

**Государственный Университет
Высшая школа экономики**

Г.И.Деев, Е.В.Родительская

**ОЦЕНКА ОШИБОК ИСЧИСЛЕНИЯ
ИНДЕКСОВ ВНЕШНЕЙ ТОРГОВ-
ЛИ С ИСПОЛЬЗОВАНИЕМ МЕТО-
ДОВ СТАТИСТИЧЕСКОГО МО-
ДЕЛИРОВАНИЯ**

Препринт WP2/2002/05

Серия WP2

Количественный анализ в экономике

Москва 2002

УДК 311:339.5
ББК 65.051
Д 26

Деев Г.И., Родительская Е.В. Оценка ошибок исчисления индексов внешней торговли с использованием методов статистического моделирования: Препринт WP2/2002/05. – М.: ГУ ВШЭ, 2002. – 24 с.

Работа посвящена анализу рекомендаций миссии Международного валютного фонда по переходу от используемого в России метода исчисления индексов внешней торговли на основе таможенных деклараций к их исчислению на основе прямого наблюдения за внешнеторговыми ценами. Проводится сравнение этих двух альтернативных методов по уровням ошибок, с использованием статистической модели. Приводится описание методов построения статистической модели.

Классификация JEL: C15, C43.

Deev G.I., Roditelskaya E.V. Estimation of Errors of External Trade Indices by the Use of Statistical Simulation Methods: Working paper WP2/2002/05. – Moscow: State University – Higher School of Economics, 2002. – 24 p. (in Russian).

The paper is focused on recommendation of the Mission of the International Monetary Fund to pass from using in Russia methodology of calculation external trade indices based on the customs declaration to a methodology based on the direct external trade price surveys. There is comparison of errors of these methodologies by means of the statistical model. Presented are methods of constructing the statistical model.

JEL Classification: C15, C43.

Деев Гелий Иванович,
Родительская Елена Викторовна
Российская таможенная академия
140009, Московская обл., г. Люберцы,
Комсомольский пр., 4
E-mail: deev34@mail.ru

Deev, Geliy,
Roditelskaya, Elena
Russian Customs Academy
4, Komsomolsky pr., Liubertsy, Mosk-
ovskaya obl., 140009, Russia
E-mail: deev34@mail.ru

© Г.И. Деев, 2002

© Е.В. Родительская, 2002

© Оформление. ГУ ВШЭ, 2002

1. Введение

Настоящее исследование является реакцией, с позиций накопленного опыта, на рекомендации миссии Международного валютного фонда по исчислению индексов внешней торговли [1].

Используемая в настоящее время в России методология расчета системы индексов внешней торговли была введена в действие в 1997 г. В этом же году в сборнике «Таможенная статистика внешней торговли Российской Федерации» были опубликованы индексы средних цен и физического объема 1997 г. относительно 1996 г., отдельно по экспорту и импорту, а также с разделением по группам стран: дальнего зарубежья и СНГ. С 1998 г. указанные индексы публикуются поквартально: квартал текущего года к соответствующему кварталу предшествующего года. Начиная с 1999 г., индексы детализируются по 10 укрупненным аналитическим товарным группам.

Основным, в методологическом смысле, индексом в действующей системе индексов внешней торговли является индекс цен, который рассчитывается по формуле Пааше. Индекс физического объема является производным от индекса стоимости и индекса цены (находится делением первого на второе). Поэтому в дальнейшем основное внимание уделено индексу цен. Выбор большинством стран индекса цен внешней торговли в форме Пааше объясняется тем, что деление стоимости экспорта (импорта) на индекс цен Пааше дает стоимость экспорта (импорта) анализируемого периода в ценах базисного периода.

Информационной базой исчисления индексов внешней торговли в России является весь массив грузовых таможенных деклараций (ГТД), которые оформляются на каждую внешнеторговую сделку. Из ГТД берутся следующие показатели:

- идентификационный код экспортируемого или импортируемого товара, соответствующий нижнему иерархическому уровню товарного классификатора ТН ВЭД СНГ (10-значные позиции);
- код экспорта/импорта;
- код страны-контрагента;
- дата оформления ГТД;
- статистическая стоимость¹;
- количество товара в кг.

В описанной базе данных предельно детализируемой товарной группой является так называемая «десятка-страна» («10-страна») – 10-значная товарная группа, дифференцированная по странам-контрагентам. 10-страны рассматриваются как конкретные товары, по которым рассчитываются цены, а, следовательно, и индивидуальные индексы. Цены рассчитываются как отношения сумм статистических стоимостей 10-стран к суммарным весам товаров данных групп. Назовем описанную базу данных базой данных таможенной службы (БДТС).

¹ Статистическая стоимость товара определяется путем корректирования фактурной стоимости товара к стандарту FOB при экспорте и CIF – при импорте, с переводом валюты контракта в доллары США.

Принятое в БДТС определение цены товара находит применение в статистике внешней торговли и других стран применительно ко всем или части товаров, например, в Норвегии [2]. Для таких цен используются термины: «средняя цена» или «стоимость единицы товара» («unit-value»). Определенные таким образом цены не являются в точном смысле ценами товаров. Собственно цена товара может быть установлена только при специально организованном наблюдении, при котором фиксируются цены конкретных товаров.

Такое наблюдение называют прямым наблюдением за ценами. В упомянутых выше рекомендациях Международного валютного фонда переход от стоимостей единиц товаров к прямому наблюдению за ценами товаров предлагается как основное направление совершенствования статистики цен внешней торговли в России. При этом высказывается понимание, что этот переход связан со значительными затратами на организацию выборочного наблюдения за товарами-представителями, так как прямое наблюдение за ценами может быть только выборочным. Необходимо отметить, что даже в продвинутых в статистическом отношении странах переход к прямому наблюдению за внешнеторговыми ценами осуществляется в течение многих лет и в большинстве случаев не закончен до настоящего времени [3, 4]. К этому нужно добавить, что общий организационно-технологический уровень выборочных наблюдений в нашей стране еще недостаточно высок [5], что создаст дополнительные проблемы с обеспечением качества и достоверности результатов наблюдения. Таким образом, переход к прямому наблюдению за внешнеторговыми ценами – не ближайшая перспектива.

Поскольку переход к прямому наблюдению за ценами внешней торговли в настоящее время затруднителен, актуальной задачей является сравнительный анализ ошибок, допускаемых в двух системах исчисления индексов внешней торговли:

- на основе базы данных, используемой в настоящее время таможенной службой (БДТС);
- на основе базы данных прямого наблюдения (БДПН) за ценами, сформированной по выборке товаров-представителей.

БДТС реально существует, а БДПН может быть генерирована посредством статистической модели. Поставленная задача может быть конкретизирована следующим образом. Как известно, для вычисления основного индекса системы индексов внешней торговли – индекса цены с применением наиболее используемых формул Пааше, Ласпейреса или Фишера достаточно по всем товарам рынка для базисного и анализируемого периодов иметь следующие данные:

- цену товара;
- количество товара.

На информации о ценах товаров следует остановиться более подробно. В условиях рынка цена одного и того же товара может быть различной для разных партий товара – лотов, поэтому цена товара даже при прямом наблюдении всегда в некоторой мере случайна в силу случайности лотовой структуры и того, что не все лоты наблюдаются. Наблюдение за ценами обычно игнорирует лотовую структуру товаров и связанный с этим элемент случайности, но при

моделировании ценообразования лотовую структуру товара можно задавать и учитывать при расчетах индексов цен. Для этих целей введем понятие «базы данных полной формы» (БДПФ), используемой для исчисления индексов внешней торговли. Эта гипотетическая база данных содержит сведения по каждому лоту товара. В отличие от БДПФ, реальная база данных таможенной службы не является полной, так как содержит только обобщенные на уровне 10-стран данные по стоимости и количеству товаров. Неполной является также и база данных прямого наблюдения за ценами, так как она формируется на выборочной основе. Задача сопоставления по точности исчисления индексов внешней торговли на основе БДТС и БДПН может быть интерпретирована как задача сопоставления по точности индексов, рассчитанных на основе этих баз, с индексами, рассчитанными на основе БДПФ. Причем индексы, рассчитанные на основе БДПФ, можно рассматривать в определенном смысле как истинные.

В практике исчисления индексов внешней торговли движение цен часто подразделяют на две составляющие: реальное движение цены и случайную флуктуацию [4]. К факторам, обуславливающим флуктуацию, относятся уже отмеченные колебания лотовой структуры товаров, ошибки репрезентативности (в случае товаров-представителей), колебания структуры 10-стран (в случае БДТС). Кроме того, между индексами цен и индексами количества товаров имеется статистическая (с элементами случайности) зависимость. Наши исследования показывают, что для большинства товарных групп внешней торговли наблюдается отрицательная корреляция между индексами цены и индексами количества товаров. Таким образом, индексы цен, вычисленные на основе всех трех указанных выше баз данных, содержат случайную вариацию, и поэтому сравнительный анализ точностей исчисления индексов по этим базам данных может вестись только в терминах сопоставления распределений вероятностей значений индексов.

Такой анализ может быть проведен с использованием статистической модели генерирования исходных данных. При этом данные генерируются в трех вариантах: БДПФ, БДПН, БДТС. Все базы данных должны быть взаимосвязаны, а именно – БДПН и БДТС формируются на основе БДПФ. Для получения распределений значений индексов необходимо достаточное число реализаций процесса формирования данных. Убедительность и обоснованность получаемых результатов может быть достигнута только в том случае, если генерируемые данные будут достаточно близки к реальным. Это требование обеспечивается тем, что практически все статистические характеристики модели получены путем обработки больших массивов реальной информации по внешней торговле. Таких характеристик около 30. К ним, в частности, относятся:

- характеристики дифференциации (децильные коэффициенты) количества товаров (весов) в товарных группах различного иерархического уровня;
- параметры вариации цен в рамках товарных групп (параметры бета-распределений);
- параметры тренда цен на уровнях отдельных товаров и 10-значных товарных групп;

- параметры вариации цен и объемов товаров анализируемого периода относительно трендов (параметры бета-распределений);
- параметр, определяющий статистическую зависимость количеств товаров анализируемого периода от изменения цен, и др.

Кроме того, для обеспечения реальных структурных соотношений в основу моделирования данных базисного периода положены реальные данные по экспорту в страны СНГ за I квартал 2000 г. товаров укрупненной аналитической группы – «минеральное сырье», охватывающей коды ТН ВЭД с 2500000000 по 2799999999. Данные охватывают 48 4-значных товарных групп, 841 10-значную группу. Эти данные моделью дезагрегируются примерно на 4000 товаров и 20000 товарных лотов. Данные анализируемого периода генерируются на основе данных базисного периода с использованием статистических зависимостей и параметров, полученных при обработке реальной информации о внешней торговле за несколько лет.

2. Методология генерирования данных, используемых для вычисления индексов внешней торговли

Методология генерирования данных, используемых для вычисления индексов внешней торговли (в дальнейшем для краткости просто – данных) представляет собой методологию генерирования данных в полной форме, так как другие формы могут быть получены из полной формы путем выборки и агрегирования. Модель генерирования данных имеет несколько функций:

- обеспечение заданной структуры данных;
- генерирование значений цен, стоимостей и количеств товаров как значений, имеющих детерминированную и случайную составляющие;
- моделирование корреляционных связей между определенными показателями;
- моделирование трендов (тенденций) изменения цен товаров.

2.1. Моделирование структуры данных

Значения сводных индексов внешней торговли существенно зависят от структуры статистических исходных данных. Под структурой данных понимается взаимоотношение элементов соответствующих матриц объект-свойство [10]. Структура отражает как соотношения различных объектов, например, соотношение численностей 10-стран в различных 10-значных товарных группах (БДТС), так и соотношения между показателями одного объекта, например, объема и цены данного товара, а также соотношения между показателями различных объектов, например, ценами различных товаров. Структура также отражает соотношения между различными иерархическими уровнями ТН ВЭД.

При моделировании структуры данных желательно обеспечить ее близость к реальной структуре. Можно просто скопировать структуру определенного фрагмента экспортного или импортного потока товаров, но при этом необходимо учесть следующие обстоятельства:

- желательно, чтобы моделировалась целая аналитическая группа, по которой ведутся публикации ГТК;
- объем структурообразующих данных должен обеспечить приемлемые затраты машинного времени генерации данных в формах, описанных выше.

Ввиду того, что реально располагаемыми данными являются данные в форме ГТК, с учетом сформулированных требований, за структурную основу была принята структура данных по экспорту в страны СНГ за I квартал 2000 г. по укрупненной аналитической группе «Минеральные продукты». Имея в виду, что при проведении модельных расчетов было принято, что каждая 10-страна содержит от 1 до 10 конкретных товаров (это число случайно) и каждый товар имеет от 1 до 10 лотов (также случайное число), общее число товаров, отличающихся по цене, будет колебаться в районе 20 тысяч.

Смысл алгоритма генерирования данных базисного периода в форме БДФ на основе БДТС по выбранной группе товаров состоит в двухступенчатой декомпозиции (разделении) сначала заданных фактических значений весов и стоимостей товарных групп с 10-значным кодом на модельные веса и стоимости отдельных товаров, а затем весов и стоимостей отдельных товаров – на веса и стоимости товаров по лотам.

2.2. Весовая (объемная) структура данных базисного периода

На основании анализа реальных статистических данных была принята следующая методология декомпозиции весов (объемов) как на уровне 10-стран, так и на уровне товаров. Предполагается, что если веса товаров, принадлежащих одной 10-стране, ранжировать по убыванию веса, то веса товаров будут убывать по показательному закону, то есть пропорционально величине $i^{-\alpha_g}$, где i – порядковый номер товара в ранжированном ряду, α_g – показатель степени, характеризующий скорость убывания весов, а значит, и дифференциацию весов товаров. Однако использование показателя степени в качестве показателя дифференциации весов и структурной характеристики неудобно, более употребительным показателем дифференциации и структуры является децильный коэффициент дифференциации $k_{g\ dec\ 10}$, определяемый как отношение суммы весов 10% товаров с наибольшими весами к сумме 10% товаров с наименьшими весами. Приближенными регрессионными методами была определена зависимость α_g от децильного коэффициента дифференциации, длины ранжированного ряда n_{10} и нормирующего коэффициента $k_{g\ 10\ H}$, который обеспечивает равенство суммы весов отдельных товаров заданному значению веса 10-страны. Веса отдельных товаров в данной 10-стране определяются из соотношений

$$(1) \quad g_{тов\ i} = i^{\alpha_{g\ 10}} ,$$

$$(2) \quad \alpha_{g\ 10} = (-2,953 + 6,82648 \cdot k_{g\ dec\ 10}^{-0,0569} - 4,79401 \cdot n_{10}^{-0,05902}) \cdot k_{g\ 10\ H} ,$$

где i – порядковый номер товара в 10-стране, n_{10} – число товаров в данной 10-стране, $k_{g dec 10}$ – децильный коэффициент дифференциации, определяющий объемную товарную структуру 10-страны, $k_{g 10 H}$ – нормирующий коэффициент, обеспечивающий равенство суммы генерированных весов заданному весу 10-страны базисного периода,

$$(3) \quad k_{g 10 H} = \frac{G_{10}}{\sum_{i=1}^{n_{тов}} g_{тов i}},$$

где G_{10} – вес данной 10-страны базисного периода.

Рассмотренная методология декомпозиции веса 10-страны на веса товаров применяется также для декомпозиции веса товара по лотам.

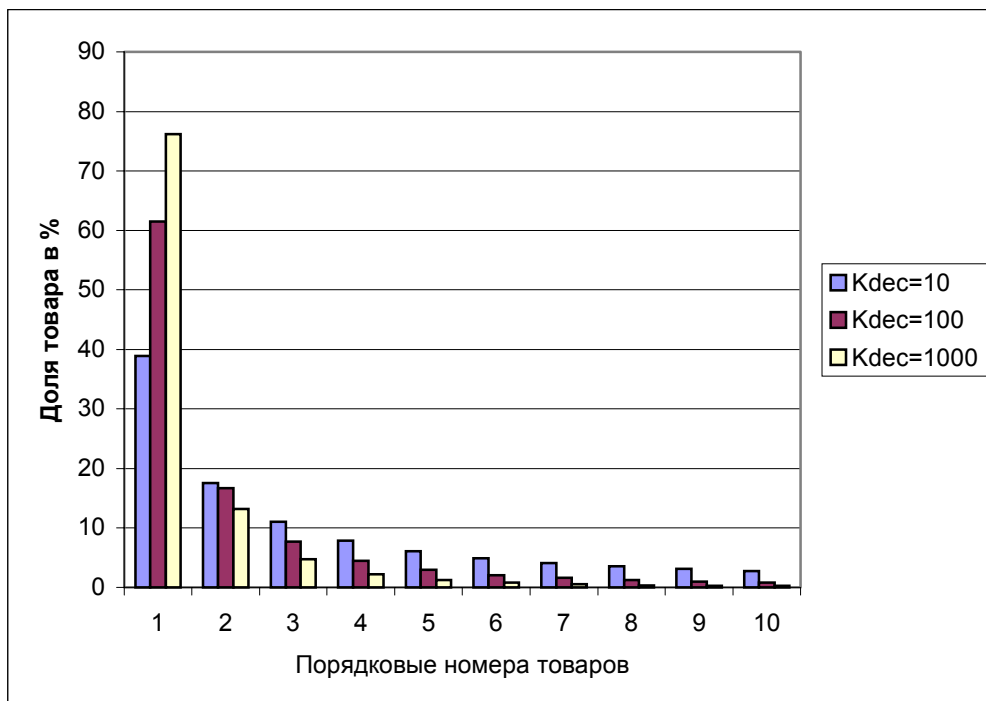


Рис. 1. Распределение весов товаров 10-страны при различных значениях децильного коэффициента дифференциации

На рис. 1 для ряда из 10 товаров представлены весовые товарные структуры при значениях децильного коэффициента, используемых в модели: 10, 100 и 1000, которые близки к реальным значениям при различных условиях. Изменяя в процессе моделирования значения децильных коэффициентов дифференциации, то есть изменяя товарные весовые структуры 10-стран и товаров, можно выявить влияние товарных структур на значения индексов.

2.3. Генерирование цен товаров и лотов базисного периода

Представляется методически обоснованным для базисного периода цены товаров, входящих в 10-страны, и цены отдельных лотов, составляющих товары, генерировать как случайные величины, коррелированные с ценами базисного периода, с распределениями, близкими к реальным.

Анализ статистических данных показал, что для генерирования цен целесообразно воспользоваться общей формой бета-распределения, плотность которого имеет вид:

$$(4) \quad f(x; \alpha, \beta, a, b) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)(b-a)^{\alpha+\beta}} (x-a)^\alpha (b-x)^\beta \\ 0 \quad \text{в остальных случаях} \end{cases},$$

$$a < x < b, \quad a < b, \quad \alpha > 0, \quad \beta > 0$$

где α и β – параметры формы бета-распределения, a и b – нижняя и верхняя границы распределения, соответственно, $\Gamma()$ – символ гамма-функции.

Выбор общей формы бета-распределения в качестве основы для генерации значений цен обусловлен несколькими обстоятельствами:

- бета-распределение обладает большой гибкостью и разнообразием форм, что позволяет выбрать форму, близкую к форме распределения реальных цен;
- бета-распределение имеет конечный, задаваемый интервал распределения, что позволяет легко и наглядно масштабировать вариацию распределения.

В алгоритме модели генерирования данных распределение (4) используется с различными значениями параметров. Параметры формы выбраны следующие: $\alpha = 1$, $\beta = 3,7$. Эти значения параметров дают коэффициент асимметрии равный 1, что является типичным для распределения цен. Математическое ожидание $M_{цены}$ и границы распределения цен a и b , а значит и вариация цены постоянно меняются в процессе работы алгоритма. Например, при генерировании цен товаров, входящих в 10-страну в базисный период, за математическое ожидание принимается средняя цена по 10-стране. Границы a и b определяются в зависимости от вариации цены, которая различна для разных случаев. В качестве показателя вариации цен выбран коэффициент масштаба вариации $K_{маш}$ – отношение максимального значения цены к минимальному. Можно показать, что задание математического ожидания $M_{цены}$ и коэффициента масштаба вариации $K_{маш}$ при известных параметрах формы α и β однозначно определяет границы распределения a и b :

$$(5) \quad a = \frac{M_{цены}}{\frac{\alpha}{\alpha + \beta}(K_{маш} - 1) + 1},$$

$$(6) \quad b = \frac{M_{цены} K_{маш}}{\frac{\alpha}{\alpha + \beta}(K_{маш} - 1) + 1}.$$

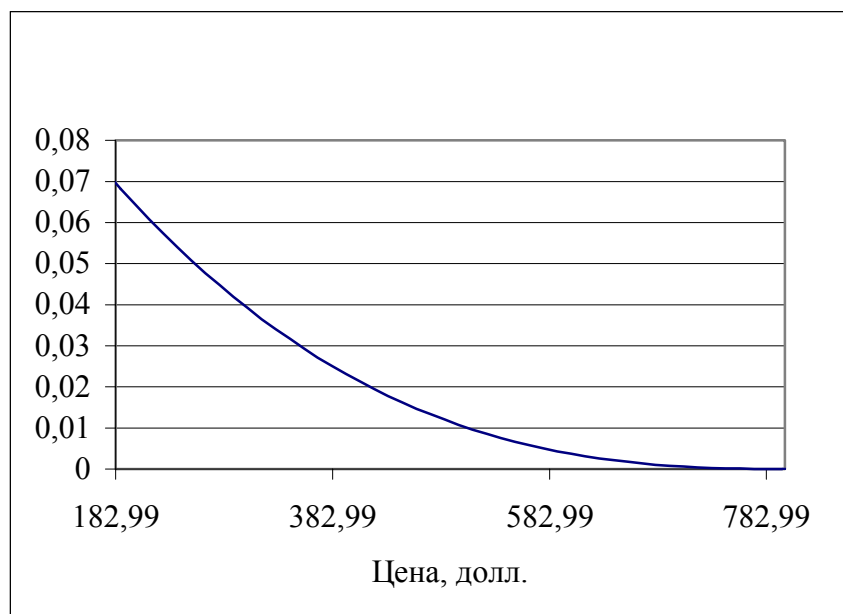


Рис. 2. Плотность распределения цены

На рис. 2 в качестве иллюстрации представлено распределение цен с параметрами $\alpha = 1$, $\beta = 3,7$, $M_{цены} = 300$ долл., $K_{маш} = 5$.

При генерировании цен, товаров и лотов для сохранения структурных соотношений на уровне 10-стран необходимо нормирование цен. Если случайное значение цены, полученное по распределению (4), обозначить как $\beta_{цен}$, то цена i -го товара определится по формуле

$$(7) \quad P_{тов i} = \beta_{цен} k_{p10n},$$

где $\beta_{цен}$ – случайная величина, имеющая бета-распределение с математическим ожиданием, равным средней цене по 10-стране, k_{p10n} – нормирующий коэффициент, обеспечивающий равенство суммы генерированных стоимостей товаров заданной стоимости 10-страны базисного периода,

$$(8) \quad k_{p10n} = \frac{C_{10}}{\sum_{i=1}^{n_{тов}} g_{тов i} \cdot P_{тов i}},$$

где C_{10} – стоимость товаров данной 10-страны базисного периода.

Методология генерирования значений цены лотов для базисного периода аналогична изложенной, с той лишь разницей, что в качестве математического ожидания цены берется средняя цена по данному товару.

2.4. Моделирование тренда цен и корреляционных связей данных анализируемого и базисного периодов

При моделировании предполагается, что каждый конкретный товар имеет свое неслучайное движение цен, обусловленное объективными факторами, –

тренд цены товара. Предполагается также, что отдельные независимые групповые тренды имеют и товарные группы на уровне 10-значных групп. При моделировании задаются два обобщенных тренда цен: тренд цены уровня 10-значных групп – $Tr10$ и тренд цены на уровне товаров – TrT . Тренды по конкретным 10-значным товарным группам – $Tr10_i$ ($i = \overline{1, n_{10}}$, n_{10} – число 10-значных товарных групп) выбираются случайным образом из равномерного распределения с математическим ожиданием, равным $Tr10$ и заданным коэффициентом масштаба вариации – $MTr10$. Аналогично формируются тренды по отдельным товарам – TrT_i ($i = \overline{1, n_{тов}}$, $n_{тов}$ – число товаров в 10-значной товарной группе). При этом используется свой коэффициент масштаба вариации – $MTrT$.

При генерировании данных по анализируемому периоду помимо трендового движения цен необходимо учесть другие аспекты статистической зависимости данных анализируемого периода от данных базисного периода. Модель должна отражать:

- корреляцию между ценами товаров базисного и анализируемого периодов, то есть ценовая структура анализируемого периода должна в основном, с точностью до тренда и случайных колебаний цен, соответствовать ценовой структуре базисного периода;
- корреляцию между количествами товаров базисного и анализируемого периодов, то есть объемная структура анализируемого периода должна в основном, с точностью до случайных колебаний объемов товаров, соответствовать объемной структуре базисного периода;
- корреляцию между стоимостями товаров, то есть стоимостная структура анализируемого периода должна в основном повторять стоимостную структуру базисного периода. Необходимо заметить, что это условие выполняется автоматически, если выполнены два предыдущих условия;
- корреляцию между индексами цены и объема.

Корреляция между ценами товаров анализируемого и базисного периода обеспечивается посредством использования соотношения:

$$(9) \quad p_{1ijk} = p_{0ijk} \times Tr10_i \times TrT_j \times \beta_p,$$

где p_{1ijk} – цена в анализируемом периоде k -ого лота, j -го товара, входящего в i -ю 10-значную товарную группу, p_{0ijk} – цена в базисном периоде k -ого лота, j -го товара, входящего в i -ую 10-значную товарную группу, $Tr10_i$ – тренд цены по i -ой 10-значной товарной группе, TrT_j – тренд по j -му товару i -ой 10-значной товарной группы, β_p – случайная мультипликативная вариация цены, имеющая бета-распределение с математическим ожиданием, равным 1.

Из вышеприведенного соотношения видно, что цена товара анализируемого периода линейно зависит от цены базисного периода, но на эту зависимость накладывается мультипликативная случайная вариация с математиче-

ским ожиданием, равным 1. Это значит, что между ценами анализируемого p_{1ijk} и базисного p_{0ijk} периодов обеспечивается корреляция. Коэффициент корреляции определяется величиной трендов $Tr10_i$, TrT_j и величиной вариации β_p . Коэффициенты корреляции в этом случае не велики, так как реальные значения трендов малы.

Корреляция между количествами товаров базисного и анализируемого периодов моделируется посредством соотношения:

$$(10) \quad g_{1grad\ ijk} = g_{0\ grad\ ijk} \cdot \left(\frac{p_{1\ ijk}}{p_{0\ ijk}} \right)^{\alpha_{инд}} \cdot \beta_g,$$

где $g_{0\ grad\ ijk}$ – количество товара по k -му лоту, j -го товара, входящего в i -ю 10-значную товарную группу, в базисном периоде, $g_{1\ grad\ ijk}$ – количество товара по k -ому лоту, j -го товара, входящего в i -ую 10-значную товарную группу, в анализируемом периоде, $\frac{p_{1\ ijk}}{p_{0\ ijk}}$ – индекс цены по соответствующему лоту

соответствующего товара, $\alpha_{инд}$ – показатель степени коррелированности индексов цены и объема, β_g – мультипликативная случайная вариация, имеющая бета-распределение с математическим ожиданием, равным 1.

Из рассмотрения вышеприведенной формулы видно, что объем товара анализируемого периода коррелирован как с объемом товара базисного периода, так и с индексом цены. При этом корреляция с индексом цены нелинейная (за исключением далекого от реальности случая, когда $\alpha_{инд} = 1$), что соответствует экономической теории. Показатель $\alpha_{инд}$ может быть как положительным, так и отрицательным. При положительном показателе имеет место так называемый «рынок производителя» [8]. Как показали статистические исследования, закономерности рынка производителя наблюдаются в экспорте продукции ТЭК. Для экспорта и импорта большинства товаров характерны закономерности «рынка потребителя», для которого $\alpha_{инд}$ отрицателен и имеет величину порядка нескольких сотых.

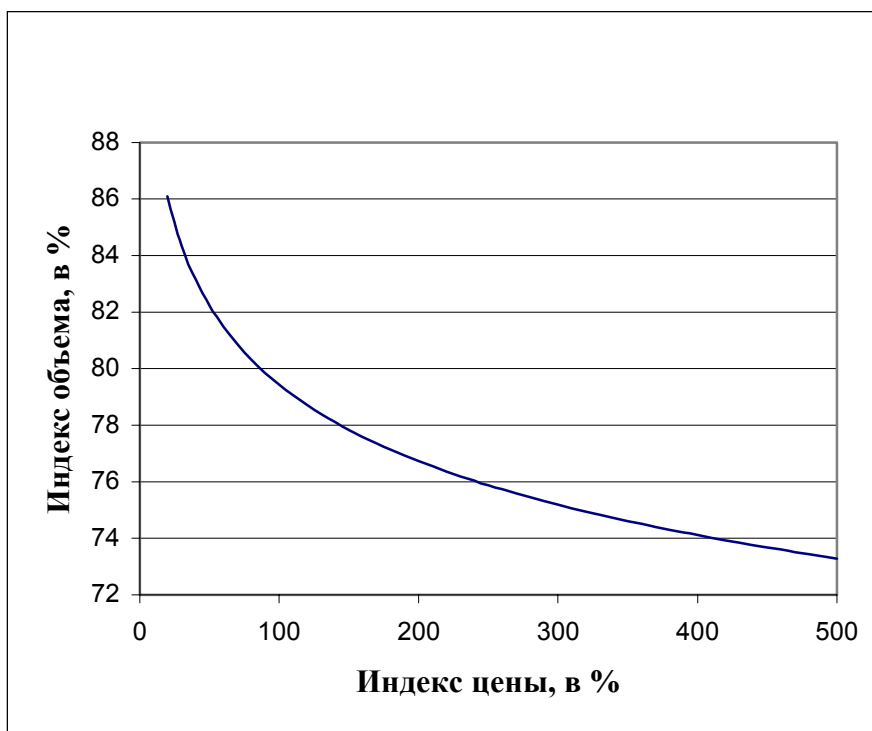


Рис. 3. Регрессия индекса объема по индексу цены

На рис. 3 представлена регрессия индекса объема по индексу цены при $\alpha_{инд} = -0,05$. Характер кривой соответствует теоретической кривой спроса – спрос падает с ростом цены.

3. Некоторые результаты использования статистической модели генерирования данных внешней торговли

Использование генерируемых моделью данных по внешней торговле дает возможность определить ошибки сравниваемых методов в вычислении индексов цен как разности между значением индекса, рассчитанного по анализируемому методу, и значением индекса, рассчитанному по базе данных полной формы (БДПФ).

С использованием описанной в разделе 2 статистической модели был проведен большой объем вычислительных работ. По каждой фиксированной совокупности параметров модели осуществлялось 100 реализаций процесса генерирования данных для вычисления индексов внешней торговли. При каждой реализации формировались три базы данных:

- база данных полной формы – БДПФ;
- база данных формы таможенной службы – БДТС;
- база данных прямого наблюдения по выборке товаров – БДПН.

Для каждой реализации проводились расчеты на иерархических уровнях ТН ВЭД:

- укрупненная аналитическая товарная группа «Минеральное сырье», экспорт в страны СНГ;
- все 4-значные группы, входящие в группу «Минеральное сырье»;
- все 10-значные группы, входящие в группу «Минеральное сырье».

Для каждой из перечисленных групп вычислялись:

- индекс цены Пааше;
- индекс цены Ласпейреса;
- индекс физического объема Ласпейреса;
- индекс стоимости;
- индекс реального объема;
- коэффициент вариации индекса цены;
- коэффициент вариации индекса объема;
- коэффициент корреляции между индексами цены и объема;
- показатели результатов фильтрации аномальных значений индексов и другие показатели.

В табл. 1-3 и на рис. 4-9 представлены результаты сопоставления по уровню ошибок индексов в процентах, вычисленных по данным в форме, используемой ГТК РФ (БДТС), и данным на выборочной основе (БДПН) или данным так называемого «прямого наблюдения». Мотивы выбора в качестве основы моделирования товарной группы экспорта в СНГ группы «минеральное сырье» приведены выше. Сопоставление по точности рассматриваемых методологий на уровне 4-значных товарных групп проводится на примере товарной группы «Цемент» (код 2523), входящей в группу «Минеральное сырье», а на уровне 10-значных групп – на примере товарной группы «Соль, употребляемая в пищу» (код 251009100), также входящей в «Минеральное сырье». Цемент и соль выбраны как типичные группы своего уровня, по которым можно судить о закономерностях на данных уровнях.

Таблица 1.

Показатели вариации ошибок (%) исчисления индексов цен импорта в страны ближнего зарубежья по товарной группе «Минеральное сырье», коды ТН ВЭД 25-27

Уровень товарной группы по ТН ВЭД	Минеральное сырье (25-27)			
	Пааше		Ласпейрес	
Тип индекса	Пааше		Ласпейрес	
Форма данных	БДТС	Выборка	БДТС	Выборка
Показатели вариации ошибки				
Среднее	4,5982	8,1049	4,0283	8,23
Стандартное отклонение	31,5436	46,4500	39,1062	43,87
Асимметричность	1,2561	0,9333	1,3798	1,14
Интервал	171,2391	238,2144	255,1774	193,48
Минимум	-41,4131	-77,4632	-61,1080	-51,36
Максимум	129,8260	160,7512	194,0695	142,13

Таблица 2

Показатели вариации ошибок (%) исчисления индексов цен импорта в страны ближнего зарубежья по товарной группе «Цемент», код ТН ВЭД 2523

Уровень товарной группы по ТН ВЭД	4-значная группа, цемент, код 2523			
	Пааше		Ласпейрес	
Тип индекса	БДТС	Выборка	БДТС	Выборка
Показатели вариации ошибки				
Среднее	8,2950	21,9336	8,2325	148365,33
Стандартное отклонение	27,8971	149,3771	26,3986	1306085,85
Асимметричность	1,0378	6,8975	0,6315	9,76
Интервал	150,9086	1386,0034	117,9587	12927029,18
Минимум	-39,6749	-80,8985	-49,8220	-99,58
Максимум	111,2337	1305,1049	68,1367	12926929,60

Таблица 3

Показатели вариации ошибок (%) исчисления индексов цен импорта в страны ближнего зарубежья по товарной группе «Соль пищевая», код ТН ВЭД 251009100

Уровень товарной группы по ТН ВЭД	10-значная группа, соль, код 251009100			
	Пааше		Ласпейрес	
Тип индекса	БДТС	Выборка	БДТС	Выборка
Показатели вариации ошибки				
Среднее, %	4,7798	19,3287	7,6069	19,32
Стандартное отклонение, %	32,5900	66,7982	32,2968	60,36
Асимметричность	1,2357	0,9701	1,2183	0,91
Интервал, %	193,2710	307,1094	163,1215	290,47
Минимум, %	-63,8040	-80,8696	-41,1721	-75,13
Максимум, %	129,4670	226,2399	121,9493	215,34

Таблица 4

Характеристики распределений значения индекса цены Пааше при различных объемах выборки

Объем выборки, %	5	10	15
Среднее значение индекса	0,9689	0,9603	0,9516
Стандартное отклонение	0,2220	0,1984	0,1727
Асимметричность	1,5275	0,6201	0,5540

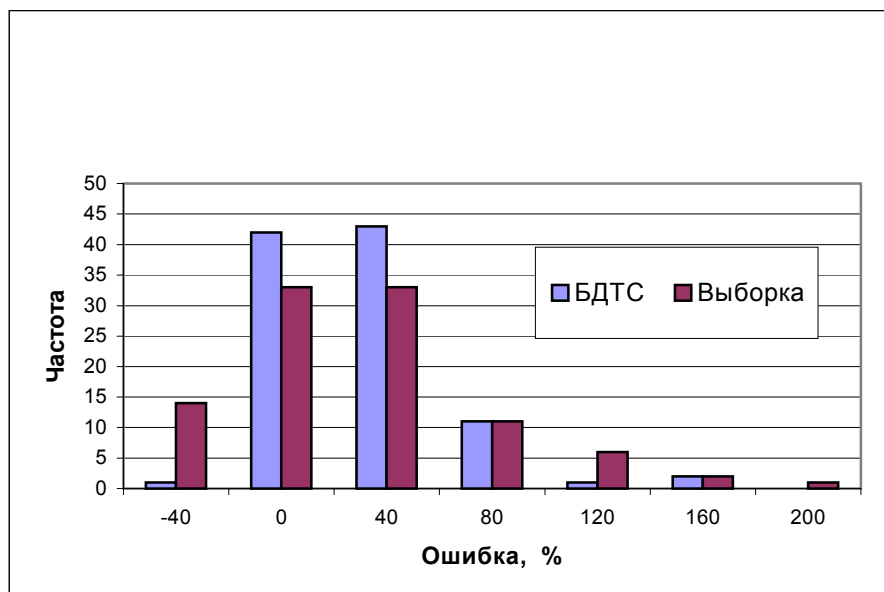


Рис. 4. Распределение ошибки исчисления индексов Пааше, вычисленных по БДТС и по 10%-й выборке. Экспорт в СНГ минерального сырья (коды 25-27)

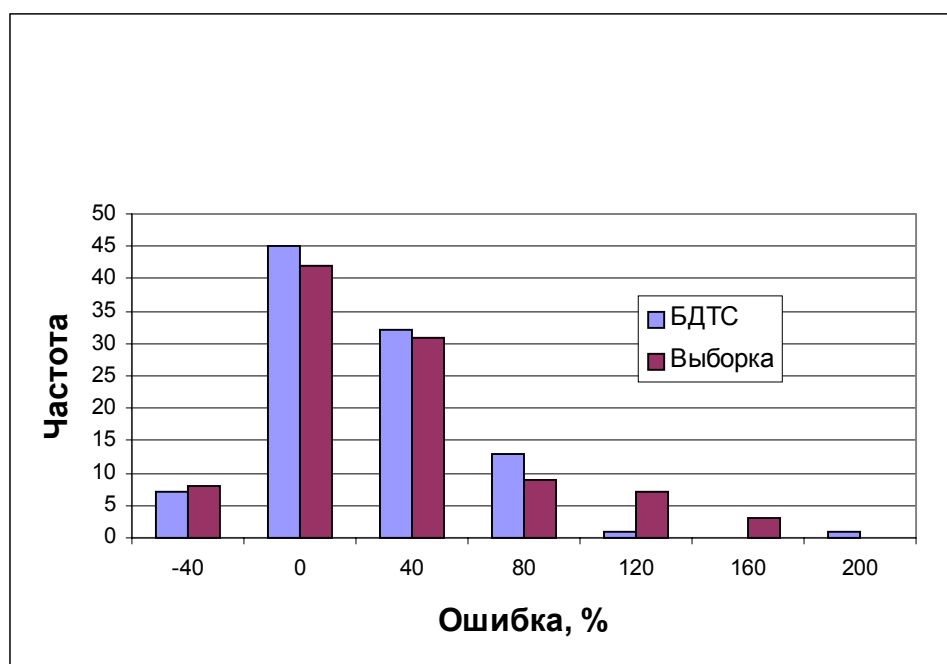


Рис.5. Распределение ошибки исчисления индекса Ласпейреса, вычисленного по БДТС и 10%-й выборке. Экспорт в СНГ минерального сырья (коды 25-27)

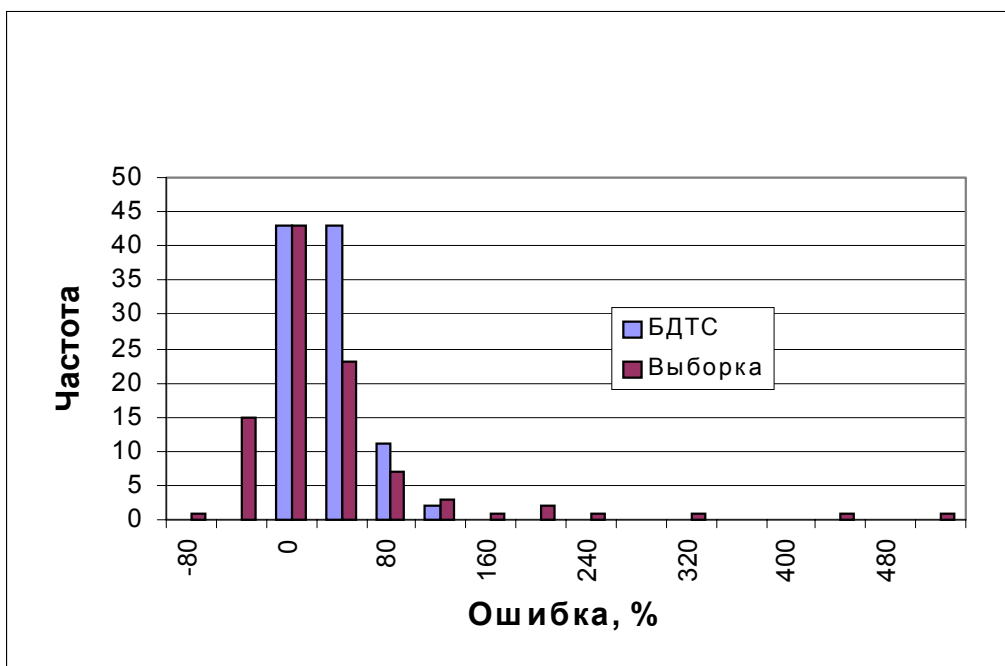


Рис.6. Распределение ошибок исчисления индекса Пааше по БДТС и 10%-й выборке. Экспорт в СНГ, цемент (код 2523)

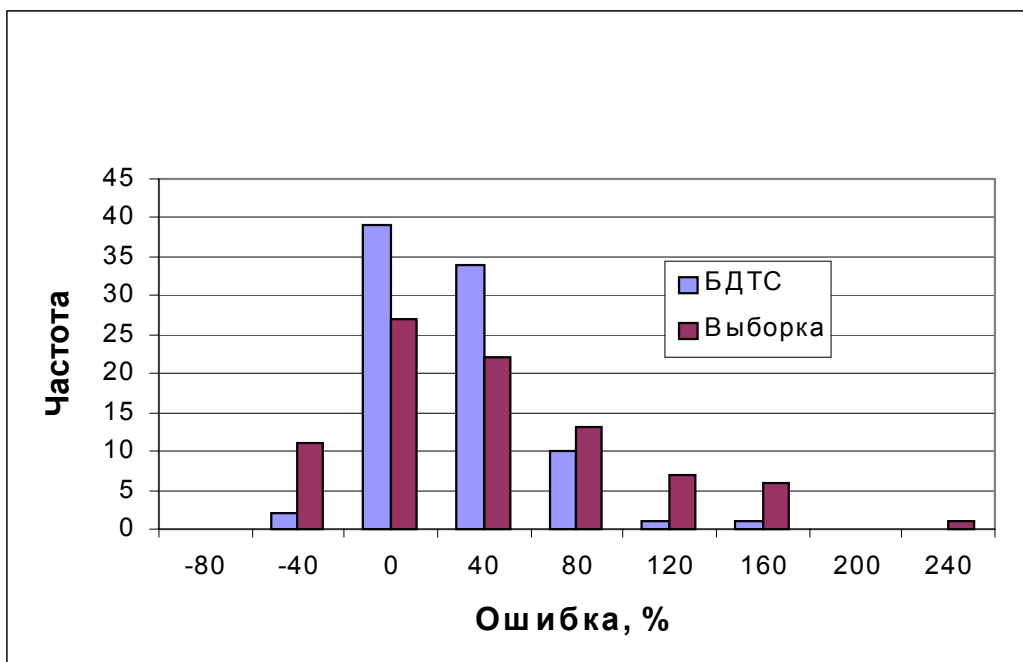


Рис.7. Распределение ошибки исчисления индекса Ласпейреса. Экспорт в СНГ, соль (код 2501009100)

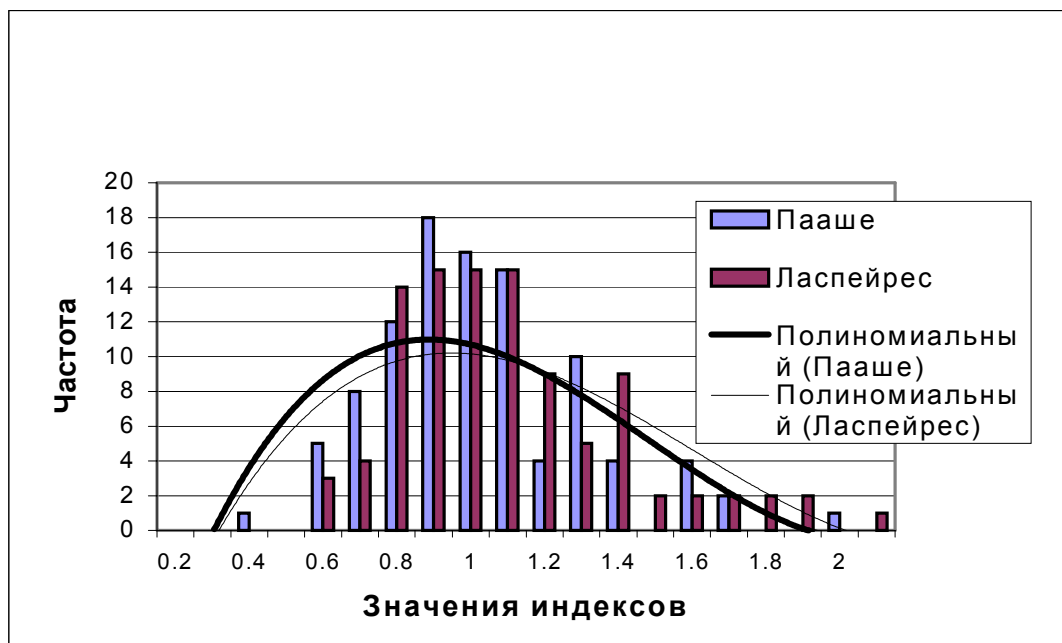


Рис.8. Распределение значений индексов Пааше и Ласпейреса по группе 25-27. Форма данных – БДТС.

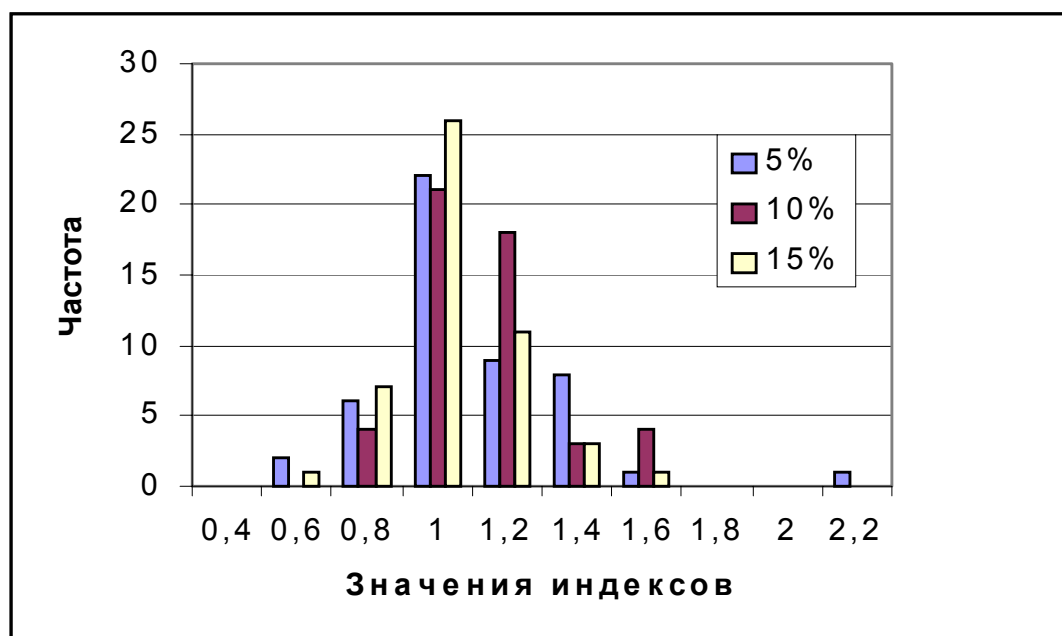


Рис.9. Сравнение распределений значений индекса Пааше при различных объемах выборки. Экспорт в СНГ минерального сырья (коды 25-27)

При рассмотрении таблиц и графиков следует отметить следующие положения.

1. Для укрупненных товарных групп и 10-значных групп ТН ВЭД случайные ошибки исчисления индексов цены как типа Пааше, так и типа Ласпейреса, при использовании данных в форме ГТК РФ (БДТС) имеют уровень в 2-3 раза меньший, чем при использовании 10% товарной выборки (БДПН). Для 4-значных товарных групп 10%-я выборка дает значительно худшие результаты.

2. Высокий уровень ошибок для 4-значных товарных групп по сравнению 10-значными и укрупненными товарными группами объясняется тем, что в 4-значных группах объединяется небольшое число 10-значных групп, каждая из которых может иметь свое движение средней цены. Уровень 4-значных товарных групп можно рассматривать как уровень нестабильности. В 10-значных группах товары имеют, как правило, общую тенденцию в движении цен. На уровне укрупненных товарных групп проявляется закон больших чисел, так как эти группы объединяют большое количество 10-значных групп. Опыт расчета индексов цен внешней торговли показывает, что наибольшие проблемы в смысле устойчивости значений индекса имеют небольшие товарные группы, например, «кожсырье». Причем величина группы определяется не стоимостью входящих в нее товаров, а количеством 10-значных групп. В этом отношении устойчивой является группа-гигант «импорт продовольствия». На фоне малых случайных ошибок групп-гигантов становятся более заметными систематические ошибки вычисления индексов. Моделируемая группа «экспорт минерального сырья» относится к малым группам, поэтому уровень случайных ошибок по ней высок.

3. 10%-я товарная выборка, которая сопоставляется с базой данных ГТК России, является достаточно крупной выборкой и вряд ли реализуема не только в нашей стране, но и в других странах, так как количество импортируемых и экспортируемых товаров составляет многие миллионы. Поэтому прямое наблюдение за ценами внешней торговли на основе простой выборки вряд ли конкурентоспособно по точности исчисления индексов с наблюдением за ценами в форме базы данных ГТК России. Однако наблюдение за ценами товаров-представителей строится не на основе простой случайной выборки, а на основе так называемой «направленной выборки» [6], которая может быть по точности более эффективной, но вряд ли способна снизить ошибки более, чем в 3 раза. Тем более, что из данных табл. 4 и рис. 9, на примере индекса цены Пааше, видно, что вариация индекса по мере увеличения объема выборки уменьшается не обратно пропорционально корню из объема выборки (как это имеет место для оценки среднего [6]), а значительно медленнее. При увеличении объема выборки с 5 до 15 % снижение ошибки (стандартное отклонение) произошло не в 1,73 раза, а только в 1,29 раза.

4. Из рассмотрения табл. 1-3 видно, что средние значения ошибок индексов, вычисленных как на основе БДТС, так и 10%-й выборки, имеют положительные значения, то есть значения индексов как Пааше, так и Ласпейреса имеют постоянные положительные смещения относительно индексов, рассчитанных по данным полной формы. Смещения для индексов, рассчитанных по

выборке, в 2-3 раза больше, а для 4-значных товарных групп могут достигать очень больших значений.

5. Как отмечалось выше, весь предшествующий анализ ошибок исчисления индексов исходил из определения ошибки как отклонения значения индекса, вычисленного на основе анализируемой базы данных, от индекса, вычисленного на основе базы данных в полной форме. Однако известно, что и индексы, рассчитанные по полным данным, могут иметь постоянные смещения. Об этом косвенно свидетельствует формула Борткевича [7], устанавливающая расхождение между индексами Пааше и Ласпейреса в зависимости от вариаций индексов цен и объемов и ковариации между ними. Проведенные на фактических данных расчеты не подтвердили достаточной точности данной формулы, что, по-видимому, связано с нелинейностью регрессии индекса объема по индексу цены, которая молчаливо предполагается при выводе формулы, а также асимметричностью распределений цен и индексов цен. Тем не менее на фактических данных видна зависимость расхождения индексов от перечисленных факторов. Так как для большинства товарных групп как импорта, так и экспорта коэффициент корреляции между индексами объема и цены имеет отрицательный знак, то в соответствии с формулой Борткевича значение индекса Пааше оказывается ниже значения индекса Ласпейреса. Нами на фактических данных внешней торговли была проведена оценка смещения по тесту транзитивности индексов цены Пааше и Ласпейреса [8]. Как правило, индексы Пааше имеют отрицательное смещение, а индексы Ласпейреса – положительное. Масштаб смещения на интервале в один квартал даже по некоторым укрупненным товарным группам доходит до 10%. Методы статистического моделирования позволяют подойти к проблеме оценки смещения с других позиций. Дело в том, что цены товаров анализируемого периода генерируются по регрессионной модели, тренд которой задается, поэтому можно полагать, что истинное движение цен определяется их трендом. Как отмечалось выше, в модели принята схема тренда цен двух уровней: на уровне отдельных товаров и на уровне 10-значных товарных групп, что позволяет вычислить общий тренд цен. В основном блоке моделирования значение тренда цен равнялось 1,0201. Математическое ожидание индекса Пааше составило 0,9723, а индекса Ласпейреса 1,0478. Таким образом, смещение индекса Пааше составило $-0,0478$, а индекса Ласпейреса $+0,0203$. Смещение индекса Фишера равно $-0,0107$. Таким образом, смещение индекса Пааше вниз превышает смещение индекса Ласпейреса вверх более чем в 2 раза. Использование индекса Фишера также не элиминирует смещение полностью. Смещения рассматриваемых индексов видны также и на рис. 5, где виден сдвиг распределений индексов.

4. Заключение

В целом на основе проведенных исследований можно сделать следующие выводы.

1. Рекомендуемый Международным валютным фондом переход от использования стоимостей единиц товаров (в качестве цен внешней торговли) к прямому наблюдению за ценами экспортируемых и импортируемых товаров

представляется в настоящее время вряд ли целесообразным. Как показывают представленные результаты исследований, такой переход не даст в реальных условиях России существенного выигрыша в точности исчисления индексов внешней торговли. Кроме того, переход к прямому наблюдению за внешнеторговыми ценами связан со значительными затратами и проведением соответствующих организационных мероприятий.

2. Исследования показали, что используемой в настоящее время в России (как и во многих других странах) методологии исчисления индексов внешней торговли присущи существенные ошибки как случайного, так и систематического характера. Основной причиной ошибок является большая товарная неоднородность многих низших классификационных групп ТН ВЭД. Специфика России здесь состоит в том, что при разработке данного классификатора, а также при модернизации его пока не учитываются интересы исчисления индексов внешней торговли и вообще интересы ведения динамических рядов внешней торговли. Кроме того, в торговле со странами СНГ существует постоянно меняющаяся система преференций, обуславливающая большие вариации цен на некоторые товары. Основным направлением снижения уровней ошибок исчисления индексов внешней торговли следует считать снижение вариации цен внутри низших классификационных групп за счет использования соответствующих фильтров и кластерных процедур.

3. Предложенная статистическая модель открывает методические перспективы по сравнительному анализу различных видов индексов и методологий их исчисления. По существу это тестирующая программа, она генерирует исходные данные с заданным движением цен, что создает возможность сопоставить оценки индексов, полученные тем или иным методом, с фактическим движением цен. В частности, модель полезна при выборе и отработке различных процедур предварительной обработки данных, в том числе фильтров аномальных значений в данных.

Литература

1. Russian Federation: Preliminary Report on the Mission to Provide Technical Assistance on Price Statistics. December 11–12, 1995. International Monetary Fund. Statistics Department.
2. External Trade Statistics. Seminar in Vilnius. Final papers, 1992.
3. Handbook of Methods. Chapter 15. International Price Indexes. Bureau of Labor Statistics, 1997.
4. Methodological Notes on Deflation of Trade Statistics. London: Central Statistical Office, 1994.
5. Методологические положения по статистике. Вып. 3. М.: Госкомстат России, 2000.
6. Деев Г.И. Несплошное статистическое наблюдение. Т. 1. М.: Госкомстат России, 2001.
7. Bortkevich V.I. Zweck und Struktur einer Preisindexzahl. Nordisk Statistisk Fidskrift. bd. II. 1923.

8. Отчет по НИР «Совершенствование методологии построения индексов внешней торговли». М.: РТА, 2002.
9. Аллен Р. Экономические индексы. М.: Статистика, 1980.
10. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. М.: ЮНИТИ, 1998.

Оглавление

1. Введение	3
2. Методология генерирования данных, используемых для вычисления индексов внешней торговли	6
2.1. Моделирование структуры данных	6
2.2. Весовая (объемная) структура данных базисного периода.....	7
2.3. Генерирование цен товаров и лотов базисного периода	8
2.4. Моделирование тренда цен и корреляционных связей данных анализируемого и базисного периодов.....	10
3. Некоторые результаты использования статистической модели генерирования данных внешней торговли	13
4. Заключение	20
Литература.....	21

Препринт WP2/2002/05
Серия WP2
Количественный анализ в экономике

Деев Гелий Иванович
Родительская Елена Викторовна
Оценка ошибок исчисления индексов внешней торговли
с использованием методов статистического моделирования

Публикуется в авторской редакции
Оформление серии *А.М. Павлов*
Корректор *Е.Е. Андреева*

ЛР № 020832 от 15 октября 1993 г.
Отпечатано в типографии ГУ ВШЭ с представленного оригинал-макета.
Формат 60×84 $\frac{1}{16}$. Бумага офсетная. Тираж 150 экз. Уч.-изд.л. 1,41. Усл.печ.л.
1,40.

Заказ №345. Изд. №263.
ГУ ВШЭ. 117312, Москва, ул. Вавилова, 7а
Типография ГУ ВШЭ. 125319, Москва, Кочновский проезд, 3

Препринты ГУ ВШЭ
Серия WP2 "Количественный анализ в экономике"

<http://stat.hse.ru>

2002

Деев Г.И., Родительская Е.В. Оценка ошибок исчисления индексов внешней торговли с использованием методов статистического моделирования: Препринт WP2/2002/05. – М.: ГУ ВШЭ, 2002. – 24 с.

Суворов Н.В. Макроэкономическое моделирование технологических изменений (теоретические, прикладные и инструментальные вопросы): Препринт WP2/2002/04. – М.: ГУ ВШЭ, 2002. – 80 с.

Ершов Э.Б. Теория ключевых и межотраслевое моделирование: Препринт WP2/2002/03. – М.: ГУ ВШЭ, 2002. – 64 с.

Писляков В.В. Анализ контента ведущих электронных ресурсов актуальной зарубежной периодики: Препринт WP2/2002/02. – М.: ГУ ВШЭ, 2002. – 33 с.

Губанов В.А. Непараметрическое выделение динамических сезонных циклов: Препринт WP2/2002/01. – М.: ГУ ВШЭ, 2002. – 33 с.

2001

Поспелов И.Г. Экономические агенты и системы балансов: Препринт WP2/2001/03. – М.: ГУ ВШЭ, 2001. – 68 с.

Бессонов В.А. Об измерении динамики российского промышленного производства переходного периода: Препринт WP2/2001/02. – М.: ГУ ВШЭ, 2001. – 34 с.

Алексеенкова М.В. Факторы отраслевого анализа для российской переходной экономики: Препринт WP2/2001/01. – М.: ГУ ВШЭ, 2001. – 34 с.