

## Валидация региональных цен ПМС и ППС на уровне первичных групп

ДЭВИД РОБЕРТС

Национальные координаторы (НКА) стран, участвующих в региональных сопоставлениях ПМС, сообщают региональным координаторам (РКА) значения национальных среднегодовых цен на товары-представители, выбранные из общего списка точно определенных товаров. Эти цены относятся к году проведения сопоставления и охватывают целый ряд товаров и услуг для конечного потребления, включенных в ВВП. Они используются для расчета паритетов покупательной способности (ППС) на уровне первичных групп, а те в свою очередь - для получения внутрирегиональных оценок отношений цен и физических объемов для стран, участвующих в сопоставлении. Эти оценки затем публикуются региональным координатором и становятся доступными для разных пользователей, включая стратегов, экономистов, исследователей, политиков, журналистов, а также широкую общественность. Постольку оценки ППС должны служить обоснованными аргументами в дискуссии, то важно, чтобы данные о ценах, на основе которых рассчитываются ППС, были тщательно проверены, а ошибки исправлены, то есть проведена валидация данных до расчета ППС. Таким образом, валидация региональных данных является приоритетной задачей при проведении региональных сопоставлений. В данной главе рассмотрено, как в рамках ПМС происходит валидация региональных данных о ценах. Такие же процедуры используются для проведения валидации данных о ценах, собранных в разных регионах, для получения результатов глобального сопоставления.

Здесь рассматриваются два типа ошибок: ошибки, связанные и не связанные с выборкой. Ошибка выборки возникает в связи с тем, что цены для расчета ППС регистрируют в отобранных, а не во всех, торговых точках. Ошибка, не связанная с выборкой, возникает в связи со сбором цен не на те товары, а также с неправильной регистрацией цен товаров-представителей и/или единиц измерения. Важное различие между двумя типами ошибок состоит в том, что ошибка выборки исчезла бы, если бы цены регистрировались во всех торговых точках, а ошибка, не связанная с выборкой, - нет; она все равно продолжала бы

существовать. Ошибку выборки контролируют до начала обследования цен при помощи дизайна выборки, а ошибку, не связанную с выборкой, - до и во время обследования при помощи хорошего плана обследования и эффективного управления, а после проведения обследования – при помощи процедуры валидации данных. Процедура валидации данных в ПМС направлена на устранение ошибок, не связанных с выборкой, а не ошибок выборки. Цель состоит в том, чтобы минимизировать, если не устранить, ошибки, не связанные с выборкой, в региональных данных после проведения обследования цен. Эта цель достигается при помощи редактирования и проверки. Редактирование состоит в рассмотрении данных о ценах на наличие возможных ошибок. Проверка состоит в подтверждении того, что значения цен, указанные в качестве возможных ошибок, являются правильными, или в исправлении их в противном случае.

Валидация данных - это итеративный процесс, требующий проведения нескольких раундов редактирования и проверки. Возможные ошибки находят путем выявления значений цен, которые имеют отклонения, превышающие заданную критическую величину, либо значений, выходящих за границы заданного диапазона приемлемых величин. Оценки отклонений обычно определяются параметрами редактируемых рядов данных, таких как средние величины и стандартное отклонение. Следовательно, если некоторые из возможных ошибок, выявленных при первоначальном редактировании, окажутся фактическими ошибками и будут исправлены, то параметры рядов ценовых данных изменятся, также как показатели отклонений для каждого значения цены, оставшегося в ряду. Второе редактирование позволит обнаружить новые возможные ошибки, которые должны быть проверены. Опять-таки, когда фактические ошибки будут исправлены, параметры рядов ценовых данных изменятся, что может привести к обнаружению новых возможных ошибок, если будет проведен третий раунд редактирования. Обычно число новых возможных ошибок снижается по мере осуществления процедуры валидации данных, которая проводится до тех пор, пока улучшения результата при следующих раундах не будет считаться незначительным, а дальнейшие действия - нецелесообразными. Время также является важным фактором: во-первых, потому, что данные необходимо выпускать своевременно, а во-вторых, потому, что чем больше времени проходит между сбором и проверкой данных, тем сложнее становится исправлять неверные данные о ценах.

Процесс валидации данных ПМС включает два отдельных процесса: первый состоит в редактировании и проверке данных, собранных страной, и называется *внутристрановой* валидацией данных; а другой – в редактировании и подтверждении данных, собранных всеми странами, и называется *межстрановой* валидацией данных. Внутристрановая валидация данных касается наблюдений за ценами *в стране* и средних цен, полученных на их основе. Цель состоит в проверке того, что сборщики цен в стране провели наблюдение цен на сопоставимые товары и правильно зарегистрировали их. Межстрановая валидация данных касается средних цен, сообщенных странами, участвующими в программе сопоставлений, и отношений средних цен между странами. Цель процесса состоит в том, чтобы проверить, что сборщики цен *в разных странах* собрали данные о ценах на сопоставимые товары и сделали это правильно. Межстрановая валидация проводится после внутристрановой валидации и осуществляется совместно РКА и НКА. В этой главе описываются оба эти процесса.

Прежде чем перейти к рассмотрению процессов валидации, следует заметить, что лучше предотвратить ошибки, чем исправлять их. Величина ошибки, не связанной с выборкой, может быть существенно уменьшена при помощи хорошего дизайна обследования и управления обследованием. Сбор цен должен быть хорошо спланирован, выполняться эффективно, и за ним следует обеспечить надлежащий надзор; спецификации товаров должны быть подробными, чтобы сборщики цен могли четко определить товары в посещаемых торговых точках; сборщики цен должны пройти обучение и получить четкие инструкции; выдаваемые им формы для заполнения данных должны быть ясными; следует проводить мониторинг полевой

работы, чтобы убедиться в том, что сборщики регистрируют цены, количества и другие необходимые данные; сотрудники статистических служб, занятые обработкой и валидацией ценовой информации, должны быть соответствующим образом обучены, и их необходимо контролировать. Процесс валидации является дополнением к хорошей практике проведения обследования. Результаты некачественного обследования нельзя улучшить посредством редактирования и проверки, как бы тщательно они ни проводились.

## Возможные ошибки, ошибки и выбросы

Для того чтобы сопоставление цен было действительным, товары, цены которых сравниваются, должны быть сопоставимыми, то есть одинаковыми или эквивалентными, а сами данные о ценах не должны содержать ошибок и искажений. Невыполнение хотя бы одного из этих требований сводит на нет сопоставления. Для того чтобы этого не произошло, при редактировании и проверке данных, выполняемых в рамках региональных сопоставлений ПМС, обращают внимание на основные виды ошибок, не связанных с выборкой: ошибку товара и ценовую ошибку.

*Ошибка товара* возникает в том случае, когда сборщики цен регистрируют цены на товары, которые не соответствуют спецификациям, а затем не сообщают об этом. Возможно, что они и не знают о несоответствии, например, когда спецификации слишком неопределенны<sup>1</sup>, или же они регистрируют цены на товар-заменитель в соответствии с инструкцией, но не сообщают об этом в форме для сбора цен. Сборщики цен обычно проинструктированы регистрировать цену аналогичного товара, если они не способны найти товар с нужными спецификациями. В соответствии с инструкциями они должны отметить товар-заменитель в форме и описать различия между заменителем и первоначально указанным товаром. Это позволяет национальным координаторам обратить внимание на замену и затем совместно с региональным координатором решить, как использовать информацию о цене. Можно корректировать цену на различия в качестве между фактически оцененным товаром и указанным товаром. Или же, если другие страны также сообщают цены тех же самых товаров-заменителей<sup>2</sup>, то можно провести сопоставление для них наряду с товарами, описанными первоначально. Если ни один из этих подходов не работает, то эту ценовую котировку следует отбросить. Замена сама по себе не вызывает ошибку. Ошибка товара возникает, когда сборщики цен не отмечают и не описывают произведенную замену<sup>3</sup>.

*Ценовая ошибка* происходит, когда сборщики цен регистрируют цены на товары, соответствующие спецификациям, но делают это неправильно, или же, когда они правильно регистрируют цену, а ошибка возникает позднее в процессе сообщения и передачи данных. Каждой цене соответствует количество: указанное количество (для сбора цены) и базисное количество (к которому должна быть приведена собранная информация о цене). Ценовая ошибка может возникнуть также в том случае, если цена была зарегистрирована правильно, но оцененное количество было записано неверно (или оно было записано правильно, а ошибка возникла позднее в процессе обработки), так что приведенная цена базисного количества – та цена, которая будет участвовать в процессе валидации, – будет неверной<sup>4</sup>.

Редактирование данных в целях устранения ошибок товара и ценовых ошибок включает выявление экстремальных значений цен, то есть цен, значения которых слишком высоки или слишком низки по отношению к средним в соответствии с выбранными критериями. Может оказаться, что цена превышает установленное заранее критическое значение или не попадает в некоторый предопределенный диапазон приемлемых значений. Это стандартные пути обнаружения ошибок в данных обследования, и оба эти метода используются в ПМС. Экстремальные значения цен не обязательно являются неверными. Но тот факт, что эти значения считаются экстремальными, говорит о том, что они представляют собой возможные

ошибки и их необходимо исследовать. В ПМС не принято немедленно отклонять цены с экстремальными значениями, здесь сначала выясняют, являются ли они результатами подлинных наблюдений. Затем можно решать, как с ними поступать. Экстремальные значения цен, оказавшиеся неверными, являются ошибками и должны быть исправлены или отброшены. Экстремальные значения цен, являющиеся точными наблюдениями, представляют собой «выбросы» (outliers)<sup>5</sup> и должны быть оставлены, если они часть совокупности, определенной остальными наблюдениями цен. На практике «выбросы», которые удовлетворяют данному критерию, часто «исправляют», то есть отбрасывают или заменяют на импутированные значения для устранения «шума», который они привносят в данные.

## «Инлаеры» и смещения

Так же как не все цены, контрольные значения для которых превышают критические или выпадают из установленного диапазона приемлемых значений, обязательно являются ошибками, так и не все цены, контрольные значения для которых ниже критических или принадлежат диапазону приемлемых значений, обязательно свободны от ошибок товара и/или ценовых ошибок. Эти значения цен, которые иногда называются «инлаерами», не выбираются при проверке на экстремальные значения, во всяком случае, вначале. Однако поскольку валидация данных является итеративным процессом, они могут быть выявлены на следующих стадиях. Другой и, возможно, более надежный путь нахождения «инлаеров» в ходе валидации данных внутри страны состоит - в дополнение к редактированию всех ценовых данных вместе – в группировке их по типам торговых точек и местоположению и редактировании каждой группы отдельно. В таком случае могут возникнуть экстремальные значения и возможные «инлаеры», которые не считались экстремальными при редактировании всех ценовых данных вместе.

Редактирование данных по типам торговых точек и местоположению не подходит межстрановой валидации данных, поскольку не существует международных принятых определений типов торговых точек или городской и сельской местности. Более того, не все страны собирают цены в сельской местности, вместо этого они используют пространственные коэффициенты для приведения цен городской местности к национальным ценам. Но даже при этом редактирование цен в группировке по типам торговых точек и местоположению в ходе внутристрановой валидации данных может помочь процессу межстрановой валидации, поскольку позволит проанализировать состав и местоположение торговых точек, где проводилось обследование цен товаров-представителей. Это особенно важно для тех товаров, для которых в процессе внутристрановой валидации данных были выявлены экстремальные средние цены. Эти средние цены могли оказаться экстремальными не из-за выбросов или не выявленных «инлаеров» в ценовых данных, а в результате смещения, вызванного тем, что цены наблюдались на выборке торговых точек и местоположений, которые не соответствуют профилю распределения.<sup>6</sup>

Предполагается, что выборка торговых точек для обследования цен отражает модели покупательского поведения потребителей для оцениваемых товаров в разных типах торговых точек. Например, если потребители покупают 60% одежды и обуви в супермаркетах, 20% в универмагах и 20% в специализированных магазинах, то выборка из 10 торговых точек для обследования цен на одежду и обувь включит шесть супермаркетов, два универмага и два специализированных магазина. А если 70% этих покупок было сделано в городской местности и 30% в сельской местности, то семь из выбранных торговых точек будут в городе, а три – в сельской местности. Осуществление выборки торговых точек таким образом привносит неявные веса для учета разных элементов обслуживания в торговых точках и их влияния на цену.<sup>7</sup> Результатом будут несмещенные значения средних цен.<sup>8</sup>

В принципе национальные координаторы пытаются выбирать торговые точки в соответствии с профилем распределения обследуемых товаров, насколько они известны. Процесс отбора не является специфическим для каждого вида обследуемых товаров, а только для группы товаров. Следовательно, хотя выборка и будет отражать профиль распределения данной группы, она не обязательно будет отражать профили распределения всех товаров в данной группе, и средняя цена тех товаров, для которых состав торговых точек неверный, будет либо завышенной, либо заниженной. Например, если имеется следующее распределение покупок нижнего белья: 80% супермаркеты, 10% - универмаги и 10% специализированные магазины, то структура выборки 60/20/20, описанная выше для одежды и обуви, приведет к систематическому завышению средних цен нижнего белья, поскольку составляющая обслуживания в универмагах и специализированных магазинах обычно выше, чем в супермаркетах, и их цены соответственно выше.

Эту проблему нельзя решить, поскольку ресурсы для проведения обследования ограничены. Выборка торговых точек, подходящая для широкой группы товаров, неизбежно не будет подходить для некоторых видов товаров, включенных в группу. Просто невозможно построить выборку торговых точек, которая будет проходящей для всех и каждого обследуемого товара. Эту проблему следует решать после сбора цен. Внутристрановая валидация средних цен по типу торговых точек и местоположению может помочь определить товары, для которых состав торговых точек непредставителен. В частности, это может быть применено для определения товаров, средние цены которых смещены в результате преобладания одного конкретного вида торговых точек. Средние цены таких товаров могут быть скорректированы путем подавления наблюдений цен из тех торговых точек, которые представлены избыточно, или дублирования наблюдений цен из тех точек, которые представлены недостаточно.

Выбор торговых точек – будь это рандомизированный или целенаправленный – может привести к отбору таких точек, которые сами представляют собой выбросы и не представительны для покупательской модели среднего потребителя. Если отбор торговых точек рандомизированный, то такие выбросы следует оставить, так как отбрасывание их подорвет теоретическое обоснование случайного отбора. Но если отбор был целенаправленным, то ответ не столь очевиден. Цель целенаправленного отбора состоит в том, чтобы выбрать множество торговых точек, которые бы были репрезентативными для тех, которые домохозяйства используют большую часть времени. Если цены в выбранной торговой точке не соответствуют уровням цен в других торговых точках того же типа, будучи постоянно выше или ниже, то эта торговая точка является выбросом и должна быть заменена на ту, где цены ближе к средним. Внутристрановая валидация средних цен по типам торговых точек и местоположению помогает выявить выбивающиеся торговые точки.

## Процесс валидации

Валидация включает два отдельных процесса: внутристрановую и межстрановую валидацию. Внутристрановая валидация предшествует межстрановой. Цель состоит в проверке того, что сборщики цен в стране провели наблюдение цен на товары, которые соответствуют спецификациям, и что цены, которые они сообщили, правильные. Те, кто проводит внутристрановое редактирование данных, сначала стараются выявить экстремальные значения среди индивидуальных цен, которые были собраны в стране для каждого обследуемого товара, а затем - экстремальные значения среди средних цен на эти товары. За редактирование данных и их последующую проверку отвечает НКЛ. Этот процесс проводится без учета ценовых данных других стран.<sup>9</sup> После завершения процесса внутристрановой валидации данных национальный координатор передает региональному координатору следующую информацию: средние цены на обследованные товары, плюс коэффициент вариации, отношение максимального и

минимального значений, и количество индивидуальных наблюдений цен для каждого сообщенного значения средней цены. Эта информация рассматривается РКА до начала процесса межстрановой валидации данных.

Цель процесса межстрановой валидации данных состоит в том, чтобы проверить, что сборщики цен в разных странах собрали данные о ценах на сопоставимые товары – иными словами, что они все проинтерпретировали спецификации товаров одинаково – и что цены, которые они собрали, правильные. Те, кто проводит межстрановое редактирование данных, смотрят на экстремальные значения среди средних цен, которые НКА региона передали в РКА, на те же самые продукты внутри первичных групп. При этом они переводят средние цены, выраженные в национальной валюте, в общую валюту. Для этого используются как обменные курсы валют, так и ППС, как будет объяснено ниже. После конвертирования средние цены каждого НКА сравниваются со средними ценами других НКА региона. Эту задачу невозможно выполнить эффективно без инициативы и активного участия РКА и НКА, согласившихся обмениваться данными о своих средних ценах друг с другом.

Внутристрановое редактирование обычно включает два раунда редактирования и проверки и занимает около двух месяцев или больше. Внутристрановая валидация данных требует проведения в среднем четырех раундов редактирования и проверки и около четырех месяцев для завершения. Межстрановая валидация занимает больше времени в связи с тем взаимодействием, которое возникает в процессе валидации между наборами данных из разных стран. Изменения, внесенные одним НКА, могут привести к изменению результатов редактирования цен других НКА. Работа с такими изменениями может занимать много времени, потому что не все страны, участвующие в сопоставлениях, бывают охвачены на ранних этапах валидации, они включаются в последующие раунды, когда информация об их средних ценах становится доступной. По этой причине, межстрановая валидация может занять больше четырех месяцев.<sup>10</sup> Процесс не является завершенным до тех пор, пока все страны региона не будут включены в межстрановые диагностические таблицы, а их НКА не утвердят официально свои пересмотренные данные о ценах.

Важной чертой процесса межстрановой валидации является его итеративная природа; при этом межстрановые диагностические таблицы передаются из РКА в НКА региона и обратно. Процесс начинается с того, что НКА посылает свои данные о средних ценах в РКА. РКА составляет диагностические таблицы, изучает их, выявляет значения цен, которые кажутся неправдоподобными, и посылает таблицы с вопросами в НКА. После того, как НКА исследует вопросы РКА, НКА пересматривает неправильные цены и посылает их обратно в РКА. После внесения корректировок в базу данных РКА пересчитывает диагностические таблицы, рассматривает их и посылает их в НКА с дальнейшими вопросами по поводу надежности конкретных цен. Процесс продолжается до тех пор, пока и РКА и НКА не посчитают пересмотренные данные окончательными. Диагностические таблицы для межстрановой валидации данных называются таблицами Куаранта или таблицами Дикханова; они будут объяснены и сопоставлены в приложении к данной главе.

## **Внутристрановая валидация**

Центральное место в процессе внутристрановой валидации занимает модуль валидации и расчета средних значений. В рамках данного модуля строятся две диагностические таблицы для использования НКА при редактировании наблюдений индивидуальных цен и средних цен, полученных на основе индивидуальных. Эти диагностические таблицы - таблица наблюдений цен и таблица средних цен (см. табл. 9.1. и 9.2). Таблица наблюдений цен строится для каждого обследованного товара внутри первичной группы. Столбцы отражают индивидуальные зарегистрированные цены, другие характеристики товара и результаты диагностических тестов, имеющих целью выявить возможное ошибки. Таблица средних цен содержит цены и

диагностические характеристики для каждого товара в первичной группе. НКА рассматривает обе таблицы, однако РКА видит только таблицу средних цен. В таблице наблюдений цен определяются экстремальные значения *индивидуальных цен*, собранных для конкретного товара, тогда как в таблице средних цен отмечаются экстремальные значения *средних цен* на обследованные товары. Общими для обеих таблиц являются поля 01-04, которые определяют товар и указывают на базисное количество (количество к которому должны быть приведены наблюдения цен, в том случае, если спецификации товара предполагают использование диапазона количества для оценки). В процессе валидации участвуют данные об индивидуальных и средних ценах на базисное количество товара.<sup>11</sup> Первый шаг процесса валидации состоит во вводе индивидуальных цен и другой информации в модуль (поля 05-15 таблицы 9.1.а). Ввод данных следует проводить как можно раньше после сбора данных, так чтобы сборщики цен и персонал торговых точек еще бы хорошо помнили обстоятельства, имевшие место во время сбора цен. В идеальном случае индивидуальные цены следует вводить на следующий день после сбора данных, что позволит супервайзерам выявить и исправить ошибки сборщиков цен в самом начале процесса.

Второй шаг – после ввода описания каждого наблюдения цен – состоит в проверке ввода данных, чтобы убедиться, что коды товаров были введены правильно и что коды и соответствующие наблюдения цен были введены в правильной последовательности. Ввод в числовом выражении осуществляется только для базисного количества, наблюдаемого количества, и цены. Наблюдаемое количество должно быть в тех же единицах измерения, что и базисное количество. Показатель важности<sup>12</sup> должен быть введен для всех наблюдений цен на важные товары.

Внутристрановое редактирование данных проводится в два этапа. Первый состоит в выявлении экстремальных значений среди индивидуальных цен, занесенных в таблицу наблюдений цен. Второй этап включает выявление экстремальных значений среди средних цен товаров в таблице средних цен. Экстремальное значение определяется как значение индивидуальной цены или средней цены, которая находится вне определенного критического значения для заданного теста.

Все диагностические тесты, основанные на теории статистики, включают измерение вариации. Если значения цен распределены нормально, то они сконцентрированы вокруг среднего значения, а стандартное отклонение является показателем среднего расстояния от него. Девяносто пять процентов наблюдений будет лежать в интервале плюс-минус два значения стандартных отклонения, а девяносто девять процентов – в интервале плюс-минус три значения стандартных отклонения. Во многих случаях, однако, распределение значений цен будет иметь форму прямоугольно треугольника (в основном, низкие цены со снижающимся количеством более высоких цен). Среднее значение и стандартное отклонение являются функциями от минимального и максимального значений. Диагностические характеристики, используемые для валидации данных о ценах, учитывают оба вида распределений:

- Экстремальные значения для наблюдений цен в таблице наблюдений цен определяются путем использования двух тестов: показатель отношения к средней цене и Т-критерия.

*Отношение к средней цене.* Отношение к средней цене – это отношение цены базисного количества наблюдаемого товара к средней цене базисного количества товара. В соответствии с критерием данного теста, это отношение должно находиться в диапазоне 0,5 – 1,5. Другими словами, ожидается, что индивидуальная цена будет не ниже половины средней цены и не выше ее удвоенной величины. Зарегистрированная цена, для которой это отношение находится вне этого диапазона, будет отмечена в таблице наблюдений цен как имеющая экстремальное значение, которое должно быть проверено. Это простой первый тест, так как он не использует оценку стандартного отклонения, на которое также оказывает влияние экстремальное значение.

Выбор диапазона 0,5 – 1,5 в качестве критического значения основан на теории статистики и практическом опыте предыдущего сбора данных. Ожидается, что подробные спецификации товаров должны обеспечить сбор цен, значения которых отличаются от среднего не более чем на половину в большую или меньшую сторону, если был оценен правильный товар.

*T-критерий.* Отношение отклонения цены базисного количества товара для наблюдения цены от средней цены базисного количества товара к стандартному отклонению для товара. Для прохождения этого теста, это отношение должно быть меньше или равно 2.0 (при величине больше 2.0 наблюдение подозрительно, так как выпадает из 95% доверительного интервала). Наблюдение цены со значением T-критерия выше 2.0 не проходит тест и должно быть отмечено в таблице наблюдений цен, как имеющее экстремальное значение, требующее исследования.

- Экстремальные значения средних цен, приведенных в таблице средних цен, также определяются по двум тестам: отношение максимального и минимального значения и коэффициент вариации.

*Отношение максимального и минимального значений.* Отношение максимального и минимального значений – это отношение максимальной цены базисного количества, зарегистрированной для товара, к минимальной цене базисного количества, зарегистрированной для товара. Средняя цена с отношением выше 2.0 не проходит тест и будет отмечена в таблице средних цен как имеющая экстремальное значение, которое следует проверить. Отношение, равное 2.0, предполагает коэффициент вариации (стандартное отклонение/среднее значение), равный 10-15%, или 95% доверительный интервал, в 20-30%.

*Коэффициент вариации.* Коэффициент вариации – это стандартное отклонение цены товара, выраженное как процент от средней цены товара. Для того чтобы удовлетворить тесту, коэффициент вариации должен быть меньше или равен 20%. Средняя цена с коэффициентом вариации выше 20% не удовлетворяет тесту и будет отмечена в таблице средних цен как имеющая экстремальное значение, которое требует проверки.

Как и в случае теста для отношения к средней цене, критические значения основаны на ожиданиях, возникающих от точности спецификаций товара. Эти два теста дают аналогичные результаты. Тест с использованием отношения максимального и минимального значений наиболее полезен в тех случаях, когда наблюдение цены существенно отличается от остальных наблюдений. Когда коэффициент вариации высок, это может указывать на то, что товар был нечетко определен.

Модуль валидации и расчета средних значений запрограммирован на отметку наблюдений цен и средних цен, для которых тестовые характеристики находятся вне критического диапазона. Следует установить надежность данных для отмеченных наблюдений цен и средних цен. Или, если быть более точным, то НКА должны исследовать надежность отмеченных наблюдений цен и тех данных, на основе которых были рассчитаны отмеченные значения средних цен.<sup>13</sup>

Сначала НКА следует проверить, были ли отмеченные наблюдения цены правильно введены в модуль. Другими словами, следует проверить, что цены в таблице наблюдений цен те же, что в формах регистрации цен. Если это не так, то цена в таблице должна быть исправлена, а изменение отмечено в поле 20. Если они совпадают, то цены были введены правильно. В этом случае НКА следует повторно посетить торговую точку, где были собраны цены, для того, чтобы выяснить, соответствуют ли оцененные товары спецификациям и правильно ли были сообщены цены. Если товар соответствует спецификациям и была сообщена правильная цена, то проверка завершается. Установлено, что экстремальное значение является точным

наблюдением и выбросом – этот результат должен быть отражен в поле 19 таблицы. Если оцененный товар не соответствует спецификациям товара, или цена была сообщена неправильно, то ситуация должна быть исправлена путем поиска товара в данной торговой точке, который соответствует спецификациям, и регистрации цены на него или установления правильной цены на товар, оцененный первоначально, если он еще имеется в продаже.

НКА может выполнять следующие действия:

- Наблюдения цен, которые были отмечены и признаны неверными, либо подавляются, либо заменяются правильными наблюдениями.
- Наблюдения цен, которые отмечены в связи с тем, что они не соответствуют тесту на отношение к средней цене, но не тесту на T-статистику, считаются правильными и представляют собой выбросы. Выброс должен быть сохранен, если это правильное наблюдение и часть совокупности, определенной другими наблюдениями цен. Ситуация может быть исправлена путем пересчета средней цены и стандартного отклонения без включения выброса и использования их для расчета T-статистики для выброса. Если значение T-статистики теперь выше 2.0, то выброс, хотя это и точное наблюдение, неприемлем и должен быть подавлен. Если значение T-статистики все еще не превышает критического значения, то наблюдение является приемлемым и должно быть оставлено, во всяком случае, первоначально. Позже, в ходе межстрановой валидации, можно будет решить, следует ли его подавить или заменить импутированным значением.
- Наблюдения цен, которые отмечены в связи с тем, что не удовлетворяют тесту на T-статистику, но оказываются правильными, также представляют собой выбросы. Несмотря на это, они должны быть подавлены, поскольку явно не являются частью той же совокупности, что другие наблюдения цен, даже когда они включены в расчет средней цены и стандартного отклонения.

Во всех случаях подавление и исправления наблюдений цен должны быть отмечены в поле 20 таблицы наблюдений цен.

Таблица 9.1.a показывает, как выглядит таблица наблюдений цен сразу после ее построения и после первого рассмотрения НКА: экстремальные значения выделены жирным шрифтом, комментарии НКА помещены в поле 19. Показано, что все наблюдения цен для риса А не прошли тест на отношение к средней цене, хотя только одно (поле 16, столбец 8) отмечено, как не прошедшее тест на T-статистику. Похоже, что все наблюдения цен для риса А требуют проверки, но так ли это? Более внимательное изучение показывает, что только одно наблюдение требует исследования на начальном этапе, это цена риса, равная 499.00, с высоким значением T-статистики. Эта цена в десять раз выше, чем в других наблюдениях, и оказывает доминирующее воздействие на среднюю цену и на стандартное отклонение, все это объясняет плохие показатели для других наблюдений по тесту на отношение к средней цене.

Таблица 9.1.b показывает таблицу наблюдений цен после проверки. При исследовании, оказалось, что 499.00 – ценовая ошибка. Цена была внесена неправильно во время ввода данных. Она была заменена правильной ценой, что было отмечено в поле 20. В результате, наблюдение соответствует тесту на T-статистику, а все остальные наблюдения соответствуют тесту на отношение к средней цене.

**Таблица 9.1а** Процесс валидации: Таблица наблюдений цен после первого рассмотрения НКА и до проверки

Область		После первого обзора и до проверки								
01	Код продукта	1101111.0111								
02	Сокращенное описание продукта	рис А, пакет 400–600 грамм								
03	Исходная величина	500								
04	Единица исходной величины	грамм								
05	Дата наблюдения	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год
06	Идентификатор сбора цен	A	A	B	B	C	C	C	D	D
07	Идентификатор торговой точки	001	009	015	019	025	036	037	048	051
08	Расположение торговой точки	сельское	сельское	сельское	сельское	городское	городское	городское	городское	городское
09	Тип торговой точки	рынок	рынок	рынок	магазин на углу	супермаркет	рынок	магазин на углу	киоск	супермаркет
10	Наблюдаемая величина	400	400	400	500	1,000	400	600	450	400
11	Единица наблюдаемой величины	грамм	грамм	грамм	грамм	грамм	грамм	грамм	грамм	грамм
12	Наблюдаемые цены	28.72	31.92	31.92	50.00	80.00	34.32	59.88	449.10	44.00
13	Важность индикатора	*	*	*	*				*	*
14	Тип цены	обменный	обменный	обменный	обычный	со скидкой	обменный	обычный	обычный	обычный
15	Дополнительная информация				продается россыпью	доступен только упакованный				
16	Исходная величина цены	35.90	39.90	39.90	50.00	40.00	42.90	49.90	499.00	55.00
17	Отношение к средней стоимости	<b>0.38</b>	<b>0.42</b>	<b>0.42</b>	<b>0.53</b>	<b>0.42</b>	<b>0.45</b>	<b>0.53</b>	<b>5.27</b>	<b>0.58</b>
18	T-значение	-0.39	-0.36	-0.36	-0.29	-0.36	-0.34	-0.30	<b>2.66</b>	-0.26
19	Комментарии НКА				не соот- ветствует	не соот- ветствует			проверить	
20	Статус ценового наблюдения	исходный	исходный	исходный	исходный	исходный	исходный	исходный	исходный	исходный
Номер столбца		1	2	3	4	5	6	7	8	9

*Примечание:* Продукт важен, если это - объем ходового товара. Важные продукты идентифицированы астериском (\*) и часто упоминаются как “продукты с астериском.”

**Table 9.1b** Процесс валидации: Наблюдения цен после проверки

Область		После проверки								
01	Код продукта	1101111.0111								
02	Сокращенное описание продукта	рис А, пакет 400–600 грамм								
03	Исходная величина	500								
04	Единица исходной величины	grams								
05	Дата наблюдения	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год	день/ месяц/ год
06	Идентификатор сбора цен	A	A	B	B	C	C	C	D	D
07	Идентификатор торговой точки	001	009	015	019	025	036	037	048	051
08	Расположение торговой точки	сельское	сельское	сельское	сельское	городское	городское	городское	городское	городское
09	Тип торговой точки	рынок	рынок	рынок	магазин на углу	супермаркет	рынок	магазин на углу	киоск	супермаркет
10	Наблюдаемая величина	400	400	400	500	1,000	400	600	450	400
11	Единица наблюдаемой величины	грамм	грамм	грамм	грамм	грамм	грамм	грамм	грамм	грамм
12	Наблюдаемые цены	28.72	31.92	31.92	50.00	80.00	34.32	59.88	44.91	44.00
13	Важность индикатора	*	*	*	*		*	*	*	*
14	Тип цены	обычный	обычный	обычный	обменянный	со скидкой	обычный	обычный	обычный	обычный
15	Дополнительная информация				продается россыпью	доступен только упакованный	добавлен астериск	добавлен астерик		
16	Исходная величина цены	35.90	39.90	39.90	50.00	40.00	42.90	49.90	49.00	55.00
17	Отношение к средней стоимости	<b>0.82</b>	<b>0.91</b>	<b>0.91</b>			<b>0.98</b>	<b>1.14</b>	<b>1.14</b>	<b>1.25</b>
18	T-значение	-1.12	-0.56	-0.56			-0.14	-0.84	-0.84	-1.55
19	Комментарии НКА				не соответствует	не соответствует			ошибка в цене	
20	Статус ценового наблюдения	исходный	исходный	исходный	погашен	погашен	исходный	исходный	исправлен 49.90	исходный
Номер столбца		1	2	3	4	5	6	7	8	9

В Табл. 9.1.b два наблюдения в 4 и 5 столбцах не соответствуют спецификациям товара. В лавке (торговая точка 19) оцениваемый рис продавался россыпью, а не в пакетах, как описано в спецификации, а в супермаркете (торговая точка 25) размер оцениваемого пакета намного превышал установленный диапазон. Сборщики цен зарегистрировали цены на эти товары-заменители, поскольку рис А не продавался в пакетах в лавке, а в меньших упаковках - в супермаркете. Поскольку цены на товары-заменители не могли быть использованы, они были подавлены, что и записано в поле 20. Заметим, что эти наблюдения не соответствуют спецификациям, а не являются ошибками товара, поскольку сборщики цен отметили произведенные замены (см. поле 15). Они были бы ошибками товара, если бы не были отмечены. Они бы также были бы «инлаерами», поскольку их значения не были бы выявлены как экстремальные. Это говорит о том, как важно рассматривать дополнительную информацию, предоставляемую сборщиками цен, наряду с самими значениями цен, а не просто полагаться на механическое определение экстремальных значений.

Как видно из приводимого ниже примера, подавление или исправление экстремальных значений среди наблюдений цен приводит к изменениям, которые оказывают влияние на возможный статус ошибки средней цены, особенно с высоким значением отношения максимального и минимального значений. Высокий коэффициент вариации, с другой стороны, может быть вызван причинами, отличными просто от ошибки товара или ценовой ошибки. Цена товара может сильно отличаться в разных типах торговых точек, или разные модификации товара могли быть оценены в разных торговых точках в связи с тем, что спецификации были слишком широкими или были поняты по-разному сборщиками цен.

При условии, что наблюдения цен верны и в разных торговых точках были оценены сопоставимые товары, вариация цен, возникающая в результате проведения наблюдений в разных торговых точках, есть экономический факт жизни. Значение средней цены должно быть сохранено, а причина вариации объяснена РКА. Возможно, что состав торговых точек, отобранных для обследования, не соответствует профилю распределения покупок для данного товара. Если это так, следует провести исследование и при необходимости скорректировать состав путем подавления наблюдений цен из тех торговых точек, которые представлены избыточно, или дублирования наблюдений цен из тех точек, которые представлены недостаточно.

Товары со средними ценами, вариация которых вызвана слишком широкой спецификацией или несогласованной оценкой в разных торговых точках, должны быть удалены, если только они не являются важными а в первичной группе недостаточное количество товаров. В этом случае эти товары и их средние цены должны быть сохранены на временной основе, и это должно быть отмечено в поле 30 таблицы средних цен. Позже можно будет решить совместно с РКА, отбросить ли эти товары, сохранить или разбить – на основании того, что сообщат другие национальные координаторы.

Таблица 9.2.a показывает, как выглядит только что построенная таблица средних цен до всяких исправлений, которые были внесены в лежащую в ее основе таблицу наблюдений цен (табл. 9.1.a). Мы видим, что цена на рис А не соответствует ни тесту на отношение максимального и минимального значений, ни тесту на коэффициент вариации. Для того чтобы понять причину, необходимо вернуться к таблице наблюдений цен и ценовой ошибке в столбце 8. Табл. 9.2.b показывает таблицу средних цен после проверки, вернее - после проверки таблицы наблюдений цен и исправления столбца 8. В результате этого исправления, средняя цена проходит тесты на отношение максимального и минимального значений и на коэффициент вариации.

**Таблица 92a** Таблица средних цен после первого редактирования НКА и до проверки

Область		После первого обзора и до проверки		
01	Код продукта	1101111.0111	1101111.0112	1101111.0113
02	Сокращенное описание продукта	рис А, пакет 400–600 грамм	рис В, пакет 250–500 грамм	рис С, продается россыпью
03	Исходная величина	500	500	1000
04	Единица исходной величины	грамм	грамм	грамм
21	Отчетный период	месяц/год	месяц/год	месяц/год
22	Количество наблюдений	9	5	10
23	№ с показателем важности	5	5	10
24	Средняя цена продукта	94.72	46.75	45.46
25	Максимальная цена на продукт	499.00	69.00	59.90
26	Минимальная цена на продукт	35.90	39.90	32.90
27	Макс-Мин соотношение	13.9	1.7	1.8
28	Стандартное отклонение	151.7	12.63	8.8
29	Коэффициент вариации	160.2	27.0	19.4
30	Комментарии НКА	проверить	проверить	принять
31	Статус средней цены	исходный	исходный	исходный

**Таблица 92b** Таблица средних цен после проверки

Область		После проверки		
01	Код продукта	1101111.0111	1101111.0112	1101111.0113
02	Сокращенное описание продукта	рис А, пакет 400–600 грамм	рис В, пакет 250–500 грамм	рис С, продается россыпью
03	Исходная величина	500	500	1000
04	Единица исходной величины	грамм	грамм	грамм
21	Отчетный период	месяц/год	месяц/год	месяц/год
22	Количество наблюдений	7	4	10
23	№ с показателем важности	7	4	10
24	Средняя цена продукта	43.92	41.19	45.46
25	Максимальная цена на продукт	55.00	45.00	59.90
26	Минимальная цена на продукт	35.90	39.90	32.90
27	Макс-Мин соотношение	1.5	1.1	1.8
28	Стандартное отклонение	7.2	2.5	8.8
29	Коэффициент вариации	16.3	6.2	19.4
30	Комментарии НКА	проверить	проверить	принять
31	Статус средней цены	поправлен	поправлен	исходный

Валидация индивидуальных цен и средних цен для товаров по первичным группам. После второго раунда валидации, большая часть, если не все, выявленные экстремальные значения будут исследованы. И среди наблюдений цен, и среди средних цен все еще могут оставаться экстремальные значения, но они должны быть зарегистрированы и могут быть удалены позже в ходе межстрановой валидации, если необходимо. По завершении внутристрановой валидации НКА отправляет таблицу средних цен РКА, где она рассматривается до введения средних цен в диагностический модуль для средних цен для проведения межстрановой валидации. Это

рассмотрение может привести к тому, что РКА отправит таблицу обратно в НКА, отметив аномальные значения средних цен, которые требуют дальнейших разъяснений. После рассмотрения вопросов, поставленных РКА, и корректировки средних цен в соответствии с требованиями НКА возвращает таблицу в РКА.

## Межстрановая валидация

Внутристрановая валидация предназначена для проверки средних цен, сообщенных НКА региона, на наличие возможных ошибок для оценки надежности ППС, получаемых на их основе. Цель процесса состоит в том, чтобы удостовериться, что средние цены характеризуют сопоставимые товары в разных странах, и что эти товары были оценены правильно; другими словами, что НКА интерпретировали спецификации товаров одинаково и точно оценили их. Для этого сравнивают средние цены на одни и те же товары в разных странах и анализируют разброс отношений средних цен между странами по товарам и по странам. Коротко говоря, межстрановое редактирование включает выявление экстремальных значений среди средних цен при помощи соответствующих ценовых отношений. Во время этого процесса окончательно отбирают те товары, которые будут включены в финальные расчеты.

Поскольку НКА сообщает цены в национальной валюте, средние цены можно сравнивать, только если они выражены в общей валюте. После перевода в общую валюту средние цены из разных стран на один и тот же товар можно сравнить друг с другом и выявить экстремальные значения в соответствии с заранее определенными критериями. Но цены товаров, даже выраженные в общей валюте, нельзя сопоставлять напрямую. С другой стороны, ценовые соотношения в странах, обследующих цены на товар, могут быть сопоставлены с эквивалентными ценовыми соотношениями для других товаров, при условии, что они сначала были «стандартизованы». Стандартизированные ценовые соотношения для товара – это отношения индивидуальных средних цен в странах, оценивающих товар, и средней геометрической средних цен всех стран, оценивающих товар, когда средние цены выражены в общей валюте.<sup>14</sup> В Таблицах 9.3.a и 9.3.b приведен числовой пример расчета и интерпретации этих соотношений. В приложении к этой главе показан анализ этих данных с использованием таблиц Куаранта и таблиц Дикханова. Данные в этих таблицах отражают результаты нескольких раундов валидации; цель состоит в объяснении концепций валидации данных.

**Таблица 9.3a** Пример валидации данных с использованием обменных курсов для четырех стран и трех товаров

		Первичная группа «Рис», информация о стране - обзор курса валют				
		Страна А	Страна В	Страна С	Страна D	Среднее геометрическое
1	NC-цена 1	5,770	2.11	1.82	1.98	
2	NC-цена 2	21,757	7.87	4.66	7.60	
3	NC-цена 3	7,075	3.10	2.59	3.09	
4	XR	3,830	1.84	0.759	1.00	
5	XR P1	1.51	1.15	2.40	1.98	1.69
6	XR P2	5.58	4.27	6.14	7.60	5.80
7	XR P3	1.85	1.68	3.40	3.09	2.39
8	XR -коэффициент	89	68	142	117	
9	XR -коэффициент	98	74	106	131	
10	XR -коэффициент	77	70	142	129	

Примечание: NC = национальная валюта; XR = обменный курс.

**Table 9.3b** Пример валидации данных с использованием ППС для четырех стран и трех товаров

		Первичная группа «Рис», информация о стране				Среднее геометрическое	Коэффициент вариации
		Страна А	Страна В	Страна С	Страна D		
1	NC-рис 1	5,770	2.11	1.82	1.98		
2	NC-рис 2	21,757	7.87	4.66	7.60		
3	NC-рис 3	7,075	3.10	2.59	3.09		
4	PPC – рис 1	2,914	1.07	0.92	1.00		
5	PPC – рис 2	2,662	1.04	0.61	1.00		
6	PPC – рис 3	2,230	1.00	0.84	1.00		
7	PPC – рис	2,673	1.04	0.779	1.00		
8	PPC – цена 1	2.16	2.04	2.34	1.98	2.13	
9	PPC – цена 2	8.14	7.60	5.98	7.60	7.28	
10	PPC – цена 3	2.65	3.00	3.32	3.09	3.00	
11	PPC – коэффициент1	102	96	110	93		7.5
12	PPC – коэффициент2	112	104	82	104		12.7
13	PPC – коэффициент3	88	100	111	103		9.3
14							
15	Коэффициент вариации	11.8	4.2	16.1	6.1		9.75

Примечание: NC = национальная валюта.

В ходе межстрановой валидации используются и ППС, и обменные курсы валют для перевода средних цен в общую валюту, и для получения стандартизированных ценовых соотношений. Стандартизированные ценовые соотношения на основе цен, полученных с использованием обменных курсов, обозначаются XR-соотношения, а стандартизированные ценовые соотношения на основе цен, полученных с использованием ППС - PPP-соотношения.<sup>15</sup> И XR-соотношения, и PPP-соотношения подвергаются редактированию и проверке. Но только PPP-соотношения используются для получения показателей дисперсии, которые будут рассмотрены ниже.

ППС для конвертирования средних цен в общую валюту рассчитываются на основе средних цен, прошедших валидацию по странам. Это означает, что редактирование начинается с ППС, основанных на ценах, которые все еще следует проверять. Эти начальные ППС скорее всего будут ненадежными, а отметка экстремальных значений среди PPP-соотношений, скорее всего, также будет ненадежной. Обменные курсы, с другой стороны, не определяются по средним ценам и последние не оказывают на них влияния. По этой причине XR-соотношения используются на начальных стадиях валидации. Может показаться парадоксальным использовать XR-соотношения для редактирования цен, на основе которых должны быть получены ППС, имея в виду тот факт, что сами ППС рассчитываются в связи с тем, что цены, конвертированные с использованием обменных курсов валют, не отражают уровни цен в странах. Но опыт показывает, что XR-соотношения позволяют лучше «почувствовать» надежность средних цен, сообщенных в начале процесса валидации. Опыт также показывает, что многие соотношения, первоначально считавшиеся экстремальными значениями среди XR-соотношений, не оказывались таковыми.

В Таблице 9.3.a приведен пример использования обменных курсов для валидации цен по странам. В примере рассмотрена первичная группа «рис» с тремя продуктами по четырем странам. В полях 1-3 показаны национальные среднегодовые цены в валюте каждой страны для каждого товара, а в поле 4 – обменный курс валюты каждой страны к валюте страны D, чей обменный курс равен 1, так как это базовая страна. Встают вопросы: оценила ли каждая страна один и тот же или подобный продукт, и зарегистрированы ли цены в одинаковых единицах? Должен ли товар быть включен? Следовательно, следующий шаг состоит в конвертировании цены каждого товара в

валюту страны D, используя соответствующие обменные курсы, для получения средней цены товара в странах. Эти цены товара – цены на основе обменных курсов или XR-цены – и соответствующие средние геометрические показаны в полях 5–7. Несмотря на то, что теперь цены выражены в общей валюте, ясно видно, что уровни цен отличаются по товарам и по странам. Товар 2 в несколько раз дороже других товаров, а все цены на товары в стране D относительно низкие. Для того чтобы проанализировать взаимоотношение страна-товар более подробно, в полях 8-10 показаны стандартизированные ценовые соотношения на основе цен, полученных с использованием обменных курсов, или XR-соотношения. Это просто XR-цена товара в каждой стране, деленная на среднюю геометрическую цену товара. В стране C наблюдается самая большая изменчивость, что говорит о необходимости рассмотрения средних цен на товар 1 и 3. Наибольшую вариацию демонстрирует товар 3, что позволяет считать, что спецификации, используемые для определения его характеристик, должны быть рассмотрены во всех странах.

Следующий шаг процесса валидации состоит в повторении только что описанных шагов, но с использованием ППС для конвертирования национальных цен на товар в общую валюту. Сначала следует рассчитать ППС для товара и агрегировать их в ППС для страны, как показано в табл. 9.3.b. Для простоты, цены в национальной валюте повторены в полях 1-3. В полях 4-6 показаны ППС для товара, которые рассчитаны просто как цена каждой страны в ее валюте, деленная на цену страны D в ее валюте. Результат – ППС для товара, где страна D является базовой. В предыдущих главах говорилось, что любая страна может быть выбрана в качестве базовой. В поле 7 показан агрегированный или средний ППС для риса для каждой страны – он может быть рассчитан с использованием любого метода, описанного в предыдущих главах. Если средние цены товара для каждой страны разделить на соответствующий ППС, то они будут конвертированы в цену каждого товара на основе ППС и ее среднюю, как показано в полях 4-6. Поскольку использование ППС устраняет различия в уровнях цен между странами, можно ожидать, что цены на каждый товар на основе ППС будут схожими с небольшими отклонениями от геометрической средней.

PPP-соотношения в полях 11–13 (PPP-цена, деленная на среднюю геометрическую цену товара) различаются по странам и по товарам. Коэффициент вариации в поле 15 есть показатель разброса значений ППС для товаров в каждой стране (коэффициент вариации для страны). Страна C демонстрирует наибольшую вариацию – в два раза выше, чем страны B и D. На практике это может означать, что в стране не были собраны цены на достаточное количество товаров-представителей или то, что была проблема с ценой на товар 2, который показал иной уровень цен по сравнению с другими товарами в этой стране. Продукт 2 в стране C должен быть исследован повторно. Коэффициент вариации, который говорит о разбросе значений цен по продуктам и по странам, приведен в поле 15. Этот коэффициент наиболее полезен при сравнении ППС по первичным группам и будет рассмотрен более подробно ниже.

Межстрановая валидация является итеративным процессом, который может включать несколько итераций или раундов. После каждого раунда, когда неправильные цены будут удалены или исправлены, показатели ППС будут становиться все более надежными, и также более надежными будут становиться отметки экстремальных величин среди PPP-соотношений. Следовательно, по мере осуществления валидации внимание переключается с экстремальных значений XR-соотношений на экстремальные значения PPP-соотношений. Цель данного упражнения состоит в том, чтобы удалить, или хотя бы уменьшить, экстремальные значения PPP-соотношений. Если это получится сделать, то экстремальные значения, оставшиеся среди XR-соотношений, можно будет проигнорировать. XR-соотношения и PPP-соотношения, которые выходят из диапазона 80–125, отмечаются как имеющие экстремальные значения и нуждаются в

проверке.

Коэффициенты вариации выше 33% являются экстремальными значениями. В этом случае НКА должны исследовать PPP-соотношения, которые отмечены среди PPP-соотношений, относящихся к этому коэффициенту. Кроме того, что коэффициент вариации является инструментом редактирования, он предоставляет возможность проводить мониторинг во время валидации и по ее завершении оценить, насколько эффективен был весь процесс редактирования и проверки для сокращения ошибок, не связанных с выборкой, в ценовых данных. Значения коэффициентов вариации должны быть существенно ниже в конце процесса валидации, чем в ее начале.

## Выбросы

Редактирование первичных групп с использованием таблиц Куаранта и агрегатов – с использованием таблиц Дикханова приводит к идентификации средних цен с экстремальными значениями или, более точно, - PPP-соотношений с экстремальными значениями. Средние цены, на основе которых были рассчитаны отмеченные экстремальные PPP-соотношения, представляют собой только возможные ошибки. Они не являются ошибками по определению, независимо от того, насколько хорошо проработаны критерии для их выявления. Они не могут быть удалены автоматически. Их следует возвращать обратно в НКА для проверки. НКА должны исследовать средние цены, возвращенные как возможные ошибки, и подтвердить, являются ли их значения верными или неверными. Если выяснится, что цены неверные, то НКА следует исправить их, в противном случае, они будут подавлены. Но если выяснится, что значения цен верные, то они представляют собой выбросы, и следует принять решение о том сохранить ли их, заменить ли импутированной величиной или отбросить – это бывает нелегким решением. Один из возможных компромиссов состоит в рассмотрении важности, для того чтобы обеспечить правильную классификацию важных товаров. Некоторые отклонения, даже значительные, могут быть легитимными. Отдельные страны могут иметь свою политику ценообразования, как, например, поддержание низких цен на энергоносители в некоторых нефтедобывающих странах. Такие цены могут быть отмечены как экстремальные значения, но они не будут неверными, и было бы неправильно удалять их, несмотря на «шум», который они вносят в набор данных.

Если, однако, нет никаких смягчающих обстоятельств, то возмущение, созданное выбросом, может оказать влияние не только на ППС для страны, которая зарегистрировала выброс, но и на ППС для других стран в региональном сопоставлении. В таких случаях следует рассмотреть возможности подавления выброса или замены его импутированным значением. Если при рассмотрении первичных групп выбивающееся значение средней цены касается товара, который очень важен для сообщившей о нем страны, то удаление может быть необоснованным, в то время как импутирование может быть таковым. Но если средняя цена относится к менее важному или неважному товару, то удаление ее может быть оправдано. Какое бы действие ни предпринималось, решение должно быть принято совместно НКА страны и РКА в каждом отдельном случае.

В приложении к данной главе более подробно рассмотрено использование таблиц Куаранта и таблиц Дикханова, которые разработаны специально для валидации данных. В следующем разделе приведен обзор процесса анализа ППС для первичных групп. Тот, кто не знаком с таблицами Куаранта и Дикханова, может захотеть ознакомиться с приложением, прежде чем переходить к следующему разделу.

## Валидация ППС для первичных групп

Таблицы Дикханова обычно строят для нескольких первичных групп, составляющих агрегат. Валидация данных на уровне агрегата образует более широкие рамки для редактирования и проверки средних цен. Другими словами, встает вопрос о том, являются ли средние цены согласованными не только внутри первичной группы, но и для более широкого круга товаров в разных первичных группах. Редактирование на уровне агрегата позволяет выявить несоответствия, которые не были бы обнаружены при редактировании только на уровне первичных групп. Например, предположим, что для алкогольных напитков страна представила все цены на все товары за кварталы, а не за литры, как положено. Ценовые соотношения будут согласованными внутри первичной группы, но не будут согласованы с ценовыми соотношениями страны в других первичных группах. Такие ошибки выявляются при редактировании по разным первичным группам. В этом отношении полезно составлять таблицы Дикханова на разных уровнях агрегирования. Например, первичные группы, относящиеся к продуктам питания, сначала могут быть проверены по таблице Дикханова для продуктов питания и безалкогольных напитков, а затем по таблице Дикханова для всех расходов домашних хозяйств на конечное потребление.

Пример таблицы Дикханова для агрегата – в данном случае для ‘расходов на конечное потребление домашних хозяйств’ – приведен в Таблице 9.4.a. Ее модифицированная версия, где остатки регрессии по методу CPD конвертированы в PPP-соотношения, а стандартное отклонение показано как коэффициент вариации, приведена в Табл. 9.4. Ни одна из таблиц не полна. Остатки регрессии для модели CPD (PPP-соотношения) показаны только для 22 из 864 обследованных товаров и только для 9 из 18 стран, включенных в сопоставление. Заметим также, что ППС для агрегата ‘расходы на конечное потребление домашних хозяйств’ не взвешены. Они были рассчитаны методом CPD, в котором используется весь набор товаров и их цен без учета расходов на уровне первичной группы.

Таблица 9.4.b состоит из двух частей. В первой части представлена суммарная информация (ППС, стандартные отклонения, и уровни цен) по странам для агрегата. Во второй части представлены первичные группы, входящие в агрегат. Для каждой первичной группы таблица также состоит из двух частей: в первой – такая же суммарная информация как по странам, но по первичной группе; во второй – остатки регрессии CPD и коэффициенты вариации для товаров, обследованных в данной первичной группе. В табл. 9.4.a остатки регрессии приведены с тремя знаками после запятой. Для помощи в определении экстремальных значений среди остатков регрессии CPD, они отображаются следующим образом:

Значения остатков регрессии CPD	Шрифт	Эквивалентное PPP-соотношение
Между -0.25 и 0.25	Обычный	Между 78 и 128
Между -0.75 и -0.25 или 0.25 и 0.75	Курсив	Между 47 и 78 или 128 и 212
Между -2.0 и -0.75 или 0.75 и 2.0	Жирный	Между 14 и 47 или 212 и 739
Меньше -2.0 или больше 2.0	Жирный курсив	Меньше 14 или больше 739

Все значения остатков, выделенные шрифтом, должны быть исследованы. Все значения должны быть тщательно проверены, но те, которые выделены жирным шрифтом и жирным курсивом – особенно. В табл. 9.4.b остатки показаны в процентной форме, что легче для интерпретации.

Первая часть таблицы 9.4a облегчает сопоставление ППС (или индексов уровней цен) на

уровне первичных групп товаров. Ожидается, что ППС будут различаться по первичным группам даже в гомогенных регионах. Тем не менее, важно проанализировать и проверить различия между ППС на уровне первичных групп, чтобы убедиться, что значения ППС правдоподобны. Например, ППС для фруктов в странах D, G, и I примерно на 40% выше, чем ППС для овощей в этих странах, тогда как в других странах эти два набора ППС одинакового порядка. Эта ситуация может отражать реальность, однако может быть, что фрукты, выбранные для обследования цен, нерепрезентативны для этих стран, но эту возможность следует проверить. В этом случае в стране I, похоже, что проблемы возникли со средними ценами зеленого сладкого перца и баклажанов. Если средние цены проверены, и они верные, то следует рассмотреть возможность отбрасывания этих двух товаров, поскольку и стандартное отклонение для перца, и стандартное отклонение для первичной группы для страны превышают критическое значение 0,30. Такие же аргументы можно привести для чеснока в стране F, если сообщенное значение средней цены является точным.

Из таблицы видно, что различные шрифты помогают увидеть продукты, чьи средние цены нуждаются в проверке. Но более важно то, что они облегчают выявление возможных проблем в странах – смотри, например, страны A, C, и D в таблице.

На этом этапе следует проверить классификацию товаров по их важности. Другими словами, цена важного товара не должна быть выбросом. Это может также указывать на возможную ошибку выбора товара.

Table94a Таблица Дикханова для агрегата (Расходы на конечное потребление домашних хозяйств). Страны А-І

Код продукта	Наименование продукта	A	B	C	D	E	F	G	H	I	STD1	товары/страны
		2005	2005	2005	2005	2005	2005	2005	2005	2005	STD3	
<b>1100000</b>	<b>HFCE</b>											
	<b>PPP</b>	4.052	16.02	0.739	75.69	43.7	7.879	96.76	1.000	5,227		
	<b>STD 2</b>	0.291	0.221	0.297	0.258	0.240	0.279	0.249	0.271	0.276	0.510	
	<b>No. оцененных товаров</b>	567	608	450	566	525	420	481	611	457		864
	<b>XR (LCU/US\$)</b>	5.779	13.01	0.709	75.55	51.88	6.359	102.3	0.571	4,464		
	<b>Rebased XR</b>	10.12	22.78	1.242	132.3	90.87	11.14	179.2	1.000	7,817		
	<b>PLI</b>	0.400	0.703	0.595	0.572	0.481	0.707	0.540	1.000	0.669		
<b>1101161</b>	<b>Свежие или охлажденные фрукты</b>											
	<b>PPP</b>	1.900	15.65	0.423	57.09	48.31	6.580	99.68	1.000	6,074		
	<b>STD</b>	0.326	0.189	0.353	0.395	0.260	0.164	0.364	0.204	0.126		
	<b>PLI</b>	0.188	0.687	0.341	0.431	0.532	0.591	0.556	1.000	0.777		
1101161.01	Яблоко: золотистое восхитительное		-0.080	0.177		-0.050	0.002	0.105	-0.290	0.092	0.145	12
1101161.02	Яблоко: красное восхитительное		-0.140	0.189		-0.180	0.039	0.273	-0.300	0.079	0.194	13
1101161.03	Яблоко: местный сорт	-0.050	-0.320	0.258			0.271		-0.030		0.248	9
1101161.04	Банан	0.285	0.457	0.581	-0.330	-0.390	0.050	-0.150	0.184	0.223	0.287	18
1101161.05	Красный виноград	-0.510	-0.010	-0.250	0.482	0.174	0.055		0.038	-0.140	0.265	15
1101161.06	Зеленый виноград	-0.610	-0.120	-0.300	0.414	0.292	-0.100	0.484	-0.130	-0.070	0.285	16
1101161.07	Грейпфрут	0.184	0.099	-0.550	0.501		-0.120	-0.160	-0.020		0.246	14
1101161.08	Лимон	0.363	0.030	-0.120	-0.420		0.270	-0.640	0.447		0.361	14
1101161.09	Мандарин		0.157	-0.380	-0.200	0.345	-0.000	-0.090	0.234		0.243	13
1101161.10	Апельсин	0.024	-0.060	0.397	-0.510	0.187	-0.110	0.451	-0.050	-0.040	0.234	16
1101161.11	Груша	0.260	-0.010			-0.390	-0.380	0.245	-0.120	-0.190	0.268	14
1101161.12	Арбуз	0.51			0.072	0.012	0.258	-0.530	0.036	0.040	0.228	13
<b>1101171</b>	<b>Свежие или охлажденные овощи</b>											
	<b>PPP</b>	1.557	14.29	0.370	32.00	44.66	5.708	63.49	1.000	3,425		
	<b>STD</b>	0.292	0.503	0.321	0.254	0.239	0.383	0.255	0.337	0.608		
	<b>PLI</b>	0.154	0.627	0.398	0.242	0.492	0.513	0.354	1.000	0.438		
1101171.01	Огурец	0.066		0.124	-0.170	-0.230	0.009	-0.290		0.220	0.272	13
1101171.02	Чеснок	0.227	0.354	-0.090	0.290	-0.300	<b>-0.930</b>	0.137	0.027		0.448	17
1101171.03	Крутой картофель	-0.390	0.391	-0.400	-0.120	0.045	0.408	-0.130	0.221	0.145	0.250	16
1101171.04	Зеленый сладкий перец	-0.410	0.479	0.256	-0.000	-0.140	0.226		0.587	<b>-1.100</b>	0.400	16
1101171.05	Зеленая капуста	0.444	-0.480	0.475	-0.310	0.515	0.256	0.105		-0.700	0.325	17
1101171.06	Морковь	-0.030	-0.640	0.367	-0.000	-0.140	0.138	0.525	-0.440		0.295	17

1101171.07	Цветная капуста	0.286	0.503	-0.440	0.351	-0.060	-0.070		-0.110		0.272	16
1101171.08	Лук	0.118	-0.610	0.059	0.394	0.302	-0.030	0.049	-0.280	0.266	0.295	18
1101171.09	Кукуруза			-0.340	-0.050	0.122		-0.070		0.424	0.203	10
1101171.10	Баклажан	-0.300			-0.370	-0.120		-0.320		0.748	0.317	13

Table 94b Модифицированная таблица Дикханова для агрегата (Расходы на конечное потребление домашних хозяйств). Страны А–I

Код продукта	Наименование продукта	A	B	C	D	E	F	G	H	I	Var. co.1	товары/страны
		2005	2005	2005	2005	2005	2005	2005	2005	2005	2005	
<b>1100000</b>	<b>HFCE</b>											
	PPP	4.052	16.02	0.739	75.69	43.7	7.879	96.76	1.000	5,227		
	Var. co. 2	29.1	22.1	29.7	25.8	24.0	27.9	24.9	27.1	27.6	51.0	
	No. оцененных товаров	567	608	450	566	525	420	481	611	457		864
	XR (LCU/US\$)	5.779	13.01	0.709	75.55	51.88	6.359	102.3	0.571	4,464		
	Rebased XR	10.12	22.78	1.242	132.3	90.87	11.14	179.2	1.000	7,817		
	PLI	40.0	70.3	59.5	57.2	48.1	70.7	54.0	100.0	66.9		
<b>1101161</b>	<b>Свежие или охлажденные фрукты</b>											
	PPP	1.900	15.65	0.423	57.09	48.31	6.580	99.68	1.000	6,074		
	Var. co. 2	32.6	18.9	35.3	39.5	26.0	16.4	36.4	20.4	12.6		
	PLI	18.8	68.7	34.1	43.1	53.2	59.1	55.6	100.0	77.7		
1101161.01	Яблоко: золотистое восхитительное		92.3	119.4		95.1	100.2	111.1	74.8	109.6	14.5	12
1101161.02	Яблоко: красное восхитительное		86.9	120.8		83.5	104.0	131.4	74.1	108.2	19.4	13
1101161.03	Яблоко: местный сорт	95.1	72.6	129.4			131.1		97.0		24.8	9
1101161.04	Банан	133.0	157.9	178.8	71.9	67.7	105.1	86.1	120.2	125.0	28.7	18
1101161.05	Красный виноград	60.0	99.0	77.9	161.9	119.0	105.7		103.9	86.9	26.2	15
1101161.06	Зеленый виноград	54.3	88.7	74.1	151.3	133.9	90.5	162.3	87.8	93.2	28.5	16
1101161.07	Грейпфрут	120.2	110.4	57.7	165.0		88.7	85.2	98.0		24.6	14
1101161.08	Лимон	143.8	103.0	88.7	65.7		102.7	52.7	156.4		36.1	14
1101161.09	Мандарин		117.0	68.4	81.9	141.2	100.0	91.4	126.4		24.3	13
1101161.10	Апельсин	102.4	94.2	148.7	60.0	120.6	89.6	157.0	95.1	96.1	23.4	16
1101161.11	Груша	126.7	99.0			67.7	68.4	127.8	88.7	82.7	26.8	14
1101161.12	Арбуз	105.2			107.5	101.2	129.4	85.9	103.7	104.1	22.8	13
<b>1101171</b>	<b>Свежие или охлажденные овощи</b>											
	PPP	1.557	14.29	0.370	32.00	44.66	5.708	63.49	1.000	3,425		
	STD	29.2	50.3	32.1	25.4	23.9	38.3	25.5	33.7	60.8		
	PLI	15.4	62.7	29.8	24.2	49.2	51.3	35.4	100.0	43.8		

1101171.01	Огурец	106.8		113.2	84.4	79.5	100.9	74.8		124.6	27.2	13
1101171.02	Чеснок	125.5	142.5	91.4	133.6	74.1	<b>39.5</b>	114.7	102.7		44.8	17
1101171.03	Круглый картофель	67.7	147.8	67.0	88.7	104.6	150.4	87.8	124.7	115.6	25.0	16
1101171.04	Зеленый сладкий перец	66.4	161.4	129.2	100.0	86.9	125.4		179.9	<b>33.3</b>	40.0	16
1101171.05	Зеленая капуста	155.9	61.9	160.8	73.3	167.4	129.2	111.1		49.7	32.5	17
1101171.06	Морковь	97.0	52.7	144.3	100.0	86.9	114.8	169.0	64.4		29.5	17
1101171.07	Цветная капуста	133.1	165.4	64.4	142.0	94.2	93.2		89.6		27.2	16
1101171.08	Лук	112.5	54.3	106.1	148.3	135.3	97.0	105.0	75.6	130.5	29.5	18
1101171.09	Кукуруза			71.2	95.1	113.0		93.2		153.8	20.3	10
1101171.10	Баклажан	74.1			69.1	88.7		72.6		211.3	31.7	13

Источник (обе таблицы): 2005 ПМС.

Примечание: STD = стандартное отклонение; XR = обменный курс; LCU = местная денежная единица, плюс = уровень индекса цен ; var.co. . = коэффициент отклонения.

## Заключение

Межстрановая валидация – этот итеративный процесс. Ее можно начинать до того, как все страны, участвующие в региональном сопоставлении, представили свои средние цены. После каждой итерации или раунда проверки, РКА внесет изменения в региональную базу данных в соответствии с результатами, сообщенными НКА стран, участвующих в раунде, добавит в базу данных цены тех стран, которые включаются в процесс валидации данных, и пересчитает диагностический модуль средних цен для построения новых таблиц Куаранта и Дикханова. Эти таблицы дадут возможность обнаружить новые экстремальные значения, появившиеся в результате изменений, произведенных РКА, которые должны быть исследованы НКА. Постепенно, после нескольких раундов проверки и после того, как цены всех стран, участвующих в сопоставлении, будут включены в базу данных, будет достигнута сходимость результатов, а результаты повторных раундов окажутся маргинальными, так что НКА и РКА посчитают продолжение процесса не стоящим усилий. Теперь процесс межстрановой валидации завершен. Прекратив участие в процессе валидации, НКА принимает ответственность за свою информацию о средних ценах. Процесс считается завершенным, когда НКА формально одобряет ценовые данные, прошедшие процесс валидации.

### Таблицы Куаранта и таблицы Дикханова

И таблицы Куаранта, и таблицы Дикханова предоставляют аналогичные инструменты для валидации цен для товаров и стран; они используют либо ППС на уровне первичных групп для редактирования данных по отдельной первичной группе, либо ППС для агрегата 16 для редактирования данных по первичным группам, составляющим агрегат. Таблица Дикханова применяются в том случае, когда используются метод товарно-страновых фиктивных переменных (CPD) или метод CPD с дополнительным учетом репрезентативности (CPRD), тогда как таблица Куаранта имеет более широкое применение, включая метод Джини-Элтэтэ-Ковеша-Шульца (GEKS) и модифицированный метод GEKS (GEKS\*), а также методы CPD и CPRD.<sup>17</sup> Регионы ПМС использовали метод CPD при проведении сопоставления в 2005 г. Таблицы Куаранта используются для редактирования цен в рамках первичной группы товаров, а таблицы Дикханова – для редактирования в рамках агрегата из нескольких первичных групп. Таблица Куаранта изначально была предназначена для редактирования цен внутри первичных групп товаров. Она предоставляет большой объем информации о ценах товаров, но ее форма является устаревшей, что делает ее неудобной для анализа большого количества продуктов, обследуемых в рамках агрегата.<sup>18</sup> Таблица Дикханова содержит большую часть той же информации, что и таблица Куаранта<sup>19</sup>, но она запрограммирована так, что можно скрывать некоторые позиции, которые могут быть открыты при необходимости, так что на дисплей выводятся только основные ряды. Более компактная форма таблицы Дикханова в сочетании с цветовым решением для идентификации разных уровней экстремальных значений делают ее более подходящей для редактирования цен по разным первичным группам и товарам, составляющим агрегат.

В рамках диагностического модуля средних цен строится множество таблиц Куаранта – по одной на каждую первичную группу, которая участвует в процессе валидации, и одна таблица Дикханова для соответствующего агрегата. РКА предоставляет НКА таблицы Куаранта и Дикханова одновременно. Это оставляет открытым вопрос, следует ли начинать процесс межстрановой валидации с таблиц Куаранта или с таблицы Дикханова. Некоторые аналитики предпочитают начинать с таблиц Куаранта и обращаться к таблице Дикханова после того, как прошло несколько раундов проверки и оценки ППС стали более надежными. Другие предпочитают сначала использовать таблицу Дикханова для выявления стран и продуктов, которые требуют исследования, и далее организуют такое исследование вокруг таблиц Куаранта. Одним из аргументов в пользу первого подхода является то, что большинство операций (если не все), относящихся к проведению сопоставления, организовано вокруг первичных групп. В частности, ППС сначала рассчитываются и усредняются на уровне первичных групп товаров, и это ППС для первичных групп, которые взвешиваются во время агрегирования. Поэтому представляется логичным и последовательным начинать процесс валидации на уровне первичных групп с использования таблиц Куаранта. Еще один аргумент состоит в том, что таблица Дикханова не уделяет такого внимания, как таблица Куаранта, XR-соотношениям, которые необходимо использовать на ранних стадиях межстрановой валидации. В таблице Дикханова они являются подтягивающими элементами, а экстремальные значения не отмечаются. В таблице Куаранта равные веса приписываются всем XR-соотношениям и PPP-соотношениям, а экстремальные значения среди тех и других отмечаются.

Table9A1 Таблица Куаранта

<b>ПОДРОБНОСТИ ТАБЛИЦЫ КУАРАНТА</b>								
[1] Регион: Азия [2] Опрос: Еда и напитки [3] Период времени: 1-й квартал 2011 года [4] Версия: Последняя [5] Дата: 30.06.11								
[6] Метод расчета: CPD [7] Пределы XR-и PPP-соотношений: 80%, 125% [8] знаменатели: LCU Страны D [9] Страница: 1								
<b>ПЕРВИЧНЫЕ ГРУППЫ И ПОДРОБНОСТИ СТРАНЫ</b>								
[10] 1101111 Рис [11] Ср. Вес: 45,1 [12] Кол-во пунктов: 3 [13] Var. co. 1: 9.5								
[14] Страна	[15] XR	[16] PPP	[17] PLI (%)	[18] Вес	[19] Кол-во товаров	[20] Var. co. 2		
A	3,830.9492	2,673.8500	69.8	23.8	3: *1	11.8		
B	1.8439	1.0353	56.1	34.6	3: *3	4.2		
C	0.7594	0.7791	102.6	58.3	3: *1	16.1		
D	1.0000	1.0000	100.0	63.9	3: *2	6.1		
<b>ДЕТАЛИ ТОВАРА ПО СТРАНАМ</b>								
1 [21] +1101111,0111 рис, длиннозерный, 500-1000 г пакет [22] Var. co. 3: 7.5								
[23] Страна	[24] NC-цена	[25] Астериксы	[26] Котируются	[27] Var. co. 4	[28] XR-цена	[30] XR-коэффициент	[31] PPP-цена	[33] PPP-коэффициент
A	5,770.58	*	15	5.1	1.51	89	2.16	102
B	2.11	*	13	4.4	1.15	68	2.04	96
C	1.82		9	21.1	2.40	142	2.34	110
D	1.98		5	16.8	1.98	117	1.98	93
[29] GM = 1.69 [32] GM = 2.13								
2 [21] +1101111,0112 Рис, короткое зерно, 500-1000 г пакет [22] Var. co. 3: 12.7								
[23] Страна	[24] NC-цена	[25] Астериксы	[26] Котируются	[27] Var. co. 4	[28] XR-цена	[30] XR-коэффициент	[31] PPP-цена	[33] PPP-коэффициент
A	21,757.60		9	14.2	5.68	98	8.14	112
B	2.11	*	10	7.7	4.27	74	7.60	104
C	1.82	*	15	9.1	6.14	106	5.98	82
D	1.98	*	8	8.9	7.60	131	7.60	104
[29] GM = 5.80 [32] GM = 7.28								
3 [21] +1101111,0113 рис, басмати, 500-1000 г пакет [22] Var. co. 3: 9.3								
[23] Страна	[24] NC-цена	[25] Астериксы	[26] Котируются	[27] Var. co. 4	[28] XR-цена	[30] XR-коэффициент	[31] PPP-цена	[33] PPP-коэффициент
A	7,075.88		6	15.3	1.85	77	2.65	88
B	3.10	*	14	10.4	1.68	70	3.00	100
C	2.59		7	24.7	3.40	142	3.32	111
D	3.09	*	10	10.2	3.09	129	3.09	103
[29] GM = 2.39 [32] GM = 3.00								

Примечания к табл. 9А.1

ОПИСАНИЕ ТАБЛИЦЫ КУАРАНТА		
[1]	<b>Region</b>	Регион, представленный в таблице.
[2]	<b>Survey</b>	Тип обследуемого товара, представленный в таблице.
[3]	<b>Time period</b>	Период – период сбора цен на товары, представленные в таблице.
[4]	<b>Version</b>	Версия таблицы – первая, вторая, третья, ... , окончательная.
[5]	<b>Date</b>	Дата расчета таблицы
[6]	<b>Calculation method</b>	Метод расчета – метод, использованный для расчета ППС на уровне первичных групп в столбце [16]. В настоящее время это CPD, но это может быть также CPRD, GEKS или GEKS*.
[7]	<b>Limits of XR and PPP-ratios</b>	<i>Пределы для XR и PPP-соотношений. Диапазон, в котором должны лежать XR – соотношения в столбце [30] и PPP-соотношения в столбце [33], чтобы не быть отмеченными в качестве экстремальных. В настоящее время: 80–125, но 65–155 также используется. Интервалы не равны, так как оцениваются относительные отклонения от средней геометрической.</i>
[8]	<b>Numeraire</b>	Валюта, выбранная в качестве расчетной. РКА может выбрать валюту любой страны своего региона.
[9]	<b>Page</b>	Номер страницы. Таблица Куаранта может занимать большое количество страниц в зависимости от количества оцениваемых продуктов и количества стран.
ИНФОРМАЦИЯ ПО ПЕРВИЧНОЙ ГРУППЕ И ПО СТРАНЕ		
[10]		Код и название первичной группы.
[11]	<b>Av. weight</b>	Средний вес расходов для группы стран по первичной группе. Невзвешенная средняя арифметическая национальных весов в столбце [18]. Как и для национальных весов, значение приведено к 100,000.
[12]	<b>No. of items</b>	Количество товаров, описанных и обследованных в первичной группе.
[13]	<b>Var.co.1</b>	Общий коэффициент вариации или – более точно – средний коэффициент вариации для товаров, обследованных в первичной группе. Рассчитывается как невзвешенная средняя арифметическая коэффициентов вариации для товаров из [22]. Измеряет среднюю вариацию PPP-соотношений из столбца [33] для всех товаров, обследованных в первичной группе. Отмечен, если превышает 33%.
[14]	<b>Country</b>	Сокращенные названия стран, охваченных таблицей.
[15]	<b>XR</b>	Рыночные обменные курсы стран, выраженные в единицах национальной валюты на единицу расчетной валюты, указанной в [8].
[16]	<b>PPP</b>	Паритет покупательной способности (ППС) на уровне первичной группы, рассчитанный как сказано в [6] и выраженный как количество единиц национальной валюты на единицу выбранной расчетной валюты, указанной в [8]. Цены, используемые для расчета ППС, являются средними ценами в национальных валютах, которые сообщили страны на продукты, обследованные в первичной группе – то есть национальные цены из столбца [24].
[17]	<b>PLI</b>	Индексы уровня цен. ППС из столбца [16], выраженные в процентном отношении от соответствующего обменного курса в столбце [15].

[18]	<b>Weight</b>	Национальные веса расходов, приведенные к 100.000. Та часть ВВП страны, которая тратится на товары первичной группы, где расходы выражены в национальной валюте при национальном уровне цен.
[19]	<b>No. of items</b>	<i>Количество товаров, обследованных каждой страной в первичной группе, которые являются важными, то есть количество товаров, отмеченных (*).</i>
[20]	<b>Var.co.2</b>	Кoeff. Вар. 2. Коэффициент вариации для страны. Стандартное отклонение PPP-соотношения для страны в столбце [33] для всех продуктов, обследованных в стране в первичной группе, выраженное в процентном отношении к средней арифметической PPP-соотношений для страны в столбце [33] для всех продуктов, обследованных в первичной группе. Отмечается при превышении значения в 33%.
<b>ХАРАКТЕРИСТИКИ ТОВАРОВ ПО СТРАНАМ</b>		
[21]		[21] Код, название и краткое определение товара, представленного в подразделе.
[22]	<b>Var.co.3</b>	Кoeff. вар. 3. Коэффициент вариации для товара. Стандартное отклонение PPP-соотношения для товара в столбце [33], выраженное в процентном отношении к средней арифметической PPP-соотношений в столбце [33]. Отмечается при превышении значения в 33%.
[23]	<b>Country</b>	Сокращенные названия стран, охваченных таблицей.
[24]	<b>NC-price</b>	Средняя цена в национальной валюте.
[25]	<b>Астериск</b>	Астериск. Индикатор важности. Обычно важные продукты отмечают звездочкой (*).
[26]	<b>Quotations</b>	Количество наблюдений цены, на основе которых рассчитаны средние цены в столбце [24].
[27]	<b>Var.co.4</b>	Кoeff. вар. 4. Коэффициент вариации для наблюдений цен. Стандартное отклонение для наблюдений цен, лежащих в основе средних цен товара, в столбце [24], выраженное в процентном отношении от средней арифметической наблюдений цен, лежащих в основе средних цен товара из столбца [24]. Отмечается при превышении значения в 20%.
[28]	<b>XR-price</b>	XR-цена. Средняя цена в национальной валюте из столбца [24], конвертированная в расчетную валюту при помощи обменных курсов из столбца [15].
[29]	<b>GM</b>	Средняя геометрическая величина значений цен, полученных конвертированием с использованием обменных курсов, из столбца [28]. Использование средней геометрической здесь и в [32] обеспечивает инвариантность в отношении выбора расчетной валюты.
[30]	<b>XR-ratio</b>	XR-соотношение. Стандартизированные ценовые соотношения, рассчитанные на основе цен, полученных конвертированием с использованием обменных курсов, из столбца [28]. XR-цены, выраженные в процентном отношении к их средней геометрической величине в [29].
[31]	<b>PPP-price</b>	PPP-цена. Средние цены в национальной валюте из столбца [24], конвертированные в расчетную валюту с использованием ППС из столбца [16].
[32]	<b>GM</b>	GM Средняя геометрическая величина значений цен, полученных конвертированием с использованием ППС, из столбца [31].
[33]	<b>PPP-ratio</b>	PPP-соотношение. Стандартизированные ценовые отношения на основе цен, полученных конвертированием с использованием ППС из столбца [31]. PPP-цены, выраженные в процентном отношении к их средней геометрической величине из [32].

Таблица 9А.1 - это пример таблицы Куаранта. Номера курсивом были добавлены для простоты ссылки, а объяснения приведены в конце таблицы. В таблице три раздела. В первом разделе приведена общая информация о таблице, из которой самой важной является метод расчета [6], диапазон допустимых значений XR-соотношений и PPP-соотношений [7], и расчетная валюта [8]. В следующем разделе представлена общая информация о первичной группе товаров, которая касается либо первичной группы в целом, либо каждой страны, где обследуется эта первичная группа. Именно в этом разделе приводится общий коэффициент вариации [13], оценки ППС [16], веса расходов [18], и коэффициенты вариации по странам [20] для первичной группы. В третьем разделе представлены товары, обследуемые в рамках первичной группы. Для каждого товара выделен подраздел, где показан коэффициент вариации для товара [22]; средние цены, сообщенные странами в национальной валюте [24]— это цены, которые проходят процесс валидации и на основе которых рассчитываются ППС, приведенные в таблице; средние цены, конвертированные в общую валюту на основе обменных курсов [28], их среднегеометрическая величина [29] и их XR-соотношения [30]; а также средние цены, конвертированные в общую валюту с использованием ППС на уровне первичной группы [31], их среднегеометрическая величина [32] и их PPP-соотношения [33].

Примерами таблиц Дикханова являются таблицы 9А.2, 9А.3, and 9А.4. Были добавлены номера ссылок курсивом и приведены объяснения после таблицы. Для иллюстрации таблица охватывает те же первичные группы, что и таблица Куаранта, несмотря на то, что обычно таблицы Дикханова строятся для агрегатов. Сопоставление таблиц Куаранта и Дикханова для одной и той же первичной группы полезно для понимания сходства и различия; имеется в виду, что ППС для обеих таблиц рассчитаны с использованием метода CPD. Таблица 9А.2 - это расширенная версия таблицы Дикханова, где показаны ряды данных, которые обычно бывают скрыты. Это – средняя цена в национальной валюте [16], количество наблюдений цены, на основе которых рассчитывается средняя цена [17], коэффициент вариации для наблюдений цен [18] и XR-соотношения [19]. Таблица 9А.3 показывает, как выглядит таблица Дикханова, когда эти ряды данных скрыты, то есть то, как обычно бывает представлена таблица Дикханова. В таблице две части. В одной представлена краткая информация о первичных группах и странах, включая ППС [4], стандартные отклонения для стран [5] и общее стандартное отклонение [6]. В другой части приведены остатки регрессии CPD [13] и их стандартные отклонения [14] для товаров по странам. Остатки регрессии CPD и их вывод объясняются в примечаниях к таблицам 9А.2, 9А.3, and 9А.4. Рассмотрение таблиц Куаранта и Дикханова показывает, что остатки регрессии CPD из таблицы Дикханова равны логарифмам PPP-соотношений из таблицы Куаранта. Остаток регрессии CPD в 0.01505 для длиннозерного риса для страны А в таблицах 9А.2 и 9А.3 равен  $\log 1.015$ , где 1.015 - значение соответствующего PPP-соотношения из таблицы 9А.1. Заметим, что это тождество выполняется только, если ППС в обеих таблицах рассчитаны методом CPD и все товары включены в расчет как одна группа, то есть либо как первичная группа, либо как агрегат. Оно не выполняется, если в таблице Дикханова для расчета использован метод CPD, а в таблице Куаранта – метод GEKS. Оно также не выполняется, если остатки регрессии CPD для товаров в первичной группе в таблице Дикханова основаны на ППС для агрегата, а PPP-соотношения в таблице Куаранта получены из ППС для первичной группы. В Табл.9А.4 показаны остатки регрессии CPD из таблиц 9А.2 и 9А.3, выраженные как PPP-соотношения.

Table9A2 Таблица Диханова для первичной группы (Расширенная версия)

		[1]A	[1]B	[1]C	[1]D	STD 1 STD 3	Товары/ Страны
Код	Название	[2] Q1 2011	[2] Q1 2011	[2] Q1 2011	[2] Q1 2011		
[3] 1101111	[3] Рис						
	[4] PPP	2,673.85	1.0353	0.7791	1.0000		
	[5] STD 2	0.0974	0.0341	0.1391	0.0505	[6] 0.0902	
	[7] No. оцененных товаров	3	3	3	3		[8] 3
	[9] XR (LCU/US\$)	1,473	0.709	0.292	0.3845		
	[10] переоснован XR	3,830.94	1.8439	0.7594	1.000		
	[11] PLI	0.698	0.561	1.026	1.000		
[12] 1101111.0111	[12] рис, длиннозерный	[13] 0.01505	[13]- 0.04068	[13] 0.09660	[13]- 0.07097	[14] 0.0637	[15] 4
	[16] Средняя цена	5,770.58	2.11	1.82	1.98		
	[17] No. of observations	15	13	9	5		
	[18] Variation coefficient	5.1	4.4	21.2	16.8		
	[19] XR-ratio	89.0	67.7	141.9	117.0		
[12] 1101111.0112	[12] Рис, короткие зерна	[13] 0.11105	[13] 0.04284	[13]- 0.19665	[13] 0.04277	[14] 0.1169	[15] 4
	[16] Средняя цена	21,757.60	7.87	4.66	7.60		
	[17] No. of observations	9	10	15	8		
	[18] Variation coeffi cient	14.2	7.7	9.1	8.9		
	[19] XR-ratio	97.9	73.6	105.8	131.1		
[12] 1101111.0113	[12] Рис, басмати	[13]- 0.12610	[13]- 0.00216	[13] 0.10005	[13] 0.02820	[14] 0.0817	[15] 4
	[16] Средняя цена	7,075.88	3.10	2.59	3.09		
	[17] No. of observations	6	14	7	10		
	[18] Variation coeffi cient	15.3	10.4	24.7	10.2		
	[19] XR-ratio	77.3	70.4	142.4	129.2		

Table9A3 Таблица Дикханова для первичной группы (сокращенная версия)

		[1]A	[1]B	[1]C	[1]D	STD 1 STD 3	Товары/ Страны
Код	Название	[2] Q1 2011	[2] Q1 2011	[2] Q1 2011	[2] Q1 2011		
[3] 1101111	[3] Рис						
	[4] PPP	2,673.85	1.0353	0.7791	1.0000		
	[5] STD 2	0.0974	0.0341	0.1391	0.0505	[6] 0.0902	
	[7] No. оцененных товаров	3	3	3	3		[8] 3
	[9] XR (LCU/US\$)	1,473	0.709	0.292	0.3845		
	[10] переоснован XR	3,830.94	1.8439	0.7594	1.000		
	[11] PLI	0.698	0.561	1.026	1.000		
[12] 1101111.0111	[12] рис, длиннозерный	[13] 0.01505	[13] -0.04068	[13] 0.09660	[13] -0.07097	[14] 0.0637	[15] 4
[12] 1101111.0112	[12] Рис, короткие зерна	[13] 0.11105	[13] 0.04284	[13] -0.19665	[13] 0.04277	[14] 0.1169	[15] 4
[12] 1101111.0113	[12] Рис, басмати	[13]- 0.12610	[13] - 0.00216	[13] 0.10005	[13] 0.02820	[14] 0.0817	[15] 4

Стандартные отклонения для остатков регрессии CPD в таблице Дикханова фактически являются коэффициентами вариации, так как математическое ожидание остатков равно 1. Они были выражены как коэффициенты вариации (то есть, умножены на 100) в табл. 9А.4 для

облегчения сравнения с коэффициентами вариации из таблицы Куаранта (9А.1). Эти два набора коэффициентов не одинаковы из-за различий в вычислении. Общий коэффициент вариации в таблице Куаранта является средним из коэффициентов вариации для товаров, обследованных в первичной группе, тогда как общий коэффициент вариации в таблице Дикханова рассчитывается на основе всех остатков регрессии в товарной части таблицы, обеспечивая согласованность общего коэффициента вариации, коэффициента вариации для товаров и для стран. Кроме того, коэффициент вариации для товаров в таблице Куаранта теоретически должен был бы рассчитываться с использованием логарифмов, поскольку PPP-соотношения основаны на средней геометрической цен на основе ППС, но из практических соображений они рассчитываются с использованием средней арифметической и стандартных отклонений PPP-соотношений. Коэффициенты вариации в таблице Дикханова рассчитываются по-другому. Они основаны на остатках регрессии CPD, которые, как мы только что заметили, равны логарифмам от PPP-соотношений. Несмотря на различия в расчете, эти два набора коэффициентов вариации являются величинами одного порядка и одинаковой надежности в смысле выявления экстремальных значений.

Версия таблицы Дикханова, представленная в табл. 9А.4 более удобна для пользователя, чем версия в табл. 9А.3, поскольку здесь используется та же терминология и такие же концепции, как в таблице Куаранта, поэтому пользователю легче переходить от одной таблицы к другой в процессе валидации данных. Возможность получать таблицу Дикханова, показывающую остатки регрессии CPD, выраженные как PPP-соотношения, и их стандартные отклонения как коэффициенты вариации, следует включить в диагностический модуль для расчета средних цен.

**Table 9A4** Модифицированная таблица Дикханова для первичной группы (сокращенная версия)

		A	B	C	D	STD1 STD3	Товары/ Страны
Код	Название	Q1 2011	Q1 2011	Q1 2011	Q1 2011		
<b>1101111</b>	<i>Рис</i>						
	<i>PPP</i>	2,673.85	1.0353	0.7791	1.0000		
	<i>STD 2</i>	0.0974	0.0341	0.1391	0.0505	0.0902	
	<i>Но. оцененных товаров</i>	3	3	3	3		3
	<i>XR (LCU/US\$)</i>	1,473	0.709	0.292	0.3845		
	<i>пересоснован XR</i>	3,830.94	1.8439	0.7594	1.000		
	<i>PLI</i>	0.698	0.561	1.026	1.000		
<b>1101111.0111</b>	<i>Рис, длиннозерный</i>	0.01505	-0.04068	0.09660	-0.07097	0.0637	4
<b>1101111.0112</b>	<i>Рис, короткие зерна</i>	0.11105	0.04284	-0.19665	0.04277	0.1169	4
<b>1101111.0113</b>	<i>Рис, басмати</i>	-0.12610	-0.00216	0.10005	0.02820	0.0817	4

Примечания к таблицам 9А.2, 9А.3, и 9А.4

[1]	Сокращенные названия стран, охваченных таблицей.
[2]	Период – период сбора цен на товары, представленные в таблице

[3]	Код и название первичной группы или агрегата.
[4]	Паритет покупательной способности на уровне первичной группы или агрегата, выраженный как количество единиц национальной валюты на единицу выбранной расчетной валюты. Цены, используемые для расчета ППС, являются средними ценами в национальных валютах, которые сообщили страны на продукты, обследованные в первичной группе или агрегате – то есть национальные цены из столбца [16]
[5]	STD 2: стандартное отклонение остатков регрессии CPD или CPRD для каждой страны для первичной группы или агрегата. Может быть конвертировано в коэффициент вариации для страны умножением на 100. Математическое ожидание остатков регрессии для каждой страны равно 1.
[6]	STD 1: стандартное отклонение остатков регрессии CPD или CPRD для всех товаров, обследованных в первичной группе или агрегате. Может быть конвертировано в общий коэффициент вариации для товаров умножением на 100. Математическое ожидание остатков регрессии для всех товаров равно 1.
[7]	Количество товаров, обследованных каждой страной
[8]	Количество товаров, определенных для первичной группы или агрегата
[9]	Рыночные обменные курсы в странах, выраженные в единицах национальной валюты на доллар США
[10]	Обменные курсы [9] к расчетной валюте. Количество единиц национальной валюты к единице расчетной валюты.
[11]	Индексы уровня цен. ППС из поля [4], как отношение к соответствующему обменному курсу к расчетной валюте из поля [10].
[12]	Код и название товара
[13]	Остатки регрессии CPD или CPRD по товарам и странам
[14]	STD 3: стандартное отклонение остатков регрессии CPD или CPRD для товара. Может быть конвертировано в коэффициент вариации для товара умножением на 100. Математическое ожидание остатков регрессии для товара равно 1.
[15]	Количество стран, обследовавших цены товара
[16]	Средняя цена в местной валюте
[17]	Количество наблюдений, на основе которых получена средняя цена в [16].
[18]	Коэффициент вариации наблюдений цен
[19]	Отношения цен, полученных конвертированием с использованием обменных курсов

Сопоставление таблиц Куаранта и таблиц Дикханова показывает, что они имеют следующие общие показатели и экстремальные значения:

- *Общий коэффициент вариации* оценивает разброс значений всех PPP-соотношений для первичной группы. При этом оценивается однородность ценовых структур в странах, где проводится обследование цен в первичной группе, и надежность оценок ППС,

рассчитанных на уровне первичной группы. Чем выше коэффициент вариации, тем менее однородны структуры цен и менее надежны оценки ППС. Значение, превышающее 33%, является экстремальным.

Во время проверки экстремальных значений в первую очередь следует уделять внимание первичным группам, где значение коэффициента вариации превышает 33%, особенно если их вес расходов высок. Первичные группы с высокими весами расходов окажут большее влияние на общую оценку ППС, чем группы с низкими весами расходов. Веса расходов представлены только в таблице Куаранта.

Значение коэффициента вариации должно снижаться по мере осуществления валидации данных, тем самым обеспечивая средство для оценки общей эффективности процесса валидации.

- *Коэффициент вариации для страны* дает оценку разброса РРР-соотношений для первичной группы в стране. Другими словами, он измеряет вариацию уровней цен в стране для товаров, обследованных в первичной группе, и надежность ее оценок ППС для первичной группы. Чем выше значение коэффициента, тем менее однородны уровни цен в стране и менее надежны ее оценки ППС. Значение, превышающее 33%, является экстремальным.

Во время проверки НКА должны в первую очередь уделять внимание первичным группам, где значение коэффициента вариации для страны превышает 33%, особенно если их вес расходов высок.

Значение коэффициента вариации должно снижаться по мере осуществления валидации данных. Это помогает НКА оценить эффективность валидации.

Коэффициент вариации для страны дополняет коэффициент вариации для товара (следующий пункт), позволяя взглянуть с другой стороны на тот же набор данных. Акцент на странах, а не продуктах может помочь обнаружить страны, чьи данные являются подозрительными. В этом отношении лучше использовать таблицу Дикханова, где товары представлены более широко, для выявления ошибок стран.

- *Коэффициент вариации для товара* – самый важный из коэффициентов вариации для целей валидации, он измеряет разброс между РРР-соотношениями для товара. Он является показателем сравнимости и точности и позволяет поставить вопрос о том, был ли обследован НКА один и тот же товар или его эквивалент и был ли он оценен правильно. Чем выше коэффициент вариации, тем менее однородны уровни цен на товар и тем больше подозрений вызывает сопоставимость товара и точность обследования цен в странах. Такие товары являются кандидатами на разбиение или удаление, и РКА должно обеспечить их тщательное исследование НКА. Во время валидации данных следует уделять первостепенное внимание товарам с коэффициентом вариации выше 33%.

- *XR-соотношения* - это стандартизированные отношения цен, полученных конвертированием средних цен с использованием обменных курсов. Этот косвенный показатель позволяет сравнивать средние цены товара в разных странах. Он облегчает выявление экстремальных значений среди отношений цен для товара в начале процесса межстрановой валидации, когда оценки ППС и РРР-соотношения еще ненадежны, так как они рассчитываются на основе средних цен, которые проходят валидацию. Сначала, в ходе первого и второго раундов валидации, следует исследовать XR-соотношения со значениями за пределами диапазона 80-125. В последующих раундах, когда оценки ППС становятся более надежными, экстремальные значения среди XR-соотношений можно игнорировать. Для анализа этого показателя лучше обратиться к таблице Куаранта, так как в ней все ряды ясно представлены и экстремальные значения отмечены.

- *PPP-соотношения* - это стандартизированные отношения цен, полученных конвертированием средних цен с использованием ППС. Это правильный показатель, с которым следует сравнивать средние цены на товар в разных странах и средние цены на разные товары в стране. В ходе межстрановой валидации данных пытаются обнаружить и проверить экстремальные значения именно среди этих ценовых соотношений для товара. Соотношения за пределами диапазона 80-125 должны быть исследованы.
- *Коэффициент вариации наблюдений цен* измеряет вариацию наблюдений цен, на основе которых рассчитана средняя цена товара, сообщенная страной. Он берется напрямую из таблицы средних цен. Он используется для обнаружения экстремальных значений среди средних цен в ходе внутристрановой валидации данных, когда средние цены с коэффициентом вариации выше 20% считаются экстремальными значениями. Если после проведения внутристрановой валидации данных коэффициент вариации остается выше 20%, НКА может потребоваться провести повторное редактирование наблюдений цен, если имеются экстремальные значения среди PPP-соотношений для товара или если коэффициент вариации для товара превышает 33%.
- *Индексы уровня цен* измеряют различия в уровнях цен между странами для первичной группы. Они полезны для сравнения согласованности относительных уровней цен по первичным группам.

---

## ПРИМЕЧАНИЯ

1. Когда спецификации продукта плохо определены, проблема состоит не столько в том, что оцененный товар не соответствует товару из списка (он, возможно, соответствует), но в том, что он не сопоставим с товарами, которые оценили другие сборщики цен, как внутри страны, так и в других странах, используя те же самые спецификации. Внутри страны эта проблема может быть уменьшена НКА и сборщиками цен, которые договорятся о том, как интерпретировать спецификации и какие товары следует оценивать до начала сбора цен. Это не устранит различий между странами, которые станут очевидными только на этапе валидации данных.
2. Это может случиться, когда спецификации товара касаются определенной модели, которая в некоторых странах находится в процессе замены на более новую модель, что нередко происходит в случае предметов длительного пользования домашних хозяйств.
3. Такая ситуация усиливает понимание того, что сборщики цен должны пройти хорошее обучение и получить четкие инструкции и формы для регистрации цен, которыми удобно пользоваться. Следует проводить мониторинг полевой работы, чтобы убедиться, что сборщики записывают цены, количества и другие необходимые данные. Обучение и мониторинг полевой работы не исключают ошибки товара, но сокращают их появление.
4. Даже если цена и указанное количество записаны правильно, ценовая ошибка может быть внесена, если цена неправильно приведена к базисному количеству.
5. Термины '*экстремальное значение*' и '*выброс*' часто используют как синонимы. В данном тексте термин '*выброс*' зарезервирован для экстремального значения, правильность которого была проверена. При обнаружении '*экстремальное значение*' является возможной ошибкой, после проверки - это либо 'ошибка', либо 'выброс'.
6. Разные товары характеризуются разными профилями распределения. Некоторые товары продаются в основном в супермаркетах, другие - в основном, в специализированных

---

магазинах.

7. Цены на один и тот же товар могут быть разными в разных торговых точках, так как они продаются в разных условиях и при разных обстоятельствах. Условия и обстоятельства торговли составляют элемент обслуживания. Если элемент обслуживания меняется по типам торговых точек, то покупаемый товар не является одним и тем же в разных торговых точках, даже если он идентичен физически. Фактически покупается составной товар – то есть сам товар плюс элемент обслуживания. Различия в элементе обслуживания вызывают разницу в качестве, которая влияет на разницу в цене. Когда те, кто занимается расчетом средних цен товара, не принимают во внимание различия элементов обслуживания в тех торговых точках, где велось наблюдение цен, средние цены, возможно, будут либо завышены, либо занижены.

8. Приведут ли эти средние цены к несмещенным ценовым отношениям зависит от того, имеют ли обследованные товары такие же профили распределения в других странах, участвующих в сопоставлении, и от того, имеют ли выбранные торговые точки аналогичные элементы обслуживания во всех странах.

9. Первоначально предполагалось, что РКА будут участвовать в процессе внутристрановой валидации для обеспечения использования общего жесткого стандарта в регионе. В некоторых регионах РКА участвует в процессе, но в других – нет, из-за строгой интерпретации конфиденциальности данных странами региона. В тех регионах, где РКА участвует во внутристрановой валидации данных, страновые наблюдения цен рассматриваются только НКА и РКА; они не предоставляются больше никому ни внутри, ни за пределами региона. Опыт показывает, что участие РКА во внутристрановой валидации данных может быть контрпродуктивным. Когда РКА участвует в процессе, НКА обычно не проявляют инициативы, ожидая действий от РКА.

10. Важность своевременного завершения валидации данных трудно переоценить. Цены на потребительские товары обычно собирают на месячной основе, а редактирование и проверка проводятся раз в квартал. Необходимо обработать большое количество наблюдений цен. Нужно приложить все усилия, чтобы предотвратить отставание, которое может привести к задержке региональных сопоставлений (и возможно, глобального сопоставления). Задержки могут оказывать отрицательное воздействие на качество. Чем больше времени занимает валидация, тем труднее становится исправлять ошибки, выявленные в ее процессе. Более того, для сокращения отставания и наверстывания упущенного времени, РКА и НКА могут иметь соблазн проводить редактирование и проверку менее тщательно.

11. Например, в табл. 9.1а базисное значение составляет 500 г (поля 03 и 04), но сборщики цен могут оценить любое количество в диапазоне 400–600 г (поле 02). Цены за обследованные количества (поля 10, 11, и 12) приведены к цене за 500 г (поле 16). Именно скорректированная цена в поле 16 участвует в валидации и в усреднении для получения средней цены (поле 24) в табл. 9.3а.

12. При расчете ППС для первичной группы проводят различия между важными и неважными товарами. Странам необходимо указать, какие из обследованных товаров в первичной группе для них являются важными. Товар является важным, если он продается в больших объемах. Важные продукты отмечаются звездочкой (\*).

13. Таблица наблюдений цен и таблица средних цен связаны. Средние цены в таблице средних цен получают на основе индивидуальных цен из таблицы наблюдений цен. Изменение индивидуальной цены немедленно приведет к изменению средней цены. Каждое значение средней цены может быть «раскрыто», чтобы показать на каких индивидуальных

ценах оно основано. Таким образом, если средняя цена отмечена как экстремальное значение, можно вызвать соответствующие наблюдения цен, чтобы посмотреть, не отмечено ли какое-то из них как экстремальное значение и причина экстремального значения средней цены. Отсюда следует, что две стадии внутристрановой валидации можно провести в обратном порядке – то есть начать со средних цен, а не с наблюдений цен. Недостаток этого подхода состоит в том, что можно пропустить дополнительную информацию, предоставляемую сборщиками цен.

14. Стандартизированное соотношение цен равно

$$\left( \frac{CC - Price_{1A}}{[CC - Price_{1A} * CC - Price_{1B} * \dots * CC - Price_{1N}]^{\frac{1}{N}}} \right) 100, \text{ где } CC - Price_{1A} - \text{средняя цена}$$

товара 1 в стране А в общей валюте.

$CC - Price_{1A}$  равно  $\frac{NC - Price_{1A}}{C C_{1A}}$ , где  $NC - Price_{1A}$  - средняя цена товара 1 в стране А в национальной валюте, а  $C C_{1A}$  - курс конвертирования валюты между валютой страны А и общей валютой. Курс конвертирования валюты – это либо обменный курс валют, либо ППС:  $CC_{1A} = XR_{1A}$  или  $PPP_{1A}$

15. Также известны как CUP-соотношения. CUP означает условные единицы для выражение паритетов. Обычно это искусственная валютная единица, например стандарт покупательной способности (PPS), используемый Евростатом.

16. Это агрегат из классификации расходов, например, продукты питания и напитки, одежда и обувь, конечное потребление домашних хозяйств или ВВП.

17. Метод CPRD следует использовать для расчета ППС для таблиц Куаранта и Дикханова только в том случае, если информация о важности товаров является надежной. Опыт подсказывает, что это не всегда так. Поэтому рекомендуется сначала получать таблицы Куаранта и Дикханова с использованием метода CPD и зарезервировать использование метода CPRD для последних итераций и для валидации важности товаров. Большие различия между остатками регрессии CPD/CPRD и стандартными отклонениями будет указывать на проблемы с данными, особенно с выбором важных продуктов. То же самое верно для метода EKS\*. ППС не должны рассчитываться для таблицы Куаранта этим методом, если индикаторы важности вызывают подозрения. Вместо этого следует применять метод GEKS, во всяком случае, на ранних стадиях валидации.

18. Формат таблицы Куаранта, которая используется с начала 1990-х, был разработан таким образом, чтобы ее можно было напечатать на бумаге формата А4.

19. Это не всегда было так. Например, первые версии таблицы Дикханова не содержали XR-соотношений, на которые направлено внимание в начале процесса валидации.