

Владимир Аркадьевич БЕССОНОВ

(Национальный исследовательский университет Высшая школа
экономики)

e-mail: bessonov@hse.ru

ЧТО СОХРАНИТ ДЛЯ ИСТОРИИ СОВРЕМЕННАЯ РОССИЙСКАЯ СТАТИСТИКА?

XV Апрельская международная научная конференция по проблемам
развития экономики и общества, НИУ ВШЭ

Сессия С-07 "Статистическое моделирование макроэкономических
процессов"

03 апреля 2014 г., 15:00–16:30, Мясницкая, 20, аудитория 238

Проблема

- Откуда возьмется будущая историческая статистика о нашем времени?

Будет по крупицам восстановлена грядущими поколениями исследователей?

Или в нее со временем превратится текущая статистика?

- Произойдет ли трансформация текущей статистики в будущую историческую сама собой?
- Что необходимо делать для обеспечения этой трансформации?

Обсуждаются вопросы, связанные лишь с той частью статистики, данные которой используются для анализа экономической динамики

- Предпосылки создания исторической статистики о нашем времени

Расцвет информационных технологий

Переход плановых экономик к рынку – одно из интереснейших событий экономической истории последнего столетия

План выступления

- Проблема
- Требования к исторической статистике о нашем времени
 - Временные ряды
 - Данные высокой частоты
 - Системы показателей
 - Метаданные
- Организация процесса трансформации
 - "Резервуар" исторической статистики
- Обсуждение

Подробнее см.:

Бессонов В.А. Что сохранит для истории современная российская статистика?
Полный текст доклада на сайте Конференции (<http://conf.hse.ru/2014/prog>)

Временные ряды

- Для решения задач анализа экономической динамики данные должны быть представлены временными рядами

Временной ряд – элементарный информационный "кирпичик"

- Альтернатива – значение показателя для некоторого периода времени

Различие – уровни временного ряда связаны между собой требованием сопоставимости

Временной ряд – единый развивающийся объект

Аналогия – дерево

- Проблема: российская экономическая статистика не вполне ориентирована на построение временных рядов

Скорее, это – статистика отдельных числовых величин

- С течением времени происходит утрата сопоставимости показателей

Данные по производству ВВП в ОКОНХ доступны ныне лишь с 1995 г. Ранее существовали оценки с 1991 г.

Месячные индексы промышленного производства в разрезе крупных отраслей ОКОНХ ныне доступны с 1995 г. Рассчитывались с 1992 г.

В 2010 г. – переход от ОКП к ОКПД. Результат – снижение степени сопоставимости многих показателей

Тогда же – полная утрата сопоставимости уровней использования производственных мощностей

Конец 2010 г. – Министерство финансов повысило лимит стоимости активов, относящихся к материально-производственным запасам, с 20 тыс.руб. до 40 тыс.руб. Это привело к скачкообразному снижению оценки инвестиций в основной капитал

- Методики описывают построение отдельных уровней показателя

Не оговариваются вопросы обеспечения сопоставимости во времени

Не рассматриваются вопросы сохранения преемственности при уточнениях весов и корзин

Не описана техника сцепления сегментов временных рядов

Не обсуждаются вопросы обеспечения гладкости досчетных составляющих

Не описаны алгоритмы обеспечения согласованности динамики месячных, квартальных и годовых данных

- Практика опубликования длинных временных рядов не развита

Статистические сборники – данные за последние годы

В ЦБСД, ЕМИСС и на официальном сайте Росстата многие ряды охватывают не всю историю построения показателя

Часть рядов разрезана с вынесением ретроспективной части в архив

ИПЦ в ЦБСД – с 1995 г., хотя они строятся с 1991 г. Участки до 2001 г. вынесены в архив

В ЦБСД данные по производству отдельных видов промышленной продукции по месяцам – лишь с 1996 г. Разрезаны на фрагменты до 2004 г. и после

Через совокупность публикаций видим российскую экономику сквозь движущееся временное окно длиной около десятилетия

- Выводы

Российская статистика не вполне ориентирована на временные ряды, т.е. на объекты, адекватные задачам анализа экономической динамики

Российская статистика подобна ящерице – время от времени она избавляется от "хвоста" ретроспективных данных

Данные высокой частоты

- Данные какой частоты (годовые, квартальные, месячные) заслуживают сохранения?

Современная российская статистика отдает предпочтение сохранению данных годовой динамики

Месячные и квартальные данные публикуются в "Краткосрочных экономических показателях Российской Федерации" не ранее 1999 г.

В ЦБСД и ЕМИСС они приведены с разной степенью детализации в разрезе показателей и территорий и за разные интервалы времени

- Для чего нужны данные высокой частоты?

Высокочастотные составляющие несут основную часть информации

Позволяют повысить разрешающую способность во временной области

Аналогия – насколько обеднеет мониторинг текущих тенденций, если месячные и квартальные данные перестанут быть доступными

- В развитых странах такие данные публикуют

В США временной ряд ИПЦ в месячном выражении – с января 1913 г.

В США временной ряд индекса промышленного производства в месячном выражении – с января 1919 г.

- В России сохранению данных высокой частоты препятствует практика их опубликования

В развитых странах – показатели в базисном виде

В России – показатели в цепном виде

Ср. Российский статистический ежегодник и Statistical Abstract of the United States

Для данных годовой динамики это не порождает проблем

Для месячных и квартальных данных проблемы возникают, если не публикуются показатели по отношению к предыдущему периоду

- **Пример: динамика производства продукции сельского хозяйства**

Публикуются только данные по отношению к соответствующему месяцу предыдущего года

Данные по отношению к предыдущему месяцу рассчитываются, но не публикуются (аналогия – купил билет, но не поехал)

В "Информации о социально-экономическом положении России" публикуют данные к предыдущему месяцу максимум за два календарных года

Если восстановить длинный ряд к предыдущему месяцу, то получим, что за 2012 г. прирост производства продукции сельского хозяйства составил 13.0%

Годовая официальная оценка дает спад на 4.8%

Расхождение между двумя оценками – 17.8 п.п. (!)

Каков статус месячных данных? "Как бы официальные данные"?

- **Вывод**

Данные высокой частоты – "нелюбимый ребенок" российской статистики

Системы показателей

- Статистические службы строят системы показателей, связанных отношениями иерархической и территориальной подчиненности

Пример – ИПЦ

- Это – более сложные структуры данных по сравнению с временными рядами

В системе на показатели накладываются дополнительные ограничения

Аналогия – лес

- Какая часть этих огромных массивов данных должна сохраняться и в каком виде?

До какого уровня агрегирования данных необходимо обеспечить доступ?

- Для чего необходим доступ к дезагрегированным данным?

Агрегированные данные – целостный взгляд ценой утраты части информации

Индивидуальные индексы – видны детали, но за деревьями не видно леса

Возможность анализа с разной разрешающей способностью

- В настоящее время в России системы показателей целиком не сохраняются

Малая глубина во времени и утрата сопоставимости данных

Не все показатели систем сохраняются

Не сохраняется информация о структуре систем показателей

В ЦБСД и ЕМИСС разные системы показателей приведены с разной степенью детализации и на разную глубину по времени

В ЦБСД и ЕМИСС системы показателей представлены отдельными показателями. Отношения иерархической и территориальной подчиненности отражены не вполне адекватно

Избыточные справочники, товарные группы – в алфавитном порядке без кодов и без информации об иерархии, территории – в странном порядке, веса отсутствуют, либо не привязаны к показателям

Система показателей не представлена в ЦБСД и ЕМИСС как единый объект

Ее нельзя целиком выгрузить из базы данных

- Все системы показателей образуют текущую статистику

Системы показателей не вполне независимы – они связаны единой логикой

Аналогия – вся экосистема

- Логика всей текущей статистики не вполне отражена в публикациях

В наибольшей мере она отражена в статистических сборниках

В ЦБСД группы показателей приведены в иной последовательности (в алфавитном порядке на разных уровнях иерархии)

В ЕМИСС показатели сгруппированы по ведомствам

- Выводы

"Верблюд" российской статистики не пролазит через "игольное ушко" системы опубликования данных

Поскольку даже текущая статистика публикуется не полностью и не вполне адекватно, то надежды на ее трансформацию в будущую историческую статистику иллюзорны

Метаданные

- Уровни и динамика агрегированных показателей определяются, в том числе, и методиками их построения

- Методики эволюционируют

Эволюция объекта исследования вынуждает адаптировать методики

Развиваются представления о том, как следует проводить измерения

Изменяются возможности проведения измерений

- Результат – разные сегменты макроэкономических временных рядов бывают построены по различающимся алгоритмам

Измеряем изменяющийся объект изменяющимся аршином

- Пользователю должна быть доступна полная временная последовательность действовавших вариантов методики

Иначе "истинные" изменения нельзя отличить от обусловленных изменением методики

Пример с инвестициями в основной капитал – выше

Пример с ИПЦ. С начала 2013 г. введена уточненная методика. Ранее веса определялись на основе структуры потребительских расходов населения за один смещенный год. Теперь – за два. Использование устаревшей информации для построения весов смещает ИПЦ вверх. "Возраст" весов увеличился. В результате увеличилось смещение. Оценка роста ИПЦ за 2013 г. – 6.5%. По прежней методике было бы меньше, скажем, 6.0%

- В настоящее время публикуется либо единственный вариант методики, либо несколько без указания интервалов времени, когда они действовали
- Методики далеко не всегда полны

Пример маловразумительной методики – ВВП в постоянных ценах

- С течением времени утрачивается понимание того, как были построены показатели

Выяснение возможно лишь при личном контакте потребителя информации с ее производителем

Для исторической статистики это невозможно

В настоящее время многие детали методик сохраняет лишь устная традиция

Статистический эпос, акыны от статистики

- Вывод

Даже если показатели будут адекватно строиться и сохраняться, все равно будут проблемы с их корректным использованием

"Резервуар" исторической статистики

- Где должна аккумулироваться историческая статистика?
- В России нет "длинных" статистических сборников

В развитых странах они есть. Пример – Historical Statistics of the United States

- Не аккумулируются данные и на официальных электронных ресурсах
- Вывод: "резервуар" отсутствует
- Как организовать процесс его создания

Начинать создавать большие программные системы принято с анализа требований

Затем – определение спецификаций, программная реализация и т.д.

Пошаговая детализация

- Сейчас делается иначе

Стала ли новая версия ЦБСД лучше?

Для чего нужны две дублирующие друг друга системы – ЦБСД и ЕМИСС?

Для чего отдельные массивы размещаются еще и на сайте Росстата?

Для чего могут быть использованы данные, появившиеся в результате деятельности Открытого правительства?

- Подключение к процессу разработки системы ее будущих пользователей – принципиально

Похоже, что сейчас процесс разработки электронных систем отдан на откуп специалистам по информационным технологиям

Они не могут заменить пользователей при определении требований к системе

Пример: заданные по умолчанию параметры в ЦБСД и ЕМИСС

Набор функциональных блоков – нужна ли графика?

- Требования к системе

Простота, аскетичность, консервативность

Минимальные системные требования, переносимость

Дешевизна, минимальные эксплуатационные расходы

Самодостаточность – пользователь должен иметь возможность работать с ней без привлечения дополнительной информации

"Прозрачность" для пользователя

Не следует стремиться к созданию системы в стиле Барокко

Адекватные структуры данных

Стандартные требования к базам данных – полнота, неизбыточность, непротиворечивость, актуальность

Аккумуляирование метаданных

Обсуждение

- Текущая статистика не трансформируется в будущую историческую статистику о нашем времени

По крайней мере – в части данных, ориентированных на анализ экономической динамики

- Причины

Наследие плановой системы

Особенности переходного периода

Система подготовки кадров

- Живем ли мы в доисторическую эпоху?

Это зависит от тех решений, которые нами будут приняты или не приняты