два аспекта интерпретируемости: во-первых, насколько легко понять сами методы составителям индексов и пользователям; и, во-вторых, легко ли понять движения цен, которые отражает каждый индекс, особенно то, какие продукты оказывают наибольшее влияние на эти движения и

почему.

10.109 Этот подход может применяться НСС для определения преимуществ и проблем, связанных с использованием методов расчета многосторонних индексов с учетом обстоятельств конкретной страны.

10.110 Методы многосторонних индексов для составления ИПЦ также можно оценить с теоретической точки зрения. В оценке могут использоваться подходы, ранее применявшиеся к двусторонним и пространственным индексам. Двусторонние индексы цен оцениваются как с использованием аксиоматических/тестовых подходов (глава 3 *Теории индексов потребительских цен*), так и экономических подходов (главы 4 и 5 *Теории индексов потребительских цен*). Подобные подходы к оценке многосторонних индексов в пространственном контексте были разработаны и представлены в нескольких работах, особенно Diewert (1999b) и Balk (2001).

10.111 Имеется описание теоретических оценок методов многосторонних индексов цен в текущем временном контексте с использованием аксиоматических/тестовых подходов и экономической теории (ABS 2016a; Zhang and others 2019). Эта оценка может быть использована НСС в качестве основы для проведения аналогичных оценок для своей местной ситуации. Всестороннее обсуждение различных многосторонних методов, использующих экономический подход к теории индексов, можно найти в главе 7 *Теории индексов потребительских цен*; см. также Diewert and Fox (2017). Самый важный результат заключается в том, что в GEKS правильно рассматриваются эффекты замещения, тогда как GK и TPD подходят только при ограничительных предположениях о потребительских предпочтениях. Однако на практике GK и TPD дают очень похожие результаты. Одно НСС разработало общий метод обработки на основе GK (см. Chessa 2016). На первом этапе этот метод был реализован только для мобильных телефонов, а в последующие годы был применен к большему количеству продуктов и розничных продавцов.

10.112 Кроме того, экспертная оценка предлагаемых многосторонних методов может быть уместной в обстоятельствах, когда пользователи ИПЦ могут ожидать, что НСС продемонстрируют наличие более широкой поддержки предлагаемых изменений.

Расчет индексов

Оперативный выбор

10.113 Свойство (негедонических) многосторонних индексов для сравнимых моделей подразумевает, что без какого-либо ручного вмешательства результаты зависят от *выбора идентификатора товара (услуги)-представителя*. Например, при использовании штрих-кода в качестве идентификатора изменение цены однородного товара, штрих-код которого изменяется одновременно — «повторный запуск» — не будет измеряться. Как упоминалось ранее, использование кодов SKU смягчает проблему, поскольку SKU обычно состоит из нескольких штрих-кодов для аналогичных товаров и является более стабильным, чем штрих-код. Тем не менее, даже код SKU может быть слишком подробным.

10.114 Если относительно небольшого количества наблюдаемых атрибутов с дискретными значениями достаточно для определения однородных товаров, товары можно определить путем перекрестной классификации наборов категориальных переменных для каждого атрибута и цен, рассчитанных как стоимость единицы по всем штрих-кодам/SKU. Скорее всего, в течение периода выборки все еще будут новые и исчезающие товары. Чтобы максимизировать степень соответствия без внесения цепного смещения, можно применить многосторонний метод (Chessa 2016). Потенциальная проблема заключается в том, что доступная информация о характеристиках может быть ограничена, особенно когда характеристики извлекаются из описаний разновидностей товаров в данных сканирования, которые часто являются довольно широкими. В этом случае, вероятно, возникнет систематическая ошибка в стоимости единицы товара. Кроме того, если информация о характеристиках считается достаточной, может быть лучше построить гедонистические индексы.

10.115 Хотя построение выборки (методом отсека), где игнорируются товары-представители, пусть даже определенные, с небольшими долями расходов во многих случаях не оказало бы существенного влияния на результаты, в этом нет необходимости при использовании взвешенного многостороннего метода. Большинство вопросов, обсуждавшихся ранее в этой главе, таких как выбор расчета стоимости единиц товара на уровне магазина или на уровне сети, а также необходимость иметь структуру индекса,

облегчающую использование данных сканирования, применимы и здесь.

Эмпирические результаты

10.116 Цель получения эмпирических результатов многосторонних методов двояка: изучить эффективность различных методов в местных условиях, а также продемонстрировать пользователям ИПЦ вероятные последствия перехода от текущих источников данных и методов для расчета ИПЦ к новым подходам. В идеале эти многосторонние методы следует сравнивать друг с другом, а также с официальным ИПЦ. Эти сравнения следует проводить на самом низком уровне агрегирования для опубликованного ИПЦ и на различных уровнях агрегирования, включая ИПЦ по всем товарам (услугам)-представителям. Некоторые эмпирические данные о многосторонних методах составления индексов цен см. в главе 5 ABS (2016a), главе 3 ABS (2017) и Chessa and others (2017).

10.117 На основании эмпирических результатов можно сделать несколько выводов. Часто эти выводы еще лучше подкрепляют теоретические аргументы в пользу использования многостороннего методов составления индекса для ИПЦ. Это может включать в себя влияние использования актуальной информации для взвешивания, которая отражает поведение потребителей, включая замещение с течением времени. Эмпирические результаты должны быть доведены до сведения пользователей ИПЦ и заинтересованных сторон.

Общение с пользователями и заинтересованными сторонами

10.118 Использование данных сканирования для составления ИПЦ потенциально представляет собой весьма существенное изменение в источниках данных и методах, используемых НСС. Эти изменения должны быть тщательно доведены до сведения пользователей ИПЦ и заинтересованных сторон. Предлагаемый комплекс мероприятий включает в себя:

- Публикация информационных документов, в которых описываются предлагаемые новые методы и источники данных.
- Проведение личных встреч с ключевыми заинтересованными сторонами (например, центральными банками, казначейством и министерствами финансов) и другими заинтересованными сторонами, включая представителей общественности.
- Использование пресс-релизов и брифингов экономических журналистов для информирования общественности о предлагаемых изменениях.
- Поощрение заинтересованных сторон и общественности к предоставлению материалов для рассмотрения в НСС.
- Взаимодействие с ведущими учеными как для рассмотрения предлагаемых изменений, так и для получения их поддержки.

10.119 После этих консультаций, которые могут занять пару лет, НСС следует опубликовать документ с изложением позиции, в котором будут даны ответы на вопросы, поднятые в ходе консультаций, и описано, как НСС будет использовать данные сканирования для составления ИПЦ, включая обоснование и эмпирические результаты, подтверждающие этот подход. В документе с изложением позиции должны быть четко указаны источники данных и используемые методы, а также график внедрения изменений.

Публикация и распространение

10.120 После публикации документа с изложением позиции НСС предлагается составлять ИПЦ с использованием как текущих, так и новых источников данных и методов параллельно в течение примерно шести месяцев. Этот переходный период позволит НСС доработать процессы и процедуры для составления ИПЦ с использованием новых методов, а также сравнить эмпирические результаты двух подходов. Этот переходный период часто является первой возможностью для НСС использовать новые источники данных и методы в режиме реального времени в соответствии с графиком обработки и публикации ИПЦ. Обнародование результатов параллельной обработки остается на усмотрение НСС.

10.121 Первый период, в течение которого новые источники данных и методы будут использоваться для составления ИПЦ, должен быть объявлен заблаговременно и должен включать подробные метаданные для СМИ и других ключевых пользователей данных. Это обеспечит правильное понимание методологических изменений, внесенных в ИПЦ. Дополнительную информацию, касающуюся распространения ИПЦ, можно найти в главе 14 настоящего Руководства.