ОРГАНИЗАЦИЯ ОБЪЕДИНЕННЫХ НАЦИЙ

E

ЭКОНОМИЧЕСКИЙ И СОЦИАЛЬНЫЙ СОВЕТ

Distr.
GENERAL

ECE/CES/GE.22/2008/5
26 February 2008

RUSSIAN
Original: ENGLISH

ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ

КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ

Совместное совещание ЕЭК ООН/МБТ по индексам потребительских цен

Девятое совещание
Женева, 8-9 мая 2008 года
Пункт 3 предварительной повестки дня

СБОР И ОБРАБОТКА ДАННЫХ О ЦЕНАХ

РОЛЬ АЛГОРИТМА ТУКИ В ПРОЦЕДУРАХ ПОДТВЕРЖДЕНИЯ ДАННЫХ О ЦЕНАХ В ИНДЕКСЕ ПОТРЕБИТЕЛЬСКИХ ЦЕН: ОПЫТ СОЕДИНЕННОГО КОРОЛЕВСТВА В ЭТОМ И БОЛЕЕ ОБЩИХ АСПЕКТАХ РЕДАКТИРОВАНИЯ ДАННЫХ

Записка Бюро национальной статистики Соединенного Королевства

Резюме

Процедуры редактирования, используемые для выявления и исправления ошибок в данных о ценах, регистрируемых в торговых точках переписчиками данных о ценах, - это немаловажный вопрос, поскольку они могут оказывать систематическое численное воздействие на измерения инфляции, которое способно привести к систематической ошибке. Используемые процедуры имеют и практические последствия. В случае их правильного применения процедуры редактирования позволяют не только повысить качество индекса цен, но и добиться улучшений в практике составления индекса потребительских цен и индекса розничных цен. В этом документе представлены результаты определенной работы, проделанной Бюро национальной статистики

GE.08-21369 (R) 150408 160408

Соединенного Королевства в связи с применением алгоритма Туки, и рассматриваются вопросы, которые возникают для разработчиков индексов потребительских цен. В нем также сообщается о более подробном исследовании, проведенном по вопросу о воздействии редактирования на индекс цен применительно к одежде. Работа по обоим вопросам была проведена несколько лет назад и привела к внедрению пересмотренных и усовершенствованных процедур редактирования. Эта работа была начата в контексте индекса розничных цен, однако в равной мере касается и индекса потребительских цен[1].

Ключевые слова: ИПЦ, Туки (Тьюки), процессы подтверждения данных, применение, эффективность, точность, индекс одежды (швейных изделий).

I. ВВЕДЕНИЕ

1. Для индекса потребительских цен (ИПЦ) и индекса розничных цен (ИРЦ) Соединенного Королевства (СК) в центральном аппарате Бюро используются два отдельных компьютерных алгоритма для выявления резко отклоняющихся значений цен, собранных на местах в магазинах, т.е. экстремумных цен, которые могли бы оказать относительно большое влияние на индексы цен соответствующего товара. Все цены взвешиваются для составления публикуемых субиндексов Индекса всех товаров. В ходе работы над этим исследованием было принято то предположение, что резко отличающееся значение неправильно, и поэтому считается недействительным, если только впоследствии оно не подтверждается путем редактирования, т.е. не будет принято в качестве действительного значения цен и определено как верное после обязательного подтверждения в результате изучения всех метаданных, направленных счетчиком, собиравшим данные о ценах, и в результате соответствующих последующих действий. К последним относятся поручение счетчику вновь собрать данные о ценах и позвонить в магазин, чтобы подтвердить правильность цены, хотя и могло быть связано со всего лишь оценкой сотрудников центрального аппарата, основанной на метаданных, представленных счетчиками.

---

[1] ИПЦ - главный показатель внутренней инфляции для макроэкономических целей и в международном контексте обозначается как гармонизованный индекс потребительских цен. Его составление начато в 1997 году и стало контрольным показателем инфляции в СК в 2003 году. ИРЦ имеет гораздо более длительную историю. Он начал свою жизнь как индекс компенсации во время первой мировой войны и только гораздо позже стал использоваться как главный внутренний показатель инфляции для макроэкономических целей. Подробнее см. “The new inflation target: the statistical perspective”, Roe and Fenwick, Economic Trends, January 2004.

2. Поскольку цены в магазинах могут меняться, а график составления ИПЦ/ИРЦ весьма жесткий, редактирование должно производиться в короткий период времени. В идеале редактирование и проверка должны быть близки к "реальному времени". Длительное и бессистемное изучение цен требует слишком больших затрат времени и не представляется возможным на практике. Проверка и редактирование "в реальном времени" на местах также призваны сыграть важную роль. В ходе сбора данных о ценах в СК оказалось полезно использование миниатюрных компьютеров, которые позволяют производить интерактивное редактирование цен в момент сбора данных о ценах[2] и до того, как данные поступают в центральный аппарат. Редактирование в центральном аппарате производится после редактирования на местах.

3. Первый алгоритм, используемый для выявления резко отклоняющихся значений, охватывает две проверки, которые проводятся по каждой сообщенной цене. Эта проверка "минимум-максимум" и проверка процентного изменения, которые определяют соответствующие цены или ценовые изменения, которые лежат за пределами заранее определенного интервала. Последний определяется подразумеваемым образом на основе средних цен предыдущего месяца для данного товара и прямо выраженным образом путем установления соответствующих параметров. При таком редактировании основное внимание уделяется уровню и изменению цен конкретного товара, цена которого определяется в данном магазине, по отношению к среднему уровню и изменению цен для "того же/сходного" товара во всех магазинах. Оно отличается от интерактивного редактирования на местах, которое производится путем сопоставления с динамикой цен на данный товар в данном конкретном магазине, а также путем выявления любых очевидных несоответствий, когда работникам магазина предлагается подтвердить правильность цены. Такая оперативная процедура в центральном аппарате - первый алгоритм - получила название "изучение".

4. Второй используемый алгоритм, алгоритм Туки, применяется в отношении тех сообщенных данных о ценах, которые не определены как явный выброс в результате процесса "изучения". Алгоритм Туки по своей сути - более изощренный вариант "изучения". Более подробно об этом сказано в следующем разделе.

5. Сообщенные данные о ценах, которые определены как резко выделяющиеся значения по любому из этих алгоритмов, проходят дальнейшую проверку и редактирование. В момент работы над исследованием в результате "изучения" в качестве

---

[2] Например, использование внешних границ применительно к изменениям цен по сравнению с предыдущим месяцем.

посторонних значений было выделено в среднем порядка 2 000 сообщенных данных о ценах из общего количества свыше 100 000, главным образом, по критерию процентного изменения, а еще около 4 000 сообщенных данных о ценах было определено в качестве явных выбросов алгоритму Туки. Все резко отличающиеся значения, определенные в результате "изучения", и половина выбросов, определенных в результате использования алгоритма Туки, подвергались дополнительной проверке, например путем перепроверки цены в магазине или звонка в магазин с просьбой подтвердить цену или же путем принятия решения после изучения метаданных.

6. Редактирование резко отличающихся значений в результате "изучения", как правило, по-прежнему проводится после проверки по алгоритму Туки, поэтому оба процесса часто используются параллельно. Если большинство резко отличающихся значений (определяемых по любому алгоритму), которые впоследствии проверяются, приняты после ручной проверки, когда составитель ИПЦ/ИРЦ изучает имеющиеся данные и принимает решение о том, что цена правильна и не должна быть пересмотрена, после этого результаты автоматического редактирования заменяются вручную. Таким образом, после этих строгих процедур отсеивается очень мало резко отличающихся значений. Число сообщенных данных, которые прямо принимаются на этом этапе в процессе редактирования, примерно в 100 раз превышает число прямо отбракованных данных о ценах.

7. Преимущество использования такого рода алгоритмов в качестве механизмов фильтрации заключается в том, что они позволяют избежать необходимости изучения аналитиком цен огромного числа результатов наблюдения цен. Однако опыт СК в общем виде поддерживает мнение, принятое в Руководстве по индексам потребительских цен МОТ, что использования автоматизированных систем отсева без возможности ручной проверки следует избегать, поскольку изменения цен между месяцами могут быть весьма значительными, например, из-за распродаж, как и в пределах месяца между различными сортами товара из-за, например, сезонности. Это в равной степени касается и статистической проверки вводимых данных там, где за данный период времени каждое изменение цен сопоставляется с изменением цен в полной выборке данного рассматриваемого товара, и менее сложных систем, основанных, например, на использовании заранее определенных пределов, которые автоматически не пересматриваются в процессе редактирования. Именно в силу этих моментов и результатов исследовательской работы, проведенной в СК, было отброшено то допущение, что резко отличающиеся значения неправильны.

II. АЛГОРИТМ ТУКИ

8. Алгоритм Туки - пример системы фильтрации, основанной на статистической проверке вводимых данных.

9. Для применения метода Туки сообщенные данные о ценах расставляются в порядке соответствующих соотношений цен и самые высшие и самые низшие 5% выделяются для последующей проверки и исключаются. Отношение цен, равное единице (т.е. цена не изменилась), также исключается. Рассчитывается арифметическая средняя оставшихся отношений цен - эквивалентная определенному усеченному среднему, - и эта средняя используется для разбиения остальных отношений цен на две группы и расчета их соответствующих средних (обозначаемых как верхняя и нижняя "среднесредние" - AM_U и AM_L). Затем следующим образом рассчитываются верхний и нижний пределы Туки, которые используются для того, чтобы пометить те наблюдения о ценах, которые требуют внимания:

$T_U = AM + 2.5 (AM_U – AM)$

$T_L = AM – 2.5 (AM – AM_L)$,

где AM_L - более низкая усеченная средняя и AM_U - более высокая усеченная средняя.

10. Особое достоинство алгоритма Туки заключается в том, что, если отвлечься от метаданных, он позволяет максимизировать использование непосредственной ценовой динамики и дает более практичные и реалистичные параметры, поскольку он исключает случаи отсутствия изменения цен. Разумеется, он основывается на достаточно большом числе наблюдений о ценах. Алгоритм Туки может использоваться применительно к любому периоду времени и поэтому может применяться для изучения как помесячных, так и годовых изменений. Верхний и нижний пределы определяются данными и могут регулярно пересчитываться по наборам текущих данных о ценах. Это и использование непосредственной ценовой динамики может быть особым преимуществом, когда инфляционная ситуация и цены меняются быстро.

III. ПРОБЛЕМЫ

11. Возникают две главные проблемы:

a) эффективность процедур редактирования;
b) воздействие на точность ИПЦ/ИРЦ.

Последнее включает возможные систематические ошибки. Эти два вопроса эффективности и воздействия, разумеется, взаимосвязаны. Рассмотрим их по очереди.

Эффективность процедур редактирования

12. Эффективность процедур редактирования есть функция:

a) общей системы обработки данных, включая полные процессы, созданные для проверки и редактирования данных;

b) роли, выполняемой алгоритмом Туки в общей системе;

c) конкретного способа применения алгоритма.

13. Если взглянуть на процесс проверки данных в целом, то следует отметить значительное перекрывание между тремя основными процессами редактирования: интерактивным редактированием на местах, "изучением" и алгоритмом Туки.

a) Интерактивное редактирование на местах реализуется как ряд проверок данных в реальном времени, встроенных в миниатюрные компьютеры. После того как счетчик вводит данные о ценах вместе с полезными метаданными (в том числе является ли цена специальной "распродажной" ценой, а также какой имеется замещающий товар в случае, если предыдущий товар уже отсутствует в запасах), программа сбора данных проверяет введенные данные с использованием ряда встроенных правил. В их числе - минимальный и максимальный ценовые диапазоны, основанные на ценах, введенных в миниатюрный компьютер в предыдущий месяц на абсолютно тот же товар в том же магазине, и ряд логистических проверок, например, что цена, установленная по окончании распродажи, заносится только после цены, введенной в предыдущий месяц в качестве цены "распродажи". Каким бы ни был объем редактирования в центральном аппарате, он не сможет заменить интерактивное редактирование на местах как наиболее эффективное средство обеспечения точности регистрируемых цен.

b) "Изучение" в отличие от "интерактивного редактирования данных" производится не в "реальном времени" и представляет собой менее совершенный процесс редактирования данных после сбора, чем алгоритм Туки. Относительные достоинства "изучения" могут зависеть отчасти от его взаимодействия с двумя другими процессами, в особенности с алгоритмом Туки. Следует отметить, что изучение стало проводиться до того, как для сбора данных стали использоваться миниатюрные компьютеры и до внедрения более совершенных процессов редактирования в центральном аппарате. Последние два нововведения стали элементом ряда мер по повышению точности и надежности ИПЦ/ИРЦ и качества управления процессами сбора для уменьшения риска ошибок. Из этих трех процессов редактирования "изучение" использует данные наименее "интеллектуальным" образом. Оно во многом продолжает существовать по инерции - цены раньше собирались на бумаге и получались "пакетами", поэтому нужна была программа выявления, которая в отличие от алгоритма Туки могла бы быть запущена без привязки к основному комплексу данных о ценах. В значительной мере данные и теперь представляются пакетами, прежде чем первоначальный набор "ценовых" данных загружается в компьютер, однако преимущества "изучения" в плане своевременности менее существенны, поэтому в момент проведения исследования вопрос заключался в том, в достаточной ли мере "изучение" повышает качество окончательных данных о ценах. Его главное преимущество, как считалось, заключалось в том, что благодаря ему можно выявлять крайние резко выделяющиеся значения и большие различия в качестве, если вместо исчезнувших товаров были отобраны товары-заместители. Он также позволяет учесть цены, перепроверенные самими счетчиками.

c) Алгоритм Туки более совершенен, чем интерактивное редактирование цен на местах с использованием миниатюрных компьютеров, равно как и процесс "изучения" в центральном аппарате, однако не позволяет использовать специфические преимущества интерактивного редактирования на местах. В отсутствие доказательств обратного есть основания считать, что редактирование в центральном аппарате должно быть связано в основном с теми резко выделяющимися значениями, которые определены с помощью алгоритма Туки. Это тем более так, поскольку значительно число сообщенных данных, определенных в результате "изучения" в качестве резко выделяющихся значений и впоследствии подтвержденных как правильные, не классифицировались бы как резко выделяющиеся значения в результате проверки с помощью алгоритма Туки. Поэтому в определенной мере деятельность по редактированию направлена не в ту сторону. Такая "неэффективность" может быть усилена из-за непропорционального числа резко выделяющихся значений, определенных в результате "изучения", которые впоследствии оказались правильными. С другой стороны, "изучение" может редактировать экстремумные сообщенные данные о ценах, которые вызывают "крен" при использовании алгоритма Туки.

14. Очевидно, что та роль, которую играет алгоритм Туки, может быть подорвана, а его результативность может быть уменьшена из-за "изучения". Прежде всего, цены, которые первоначально были определены как резко выделяющиеся в результате "изучения" и тем не менее были подтверждены в последующий период, исключаются из набора данных, по которому проводится проверка Туки, даже если их включение позволило бы сделать более содержательным показатель изменения цен (или цены) по каждому товару, из которых выводятся параметры допустимости для Туки и тем самым резко выделяющиеся значения по Туки. Кроме того, резко выделяющиеся значения, определенные в результате "изучения", которые были подтверждены как недействительные, вполне могли бы быть включены в набор данных, используемых для алгоритма Туки, поскольку алгоритм Туки, применяемый при редактировании ИПЦ/ИРЦ, в любом случае отбрасывает 10% данных. Таким образом, можно было бы утверждать, что в практическом плане мало что теряется при отказе от "изучения", хотя это означало бы, что проверка данных и их редактирование должны были бы быть отложены на несколько дней до получения достаточного числа данных о ценах для расчета параметров Туки. Этот аргумент был рассмотрен, но отвергнут. Была принята та точка зрения, что, несмотря на снижение эффективности алгоритма Туки, "изучение" в целом дает положительный эффект, поскольку благодаря ему можно очень быстро и просто выявить крайние резко выделяющиеся значения. Однако также было принято мнение, что эффективность "изучения" могла бы быть повышена путем указания для данных о ценах весового коэффициента "расходов"вместе с другой полезной информацией, которой мог бы руководствоваться редактор, принимая свои решения в процессе "изучения".

15. Что касается вопроса о применении, что та посылка, что резко выделяющееся значение "неправильно, пока не будет доказано обратное", может быть оспорена и она была оспорена как ошибочная. Тот факт, что сообщенные значения, определенные в качестве резко выделяющихся с использованием любого из этих алгоритмов, были исключены из индекса, если только впоследствии они не были подтверждены в результате редактирования, обоснованы только в той мере, в какой набор резко выделяющихся значений перекрывается с набором неправильных цен. Для резко выделяющихся значений, определяемых в процессе"изучения", это неважно, поскольку они проверяются все. В случае резко выделяющихся значений, определяемых по алгоритму Туки, предполагаемое соотнесение соответствующей динамики цены (или цены) с показателем общего изменения цен (или ценами) данного товара, в результате которого это значение цены было определено в качестве резко отличающегося, могло бы дать достаточно дополнительной информацией, дополняющей ту, которая имеется у счетчика, чтобы с реальным основанием считать цену недействительной. При этом изучение данных показывает, что доля резко выделяющихся значений, определенных с использованием

алгоритма Туки (и "изучения"), которые были прямо отброшены в результате перепроверки цен или использования обоснованных суждений, весьма мала. Если предположить, что редактирование дает правильные решения, то предположение об исключении резко выделяющегося значения представлялось бы неправильным. Ввиду этого набор резко выделяющихся значений может не перекрываться надлежащим образом с набором неправильных цен[3], и в этом случае следовало бы уделить внимание скорее сообщенным данным о ценах на данный товар в данных торговых точках, которые остались неизменными на необычно длинные периоды, а также более экстремальным очевидным выбросам.

Влияние на точность ИПЦ/ИРЦ

16. Исследования БНС показали, что теоретические последствия для ИПЦ/ИРЦ использования алгоритма Туки, в том виде как он ранее применялся в период проведения исследования, могли бы заключаться в занижении индекса из-за отсева изменений цен, скорее в сторону повышения, а не понижения, и, что если только это не будет скорректировано процессом редактирования, может привести к систематической ошибке публикуемого индекса. Это результат двух вещей:

a) смещенного распределения ценовых изменений;

b) того, что в алгоритме не используются метаданные, указывающие на то, не являются ли размеры изменения цен необычными, например из-за продаж и специальных предложений. Эта информация хранилась в базе данных о ценах в виде легко выводимых обозначений.

17. Потенциальная систематическая ошибка может быть усугублена неиспользованием метаданных, представленных вместе с ценами. В частности, это касается того случая, когда цены на данный товар подвержены крайне резким колебаниям, вызванным, например, распродажами или сезонностью.

18. Та гипотеза, что резко выделяющееся значение неправильно и поэтому признается недействительным, если только впоследствии оно не подтверждается в результате редактирования, была не только необоснованна и могла привести к появлению ошибок и систематического смещения, как об этом говорилось выше, но и противоречила

---

[3] Этот вопрос можно было бы дополнительно изучить, сопоставив наборы выбросов с базой данных собранных на местах цен, которые не прошли перепроверку на точность.

регламентациям ЕС, касающимся составления Гармонизованного индекса потребительских цен (ГИПЦ). Сотрудники центрального аппарата были переобучены, чтобы всегда перепроверять цену, если только не имеется явных признаков того, что это не нужно.

19. Кроме того, в ходе исследования его масштабы были расширены и внимание было заострено на следующих моментах:

a) автоматизированной компьютерной процедуре в рамках программы обработки цен ИПЦ/ИРЦ, которая меняет пометки всех сообщенных цен с несопоставимых на сопоставимые, если изменение цены при занижении одного товара другим считается достаточно небольшим;

b) предоставлении сверщикам цен в центральном аппарате права отменять решения, принятые счетчиками на местах по таким вопросам, как сопоставимость замещающего товара. Решение сверщиков цен основывается на метаданных, предоставленных счетчиками, и там, где имеются сомнения в том, совместим ли замещающий товар, еще и на изучении сопоставимой цены и использования знания рынка.

Эти два момента и находились в центре второго исследования.

20. Резюмируем уже изложенное: как и многие процессы обработки данных, используемые национальными статистическими ведомствами, система ИПЦ/ИРЦ - продукт истории, эволюционирующий многие годы. Можно считать, что прогресс технологии, в частности использование миниатюрных компьютеров для сбора данных о ценах и электронной передачи файлов, содержащих цены, в центральный аппарат, означает, что те же проверки данных предусмотрены на различных этапах редактирования, что ведет к повторам. Такое отсутствие системного подхода может на практике привести к неэффективности. Однако еще большее значение имеет возможное влияние на измерение инфляции.

21. В следующем ниже разделе представлены результаты второго исследования, о котором говорилось выше. Речь идет о более подробном исследовании потенциального воздействия редактирования на индекс цен на одежду, когда распродажи и сезонность могут привести к большим краткосрочным колебаниям цен. В нем дан более подробный анализ, на основе которого можно получить полезное представление о некоторых вопросах, которые могут возникать.

IV. ОДЕЖДА

22. Исследование, проведенное БНС по вопросу о субиндексе одежды индекса всех товаров, было первоначально вызвано интересом к этому индексу из-за очевидной понижательной динамики цен, которая была зафиксирована и в некоторых других, хотя и не во всех, странах Европейского союза.

Фундаментальный анализ

23. В июле 2002 года индекс одежды и обуви упал ниже 100 - впервые с летних распродаж 1987 года. На диаграмме 1 ниже показан уровень индекса с января 1987 года, при этом линия скользящей средней усилила свое резкое снижение с 1998 года, которое последовало за семью годами устойчивого роста цен. С января 1998 года по январь 2002 года индекс скользящей средней[4], показанный кривой на диаграмме 1, снизился примерно на 14%. Возник вопрос, действительно ли цены на одежду стали ниже, чем 15 лет назад на сходные товары.

Индекс группы одежды (база - январь 1987 = 100)

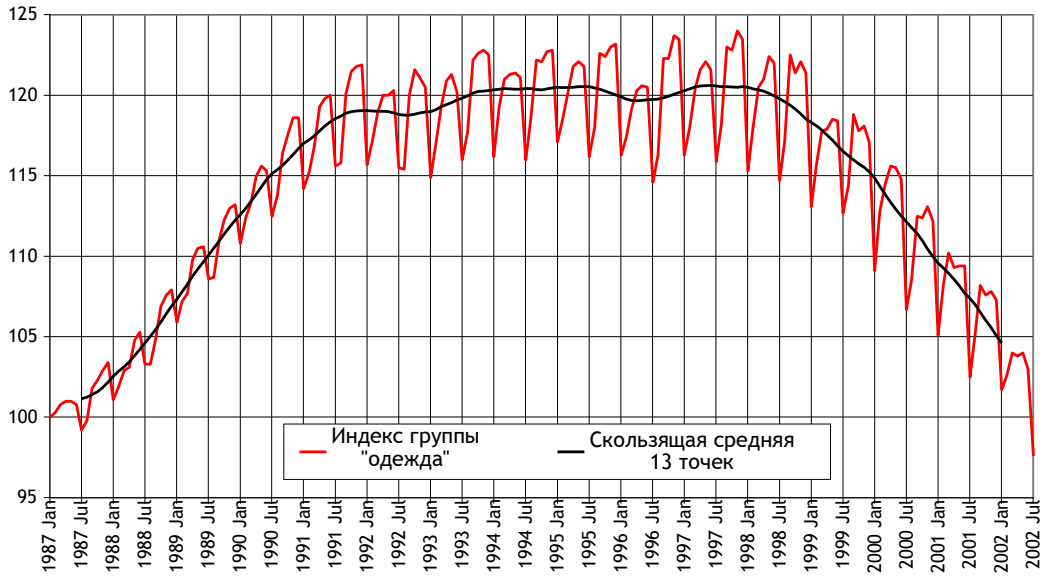


Диаграмма 1

---

[4] Скользящая средняя - *центрированный* вариант из 13 точек, когда обе краевые точки составляют половину суммы весового коэффициента других точек.

24. На диаграмме 2 ниже проиллюстрирован индекс одежды в США. Это - ряд, наиболее сопоставимый с индексом одежды и обуви СК. Он охватывает все США, представляет собой версию без поправки на сезонность и включает как мужскую, так и женскую одежду и обувь. Ряд был пересчитан к базе января 1987 года, чтобы его можно было прямо сопоставлять с данными СК. Диаграмма показывает сходную динамику индекса СК с сезонными распродажами и возвращением к нормальным ценам. Рост цен начинает тормозиться в начале 1990-х годов, а к концу десятилетия цены начинают падать. В последнем случае снижение выражено гораздо меньше, чем индекс СК, когда скользящая средняя с января 1998 года по январь 2002 года снизилась примерно на 6% по сравнению с 13% по ряду СК. Разница в размере падения цен остается, даже если учесть экономические факторы, такие, как значительное снижение курса фунта стерлинга к концу этого периода. Этот вопрос еще не удалось прояснить, особенно в свете того, что в некоторых других странах, например в Швеции, какого-либо снижения цен не отмечалось.

Индекс группы "одежда" США (приведено к новой базе, январь 1987 = 100)

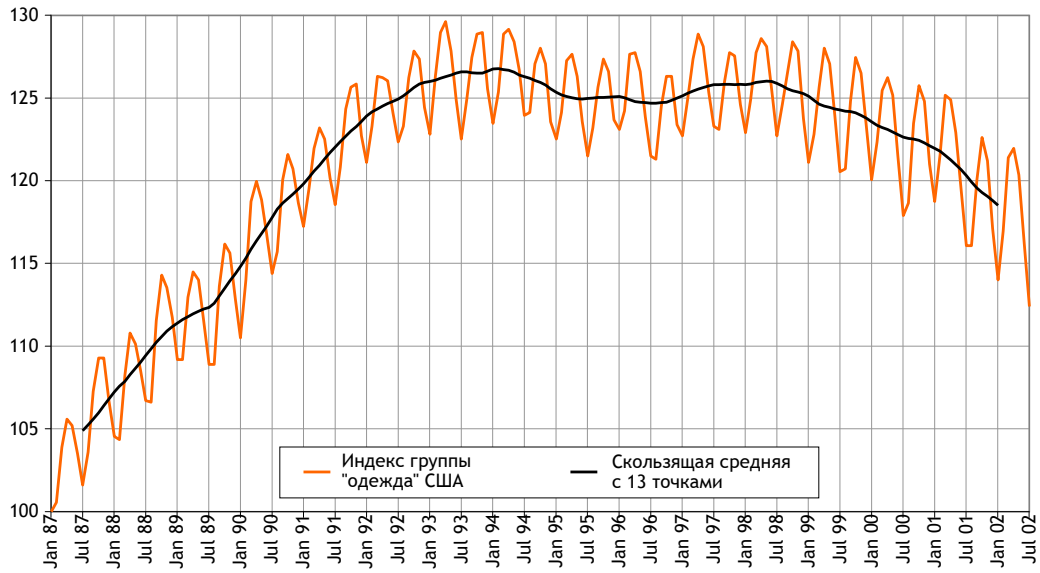


Диаграмма 2

25. На диаграмме 3 проиллюстрирован временной ряд для всех разделов группы "одежда и обувь" СК. Четыре из них демонстрируют схожую динамику - первоначальный рост цен, за которым следует период ценовой стабильности в середине 1990-х годов до ускоряющегося падения цен к концу десятилетия. Исключение - "другие предметы одежды", крайне разнородная группа, где рост продолжался и в конце 1990-х годов, и снижение произошло только в последние годы.

26. В частности, следует отметить периоды распродаж, в СК - традиционно примерно в январе и июле, и восстановление нормальных цен после них. Как видно из диаграммы 3, в более недавние годы снижение в ходе сезонных распродаж было более глубоким, а восстановление цен по большинству разделов - менее значительным. В особенности в разделе женской верхней одежды даже между циклами распродаж/восстановления цен цены за последние несколько лет рассматриваемого периода существенно снижались, указывая на то, что снижение индекса вызвано не только сезонностью и "модой".

27. Если посмотреть на отдельные позиции в разделе мужской и женской верхней одежды, то очевидно, что имеются некоторые товары, которые демонстрировали более существенное понижательное движение цен по сравнению с другими товарами. Например, предварительные исследования по женской верхней одежде показали, что индексы ИПЦ/ИРЦ для таких товаров, как жакеты, блузы и строгие платья, снизились в период с 1995 года по 2001/02 год примерно на 40%. С другой стороны, цены на некоторые товары, по-видимому, были стабильными или даже демонстрировали небольшой рост.

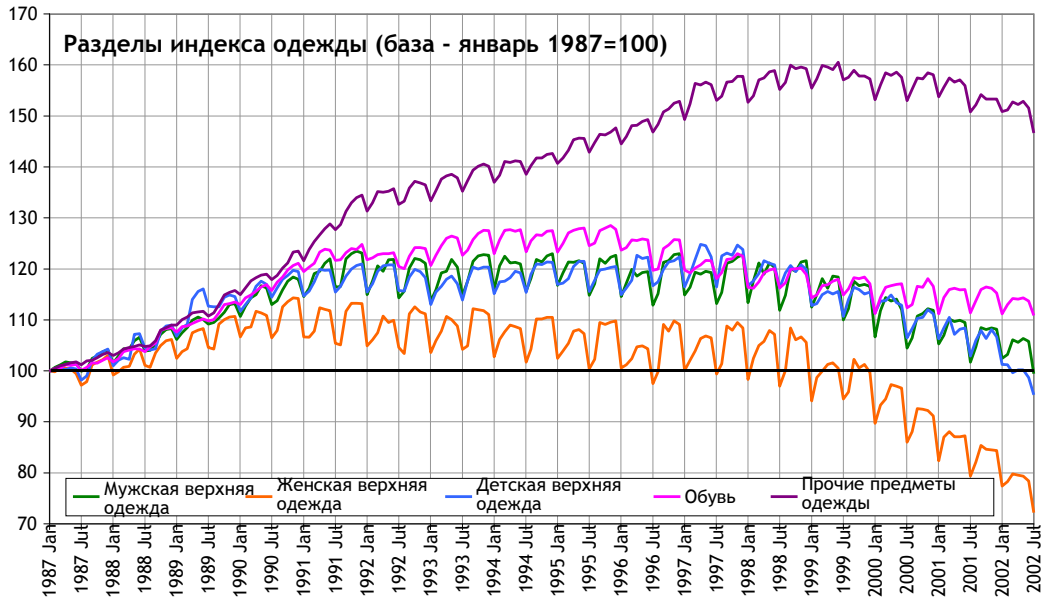


Диаграмма 3

Исследования

28. Для того, чтобы попытаться понять, что же происходило на самом деле и идет ли речь о проблеме редактирования и возможного вызванного им смещения, был проведен ряд более глубоких исследований. В ходе исследований была предпринята попытка рассмотреть ряд конкретных вопросов.

29. Является ли оценка редактора данных в центральном аппарате более надежной, чем счетчика, собиравшего цены?

30. Разумеется, это - законный вопрос, какой бы алгоритм не использовался для выявления резко выделяющихся значений, однако он, видимо, наиболее актуален для процесса автоматического переправления цен в случае замещающих товаров с "несопоставимых" на "сопоставимые", если разница в ценах невелика, и "изучения", когда в меньшей степени используется информация и общие указания, которые могут быть получены из данных для содействия в установлении координат для алгоритмов. Возможно, "изучение" требует большего внимания в том смысле, что аналитик цен в центральном аппарате может произвести замену за секунду, при том, что счетчик, возможно, затратил на такое решение несколько минут, при этом изменение цен в большой степени влияет на такое решение. Этот вопрос также актуален в контексте средств автоматизированного редактирования данных на местах.

31. Первоначальный анализ показал масштабы потенциальной проблемы. Например, только по женской верхней одежде в 2002 году восстановленные цены составляли в среднем лишь 10% распродажных цен. В период с февраля по декабрь в среднем ежемесячно фиксировалось 812 распродажных цен по сравнению с только 81 нормальной ценой после окончания распродажи. Трудность возникает там, где "несопоставимый" или новый товар заменяет товар на распродаже. В этих случаях должна быть рассчитана условная базовая цена, чтобы рост цен реально не влиял на индекс, однако в результате не оказывается восстановленной цены, которая соотносилась бы с распродажной ценой. Таким образом, можно с полным основанием утверждать, что правильное выявление того, является ли товар, замещающий данный товар в конце распродажи, сопоставимым или несопоставимым, не менее важно, чем правильная регистрация цены. Автоматизированные алгоритмы редактирования данных, такие, как алгоритмы Туки, не смогут провести проверку по этому признаку, если они только не входят в более общий пакет процедур редактирования.

32. Диаграмма ниже иллюстрирует процесс "изучения", описанный выше применительно к женской верхней одежде. Буквой "S" обозначается распродажная цена, а буквой "R" - восстановленная нормальная цена[5].

33. В верхнем разделе всего лишь показано число данных о ценах, действительных данных и данных, помеченных "продажная" или "нормальная восстановленная" цена. Ниже слева дается разбивка окончательных обозначений, т.е. обозначений после "изучения" и после принятия ценовыми аналитиками своих решений. Справа показан анализ первоначальных обозначений, присвоенных счетчиками, и обозначений, присвоенных в центральном аппарате. Линии со стрелками всего лишь представляют сходные показатели, например число окончательных обозначений N, переправленных с C, должно равняться числу первоначальных обозначений C, ставших N. Следует отметить, что в таблице "новый" означает замещающий товар, который "несопоставим".

34. Любопытно отметить, что во время работы над исследованием больше половины поступивших данных о ценах, помеченных N, были изменены в центральном аппарате, при этом половина из них были переправлены на C. Примерно 10% этих данных о ценах, поступивших с пометой C, были позднее изменены на N, и таким образом для них надо было вывести новую условную базовую цену.

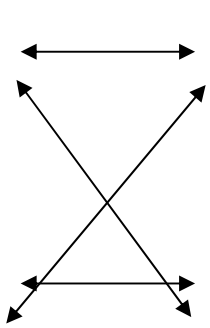
---

[5] Буквенное обозначение присваивается данным о ценах, чтобы облегчить процедуру проверки и дать дополнительную информацию о цене, показывая ее "статус" (например, сопоставимая, несопоставимая, распродажная, восстановленная и т.п.). Первоначально они присваиваются на местах, хотя автоматизированные компьютерные процедуры и сотрудники офиса могут изменять их при изучении цен. Автоматизированные компьютерная процедура заменяет все пометы "несопоставимая цена" на "сопоставимая цена", если увеличение цены на товар-заместитель "достаточно мало". Первоначальное обозначение, присваиваемое на местах, регистрируется таким образом, чтобы можно было видеть, где они были изменены. Обозначение "C" присваивается замещаемому товару, отобранному вместо товара, переставшего поступать в продажу. Он должен быть того же качества, что и предыдущий товар. Как "N" обозначается замещающий товар, который отличается по качеству от товара, который он замещает. Их следует выбирать, только если сопоставимого товара не имеется. Затем условно рассчитывается новая базовая цена для несопоставимых товаров.

Диаграмма 1

Мартовские цены - женская верхняя одежда

Общее число цен	**701**
их них действительных	**596**
из них S	**51**
из них R	**53**



ОКОНЧАТЕЛЬНЫЕ		**ПЕРВОНАЧАЛЬНЫЕ**		
Новые	71	**Новые**	47	
из них первоначально N	22	из них остаются N	22	
из них первоначально C	10	из них становятся C	24	
из них первоначально без обозначения	35	из них становятся без обозначения	14	
из них первоначально S	8			
из них первоначально R	1			
Сопоставимые	122	**Сопоставимые**	108	
из них первоначально C	97	из них остаются C	97	
из них первоначально N	24	из них становятся N	10	
из них первоначально без обозначения	8	из них становятся без обозначения	1	
из них первоначально S	0			
из них первоначально R	0			
Без обозначения	332	**Без обозначения**	367	
из них первоначально N	14	из них становятся N	35	
из них первоначально C	1	из них становятся C	8	(центр)
из них первоначально без обозначения	331	из них остаются без обозначения	331	

35. Таблица ниже выводит анализ на другой уровень. В ней рассматриваются ценовые изменения по сравнению с предыдущим месяцем (или, если цены прошлого месяца не имеется, позапрошлого месяца для каждой из сообщенных цен, помеченных N (несопоставимые) или C (сопоставимые). Например, в верхней строке рассматриваются изменения цен с прошлого месяца для всех сообщенных цен по женской верхней одежде, которые счетчики пометили N, а сотрудники в центральном аппарате оставили это обозначение без изменений. Любопытно отметить, что среднее увеличение цен по сравнению с предыдущим месяцем составило 18,30 фунта стерлингов, при этом выросло 150 цен и уменьшилось 60 цен. В этой таблице самое интересное - сопоставление двух строк, которые имеют первоначальные обозначения C. Можно отметить, что из тех, которые:

a) Были изменены на N

i) средняя цена составляла 75% по сравнению с ценой предыдущего месяца;

ii) среднее увеличение составило порядка 10 фунтов стерлингов;

iii) примерно 80% цен, сопоставленных с предыдущими ценами, продемонстрировали рост.

b) Остались C

i) средняя относительная цена составила только 12% по сравнению с ценой предыдущего месяца;

ii) среднее увеличение составило порядка 2,40 фунта;

iii) примерно 70% цен, сопоставленных с предыдущими ценами, продемонстрировали рост.

36. Если сравнить две рубрики выше, то можно заключить, что в пределах обеих категорий большинство цен демонстрируют рост, указывая на то, что, исходя из прежнего обсуждения, это, возможно, цены на "сопоставимые товары с нормальной послераспродажной ценой". Однако те, которые были изменены на **N**, показывают в целом большее увеличение цены. Хотя возможно, что такое увеличение цены вызвано ростом качества при принципиальном отличии данного товара, вполне возможно, что по крайней мере в некоторых случаях изменение цены повлияло на "изучение" и сопоставимый товар фактически был исключен из индекса, а рост цены не был зафиксирован. Если такое явление происходит постоянно, его воздействие на индекс будет нарастающим.

Таблица 1. Изменения цен в разбивке по обозначениям

Первоначальное	Окончательное	Число	Сопоставление за последние два месяца	Средний рост цены	Миним. рост цены	Максим. рост цены	Число увелич. цен	Число снижений цен	Среднее отношение цен (%)
Новый	Новый	224	214	18,30	-60,00	417,00	150	60	73
Сопоставимый	Новый	109	102	9,99	-119,01	92,95	67	17	75
Без обозначения	Новый	354	353	-0,33	-30,00	0,00	0	7	-1
Сопоставимый	Сопостав.	978	892	2,43	-49,00	145,50	288	120	12
Новый	Сопостав.	240	161	1,46	-81,00	45,00	74	43	10
Без обозначения	Сопостав.	8	8	-1,25	-10,00	0,00	0	1	-3

37. Оставляя в стороне вопрос о том, правильны ли они, можно проверить воздействие этих решений, фиксируя обозначение, занесенное счетчиком, и окончательное обозначение, присвоенное товару в момент его включения в индекс. Это было сделано БНС для группы швейных изделий. Как отмечалось выше, очевидно, что, когда сообщаемые данные о ценах включаются в публикуемый индекс (окончательное обозначение), в наборе данных будет меньше несопоставимых позиций по сравнению с периодом времени, следующим сразу после сбора цен (первоначальное обозначение), в силу использования в это время автоматизированных процессов. Поскольку многие сообщаемые данные о сопоставимых товарах с более высокой ценой получают новое обозначение несопоставимых товаров, первоначальные обозначения могли бы дать более высокий индекс.

38. Для сопоставления было составлено два отдельных индекса одежды за четырехлетний период:

a) первый - с использованием обозначений, присвоенных на местах;

b) второй - с использованием окончательных обозначений после компьютерной проверки, "изучения" и т.п.

39. Результаты представлены на диаграммах 3-5 ниже соответственно для 2000, 2001 и 2002 годов. Как и ожидалось, все три диаграммы показывают более высокий индекс в случае использования первоначальных обозначений, а не обозначений, присвоенных после "изучения". В целом различия за год между индексами составляли 3-4 индексных пункта, хотя в 2000 году оно было меньше.

Диаграмма 3

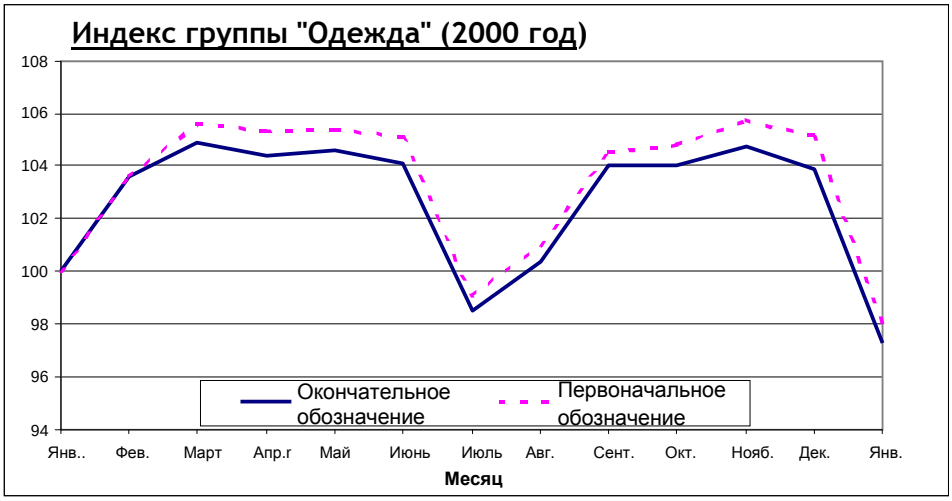


Диаграмма 4

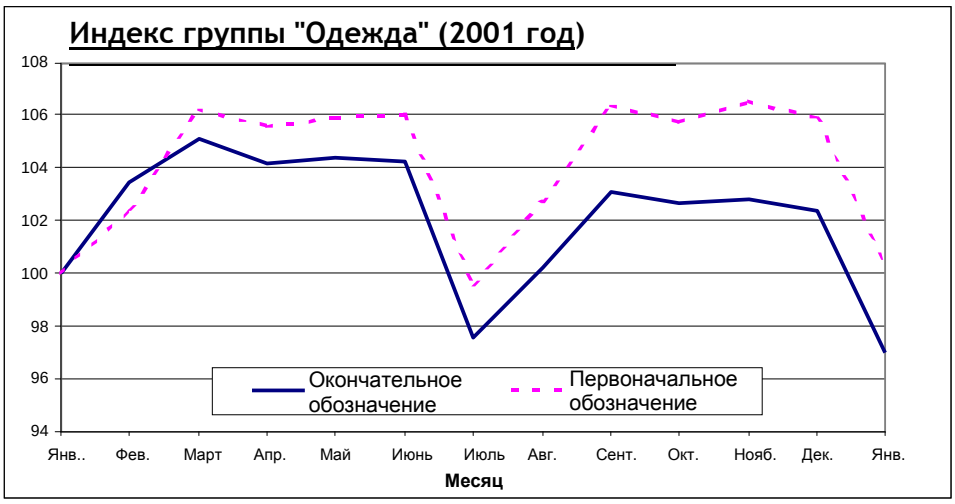
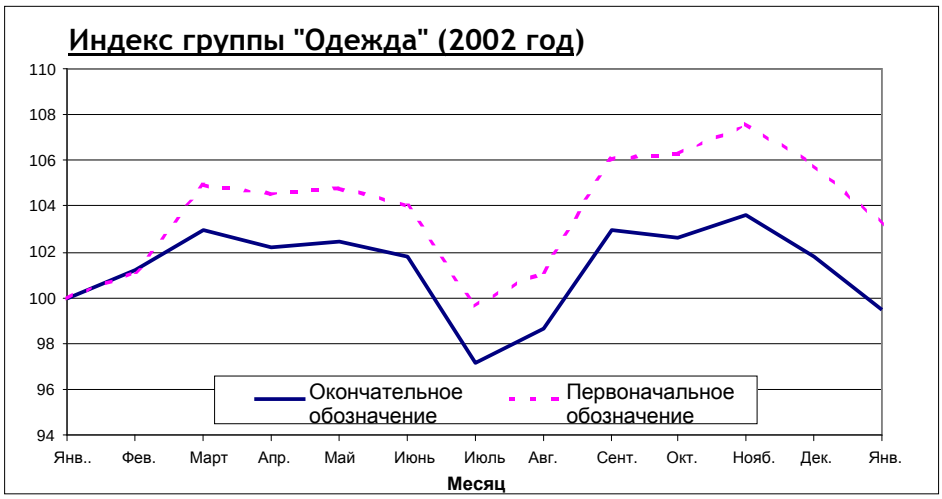
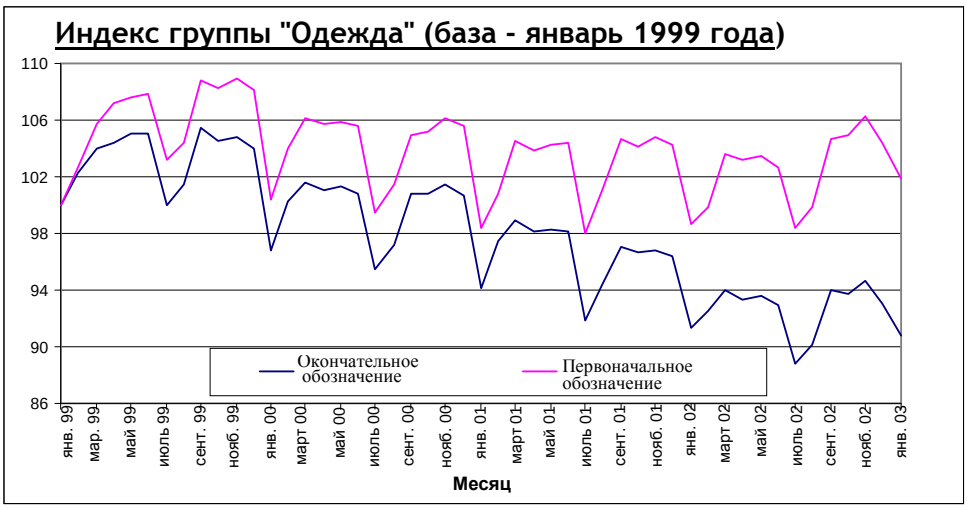


Диаграмма 5



40. Среднюю результирующую всех диаграмм можно видеть на диаграмме 6 ниже со сцепленными индексами с января 1999 года. На ней показаны различия между обоими индексами за четырехлетний период: 11 индексных пунктов - 101,9 вместо 90,9.

Диаграмма 6



41. Хотя данные моделирования за четырехлетний период показывают, что переправленные обозначения могут вызывать понижательное давление на индекс изделий швейной группы, это не вызывает трудностей, если новые обозначения даны правильно, т.е. решения, приняты в центральном аппарате, лучше, чем решения счетчиков,

собиравших данные о ценах на местах. Для проверки того, так ли это на самом деле, была проведена маломасштабная проверка.

Результативность счетчиков цен: последующая проверка обозначений

42. Проверяющим БНС было поручено перепроверить точность обозначений N (несопоставимые) и C (сопоставимые) по выборке специально отобранных мест. Хотя проверка охватывала только трехмесячный период, что было вызвано соответствующими расходами, она не дала четкого указания относительно надежности решений счетчиков. Через несколько дней после первого сбора цен была проведена перепроверка. На основании как имеющихся у них описаний товара, так и его изучения проверяющим было предложено решить, сопоставим ли данный товар или нет, и представить информацию, которая повлияла на их решение. Выдержка из такого бланка перепроверки представлена в приложении.

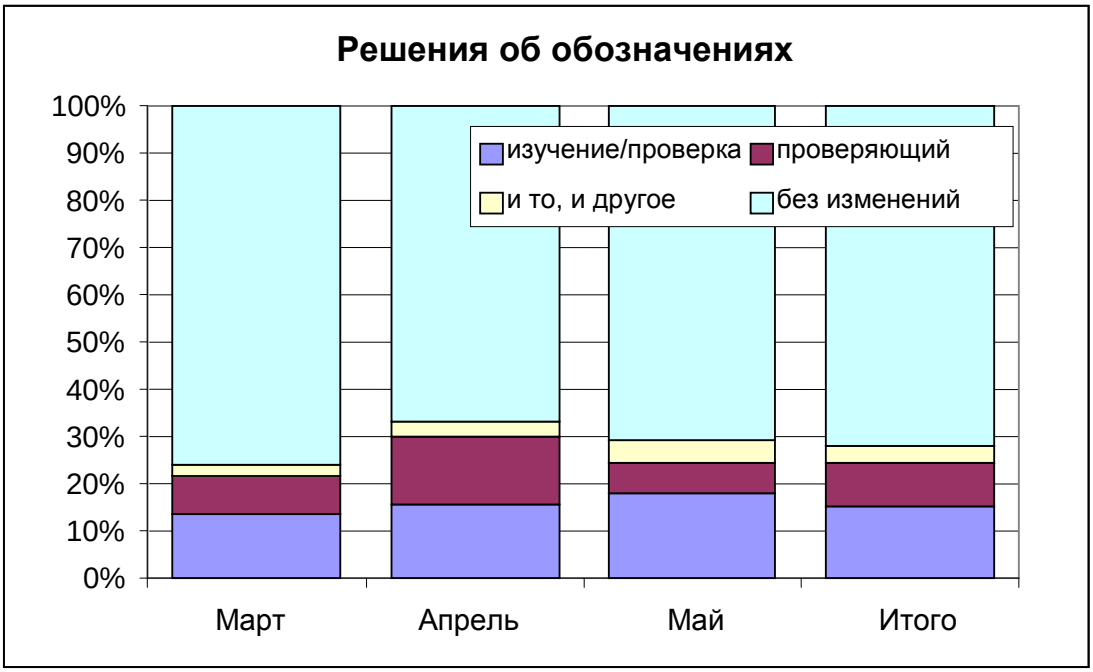
43. Как показывают результаты, в целом решения счетчиков о присвоении обозначений были точны. Хотя проверяющие изменили некоторые обозначения, можно считать, что это скорее вопрос неоднозначного толкования при неизбежном элементе субъективности, хотя небольшая часть из них неизбежно была вызвана неправильными решениями счетчика, и по этой причине необходима, как это делается в СК, постоянно действующая программа подготовки счетчиков.

44. В таблице 2 ниже показаны общие результаты. Они также представлены в графической форме на диаграмме 7. Хотя большинство сообщенных данных о ценах, рассмотренных проверяющими, остались неизменными, возникло то затруднение, что в тех случаях, когда проверяющие не указали, что обозначения, присвоенные счетчиками, были неправильными, изменение обозначений, рекомендованное ими, как правило, не совпадало с редакционными изменениями в ходе процедуры проверки в центральном аппарате.

Таблица 2

	Март	Апрель	Май	ИТОГО
Общее число сообщенных цен	256	147	155	588
Общее число изменений по результатам "изучения"	35	23	28	86
Общее число изменений проверяющими	20	21	10	51
Общее число изменений, внесенных и в результате "изучения", и проверяющими *(не показано выше)*	7	5	7	19

Диаграмма 7



45. В двух таблицах ниже представлена разбивка результатов по первоначальным обозначениям (обозначениям, присвоенным счетчиком) с уделением особого внимания изменениям обозначения с "сопоставимый" и "несопоставимый" или наоборот.

46. Первая из этих таблиц, таблица 3, показывает, что 425 цен на одежду, помеченных счетчиками "сопоставимые", только 4% были изменены в ходе "изучения"/проверки по сравнению с 14% замен проверяющими. Любопытнее всего то, что из 4% замен в центральном аппарате среднее изменение цен на данный товар по сравнению с прошлым месяцем составило увеличение на 17,87 фунта. При этом применительно к заменам проверяющими средняя цена выросла только 2,69 фунта.

Таблица 3

ПЕРВОНАЧАЛЬНОЕ ОБОЗНАЧЕНИЕ C	425	*Среднее изменение цен* (по сравнению с предыдущим месяцем)
в том числе изменено на N в ходе "изучения"	4%	+ 17,86 фунта
в том числе изменено на N проверяющими	14%	+ 2,69 фунта

47. В таблице ниже показано, что из 134 данных о ценах, помеченных как "несопоставимые", 63% были изменены в ходе проверки на сопоставимые. Среднее увеличение цен, связанное с этим, составило всего лишь 1,28 фунта (указывая на то, что значительная часть этих 63% замен была произведена автоматизированными процедурами, упоминавшимися выше, которые изменили "несопоставимые" цены только с небольшим изменением цен). Наоборот, проверяющие изменили только 9% из этих 134 сообщенных данных.

Таблица 4

ПЕРВОНАЧАЛЬНОЕ ОБОЗНАЧЕНИЕ N	134	
в том числе изменено на C в ходе "изучения"	63%	+ 1,28 фунта
в том числе изменено на C проверяющими	9%	+ 2,13 фунта

48. Хотя пределы возможных обобщений будут зависеть от ситуации на месте, в том числе от качества обучения, полученного счетчиками, следует сделать тот вывод, что при обычных обстоятельствах решения, принятые счетчиками и проверяющими на местах, в большей степени основываются на информации и предпочтительны по отношению к тем, которые принимаются в центральном аппарате.

V. НЕКОТОРЫЕ ДОПОЛНИТЕЛЬНЫЕ СООБРАЖЕНИЯ ПО АЛГОРИТМУ ТУКИ

49. До сих пор здесь говорилось о том, что больше внимания необходимо уделять усилению согласованности и комплексности процедур выявления резко выпадающих значений и редактирования, а также более широкому использованию метаданных в сочетании с приданием большего веса решениям, принятым переписчиками цен. Здесь

также проводится тот тезис, что алгоритм Туки - наиболее подходящий алгоритм, при условии что он используется эффективно и грамотно.

50. В идеале составитель ИПЦ получает все данные одномоментно и своевременно и проводит проверку по алгоритму Туки по всем данным, поскольку в ином случае его результативность ограничивается отсутствием (действительных) данных. Однако большинство составителей индексов находятся в иной ситуации. Таким образом, возникает вопрос о том, можно ли получить суррогат для выявления по первой подборке данных тех цен, которые скорее всего были бы определены как выбросы, если бы алгоритм Туки был просчитан ко всем данным. Как мы видели в случае ИПЦ/ИРЦ, это не обеспечивается процессом "изучения", хотя он и имеет свои практические преимущества.

51. Алгоритм Туки надежен в том смысле, что предполагаемые пороги, которые определяют резко отличающиеся значения, не меняются в сколь-нибудь значительной степени после добавления относительно небольшого набора данных. Таким образом, большинство сообщенных данных, временно определяемых как выбросы в результате первого выполнения алгоритма Туки, скорее всего будут подтверждены как выбросы после второго выполнения алгоритма Туки. Кроме того, число значений, временно определенных как выбросы, может быть ограничено, и внимание автоматически может быть сконцентрировано на более экстремальных движениях цен в первом наборе данных, при коррекции прямо заданных параметров, использованных при первом выполнении алгоритма Туки. Кроме того, для разных товаров могли бы использоваться различные параметры, чтобы отразить соответствующую долю данных о ценах, которые были собраны в централизованном порядке (и поэтому включены во второй набор данных).

52. Двукратное выполнение алгоритма Туки - для первого берется первоначальный главный набор данных, собранных на местах, а для второго берутся все данные, собранные как на местах, так и в центре, - имеет свои преимущества. Любые решения о подтверждении, принимаемые после первого выполнения алгоритма Туки, могут быть временно отброшены и реализованы только после того, как данные определяются как резко отклоняющиеся после второй прогонки. Сохраняется возможность оставить без изменения или отменить решение по любым данным, которые не определены как выбросы после второго выполнения алгоритма.

VI. ВОПРОС О ЦЕНТРАЛИЗОВАННОМ СБОРЕ ЦЕН И ЦЕНТРАЛИЗОВАННО РАССЧИТЫВАЕМЫХ ИНДЕКСАХ

53. Замечания, изложенные в настоящем документе, касаются сбора цен на местах применительно к большинству товаров. Сбор цен ведется в торговых точках примерно в 150 городах страны. Такой метод позволяет получить данные о ценах примерно по 110 000 позиций. Обычно счетчики должны прийти в магазин, хотя цены на некоторые товары могут быть получены по телефону.

54. Обсуждение процедур редактирования неполно без упоминания централизованного сбора. Централизованный сбор применяется в отношении товаров, цены на которые могут быть получены в централизованном порядке БНС без какой-либо работы на местах. Эти цены можно подразделить на две категории в зависимости от их последующего использования:

a) централизованные магазины, цены которых совмещаются с ценами, полученными на местах, и

b) товары с централизованно устанавливаемыми ценами, используемыми самостоятельно для составления централизованно рассчитываемых индексов.

55. Цены централизованных магазинов получают от крупных розничных сетей, где цены устанавливаются в масштабах всей страны. Местные филиалы этих систем исключаются из местного сбора данных. Некоторые сети заполняют бумажные бланки, другие заполняют данные о ценах в электронной таблице, отсылая их по электронной почте, или, как вариант, данные берут с сайта компании в Интернете. Каталоги почтовой рассылки также используются как централизованные магазины: цены регистрируются в том виде, как они проставлены в каталоге в момент выпуска каталога (как правило, дважды в год). Эти цены совмещаются с ценами на те же товары, собираемыми на местах.

56. В большинстве случаев розничная торговля выбирает товары с заданными характеристиками, в отношении которых они сообщают цены, однако в некоторых случаях они сообщают полный прейскурант, из которого сотрудники БНС выбирают товар для определения цены. Выбор зависит от опыта того, какой показатель подходит больше всего.

57. Централизованный сбор цен ведется примерно по 130 товарам, причем расчет индекса производится отдельно от расчета главного индекса. Примерно 110 из этих товаров используются при составлении ИРЦ, а остальные - ИПЦ (например,

комиссионные паевых фондов и, в новом ИПЦ, новый индекс автомобилей). Выбор такого вида сбора данных и расчета обычно зависит от одного или нескольких из следующих соображений: источники данных, представление данных, частота изменения цен, национальная ценовая политика и возможность будущих принципиальных изменений методов ценообразования.

58. Применительно к большинству из этих товаров метод сбора и расчета данных основан на общей модели. Индексы агрегируются с самого низкого уровня вверх, при этом веса часто имеются на уровне цен конкретных товаров. Данные о весах, используемые при централизованном расчете индекса, поступают из самых разных источников, которые обычно специфичны по отношению к данному индексу. По мере возможности, сбор данных о ценах ведется в Интернете. В противном случае данные о ценах собираются по мере возможности из одного центрального источника (торговые ассоциации, государственные ведомства и т.п.), хотя чаще всего сбор данных требует контакта с региональными компаниями или с рядом конкурирующих компаний. Цены могут быть запрошены письмом, по телефону или по электронной почте или же могут поступать автоматически, если БНС включено в список рассылки поставщика. Поставщики могут направлять либо полный прейскурант, либо тариф, из которых можно вывести соответствующие цены. В некоторых случаях транспортные тарифы предоставляются в виде индексов цен. Все сообщаемые цены должны подтверждаться той или иной письменной документацией. То, насколько часто направляются запросы, зависит от товаров, которые могут быть весьма различны, а также от того, известны ли и ожидаются ли изменения цен. Самая частая регулярность - раз в месяц или квартал, а также три раза (например, в случае некоторых транспортных тарифов), два раза (например, арендные ставки местных органов власти) и один раз в год (например, билеты на футбольные матчи), а также "при необходимости" (например, когда сообщается об изменении национальных железнодорожных тарифов), когда запросы также включаются в график.

59. Важность этих централизованно собираемых или централизованно рассчитываемых индексов в связи с вопросом имеет три аспекта:

a) хотя их число относительно невелико, на них приходится 40-50% ИПЦ/ИРЦ в плане расходов домохозяйств;

b) в силу относительно высокого веса, определяемого для любого конкретного ценового параметра в централизованно собираемых или централизованно рассчитываемых индексах, имеется более чем реальный риск того, что ошибка в установлении цены может привести к ошибке в публикуемых индексах;

c) неизбежно то, что чем больше времени расходуется на местном уровне для редактирования собранных данных о ценах, тем меньше времени будет израсходовано на редактирование и перепроверку централизованно собираемых данных о ценах и централизованно рассчитываемых индексов.

60. Алгоритм Туки не может применяться в отношении централизованно собираемых цен и централизованно рассчитываемых индексов в изоляции ввиду ограниченного числа сообщаемых данных о ценах. Упор делается на процедурах изучения в центральном аппарате на основе сходных процессов, аналогичных процессам, используемым на местах в отношении местного сбора данных о ценах, т.е. установлении диапазона минимальных и максимальных цен на основе сбора цен в предыдущий период на те же товары и ряде логистических проверок, например того, что "нормальная" послераспродажная цена следует только после цены, введенной в предыдущий месяц как "распродажная" цена. Кроме того, полученные цены перепроверяются в свете ожиданий; например, они проверяются, когда можно ожидать ежегодного пересмотра цен и после получения сообщений о запланированном увеличении цен. В отношении централизованно собранных данных используется алгоритм Туки, после того как цены совмещаются с ценами на те же товары, полученными на местах.

61. Из этого можно заключить, что централизованно рассчитываемые индексы связаны с наибольшим специфическим риском. Это тем более так, если иметь в виду, что многие такие индексы связаны с услугами, в случае которых имеется дополнительная проблема учета изменений в сложных тарифных структурах, таких, как коммунальные и транспортные тарифы и тарифы сотовой связи. Для проверки правильности расчета индекса необходимо следующее:

a) собрать предыдущие цены и отметить причины изменения цен;

b) сверить полученную информацию о ценах с ранее объявленными изменениями цен;

c) постоянно проверять адекватность методики, прежде всего для обеспечения того, что средние изменения цен в результате пересмотра тарифа вместе с коррективами на изменения в оказываемых услугах должным образом отражаются в расчетах;

d) обеспечить проведения независимой проверки другим составителем индекса.

62. Реализация процедуры, имеющей оптимальную эффективность, означает расходование меньшего времени на единицу сообщаемых данных в случае централизованного сбора данных о ценах и централизованного расчета индексов по сравнению со сбором данных о ценах на местах.

63. В БНС внедрен трехэтапный процесс проверки централизованно собираемых и рассчитываемых индексов цен:

a) составитель индекса в центральном аппарате, собирающий цены, проверяет их полноту и точность, например, сопоставляя данные с предыдущими ценами и с заранее опубликованными увеличениями цен;

b) другой составитель индексов проверяет индекс на предмет того, выглядит ли тот разумным образом и поддается ли он рациональному толкованию. Как минимум, если индекс "не выглядит разумным образом", все цены перепроверяются, хотя цель будет заключаться в выверке всех цен;

c) статистик по ценам проверит все соответствующие электронные таблицы, лежащие в основе логики расчетов и расчетов ошибок и странно выглядящих цен. Это - последняя проверка перед публикацией.

VII. РЕЗЮМЕ И ВЫВОДЫ

64. Выше в этом документе выдвигался тот тезис, что преимущества использования алгоритмов, таких как алгоритм Туки, как "фильтрующих" механизмов редактирования заключаются в том, что они позволяют избежать необходимости изучения аналитиком в центральном аппарате исключительно большого числа наблюдаемых цен за очень короткий период времени и принятия моментальных решений. Однако на практике их результативность в этом плане, а также как инструмента, обеспечивающего эффективный процесс выявления и исправления истинных ошибок в данных оценок, которые заносятся в ИПЦ, не гарантирована и зависит от целого ряда факторов, связанных с конкретным характером их применения.

65. Исследования, проведенные в Соединенном Королевстве несколько лет назад, показывают следующее:

a) Хотя алгоритм Туки может считаться лучше других алгоритмов выявления резко выделяющихся знаний, его результативность может быть снижена из-за

исключения "действительных" данных в результате проведенного ранее редактирования (процесса "изучения").

i) Эффективность применения алгоритма Туки еще более снижается в том случае, если предполагается, что резко отличающееся значение неправильно (т.е. недействительно), если только впоследствии оно не подтверждается редактированием.

a. Правильность предположения о необходимости исключения выброса подвергается сомнению из-за очень небольшого числа резко отличающихся значений, статус которых как таковых прямо подтверждается в ходе активного редактирования. Это предположение могло бы на практике привести к систематической ошибке.

ii) Результативность применения алгоритма Туки может быть также снижена из-за автоматической замены обозначения несопоставимых "замещений" на сопоставимые "замещения" в тех случаях, когда ценовая разница весьма невелика, а также из-за отсутствия отсылки к соответствующим "обозначениям" (C - для сопоставимых, N - для несопоставимых, S - для распродажных и т.п.).

b) Решение счетчика, собирающего данные о ценах, как правило, более надежно, чем решение специалиста в центральном аппарате (хотя здесь есть и исключения).

c) Централизованно рассчитываемые индексы относительно сложны, и на них приходится приблизительно 40% "корзины" ИПЦ/ИРЦ по весовым коэффициентам расходов. Они сопряжены с более высоким специфическим риском ошибки в ценах, ведущей к ошибке в публикуемом индексе.

66. В результате исследований, описанных в этом документе, был принят ряд мер:

a) "Изучение" было сохранено в силу того, что оно дает возможность выявить на раннем этапе крайне резко отличающиеся значения и большие различия в качестве по сравнению с ценой товара-заместителя и к тому же было улучшено благодаря указанию "весов" сообщаемых данных вместе с другой полезной информацией.

b) Предположение, что резко отличающееся значение неправильно, пока не будет доказана его правильность, было скорректировано: в настоящее время специалисты, проводящие "изучение", применяют метод перепроверки сообщаемых данных о ценах, первоначально определенных в качестве выбросов, если только уже не имеются данные, доказывающие обратное.

c) Решение счетчика отменяется только в том случае, когда этого требуют имеющиеся данные. Часть этого пересмотренного процесса - уделение большего внимания кодам "обозначений", например тому, представляет ли цена нормальную цену после распродажной цены.

d) Не производится автоматического перекодирования "несовместимых" замещений на "совместимые" замещения в случае, если различия в ценах невелики.

e) Больше внимания стало уделяться обеспечению качества и редактированию сообщаемых данных о ценах, связанных с централизованно собираемыми данными о ценах и централизованно рассчитываемыми индексами в тех случаях, когда имеется более высокий специфический риск ошибок, влияющих на публикуемые индексы.

67. Исследование также привело к следующим обобщенным выводам:

a) Редактирование важно, поскольку оно способно влиять на измеряемые темпы инфляции, и это немаловажный вопрос, который требует внимательного отношения. Процессы редактирования должны регулярно рассматриваться, в частности для того, чтобы они отражали изменения в методах сбора цен и в процессах компиляции и структуре индекса.

b) Эффективность редактирования, определяемая как потенциальное воздействие на точность ИПЦ/ИРЦ при данном объеме редактирования, будет зависеть от использования оптимально эффективных процедур:

   i) не имеется замены редактирования цен на местах "в реальном времени";

   ii) счетчик, собирающий данные о ценах, обычно находится в предпочтительном положении для того, чтобы принимать взвешенное решение;

iii) необходимо уделять относительно большее внимание обеспечению качества и редактированию цен, которые оказывают наибольшее влияние на публикуемый индекс либо из-за весовых коэффициентов расходов, либо из-за числа сообщаемых данных о ценах. Необходимо также учитывать сложность расчетов и соответствующий специфический риск ошибки.

68. Исследования, описанные в этом документе, были проведены группой исследователей ИПЦ в БНС. Этот проект осуществлялся под руководством Деймона Уингфилда, которому мы признательны за плодотворное обсуждение, которое мы провели в этом время, и за его замечания по предыдущему варианту этого документа. Мнения, выраженные в этом документе, принадлежат автору.

ANNEX

[ENGLISH ONLY]

Location 1552 - Bluewater **CONFIDENTIAL** **Clothing Indicator Backcheck**

Shop	February Last month		March This month
HOBBS LTD, L101 & L102 GUILDHALL (LOWER)	Women's skirt: casual HOBBS/NCO BLACK/PLAIN"18.5 SAIGON SK STRETCH CORD/2 FRONT/2 SLIT BACK POCKS/BELT LOOPS 97 COTTON 3 LYCRA £49.00 Notes: _____	Item 1 2 3 4 5 Price	Women's skirt - casual HOBBS/ITALY BLUE DISTRESS DENIM/"19" CATH" PANELLED A LINE WITH SEWN DOWN FRONT FLAP/NO WAISTBAND 98.5 COTTON 1.5 ELASTANE £59.00 **Comparable** ☐
CECIL GEE U137 GUILD HALL	MEN'S CASUAL S/SLEEVE SHIRT CECIL GEE/NCO LT BLUE/PLAIN KNITTED JERSEY COLLAR/6 BUTT FRONT/TURN UP ON SLEEVE 50 COTTON 50 POLY £0.00 out of stock Notes: _____	Item 1 2 3 4 5 Price	Men's casual s/sleeve shirt with collar ARMANI/ITALY NAVY/POLO SHIRT COLLAR/5 BUTTONS/CHEST POCKET WITH LOGO/WHITE STRIPE ON COLL 100 COTTON £69.00 **New** ☐
CECIL GEE	MEN'S CASUAL SHIRT-LONG SLEEVE ARMANI JEANS/ITALY LIGHT DENIM BLUE/PLAIN COLLAR,ONE CHEST POCKET WITH LOGO/FINE BLUE STITCHING 80 COTTON 20 LINEN £79.00 Notes: _____	Item 1 2 3 4 5 Price	Men's Casual Shirt, long sleeved ARMANI JEANS/ROMANIA WHITE/LIGHT BLUE 1 CM CHECK BUTTON COLLAR/CHEST POCKET WITH LOGO 100 COTTON £85.00 **New** ☐
RIVER ISLAND U32 ROSE GALLERY	MENS SUIT-READY MADE -1 RIVER ISLAND/ROMANIA/6898 MID GREY/LINEY RIBBED FEEL S/B 4 BUTTONS/3 FRONT POCKETS 4 BUTT CUFF/FLAT FRNT BOOT TRS 67 POLY/33 VISCOSE £99.00 Notes: _____	Item 1 2 3 4 5 Price	Mens suit, ready made RIVER ISLAND/ROMANIA/6900 MID GREY/LINEY RIBBED FEEL S/B 4 BUTTONS/3 FRONT POCKETS 4 BUTT CUFF/FLAT FRNT BOOT TRS 67 POLY/33 VISCOSE £99.00 **Comparable** ☐
RIVER ISLAND	MEN'S FORMAL SHIRT-LONG SLEEVE RIVER ISLAND/NCO/E 7107 WHITE WITH BLUE/BLACK LINE CHK CUFF/1BREAST POCKET/WHITE BUTT IN CUBBY HOLE ON WALL 100 COTTON £0.00 out of stock Notes: _____	Item 1 2 3 4 5 Price	Men's Formal Shirt, long sleeved RIVER ISLAND/HONG KONG/E 0017 WHITE WITH THIN GREY 2 CM STRI COLLAR/SINGLE CUFF/BUTTON FRONT/ON HANGER BY SUITS 100 COTTON £24.99 **Comparable** ☐