

***ИНСТИТУТ ЭКОНОМИКИ  
ПЕРЕХОДНОГО ПЕРИОДА***

*Научные труды  
№ 46Р*

**ПРОБЛЕМЫ ПРОГНОЗИРОВАНИЯ  
НЕКОТОРЫХ  
МАКРОЭКОНОМИЧЕСКИХ  
ПОКАЗАТЕЛЕЙ**

**Москва  
2002**

## **Институт экономики переходного периода**

В работе проведен сравнительный анализ качества различных методов прогнозирования временных рядов в применении к рядам, отражающим динамику развития основных макроэкономических показателей, характеризующих состояние экономики России. В применяемых в работе подходах выделены две составляющие – традиционное эконометрическое моделирование и моделирование с использованием структур случайных векторов.

**Авторский коллектив:** Р.М. Энтов, В.П. Носко, А.Д. Юдин,  
П.А. Кадочников, С.С. Пономаренко

Редактор: Н. Главацкая  
Корректор: С. Хорошкина  
Компьютерный дизайн: В. Юдичев

*Настоящее издание подготовлено по материалам исследовательского проекта Института экономики переходного периода, выполненного в рамках гранта, предоставленного Агентством международного развития США.*

ISBN 5-93255-093-7

**Лицензия на издательскую деятельность Серия ИД № 02079  
от 19 июня 2000 г.**

125993, Москва, Газетный пер., 5.

Тел. (095) 229-6413, FAX (095) 203-8816

**E-MAIL** – root @iet.ru, **WEB Site** – <http://www.iet.ru>

© **Институт экономики переходного периода 2002**

## Содержание

Предисловие .....	7
<b>Часть I. Прогнозирование временных рядов традиционными эконометрическими методами и их модификациями .....</b>	<b>9</b>
<b>Глава 1. Постановка проблемы и инструментарий исследования .....</b>	<b>9</b>
1.1. Общие соображения .....	9
1.2. Постановка задачи .....	12
1.3. Результаты некоторых предыдущих исследований .....	15
1.3.1. Сравнение линейных и нелинейных моделей. Комбинирование прогнозов .....	15
1.3.2. Сравнение одномерных и многомерных неадаптивных и адаптивных моделей .....	20
1.3.3. Сравнение TS- и DS-моделей .....	22
1.3.4. Моделирование и оценка стабильности соотношений между макроэкономическими показателями .....	29
<b>Глава 2. Исследование сравнительного качества прогнозов некоторых макроэкономических рядов РФ, получаемых по фиксированным и рекурсивным TS- и DS-моделям .....</b>	<b>35</b>
2.1. Денежные ряды .....	35
2.1.1. M0 .....	35
2.1.2. M1 .....	44
2.1.3. M2 .....	54
2.2. Экспорт .....	63
2.3. Налоговые доходы федерального бюджета .....	67
2.4. Безработица .....	77
2.5. Индекс интенсивности промышленного производства .....	80
2.6. Индекс интенсивности производства цветных металлов .....	84
2.7. Предпочтительность модели на временном интервале оценивания и качество прогнозов .....	93

2.8. Сравнение прогнозов, полученных по выбранным моделям, с «наивными» прогнозами	96
<b>Глава 3. Влияние на сравнительное качество последовательностей прогнозов длины интервала, на котором это сравнение производится</b>	103
3.1. Денежные ряды	103
3.1.1. M0	103
3.1.2. M1	108
3.1.3. M2	115
3.2. Экспорт	119
3.3. Налоговые доходы федерального бюджета	121
Общий вывод	126
<b>Глава 4. Источники ошибок прогнозов и некоторые методы их коррекции</b>	127
4.1. Источники ошибок прогнозов	127
4.2. Календарный эффект	130
4.3. Коррекция прогнозов методами «back-on-track» и «back-on-average»	131
4.3.1. M0	131
4.3.2. M1	136
4.3.3. M2	145
4.3.4. Налоговые доходы федерального бюджета	148
4.3.5. Экспорт	153
4.3.6. Индекс интенсивности производства цветных металлов	157
4.3.7. Индекс интенсивности промышленного производства	164
4.3.8. Безработица	168
<b>Глава 5. Моделирование и прогноз налоговых поступлений в консолидированный и федеральный бюджеты РФ</b>	171
5.1. Поступления подоходного налога	172
5.2. Поступления налога на прибыль	176
5.3. Поступления налога на прибыль (в федеральный бюджет РФ)	179

5.4. Поступления налога на добавленную стоимость .....	182
5.5. Суммарные налоговые поступления .....	186
5.6. Суммарные налоговые поступления (в федеральный бюджет РФ) .....	189
5.7. Сравнение результатов многошаговых прогнозов эконометрических моделей и модели оценки поступлений (Revenue Estimating Model) .....	192
5.7.1. Поступления подоходного налога .....	193
5.7.2. Поступления налога на прибыль .....	194
5.7.3. Поступления налога на прибыль (в федеральный бюджет РФ) .....	195
5.7.4. Поступления налога на добавленную стоимость .....	196
5.7.5. Суммарные налоговые поступления .....	197
5.7.6. Суммарные налоговые поступления (в федеральный бюджет РФ) .....	198
5.8. Основные результаты и выводы .....	199
<b>ЧАСТЬ II. Структуры случайных векторов .....</b>	<b>201</b>
<b>Глава 6. Прогнозирование с использованием структур случайных векторов .....</b>	<b>201</b>
6.1. Предварительные замечания .....	201
6.1.1. Непосредственные связи между элементами статистической системы .....	201
6.1.2. Структуры случайных векторов .....	203
6.1.3. Существенная размерность и информативные структуры. ....	205
6.1.4. Прогнозирование с использованием структур случайных векторов .....	207
6.2. Постановка проблемы и алгоритмы .....	208
6.2.1. Постановка проблемы .....	208
6.2.2. Алгоритмы моделирования и прогнозирования .....	213
6.3. Результаты расчетов .....	221
6.3.1. Инфляция .....	221
6.3.2. Денежные агрегаты .....	227
6.3.3. Динамика внешнеторговых характеристик ..	245

6.3.4. Безработица	257
6.3.5. Доходы федерального бюджета	263
6.3.6. Валовой внутренний продукт	276
<b>Заключение. Некоторые выводы</b>	
<b>из полученных результатов</b>	286
<b>Приложения</b>	301
<b>Приложение 1. Макросы</b>	301
П1.1. Построение трехмерных информаций	301
П1.2. Построение прогнозов	307
<b>Приложение 2. Исходные данные.</b>	309
П2.1. Инфляция	309
П2.2. Денежные агрегаты (M0, M1, M2)	310
П2.3. Экспорт и импорт	311
П2.4. Безработица	312
П2.5. Доходы федерального бюджета	312
П2.6. Валовой внутренний продукт	313
П2.7. Индекс интенсивности промышленного производства	313
П2.8. Индекс интенсивности производства цветных металлов (SA)	314
П2.9. Индекс интенсивности производства цветных металлов (NSA)	314
<b>Литература</b>	315

## Предисловие

Настоящая работа посвящена анализу временных рядов некоторых макроэкономических показателей, характеризующих экономическую ситуацию в Российской Федерации.

Разнообразные содержательные задачи экономического анализа требуют использования статистических данных, характеризующих исследуемые экономические процессы и развернутых во времени в форме временных рядов. При этом нередко одни и те же временные ряды используются для решения разных содержательных проблем. Одной из важных задач макроэкономической политики, включая исследование динамики фундаментальных переменных, является построение моделей, позволяющих осуществлять прогнозирование макроэкономических показателей на будущие периоды.

Целью настоящего исследования является сравнительный анализ качества различных методов прогнозирования временных рядов в применении к рядам, отражающим динамику развития основных макроэкономических показателей, характеризующих состояние экономики России. В используемых в работе подходах можно выделить две составляющие – традиционное эконометрическое моделирование и моделирование с использованием структур случайных векторов. Для того чтобы разграничить эти подходы и получаемые результаты, работа разделена на две части.

Первая часть состоит из 5 глав и посвящена сравнительному исследованию качества традиционных эконометрических методов прогнозирования временных рядов и некоторых их модификаций.

В первой главе излагается общая постановка проблемы и инструментарий исследования, применяемый в первой части работы.

Вторая глава посвящена предварительному исследованию сравнительного качества последовательности одношаговых прогнозов для некоторых рядов макроэкономических показателей РФ.

В третьей главе исследуется устойчивость выводов сравнительного анализа качества прогнозов по альтернативным моделям при расширении интервала, на котором производится сравнение качества прогнозов.

В четвертой главе обсуждаются источники ошибок прогнозов и некоторые методы их коррекции. Анализируются последствия коррекции прогнозов в применении к некоторым из рассматриваемых макроэкономических рядов.

В пятой главе в качестве примера практического применения моделей

прогнозирования производится сравнение различных моделей для прогноза поступлений отдельных налогов и суммарных налоговых поступлений в консолидированный и федеральный бюджеты РФ.

Во второй части исследования рассматривается альтернативный подход к построению прогнозов временных рядов, основанный на использовании структур случайных векторов. Этот подход, в отличие от традиционного эконометрического подхода, используемого в первой части, не предполагает каких-либо гипотез о функциональном виде случайной составляющей временного ряда. Данная часть работы состоит из основных определений, постановки задачи, описания алгоритмов построения прогнозов, а также включает результаты расчетов.

В заключении к работе приводятся краткое описание полученных результатов исследования, сравнение различных методов прогнозирования и основные выводы.

В приложениях приведены макросы (программы) построения моделей и прогнозов на основе структур случайных векторов, а также исходные данные для расчетов.



# **Часть I. Прогнозирование временных рядов традиционными эконометрическими методами и их модификациями**

## **Глава 1. Постановка проблемы и инструментарий исследования**

### **1.1. Общие соображения**

Проблема точности макроэкономических прогнозов всегда привлекала и привлекает к себе большое внимание. При этом возникает целый ряд вопросов, на которые совсем не просто ответить. После получения данных о том, что же в действительности произошло за период действия прогноза, заявленного ранее, естественно возникает желание (1) измерить ошибки прогнозов, (2) объяснить их и (3) научиться уменьшать эти ошибки в будущем.

Макроэкономическое прогнозирование является важным инструментом для проведения и мониторинга экономической политики. Достаточно точная оценка экономического роста, инфляции, безработицы, доходов бюджета и других макроэкономических показателей в будущие периоды позволяет своевременно принимать те или иные меры, осуществлять более эффективное регулирование экономики. Построение относительно точных прогнозов может быть периодическим или случайным, при этом прогресс в прогнозировании предполагает наличие некоторых не меняющихся зависимостей, которые можно надлежащим образом идентифицировать и использовать.

Экономические прогнозы обычно охватывают короткие периоды времени. Длинные последовательности точных прогнозов можно встретить чрезвычайно редко. Более того, мало кто из прогнозистов оставляет свои модели и методы неизменными в течение долгого периода времени, поскольку они ищут улучшения и пытаются адаптировать прогнозы к новым тенденциям в экономике. Поэтому данные о предшествующих результатах конкретного прогнозиста часто являются весьма ненадежной базой для выводов о том, как он будет действовать в будущем.

Еще более рискованно пытаться ранжировать прогнозистов по тому, сколь хорошо они предсказали изменение в частном коротком периоде, скажем, за год или за несколько лет. В каждом таком случае некоторые прогнозисты окажутся лучше других просто в силу случая или других причин. Говорить о таком превосходстве можно только при устойчивой разнице в результатах прогнозов во времени.

Сравнение многопериодных прогнозов с точки зрения статистических выводов осложняется коррелированностью таких прогнозов, а значит и их ошибок, по крайней мере, в двух отношениях: (а) внутри каждой последовательности прогнозов с заданным базовым периодом, (б) между следующими друг за другом последовательностями многопериодных прогнозов, которые частично относятся к одному и тому же периоду прогнозирования. Каждый многопериодный прогноз зависит от предыдущих прогнозов, являясь до некоторой степени их ревизией. Получающаяся в итоге сложная корреляционная структура препятствует сравнительному оцениванию качества прогнозов, затрудняя интерпретацию полученных мер точности прогнозов (среднее, смещение и т.п.).

Успешность прогнозирования в весьма высокой степени зависит от стабильности общей экономической ситуации. В этой связи можно отметить очень большие ошибки документированных прогнозов ВВП в США, сделанных непосредственно после Второй мировой войны. Так (см. [Zarnowitz (1978)]), одна весьма уважаемая группа частных прогнозистов прогнозировала на 1947 г. 6%-ное убывание ВВП, в то время как в действительности имел место рост ВВП примерно на 11%. Та же группа затем предсказала на 1948 г. незначительное уменьшение ВВП, хотя в действительности он опять показал столь же бурный рост. И такое положение было общим в те годы: прогнозы опирались на данные и соотношения, имевшиеся в 30-е гг., и ложные аналогии с периодом, непосредственно следовавшим за Первой мировой войной. Напротив, начиная с 1953 г., прогнозы ВВП становятся более точными. Следует также отметить, что в большинстве случаев завышенные прогнозы соответствовали годам экономического спада, а заниженные прогнозы – годам, когда ВВП существенно возрастал. К сожалению, «обучение на прошлых ошибках» дает здесь не очень много, т.к. ошибки прогнозов на коротком временном интервале либо имеют малую автокоррелированность, либо вовсе не коррелированы (при прогнозах на один шаг вперед).

Средняя точность прогнозов обычно убывает с увеличением горизонта прогнозирования. Так, ВВП прогнозируется на один квартал вперед лучше, чем на два квартала, на два квартала вперед лучше, чем на три квартала, и

т.д., но эта разница постепенно сглаживается с удалением в будущее. В определенном смысле можно говорить о правиле «длиннее прогнозы – больше ошибки». Каждая из потенциальных составляющих прогноза – экстраполяция временных рядов и связей между ними, внешняя информация, а также мнения экспертов – подвержена ухудшению при удлинении горизонта прогнозирования.

Весьма большие ошибки прогнозов характерны для поворотных точек экономических циклов. Можно сказать, что наличие небольшого количества ошибок этого типа является необходимой платой за то, чтобы избежать большого числа существенных ошибок в промежутках между поворотными точками путем применения процедур оптимального оценивания, таких, как метод наименьших квадратов. Разумеется, было бы желательным каким-то образом предугадывать поворотные точки, чтобы уменьшить ошибки прогнозов.

Качество прогнозов зависит и от того, рассматриваются ли прогнозы для номинальных величин или прогнозы для реальных величин. Так, в том же исследовании [Zarnowitz (1978)] указывалось на существенно лучшее качество прогнозов для номинального, а не реального ВВП, что объяснялось значительно большей вариабельностью темпов роста ВВП в долларах в постоянных ценах.

Методы автопрогноза, основанные на анализе временных рядов, экстраполируют имеющийся в наличии ряд только на основании информации, содержащейся в нем самом. Такого рода прогноз может оказаться эффективным лишь в кратко- и, максимум, в среднесрочной перспективе. Серьезный подход к решению задач долгосрочного прогнозирования требует использования комплексных подходов, и в первую очередь привлечения различных (в том числе статистических) технологий сбора и анализа экспертных оценок. Как бы то ни было, как отмечалось в работе [Klein (1984)], среди прочих субъектов науки экономика отличается высоким отношением «шум/сигнал», и это ее внутреннее свойство, так что «следует учиться жить в такой ситуации, делая по возможности все, чтобы улучшать точечные прогнозы и указывать границы возможных ошибок».

Точность прогнозов зависит как от объективных условий, таких как природа прогнозируемой переменной и длина горизонта прогнозирования, так и от атрибутов самого прогнозиста, таких как теория (идеология), которой он следует (кейнсианский или монетаристский подход и т.п.), и методика, посредством которой эта теория используется для построения количественного прогноза. (Исследование совместного влияния последних факторов проводилось в работе [Batchelor, Dua (1990)].)

Исследования, проведенные еще в работах [McNees (1979)] и [McNees, Ries (1983)] на базе основных макроэкономических рядов США, ясно показали, что (а) точность прогноза оказывается наилучшей для медленно меняющихся переменных и является наихудшей для быстро меняющихся переменных (цены активов); (б) точность краткосрочных прогнозов выше, чем точность долгосрочных прогнозов.

Эффективный подход к решению задач кратко- и среднесрочного автопрогноза это прогнозирование, основанное на использовании «подогнанных» (идентифицированных) моделей типа  $ARIMA(p, k, q)$ , включая, в качестве частных случаев, и модели AR-, MA- и ARMA. (См., например, работы [Айвазян, Мхитарян (1998)] и Приложение П2 в работе [*Эконометрический анализ динамических рядов основных макроэкономических показателей (2001)*]).

Хотя публикуемые прогнозы достаточно часто превосходят по качеству прогнозы, сделанные по одномерным моделям ARIMA, последние дают хорошие стандарты, на которые следует ориентироваться при построении прогноза, учитывающего и дополнительную информацию.

В решении прикладных задач кратко- и среднесрочного автопрогноза весьма широко распространены так называемые адаптивные методы, позволяющие по мере поступления новых данных обновлять ранее сделанные прогнозы с минимальной задержкой и с помощью относительно несложных математических процедур.

## 1.2. Постановка задачи

При анализе временных рядов основное внимание уделяется описанию или моделированию их структуры. Построенная модель часто используется для экстраполяции или прогнозирования временного ряда, и тогда качество прогноза может служить полезным критерием при выборе среди нескольких альтернативных моделей.

В работе [*Эконометрический анализ динамических рядов основных макроэкономических показателей (2001)*] детально рассматривался вопрос об отнесении конкретных макроэкономических рядов к классу рядов, стационарных относительно детерминированного тренда (или просто стационарных) –  $TS$  (trend stationary) ряды, или к классу рядов, имеющих стохастический тренд (возможно, наряду с детерминированным трендом) и приводящихся к стационарному (или стационарному относительно детерминиро-

ванного тренда) ряду только путем однократного или  $k$ -кратного<sup>1</sup> дифференцирования ряда –  $DS$  (difference stationary) ряды.

Принципиальное различие между этими двумя классами рядов выражается в том, что в случае  $TS$ -ряда вычитание из ряда соответствующего детерминированного тренда приводит к стационарному ряду, тогда как в случае  $DS$ -ряда вычитание детерминированной составляющей ряда оставляет ряд нестационарным из-за наличия у него стохастического тренда.

Траектории  $TS$ - и  $DS$ -рядов отличаются друг от друга кардинальным образом.  $TS$ -ряды имеют линию тренда в качестве некоторой «центральной линии», которой следует траектория ряда, находясь то выше, то ниже этой линии, с достаточно частой сменой положений выше–ниже.  $DS$ -ряды помимо детерминированного тренда (если таковой имеется) имеют еще и так называемый стохастический тренд, из-за присутствия которого траектория  $DS$ -ряда весьма долго пребывает по одну сторону от линии детерминированного тренда (выше или ниже соответствующей прямой), удаляясь от нее на значительные расстояния, так что по существу в этом случае линия детерминированного тренда перестает играть роль «центральной» линии, вокруг которой колеблется траектория процесса.

В  $TS$ -рядах влияние предыдущих шоковых воздействий затухает с течением времени, а в  $DS$ -рядах такое затухание отсутствует и каждый отдельный шок влияет с одинаковой силой на все последующие значения ряда. Поэтому наличие стохастического тренда может потребовать проведения определенной экономической политики для возвращения макроэкономической переменной к ее долговременной перспективе, тогда как при отсутствии стохастического тренда возвращение к долгосрочному значению осуществляется благодаря внутренним свойствам макроэкономических процессов и самой переменной, – в этом случае макроэкономическая переменная «скользит» вдоль линии тренда как направляющей, пересекая ее достаточно часто и не уклоняясь от этой линии сколько-нибудь далеко.

$DS$ - и  $TS$ -модели одних и тех же временных рядов могут приводить к совершенно различным прогнозам (см., например, [Diebold и Senhadji (1996)]). Поэтому решение о том, какую из этих моделей использовать, чрезвычайно важно для прикладных прогнозистов. Вместо того чтобы употреблять одну из этих моделей по умолчанию, можно использовать критерий единичного корня как диагностический инструмент для выбора решения. Фактически, одной из ранних мотиваций построения критериев единично-

---

<sup>1</sup> Мы не затрагиваем здесь вопрос о возможной дробной интегрированности рядов.

го корня было именно оказание помощи в определении того, использовать ли прогностические модели в разностях или в уровнях в конкретных приложениях (см., например, [Dickey, Bell, Miller (1986)]).

Значительная часть публикаций, касающихся проблемы единичного корня, была сфокусирована на неспособности критериев единичного корня отличать при конечных выборках нулевую гипотезу о наличии единичного корня от близких стационарных альтернатив (см., например, [Christiano, Eichenbaum (1990)], [Rudebusch (1993)]). Однако низкая мощность не обязательно является проблемой для прогнозирования. Например, долгое время утверждалось, что если корень процесса близок к единице, но все же меньше единицы, то точность прогнозов можно улучшить, используя модель в разностях, а не в уровнях [Box, Jenkins (1976)]. В конечном счете, интерес для прогнозирования представляет скорее не вопрос о том, выбирают ли критерии единичного корня «истинную» модель, а вопрос о том, выбирают ли эти критерии модели, которые дают более качественные прогнозы.

Сравнительные достоинства таких стратегий, как «всегда дифференцировать», «никогда не дифференцировать» или «иногда дифференцировать, следуя результатам предварительного теста на единичный корень», в общем случае зависят от степени инерционности (последствия, «persistence») истинного процесса, горизонта прогнозирования, размера выборки и свойств теста. Одной из основных целей работы является исследование того, до какой степени выбор *TS*- или *DS*-модели влияет на качество прогноза некоторых российских макроэкономических рядов.

Ошибки прогнозов возникают вследствие целого ряда факторов. Как показывают исследования, приведенные в работах [Clements, Hendry (1998b)] и [Clements, Hendry (2000)], наиболее драматическое ухудшение качества прогнозов наблюдается при изменении параметров процесса порождения данных, особенно при изменении детерминированных составляющих этого процесса. В связи с этим еще одной целью работы является исследование возможностей некоторых методов коррекции прогнозов, направленных на улучшение качества прогнозов в условиях нестабильности процесса порождения данных.

### **1.3. Результаты некоторых предыдущих исследований**

#### *1.3.1. Сравнение линейных и нелинейных моделей. Комбинирование прогнозов*

В статье [Stock, Watson (1996)] обсуждались пять вопросов, относящихся к прогнозированию макроэкономических рядов.

1. Могут ли нелинейные модели временных рядов давать в реальном времени прогнозы, лучшие по сравнению с линейными моделями?
2. Можно ли улучшить прогнозы на полгода или год вперед, используя предварительные данные об инерционности временных рядов для выбора модели прогнозирования?
3. Могут ли комбинации прогнозов превосходить по точности прогнозы, основанные на каком-то одном методе, для целого спектра временных рядов, и если да, то в каком отношении эти комбинации должны взвешивать наилучшие на данный момент методы прогнозирования?
4. Есть ли выгода от использования этих сложных методов по сравнению с простыми авторегрессионными прогнозами, достаточная для того, чтобы оправдать их использование даже прогнозистом, не склонным к риску?

Для ответа на эти вопросы в работе проводится соответствующий эксперимент. В этом эксперименте производится сравнение различных прогнозов на одномесечном, шестимесечном и двенадцатимесечном горизонтах для 215 экономических временных рядов США. Эксперимент симулирует применение этих методов в реальном времени, т.е. все прогнозы (включая оценки всех параметров, все правила выбора модели, все предварительные тесты и т.п.) базируются исключительно на наблюдениях до даты построения прогноза, включая эту дату. Оценки параметров, статистики для выбора модели, предварительные тесты и веса в комбинации прогнозов обновляются для всех моделей ежемесячно, и эти обновленные статистики используются для построения прогнозов на будущие месяцы.

Рассматриваются прогнозы, получаемые 49 методами прогнозирования. Они называются именно методами, а не моделями, поскольку многие из этих прогнозов базируются не на единственной оцененной модели, а на результатах выбора одной из нескольких моделей в результате предвари-

тельного тестирования или применения критериев выбора модели. Индивидуальные модели, используемые этими методами прогнозирования, называются здесь примитивными (первичными) моделями, и таких моделей всего 121. Например, одним из методов прогнозирования является авторегрессия в уровнях с постоянной составляющей и выбором максимального порядка запаздывания (порядка модели авторегрессии) на основе информационного критерия Акаике (AIC), с порядком, изменяющимся от 0 до 12; в терминологии настоящей работы этот метод прогнозирования комбинирует информацию от 13 примитивных моделей. Примитивные модели подразделяются на четыре класса: авторегрессии (AR), экспоненциальное сглаживание (EX), искусственные нейронные сети (ANN) и логистические гладкие переходные авторегрессии (LSTAR). В качестве дополнительной «опорной метки» рассматривается также прогноз «отсутствие изменений».

Обсуждаются также различные процедуры комбинирования информации от этих 49 методов прогнозирования (процедуры объединения прогнозов). В работах [Bates, Granger (1969)], [Granger, Newbold (1977)], [Granger, Ramanathan (1984)] было показано, что усреднение прогнозов по различным моделям может улучшить качество прогноза, когда все модели являются только аппроксимациями. Процедуры объединения, рассматриваемые в статье, различаются весом, приписываемым модели с наилучшим на данный момент качеством, включая равные веса для всех прогнозов, взвешивание прогнозов в обратной пропорции к текущей среднеквадратической ошибке (MSE), использование медианных прогнозов и приписывание веса только методу, на настоящий момент имеющему наименьшую симулированную в реальном времени MSE; последняя процедура состоит в симулировании выбора в реальном времени модели на основе наименьших квадратов предсказаний (PLS).

[Markidakis et al. (1982)] исследовали качество одномерных методов на многих рядах, в том числе и экономических, и сделали вывод об успешности во многих случаях экспоненциального сглаживания. [Meese, Geveke (1984)] сравнивали различные линейные модели, используя 150 макроэкономических рядов, и нашли, что обычно хорошо работают AR модели с длинной запаздываний, определяемой на основании AIC. Они обнаружили также, что линейные комбинации прогнозов улучшают качество прогнозов не существенно. [Weigland, Gershenfeld (1994)] сравнивали линейные модели с большим количеством нелинейных моделей. Хотя они обнаружили нелинейную динамику в некоторых неэкономических временных рядах, нели-



нейные прогностические модели оказались относительно плохими для экономических рядов, которые они рассматривали (обменные курсы). [Swanson, White (1995, 1997)] сравнивали многомерные ANN модели с линейными векторными авторегрессиями и нашли, что векторные авторегрессии обычно имеют меньшие MSE, чем ANN модели в симулированном реальном времени. В сравнении с указанными работами, работа [Stock J.H., Watson M.W. (1996)] отличается использованием большого количества макроэкономических временных рядов и нелинейных моделей, анализом методов с тестированием на единичный корень и интенсивным исследованием процедур объединения прогнозов.

Все модели в данном эксперименте имеют вид:

$$y_{t+h} = f_i(Z_t; \theta_{ih}) + u_{t+h}$$

где  $y_t$  – ряд, для которого строится прогноз,  $h$  – горизонт прогнозирования,  $i$  – индекс модели прогнозирования ( $i=1, \dots, 121$ ),  $\theta_{ih}$  – вектор неизвестных параметров,  $u_{it}$  – ошибка, а  $Z_t$  – вектор предикторных переменных. Обычно  $Z_t = (y_t, \dots, y_{t-p}, \Delta y_t, \dots, \Delta y_{t-p}, 1, t)$  где  $p$  – максимальное запаздывание. Типичным является использование индивидуальной прогностической моделью только некоторого подмножества элементов  $Z_t$ .

Все прогнозы выполняются полностью рекурсивным способом, т.е. прогнозы значения  $y_{t+h}$  используют информацию для моментов  $1, 2, \dots, t$ . Для прогнозирования  $y_{t+h}$  оценка вектора параметров  $\theta_{ih}$  производится по наблюдениям  $y_1, y_2, \dots, y_t$ . Во всех моделях вектор параметров оценивается путем минимизации суммы квадратов остатков прогноза на  $h$  шагов вперед.

Заметим, что обычно каждый метод прогнозирования в применении к конкретному ряду имеет различные значения параметров при различных горизонтах, т.е. прогноз на  $h$  шагов вперед не вычисляется как итерация на  $h$  периодов вперед прогностической модели для одношагового прогноза. (Прогноз на  $h$  шагов вперед и ошибка прогноза равны, соответственно,

$$y_{t+h/t,ih} = f_i(Z_t; \theta_{ih}^*), e_{t+h/t,ih} = y_{t+h} - y_{t+h/t,ih}$$

Этот подход имеет и преимущества, и недостатки. Если модель для одношагового прогноза правильна, то оценивание ее на однопериодном горизонте с последующей итерацией вперед более эффективно, чем непосредственное оценивание прогностической модели на  $h$  шагов вперед. Вместе с тем, имея в виду неправильность спецификации моделей, прямое оценивание модели для прогноза на  $h$  шагов вперед дает возможность уменьшения влияния неправильной спецификации на этом горизонте. Обсуждению вопроса о том, в каких ситуациях второй подход дает лучшие результаты, пос-

выщена работа [Clements, Hendry (1996)]. С практической точки зрения, предсказание непосредственно по модели, рассчитываемой на  $h$  шагов вперед, требует значительных затрат машинного времени для оценивания параметров, но зато существенно упрощает вычисление многошаговых прогнозов по нелинейным моделям.

Для получения основных результатов все прогнозы в работе [Stock, Watson (1996)] автоматически «подстригаются», так что если прогнозируемое изменение превосходит по абсолютной величине все ранее наблюдавшиеся изменения для данного ряда, то применяется прогноз «без изменений». Такая коррекция применяется для симуляции вовлеченности в прогнозирование человеческого фактора, отсутствующего в проводимом компьютеризованном эксперименте. Поскольку прогнозы в этом эксперименте выполняются автоматически, некоторые модели могут давать (и фактически дают) экстремальные маловероятные прогнозы. Возможные источники таких экстремальных прогнозов включают оценки параметров, соответствующие локальным, а не глобальным максимумам в нелинейных моделях, скачкообразное изменение параметров, ошибки, возникающие из-за необоснованного включения в модели детерминированных трендов. В действительности реальном времени такие «невероятные» прогнозы были бы замечены и скорректированы вмешательством человека. Соответственно, используемый алгоритм «выстригания» можно понимать как некоторое практическое правило, которое прогнозист может использовать в реальном времени для обнаружения такой проблемы и принятия соответствующего решения. Хотя основное внимание уделяется подстриженным прогнозам, для сравнения представлены также некоторые результаты для «неподстриженных» прогнозов.

*Стартовый и прогнозный периоды.* Для каждого ряда имеется три отдельных периода: стартовый, на котором производятся начальные оценки модели; промежуточный, на котором производятся прогнозы по 121 примитивным моделям 49 прогностическими методами, но без использования объединения прогнозов, и период симулированного прогноза в реальном времени, на котором получаются рекурсивные прогнозы всеми моделями, всеми методами и процедурами объединения. Пусть  $T_0$  дата первого наблюдения, используемого в этом исследовании. Тогда стартовый период оценивания длится с  $T_0$  до  $T_1$ , где  $T_1 = T_0 + 120$  (месяцев). Промежуточный период – с  $T_1$  до  $T_2 - 1$ , где  $T_2 = T_1 + 24$  (месяцев). Прогнозный период – с  $T_2$  до  $T_3$ , где  $T_3$  – дата последнего наблюдения (12.1996) минус горизонт прогноза  $h$ .

Все результаты, касающиеся качества прогнозов, приводимые в таблицах, соответствуют периоду прогнозирования в симулированном реальном времени (от  $T_2$  до  $T_3$  включительно). Для большинства рядов датой начала наблюдений является 01.1959; в этом случае  $T_0=01.1959$ ,  $T_1=01.1970$ ,  $T_2=01.1972$ ,  $T_3=12.1996 - h$ .

#### Прогностические модели и методы

Авторегрессионные (AR) модели. Результаты приводятся для 18 различных авторегрессионных методов прогнозирования. Они отличаются глубиной запаздываний (3 варианта), тем, включаются ли в модель только постоянная или постоянная и временной тренд (2 варианта), а также степенью инерционности (3 варианта).

В отношении глубины запаздываний берутся три альтернативных варианта: фиксированная глубина, равная 4; глубина, определяемая BIC ( $0 \leq p \leq 12$ ); глубина, определяемая AIC ( $0 \leq p \leq 12$ ).

В отношении инерционности временных рядов также берутся три альтернативных варианта. В первом авторегрессия специфицируется в уровнях, т.е. производится прогноз  $y_{t+h}$  по значениям  $y_t, \dots, y_{t-p+1}$ , без всяких ограничений на коэффициенты. Во второй накладывается условие единичного корня, зависимой переменной является  $y_{t+h} - y_t$ , а предикторами  $y_t, \dots, y_{t-p+1}$ . В третьей используется рекурсивный претест на единичный корень, по результатам которого производится выбор между спецификациями в уровнях или в разностях. Для проверки гипотезы о наличии единичного корня существует большое количество критериев. В работе Stock (1996) методом Монте-Карло исследовались авторегрессионные прогнозы с предварительным тестированием на единичный корень. Stock сравнил несколько различных претестовых методов с различными уровнями значимости и нашел, что наилучшее качество прогнозов при различных значениях авторегрессионного корня давало использование DF-GLS критерия [Elliot, Rothenberg, Stock (1996)] с малым уровнем значимости. Соответственно, в работе Stock, Watson используется претестовая статистика DF-GLS<sup>u</sup> для различения моделей с константой, но без тренда, и претестовая статистика DF-GLS<sup>t</sup> – для различения моделей с включением линейного тренда.

Всего оценивалось 52 примитивных модели (2 спецификации детерминированных составляющих, 13 выборов глубины запаздываний, модели в уровнях или в разностях). Основанные на этих моделях 18 методов прогнозирования включают рекурсивный выбор модели, использующий информационный критерий, и/или рекурсивные претесты на единичный корень.

Основные выводы:

1. Хотя для некоторых рядов нелинейные прогнозы оказались лучше линейных, большинство нелинейных методов прогнозирования и все методы, основанные на нейронных сетях, приводят к прогнозам, имеющим худшее качество, чем прогнозы, получаемые линейными методами.
2. Улучшению прогнозов для всех рассмотренных горизонтов прогнозирования способствует предварительное тестирование ряда на наличие единичного корня.
3. Комбинирование прогнозов, получаемых различными методами, служит защитой от слишком больших ошибок прогнозов и может быть рекомендовано для использования в практике прогнозирования.
4. Если прогнозист предполагает использовать какой-то один метод прогнозирования, то в этом случае следует рекомендовать использование модели авторегрессии с предварительным тестированием на наличие единичного корня и выбором порядка авторегрессии на основании имеющихся статистических данных.

*1.3.2. Сравнение одномерных и многомерных неадаптивных и адаптивных моделей*

Исследование на эту тему было проведено в работе [Stock, Watson (1996)], в которой сравнивались результаты прогнозирования большого количества различных экономических рядов с использованием одномерных и двумерных (векторных) моделей авторегрессии, динамических моделей с авторегрессионно распределенными разностями и адаптивных вариантов этих моделей, в которых коэффициенты изменяются с течением времени.

Для прогнозирования были использованы четыре группы моделей по четыре модели в каждой группе.

Одномерная авторегрессия.

AR(p) – оценивание параметров по методу наименьших квадратов, выбор порядка модели по BIC – байесовскому информационному критерию Шварца, прогнозирование по фиксированной оцененной модели.

RRA1 – оценивание параметров посредством скользящей регрессии по 120 наблюдениям, выбор порядка модели по сколь-

заящему BIC, прогнозирование по переоцениваемым моделям.  
RRA2 – оценивание параметров посредством скользящей регрессии по 240 наблюдениям, выбор порядка модели по скользящему BIC, прогнозирование по переоцениваемым моделям.  
RLSA – оценивание параметров посредством рекурсивной регрессии, выбор порядка модели по рекурсивному BIC, прогнозирование по переоцениваемым моделям.

Динамическая модель ADL( p,p) с изменяющимися параметрами, следующими модели случайного блуждания  $\theta_t = \theta_{t-1} + \eta_t$ , с независимыми и одинаково распределенными  $\eta_t$ , ковариационные матрицы которых различаются для различных моделей коэффициентом пропорциональности  $\lambda$ .

ATVP1 – оценивание параметров по методу TVP (Time varied parameters) с  $\lambda = 0.0025$ , выбор порядка модели по минимуму суммы квадратов ошибок прогнозов на один шаг.

ATVP2 – оценивание параметров по методу TVP с  $\lambda = 0.0075$ , выбор порядка модели по минимуму суммы квадратов ошибок прогнозов на один шаг.

ATVP3 – оценивание параметров по методу TVP с  $\lambda = 0.015$ , выбор порядка модели по минимуму суммы квадратов ошибок прогнозов на один шаг.

ATVP4 – оценивание параметров по методу TVP с  $\lambda$  и  $p$ , выбираемыми рекурсивно по минимуму суммы квадратов ошибок прогнозов на один шаг.

Векторная (двумерная) авторегрессия.

VAR – оценивание параметров по методу наименьших квадратов, выбор порядка модели по BIC – байесовскому информационному критерию Шварца, прогнозирование по фиксированной оцененной модели.

RRA1 – оценивание параметров посредством скользящей регрессии по 120 наблюдениям, выбор порядка модели по скользящему BIC, прогнозирование по переоцениваемым моделям.

RRA2 – оценивание параметров посредством скользящей регрессии по 240 наблюдениям, выбор порядка модели по скользящему BIC, прогнозирование по переоцениваемым моделям.

RLSA – оценивание параметров посредством рекурсивной регрессии, выбор порядка модели по рекурсивному BIC, прогнозирование по переоцениваемым моделям.

Динамическая модель ADL(p,p).

VTVP1 – аналог VAR.

VTVP2 – аналог RRA1.

VTVP3 – аналог RRA2.

VTVP4 – аналог RLSV.

Величина параметра  $\lambda$  представляет в рамках TVP модели меру нестабильности в соответствующих соотношениях.

Исходный период для оценивания модели (in-sample) – [01.1959–12.1978].

Период, на котором сравнивались прогнозы на один шаг вперед – [01.1979–12.1993].

Адаптивные модели часто давали более точные прогнозы, чем фиксированные модели (оцененные по исходному периоду), что говорит о структурной нестабильности.

Всего было рассмотрено 5700 пар рядов, и для каждого из указанных выше 16 методов были найдены проценты случаев, когда этот метод оказывался лучшим по отношению к другим методам. Результаты оказались неоднозначными.

После этого для каждого ряда были отобраны пары с 10 наименьшими значениями BIC, построенными по исходному периоду. При этом оказалось, что для разных рядов лучшими оказывались разные методы, причем для многих рядов одномерные прогнозы были лучше VAR и ADL прогнозов.

При сравнении прогнозов по всем парам среди одномерных лучшей оказалась модель ATVPI, а среди двумерных – VTVP1. При этом ATVPI обычно превосходила двумерные модели, но это, возможно, объясняется тем, что пары формировались без какого бы то ни было обоснования.

### 1.3.3. Сравнение TS- и DS-моделей

В работе [Franses, Kleibergen (1996)] проведено сравнение TS- и DS-прогнозов для 14 макроэкономических рядов, рассматривавшихся в известной работе [Nelson, Plosser (1982)], продленных до 1988 г. Для каждого из этих рядов подбирались авторегрессионные TS- и DS-модели одинакового порядка. Эти модели оценивались для трех периодов: до 1976 г., до 1970 г. и до 1952 г. На основании оцененных моделей строились прогнозы на один и на несколько шагов вперед, соответственно, для периодов 1977–1988, 1971–1988 и 1953–1988 гг., что давало, соответственно, 12, 18 и 36 пар ошибок прогнозов.

Далее для каждой из моделей вычислялись MSEP (средний квадрат ошибок предсказания), MAE (средняя абсолютная ошибка прогноза) и разности между значениями этих показателей для TS- и DS-моделей. Для про-

верки статистической значимости этих разностей в случае одношаговых прогнозов использовались непараметрические статистические критерии: критерий знаков и критерий Уилкоксона. При этом, несмотря на малое количество прогнозов, использовалась нормальная аппроксимация статистик этих критериев.

Основной вывод: там, где различия между точечными прогнозами оказались статистически значимыми, *DS*-прогнозы работали лучше.

Помимо указанного исследования в статье [Franses, Kleibergen (1996)] строились также одношаговые прогнозы по рекурсивным регрессиям. Точнее говоря, сначала для каждого ряда оценивалась модель по данным до 1952 г. включительно, по оцененной модели строился прогноз на 1953 г. Затем модель переоценивалась по данным, включающим уже и 1953 г., и по новой оцененной модели строился прогноз на 1954 г. и т.д. Такой подход допускает непостоянство коэффициентов модели авторегрессии, описывающей ряд. При таком подходе для 5 из 14 рядов *DS*-модель работала значимо лучше, для 7 из 14 рядов различие в результатах прогнозирования было статистически незначимым, и только лишь для одного ряда (Consumer Prices) значимо лучшей оказалась *TS*-модель.

Обсуждая полученные результаты, авторы отмечают, что изменение в параметре тренда *TS*-процесса в послевыборочный (out-of-sample) период может приводить к тому, что ошибки прогнозов по «правильной» *TS*-модели будут больше, чем ошибки прогнозов по «неправильной» *DS*-модели. *DS*-процесс способен быстрее адаптироваться к изменениям структурного параметра, по крайней мере на одношаговых прогнозах.

Если изменений структуры нет и ряд порождается *TS*-моделью, то тогда, по мнению Franses, Kleibergen, прогнозы по *TS*-модели должны быть лучше прогнозов по неправильной *DS*-модели, по крайней мере теоретически.

Между тем в работе [Campbell, Perron (1991)] моделировались стационарные ряды, близкие к интегрированным, затем производился подбор *TS*- и *DS*-моделей по смоделированным данным, и по этим моделям строились прогнозы поведения рядов на один и 20 шагов вперед. Альтернативные модели сравнивались по величине среднеквадратической ошибки прогноза. Полученные результаты говорили о том, что стационарные ряды, близкие к интегрированным, лучше прогнозировать, считая их интегрированными, т.е. пренебрегая результатами соответствующих тестов.

В работе [Clements, Hendry (2000)] предпринято исследование качества прогнозов по *DS*- и *TS*-моделям в зависимости от того, соответствует ли используемая для прогнозирования статистическая модель (SM) истинному

процессу порождения данных (DGP – data generating process). Эта задача достаточно сложна, и поэтому авторы ограничились только простейшими моделями  $DS$ - и  $TS$ -рядов.

Первоначально в работе в качестве представителя  $DS$  рядов используется процесс случайного блуждания со сносом

$$y_t = y_{t-1} + \mu + \varepsilon_t, \text{ где } \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \text{ и } y_0 = 0,$$

а в качестве представителя  $TS$ -рядов используется процесс

$$y_t = \varphi + \gamma + u_t, \text{ где } u_t \sim NID(0, \sigma_u^2).$$

Исследуется поведение дисперсий ошибок прогнозов, сделанных в момент  $T$  на  $h$  периодов вперед, в четырех ситуациях:

DGP –  $DS$ -ряд; SM –  $DS$ -ряд;

DGP –  $DS$ -ряд; SM –  $TS$ -ряд;

DGP –  $TS$ -ряд; SM –  $TS$ -ряд;

DGP –  $TS$ -ряд; SM –  $DS$ -ряд.

Соответствующие дисперсии обозначаются как  $V_{DS/DS}$ ,  $V_{TS/DS}$ ,  $V_{TS/TS}$  и  $V_{DS/TS}$  соответственно.

При этом отдельно разбираются случаи, когда неопределенность, проистекающая от оценивания параметров статистической модели, игнорируется, и случаи, когда такая неопределенность учитывается. В случаях, когда такая неопределенность учитывается, оценки параметра  $\mu$  статистической  $DS$ -модели зависят от значения  $u_T$ , которое, в свою очередь, является случайной величиной. Производя усреднение по всем возможным значениям  $u_T$  (вычисляя соответствующие математические ожидания), авторы получают явные выражения для ожидаемых значений отношений дисперсий прогнозов, строящихся по альтернативным статистическим моделям при фиксированном  $DGP$ , т.е. для  $E_{DS}[V_{TS/DS}/V_{DS/DS}]$  и  $E_{TS}[V_{DS/TS}/V_{TS/TS}]$ .

Полученные явные выражения для последних характеристик зависят от  $T$  и  $h$  довольно запутанным образом, что затрудняет их интерпретацию при конечных значениях  $T$  и  $h$ . Поэтому авторы исследуют асимптотическое поведение указанных характеристик при неограниченном увеличении  $h$  и конечных или также неограниченно возрастающих значениях  $T$ . Поскольку практически приходится иметь дело все же с умеренными значениями  $T$  и малыми значениями  $h$ , для нас больший интерес представляют приведенные в работе результаты статистического моделирования (методом Монте-Карло).

Результаты статистического моделирования указывают на то, что для умеренных значений  $T$  (не превосходящих 200) и коротких горизонтов прог-



нозирования (на несколько периодов вперед) значения  $E_{TS}[V_{DS/TS} / V_{TS/TS}]$  близки к 2. Что касается значений  $E_{DS}[V_{TS/DS} / V_{DS/DS}]$ , то при коротких горизонтах прогнозирования наблюдается их линейное возрастание с увеличением  $T$ , что приводит, например, при  $h = 1$  и  $T = 200$  к значению  $E_{DS}[V_{TS/DS} / V_{DS/DS}]$  порядка 28. Это означает фактически, что если истинный DGP нам не известен, то при коротких горизонтах прогнозирования в определенном смысле «безопаснее» использовать для прогнозирования оцененную DS-модель.

Последний вывод, впрочем, относится к простейшим моделям TS- и DS-рядов, указанным выше. Для проверки устойчивости полученных результатов авторы предпринимает далее статистическое моделирование, использующее несколько более сложные TS- и DS-модели процесса порождения данных.

В качестве TS DGP используется процесс  $y_t = \phi + \gamma + u_t$ , где  $u_t = \rho u_{t-1} + a_t$ ,  $a_t \sim NID(0, \sigma_a^2)$ ,  $0 \leq \rho \leq 1$ .

В этом случае та же, что и раньше, статистическая DS-модель может более успешно аппроксимировать TS DGP при значениях  $\rho$ , близких к 1.

В качестве DS DGP берется процесс  $y_t = y_{t-1} + \mu + \varepsilon_t$ , где  $\varepsilon_t = b_t - \psi b_{t-1}$ ,  $b_t \sim NID(0, \sigma_b)$ ,  $0 \leq \psi < 1$ .

В этом случае статистическая TS-модель может более успешно аппроксимировать DS DGP при значениях  $\psi$ , близких к 1 (напомним, что при  $\psi = 1$  мы получаем процесс  $y_t$ , стационарный относительно тренда).

Статистическое моделирование с указанными (более общими) TS- и DS-моделями DGP приводит к следующим результатам.

При  $\rho = 0.9$  статистическая DS-модель является хорошей аппроксимацией для TS DGP, особенно для малых значений  $T$  и  $h$ . Более того, для  $T < 50$  однопараметрическая статистическая DS-модель дает более точные прогнозы, чем трехпараметрическая статистическая TS-модель (соответствующая TS DGP). Этот результат находится в соответствии с указанными выше результатами работы [Campbell, Perron (1991)].

При  $\psi = 0.9$  и  $T < 100$  статистическая TS-модель дает более точные прогнозы, чем статистическая DS-модель (соответствующая DS DGP), при любых  $h$ .

Таким образом, изменение DGP в направлении «другой» модели (DS или TS) может приводить при небольших значениях  $T$  к тому, что прогнозы по статистической модели, соответствующей DGP (т.е. по «своей» модели), оказываются менее точными. Это особенно важно для практических ситуаций, в которых различие между TS- и DS-моделями бывает довольно затруднительным.

В работе [Diebold, Kilian (2000)] проводится систематическое исследование того, до какой степени предварительное тестирование на наличие единичного корня влияет на качество прогноза при различных степенях инерционности процесса, различных горизонтах прогнозирования и различных объемах выборок. Внимание концентрируется на случае одномерного процесса авторегрессии с трендом и высокой инерционностью, особенно интересном для экономики и финансов. Поскольку о точных аналитических результатах при конечных выборках в этом случае не может быть и речи, приходится действовать, используя метод статистического моделирования Монте-Карло.

При планировании эксперимента, как всегда, возникают проблемы. С одной стороны, хотелось бы исследовать достаточно широкий спектр процессов порождения данных (*DGP*), так чтобы результаты пролили свет на поведение альтернативных методов в широком спектре встречающихся на практике ситуаций. С другой стороны, исследуемые *DGP* должны быть достаточно простыми, чтобы обеспечить реализацию Монте-Карло анализа и интерпретируемость полученных результатов.

Использование *DGP* в виде процесса авторегрессии первого порядка с различными степенями инерционности, соответствующими различным значениям параметра авторегрессии, представляет привлекательный компромисс. Если, однако, целью анализа служит получение рекомендаций для практической работы, то для этой цели необходимо включение в модель и временного тренда.

Поэтому в работе исследуется *AR*(1) процесс с трендом, имеющий вид  $(y_t - a - bt) = \rho(y_{t-1} - a - b(t-1)) + \varepsilon_t$ , где  $\varepsilon_t \sim iid N(0, \sigma^2)$ ,  $t = 1, \dots, T$ . Этот процесс можно записать также в виде:

$$y_t = k_1 + k_2 t + \rho y_{t-1} + \varepsilon_t$$

где

$$k_1 = a(1 - \rho) + \rho b,$$

$$k_2 = b(1 - \rho).$$

Часто бывает удобнее представить этот процесс в компонентной форме как сумму линейного тренда и процесса *AR*(1):

$$y_t = T_t + x_t,$$

где

$$T_t = a + bt, \quad x_t = \rho x_{t-1} + \varepsilon_t.$$

При  $\rho = 1$  рассматриваемый процесс является случайным блужданием

со сносом  $b$ ; при  $\rho < 1$  имеем ковариационно стационарные  $AR(1)$  отклонения от линейного тренда с угловым коэффициентом  $b$ .

В работе процесс параметризуется в соответствии с квартальными данными о послевоенном реальном ВВП в США, полагая  $a = 7$ ,  $b = 0.0065$  и  $\sigma = 0.0099$ . Рассматриваются значения  $\rho \in \{0.5, 0.9, 0.97, 0.99, 1\}$  и  $T \in \{25, 30:10:80; 100:20:180; 200:40:1000\}$ , включающие подходящие степени инерционности и объемы выборок для годовых, квартальных, месячных, недельных и дневных данных.

Сравниваются три модели прогнозов:  $AR(1)$  в уровнях с линейным детерминированным трендом ( $L$  – для «уровней»), случайное блуждание со сносом ( $D$  – для «разностей») и модель, предлагаемая претестами Дикки–Фуллера, использующими 5%-ные критические значения для конечных выборок ( $P$  – для «претестовых» моделей). Для всех моделей оценивание производится методом *OLS*. Общей задачей является предсказание уровней рядов на горизонтах  $h$  от 1 до 100 периодов вперед. В качестве характеристики прогноза каждой модели берется безусловная среднеквадратическая ошибка прогноза (*PMSE*), вычисляемая по 20000 реализаций алгоритма Монте-Карло. Для каждого значения  $\rho$  вычисляются отношения  $PMSE(D)/PMSE(L)$ ,  $PMSE(D)/PMSE(P)$  и  $PMSE(P)/PMSE(L)$  для всех комбинаций значений  $h$  и  $T$ .

#### **Результаты**

##### **D&L (разности или уровни)**

В этой связке полученные результаты указывают на предпочтительность дифференцирования в случаях, когда размер выборки небольшой или умеренный, а процесс имеет высокую инерционность, и на предпочтительность модели в уровнях в случае малоинерционных процессов или при больших выборках в случае высокоинерционного процесса.

##### **D&P**

В этой связке отношение  $PMSE(D)/PMSE(P)$  либо близко к единице, либо превосходит единицу, так что использование претеста предпочтительнее, чем безоглядное дифференцирование. Это означает, что при построении моделей с целью прогнозирования не рекомендуется следовать стратегии Бокса–Дженкинса неперменного дифференцирования ряда для достижения стационарности.

##### **P&L**

В этой связке отношение  $PMSE(P)/PMSE(L)$  не везде меньше единицы, однако преимущество претестового подхода проявляется на большей части пространства эксперимента. Это также ставит под сомнение опирающуюся на асимптотические результаты стратегию неперменного построения прогнозирующих моделей в уровнях.

### Некоторые практические советы

При широком диапазоне объемов выборок и горизонтов прогноза достаточно трудно сформулировать вполне надежные рекомендации по выбору одного из трех методов, рассмотренных выше, тем более что параметры DGP, выбранные в соответствии с квартальными данными о величине ВВП в США, вовсе не обязаны повторяться в других рядах. В этой связи в работе было проведено дополнительное моделирование, основанное на других рядах, имеющих характерные частоты наблюдений. Ниже приведена таблица параметров DGP для рядов с различной частотой наблюдений и указано, на основании какого ряда выбраны эти параметры.

Частота	$a$	$b$	$\sigma$	Базовый ряд
Год	-6.0674	0.0173	0.0500	ВВП на душу населения США
Квартал	7.3707	0.0065	0.0099	Реальный ВВП США
Месяц	3.3653	0.0024	0.0105	Индекс промышленного производства США
День	5.1126	0.0004	0.0095	Индекс Доу–Джонса

Поскольку предварительное тестирование явно лучше, чем обязательное дифференцирование, упор делается на сравнении метода предварительного тестирования (метод Р) и метода обязательного использования модели в уровнях (метод L).

Для годовых данных (для  $T=40-160$  и  $h=1-100$ ) предварительное тестирование недвусмысленно улучшает точность прогноза для всех горизонтов прогнозирования и объемов выборок, если корень DGP равен или больше 0.97. Для  $\rho = 0.9$  предварительное тестирование улучшает точность прогноза для объемов выборки, не превышающих 70, но дает меньшую точность по сравнению с методом L при большем количестве наблюдений. Для  $\rho = 0.5$  оба метода дают близкие результаты. Фактически, отсюда вытекает рекомендация использовать предварительное тестирование при наличии не более 70 годовых данных, а для большего количества наблюдений – при очень высокой инерционности ряда.

Для квартальных данных ( $T=80-200$  и  $h=1-16$ ) метод Р предпочтительнее для всех горизонтов прогноза и объемов выборки, если корень процесса имеет значение 0.97 или выше. Для  $\rho = 0.9$  метод L равномерно предпочтительнее. Для  $\rho = 0.5$  оба метода дают близкие результаты. Отсюда вытекает рекомендация использовать предварительное тестирование для процессов с корнями 0.97 и выше и метод L при меньших значениях корней.

Для месячных данных ( $T=240-480$  и  $h=1-48$ ) метод Р предпочтительнее для всех горизонтов прогноза и объемов выборки, если корень процесса ра-

вен 1 или 0.99. Для  $\rho = 0.97$  и  $\rho = 0.9$  метод L по крайней мере столь же точен, как и P. Для  $\rho = 0.5$  оба метода дают близкие результаты. Соответственно, для месячных данных предварительное тестирование имеет смысл только для процессов с корнями, равными 0.99 или выше. Во всех других случаях метод L предпочтительнее.

Для дневных данных ( $T=360-720$  и  $h=1-90$ ) предварительное тестирование предпочтительнее только для  $\rho=1$ . Для  $\rho=0.99$  предварительное тестирование предпочтительнее, если объем выборки не превосходит 600 дней (при всех горизонтах). При более длинных выборках метод L предпочтительнее на протяженных горизонтах прогнозирования и сравним с методом P на коротких горизонтах. Для  $\rho = 0.97$  метод L равномерно лучше. Для  $\rho = 0.9$  метод L несколько лучше при  $T < 500$ , за исключением очень коротких горизонтов. При возрастании объема выборки это различие исчезает. Для  $\rho = 0.5$  оба метода дают близкие результаты. Таким образом, для дневных данных предварительное тестирование имеет смысл, только если ряд в высшей степени инерционен, а именно, значения корней 0.99 и выше. В противном случае следует предпочесть метод L.

Указанные рекомендации могут показаться порочным кругом, поскольку они предполагают знание корня *DGP*. На практике точечные *OLS*-оценки корней макроэкономических рядов обычно превышают 0.97 для годовых данных, 0.99 для квартальных данных и значительно превышают 0.99 для дневных данных. Более того, наличие смещения *OLS*-оценок при малых выборках если и приводит, то только к недооценке истинных значений корней. Отсюда можно сделать вывод, что фактически предварительное тестирование рекомендуется при проведении реального прогнозирования макроэкономических рядов, имеющих трендовое поведение.

#### **1.3.4. Моделирование и оценка стабильности соотношений между макроэкономическими показателями**

Различным вопросам прогнозирования основных макроэкономических показателей посвящены и другие статьи Stock и Watson. Так, например, в работе [Stock, Watson, (1994)] авторы оценивают авторегрессионные и векторные авторегрессионные зависимости между макроэкономическими переменными (76 макроэкономических показателей США послевоенного периода, месячные данные), а также проверяют, насколько стабильны эти соотношения во времени. Кроме того, в работе проверяется, как возможное наличие нестабильности в связях между переменными отражается на возможности прогнозирования макроэкономических переменных, в том числе и с

учетом адаптивных моделей с изменяющимися во времени параметрами (TVP-модели). Для проверки на стабильность используются тесты Ниблома (см., [Nyblom (1989)]), *CUSUM*-тесты и тест Куандта на наличие излома (см. [Quandt (1960)]).

Во второй части работы [Stock, Watson (1994)] авторы проверяют, могут ли современные, часто используемые адаптивные эконометрические модели уловить эту нестабильность. Проверка осуществляется на основе сравнения прогнозов по различным моделям, в числе которых как неадаптивные модели (обычные *OLS*-модели), так и слабоадаптивные модели (рекурсивные модели, модели со сдвигом интервала оценок, *TVP*-модели со слабо меняющимися параметрами) и сильноадаптивные модели (*TVP*-модели с сильно меняющимися параметрами).

Полученные в работе результаты указывают на наиболее вероятную нестабильность уравнений для индекса промышленного производства, доходов населения, занятости, цен и процентных ставок. При этом при моделировании и прогнозировании худшими оказались модели с фиксированными параметрами, в то время как слабоадаптивные модели показали достаточно хорошие результаты. В частности, рекурсивная модель оказывается лучшей или близка к лучшей почти во всех случаях, векторные авторегрессионные модели лучше авторегрессионных моделей для одной переменной примерно в трети (38%) случаев. Исследование сильноадаптивных моделей показало, что такие модели, как *TVP*-модели с сильно меняющимися параметрами, оказались в большинстве случаев хуже рекурсивных моделей. В этом смысле такие модели плохо учитывают нестабильность соотношений между макроэкономическими переменными.

В работе [Stock, Watson, (1998a)] авторы исследуют эмпирические зависимости между бизнес-циклом (динамикой агрегированного выпуска) и различными макроэкономическими показателями, такими, как производство, цены, процентные ставки, производительность, занятость по секторам, доходы населения, инвестиции, потребление и др. Проверка таких зависимостей осуществляется через сравнение циклических компонент агрегированного выпуска (основного бизнес-цикла) и указанных переменных. Особое внимание уделяется тому, опережает или отстает циклическая компонента рассматриваемых показателей от основного цикла. В случае, когда она опережает цикл, это означает, что можно прогнозировать динамику агрегированного выпуска на основе данного показателя, и наоборот, если циклическая компонента показателя повторяет динамику бизнес-цикла с лагом, то эту информацию можно использовать для прогнозирования рассматрива-

емого показателя. Для выделения бизнес-циклов была использована стандартная методика, применяемая National Bureau of Economic Research (NBER); исследование проводилось на квартальных данных.

Полученные результаты говорят о том, что производство и использование ресурсов, в том числе и трудовых ресурсов (занятость), в различных отраслях меняется практически одновременно и сонаправленно с общим бизнес-циклом. Исключение составляют финансовый и государственный секторы, корреляция которых с бизнес-циклом существенно ниже по сравнению, например, со строительством или промышленностью. Кроме того, отмечается, что добывающие секторы также подвержены циклической динамике, которая, однако, не связана с бизнес-циклами, а определяется другими причинами (мировыми ценами на ресурсы, открытием и закрытием шахт и месторождений и др.). Потребление, сбережения, запасы и импорт изменяются в соответствии с бизнес-циклом; при этом инвестиции в оборудование, здания и сооружения несколько отстают от основного цикла. В отличие от импорта, экспорт не так значительно колеблется вместе с бизнес-циклом; это приводит к тому, что торговый баланс изменяется в противофазе с циклом.

В этой же работе авторы проверяют классические зависимости между макроэкономическими показателями – кривую Филлипса (зависимость между инфляцией и безработицей), спрос на деньги в долгосрочном периоде, соотношения сбалансированного роста (стационарность отношений макроэкономических показателей) и спреда между долгосрочными и краткосрочными процентными ставками (нестационарность ставок и стационарность спредов). Проведенный эконометрический анализ этих соотношений указывает на то, что большинство из таких зависимостей действительно имеет место в долгосрочном периоде, хотя надо быть достаточно осторожным в отношении стационарности рассматриваемых рядов и стабильности оцениваемых коэффициентов.

Достаточно интересной и полезной для моделирования и прогнозирования является работа [Stock, Watson, (1998b)], в которой авторы используют для прогнозирования конструирование агрегированных индексов макроэкономических показателей. В случае, когда количество объясняющих переменных превышает количество наблюдений прогнозируемого временного ряда, в работе предлагается строить специальные диффузионные индексы, которые затем можно использовать для оценки и прогнозирования. Здесь необходимо отметить, что одной из основных проблем, с которой сталкиваются авторы при расчете подобных вероятностных индексов, является со-

держательная интерпретация оцененных зависимостей и полученных в индексах коэффициентов – весов макроэкономических переменных.

В работе [Stock, Watson (1999)] авторы анализируют возможность прогнозирования инфляции на основе кривой Филлипса. Кривую Филлипса в общем смысле можно интерпретировать как зависимость инфляции от показателей экономической активности. В работе рассматриваются различные показатели; при этом дополнительное внимание уделяется тому, какие из этих показателей лучше подходят для прогнозирования инфляции. Кроме этого, проводится сравнение прогнозов инфляции, полученных с использованием показателей экономической активности, с прогнозами инфляции на основе других макроэкономических показателей. Оценки проводились на месячных данных США 1959–1997 гг.

Тестирование коэффициентов модели, описывающей кривую Филлипса или обобщенную кривую Филлипса, на стабильность (тесты Куандта) показало, что коэффициенты не являются стабильными (статистически значимый результат), однако при дальнейших оценках авторы не учитывают эту нестабильность по двум основным причинам. Во-первых, как показано в работе, эта нестабильность невелика. Во-вторых, применение методов с более высокой степенью адаптации не улучшает, а иногда и существенно ухудшает качество прогнозирования, что согласуется с результатами предыдущих работ. Оценка прогностической способности различных показателей экономической активности показала, что по сравнению с безработицей лучшие свойства имеют показатели загрузки мощностей и объемы продаж. При этом добавление прочих макроэкономических факторов не дает улучшения прогнозов.

Дополнительно в работе исследуется вопрос о том, как влияет на полученные результаты предположение о том, что ряд инфляции является интегрированным рядом первого порядка. Приведенные в работе оценки указывают на то, что с точки зрения прогнозирования предпочтительнее использовать модели, предполагающие, что инфляция – это процесс  $I(1)$ .

В работе [Stock, Watson (2001)] авторы исследуют качество прогнозов агрегированного выпуска и инфляции с использованием показателей финансового рынка. В работе рассматриваются прогностические свойства таких показателей, как процентные ставки и спреды между ними, цены и доходность акций, обменные курсы и др. Моделирование и прогнозирование осуществлялось на основе авторегрессионных моделей с добавлением объясняющих переменных и их лаговых значений; при этом отдельное внимание уделялось проверке стабильности прогнозов. Оценки производились на квартальных данных по семи странам OECD за 1959–1999 гг.



Полученные оценки показывают, что финансовые показатели плохо подходят для прогнозирования агрегированного выпуска; относительно лучше для этих целей подходит спрэд между краткосрочными и долгосрочными процентными ставками. При этом в работе отмечается, что наилучший с точки зрения прогнозирования показатель может быть различным для разных стран. При прогнозировании инфляции оказалось, что ни один из финансовых показателей не позволяет прогнозировать инфляцию на длительные горизонты (1–2 года) вперед, что, по-видимому, связано с тем, что большинство зависимостей являются нестабильными.

Работа [Staiger, Stock, Watson (2001)] посвящена выяснению причин того, что в конце 1990-х гг. в США наблюдались одновременно невысокая безработица, низкая инфляция и стабильно растущая реальная заработная плата, что противоречило оценкам кривой Филлипса, полученным ранее. В частности, в работе проверялась гипотеза о наличии сдвига или изменения наклона кривой Филлипса. Кроме того, отдельной задачей работы была оценка значения безработицы, не ускоряющего инфляцию (non-accelerating inflation rate of unemployment, NAIRU).

Полученные результаты указывают на то, что классическая зависимость между инфляцией и безработицей в виде кривой Филлипса сохраняется (отрицательная зависимость не изменила знак), однако происходит ее сдвиг к осям инфляция-безработица. Среди причин такого сдвига отмечают изменения в ценообразовании фирм, изменение NAIRU и др. Достаточно точных оценок значения NAIRU в рамках проведенных оценок получить не удалось, вследствие того, что оцененные доверительные интервалы достаточно широки и включают, в том числе, и фактически наблюдаемые значения безработицы.



## Глава 2. Исследование сравнительного качества прогнозов некоторых макроэкономических рядов РФ, получаемых по фиксированным и рекурсивным TS- и DS-моделям

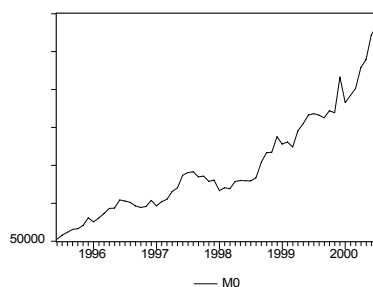
Для исследования использовались ежемесячные данные Госкомстата РФ, Центрального банка РФ, Министерства финансов РФ и Центра экономической конъюнктуры по макроэкономическим показателям Российской Федерации (денежные агрегаты M0, M1, M2, экспорт, налоговые доходы федерального и консолидированного бюджетов РФ, поступления отдельных налогов, индексы промышленного производства) за период 1994–2001 гг. При этом моделирование в большинстве случаев осуществлялось на более коротком интервале, чтобы иметь возможность анализировать ошибки прогнозирования вперед на различные периоды. Полные таблицы исходных данных приведены в Приложении 2.

### 2.1. Денежные ряды

#### 2.1.1. M0

При исследовании этого ряда, а также рядов M1 и M2 мы будем рассматривать месячные данные за период с 06.1995 по 02.2001, используя результаты, полученные в отношении этих трех рядов на временном интервале с 06.1995 по 07.2000 в работе [Эконометрический анализ динамических рядов основных макроэкономических показателей (2001)].

На временном интервале 06.1995–02.2001 ряд M0 (в номинальных значениях) имеет вид:



При использовании критериев Перрона на наличие у этого ряда единичного корня с эндогенным выбором даты излома в упомянутой работе были получены следующие результаты (относящиеся к периоду с 06.1995 по 07.2000).

Для модели, допускающей сдвиг траектории и изменение наклона тренда в форме инновационного выброса (IO), в качестве даты излома указан 01.1999.

Для модели, допускающей только изменение наклона тренда (без сдвига траектории) в форме аддитивного выброса (AO), в качестве даты излома также указан 01.1999.

В обоих случаях гипотеза единичного корня не отвергается, что, в согласии с результатами применения других критериев, приводит к выводу о том, что поведение траектории ряда  $M0$  за период с 06.1995 по 07.2000 соответствует поведению  $DS$ -ряда (т.е. поведению интегрированного ряда, приведение которого к стационарному виду требует дифференцирования).

Содержательные соображения указывают на то, что излом этого и других временных рядов может приходиться на вторую половину 1998 г. (финансовый кризис и сразу после него), а также на начало 1999 г. (оживление экономической активности и начало экономического роста). В этом смысле полученные результаты согласуются с содержательными аргументами в пользу выбора даты излома.

Для сравнения прогнозов, получаемых по моделям  $TS$ - и  $DS$ -рядов, построим для периода 06.1995–07.2000 подходящие модели обоих типов –  $DS$  и  $TS$ .

#### **DS-модели**

Сначала построим  $DS$ -модели с инновационным и аддитивным выбросами (дата излома 01.1999). Оцененные модели имеют следующий вид.

#### **DS-модель с инновационным выбросом**

Dependent Variable: D(X)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7957.736	3094.706	2.571403	0.0150
DTB	-18161.51	11183.91	-1.623897	0.1142
DU	14680.75	5592.419	2.625115	0.0132
D(X(-1))	0.035399	0.159275	0.222248	0.8255
D(X(-2))	-0.105155	0.156321	-0.672687	0.5060
D(X(-3))	-0.056082	0.155705	-0.360181	0.7211
D(X(-4))	-0.262292	0.162524	-1.613874	0.1164
D(X(-5))	-0.274312	0.169334	-1.619945	0.1151
D(X(-6))	-0.104990	0.174786	-0.600679	0.5523
D(X(-7))	-0.030949	0.187136	-0.165381	0.8697
D(X(-8))	-0.264658	0.218019	-1.213925	0.2337

Продолжение таблицы

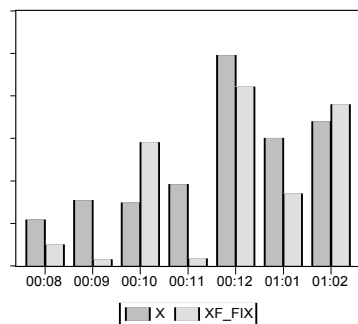
Dependent Variable: D(X)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D( X( -9))	-0.660972	0.207564	-3.184427	0.0032
D( X( -10))	-0.190682	0.231654	-0.823134	0.4165
D( X( -11))	-0.486314	0.232308	-2.093403	0.0443
D( X( -12))	0.625239	0.242277	2.580681	0.0147
D( X( -13))	-0.722586	0.263874	-2.738378	0.0100
R-squared	0.639036	Mean dependent var		4816.396
Adjusted R-squared	0.469835	S.D. dependent var		12589.30
S.E. of regression	9166.580	Akaike info criterion		21.34572
Sum squared resid	2.69E+09	Schwarz criterion		21.96945
Log likelihood	-496.2972	F-statistic		3.776772
Durbin-Watson stat	2.141405	Prob( F-statistic)		0.000779

#### DS-модель с аддитивным выбросом

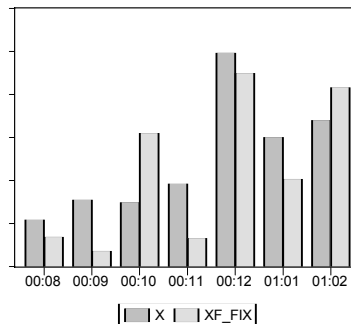
Dependent Variable: D(X)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2323.949	434.9662	5.342828	0.0000
DU	5718.042	1026.329	5.571356	0.0000
AR( 1)	-0.052317	0.153159	-0.341585	0.7348
AR( 2)	-0.194266	0.150285	-1.292652	0.2051
AR( 3)	-0.104782	0.146957	-0.713011	0.4809
AR( 4)	-0.281084	0.150118	-1.872425	0.0700
AR( 5)	-0.262538	0.139072	-1.887782	0.0679
AR( 6)	-0.046242	0.142080	-0.325462	0.7469
AR( 7)	0.018807	0.159481	0.117925	0.9068
AR( 8)	-0.193991	0.190853	-1.016441	0.3168
AR( 9)	-0.623240	0.186508	-3.341627	0.0021
AR( 10)	-0.206942	0.209388	-0.988317	0.3302
AR( 11)	-0.475055	0.209853	-2.263750	0.0303
AR( 12)	0.539607	0.225409	2.393907	0.0225
AR( 13)	-0.737280	0.235544	-3.130117	0.0036
R-squared	0.648283	Mean dependent var		4816.396
Adjusted R-squared	0.499070	S.D. dependent var		12589.30
S.E. of regression	8910.253	Akaike info criterion		21.27810
Sum squared resid	2.62E+09	Schwarz criterion		21.86285
Log likelihood	-495.6744	F-statistic		4.344682
Durbin-Watson stat	2.058355	Prob( F-statistic)		0.000255

Если зафиксировать обе модели и производить по ним прогнозы на один шаг вперед, то получаются следующие результаты:

Инновационный выброс:

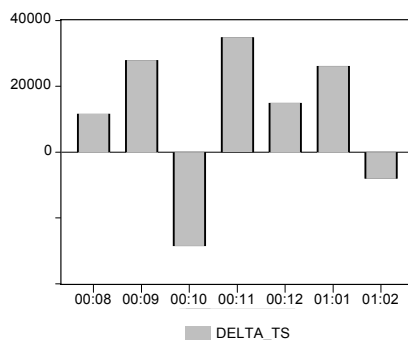


Аддитивный выброс:

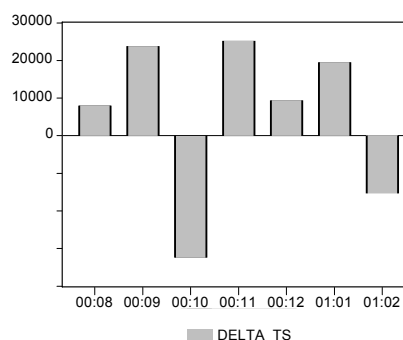


Ошибки прогнозов:

Инновационный выброс:



Аддитивный выброс:



Характеристики прогнозов по подобранным DS-моделям:

	DS (инновац.)	DS (аддитивн.)
<b>Root Mean Squared Error</b>	23660.00	20762.60
<b>Mean Absolute Error</b>	21726.06	19084.87
<b>Mean Absolute Percent Error</b>	5.966785	5.247909

По всем трем характеристикам здесь предпочтительнее оказывается модель с аддитивным выбросом.

### TS-модели

#### TS-модель, допускающая сдвиг траектории и изменение наклона тренда в форме инновационного выброса с датой излома 01.1999

Dependent Variable: X				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	83889.10	30383.71	2.760990	0.0097
DU	-11404.73	9892.396	-1.152878	0.2581
T	3349.482	1335.656	2.507742	0.0178
DT	6891.047	2317.117	2.973975	0.0058
X(-1)	0.750891	0.167935	4.471320	0.0001
X(-2)	-0.184634	0.185574	-0.994934	0.3277
X(-3)	0.011089	0.166217	0.066714	0.9473
X(-4)	-0.286183	0.169492	-1.688473	0.1017
X(-5)	-0.134868	0.169320	-0.796525	0.4320
X(-6)	0.131024	0.168245	0.778771	0.4422
X(-7)	0.059066	0.172426	0.342559	0.7343
X(-8)	-0.378450	0.255431	-1.481612	0.1489
X(-9)	-0.535159	0.263414	-2.031628	0.0511
X(-10)	0.282439	0.272772	1.035438	0.3087
X(-11)	-0.388248	0.278077	-1.396188	0.1729
X(-12)	0.930807	0.284523	3.271462	0.0027
X(-13)	-1.291862	0.329794	-3.917180	0.0005
X(-14)	0.459354	0.259166	1.772429	0.0865
R-squared	0.989083	Mean dependent var	169119.2	
Adjusted R-squared	0.982896	S.D. dependent var	63099.49	
S.E. of regression	8252.271	Akaike info criterion	21.15436	
Sum squared resid	2.04E+09	Schwarz criterion	21.85606	
Log likelihood	-489.7047	F-statistic	159.8771	
Durbin-Watson stat	2.000861	Prob( F-statistic)	0.000000	

#### TS-модель, допускающая изменение наклона тренда в форме аддитивного выброса с датой излома 01.1999

Dependent Variable: X				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	62114.74	2395.813	25.92637	0.0000
T	2238.796	86.16029	25.98408	0.0000
DT	6047.931	392.9049	15.39286	0.0000
AR(1)	0.564688	0.148698	3.797544	0.0006
AR(2)	-0.021091	0.165916	-0.127117	0.8996
AR(3)	-0.030859	0.164054	-0.188101	0.8519
AR(4)	-0.177696	0.163864	-1.084412	0.2860
AR(5)	-0.185941	0.152462	-1.219589	0.2313
AR(6)	0.089144	0.163710	0.544524	0.5897
AR(7)	0.033957	0.171104	0.198459	0.8439
AR(8)	-0.157527	0.239983	-0.656410	0.5161
AR(9)	-0.564313	0.258719	-2.181178	0.0364

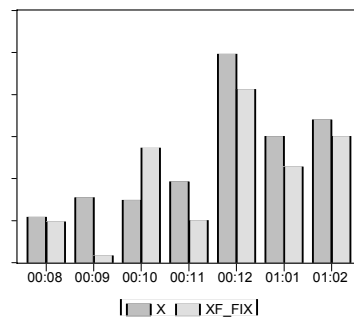
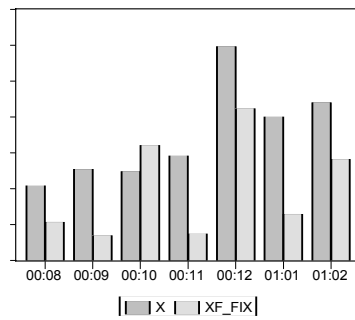
Продолжение таблицы

Dependent Variable: X				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR( 9)	-0.564313	0.258719	-2.181178	0.0364
AR( 10)	0.205723	0.274590	0.749201	0.4590
AR( 11)	-0.361187	0.276525	-1.306165	0.2005
AR( 12)	0.818430	0.284241	2.879354	0.0069
AR( 13)	-0.790364	0.225473	-3.505352	0.0013
R-squared	0.987803	Mean dependent var		167766.8
Adjusted R-squared	0.982259	S.D. dependent var		63152.37
S.E. of regression	8411.560	Akaike info criterion		21.17035
Sum squared resid	2.33E+09	Schwarz criterion		21.78809
Log likelihood	-502.6736	F-statistic		178.1749
Durbin-Watson stat	1.716294	Prob(F-statistic)		0.000000

Если зафиксировать обе модели и производить по ним последовательные прогнозы на один шаг вперед, то получаются следующие результаты:

Инновационный выброс:

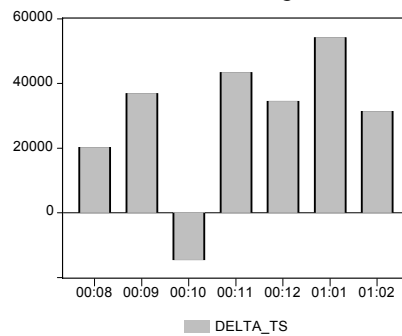
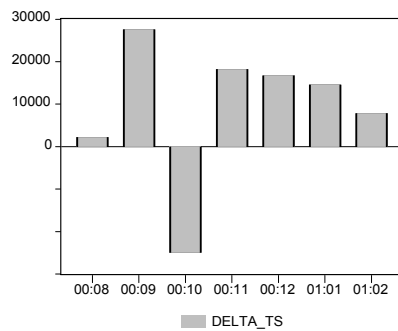
Аддитивный выброс:



Соответственно, ошибки прогнозов имеют вид:

Инновационный выброс:

Аддитивный выброс:





Характеристики прогнозов по подобранным *TS*-моделям:

	<b>TS (иннов.)</b>	<b>TS (аддитив.)</b>
<b>Root Mean Squared Error</b>	35901.47	18035.41
<b>Mean Absolute Error</b>	33681.18	16031.46
<b>Mean Absolute Percent Error</b>	9.064165	4.371771

Прогноз по второй модели оказывается предпочтительнее по всем трем показателям.

Сведем полученные результаты в одну таблицу.

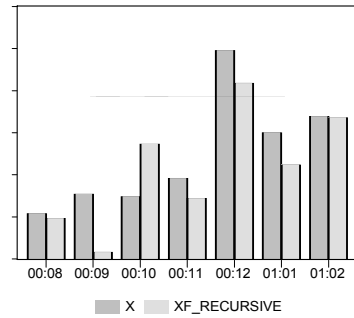
	<b>DS иннов.</b>	<b>TS иннов.</b>	<b>DS аддитив.</b>	<b>TS аддитив.</b>
<b>RMSE</b>	23660.00	35901.47	20762.60	18035.41
<b>MAE</b>	21726.06	33681.18	19084.87	16031.46
<b>MAPE</b>	5.966785	9.064165	5.247909	4.371771

Из таблицы следует, что хотя обе *DS*-модели лучше *TS*-модели с инновационным выбросом, наилучшей оказалась *TS*-модель с аддитивным выбросом.

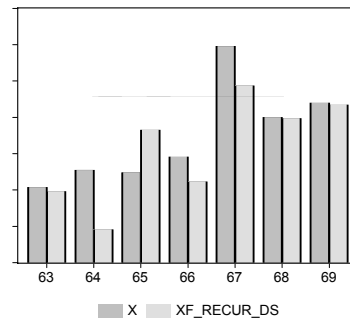
Приведем теперь результаты, получаемые при использовании рекурсивных моделей. Сравнение будем производить для моделей с аддитивным выбросом, поскольку и среди *TS*-моделей, и среди *DS*-моделей такая модель превосходит модель с инновационным выбросом по всем трем показателям.

Прогнозы:

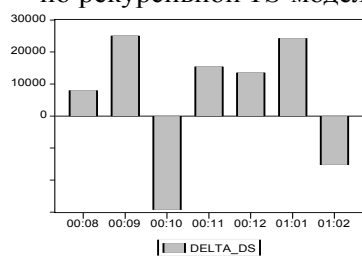
по рекурсивной *TS*-модели



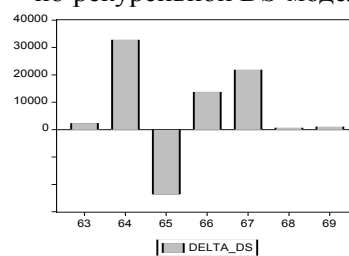
по рекурсивной *DS*-модели



Ошибки прогнозов:  
по рекурсивной *TS*-модели



по рекурсивной *DS*-модели



Характеристики точности прогнозов по рекурсивным моделям с аддитивным выбросом:

	Рекурсивная DS	Рекурсивная TS
<b>Root Mean Squared Error</b>	18103.33	16733.83
<b>Mean Absolute Error</b>	13663.46	13698.90
<b>Mean Absolute Percent Error</b>	3.737642	3.747511

Сведем результаты в отношении ряда  $M_0$ , полученные по моделям с аддитивным выбросом, в одну таблицу.

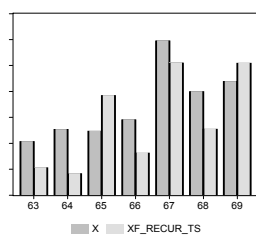
	DS рекурс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	18103.33	16733.83	20762.60	18035.41
<b>Mean Absolute Error</b>	13663.46	13698.90	19084.87	16031.46
<b>Mean Absolute Percent Error</b>	3.737642	3.747511	5.247909	4.371771

Среди фиксированных моделей с аддитивным выбросом лучшей по всем трем характеристикам точности прогнозов оказывается *TS*-модель. Среди рекурсивных моделей ни одна из двух не является более предпочтительной. При этом и среди *TS*-моделей, и среди *DS*-моделей рекурсивные модели оказались лучше фиксированных.

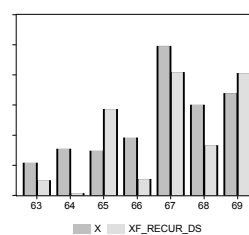
Инновационный выброс

Прогнозы (по оси абсцисс номера наблюдений):

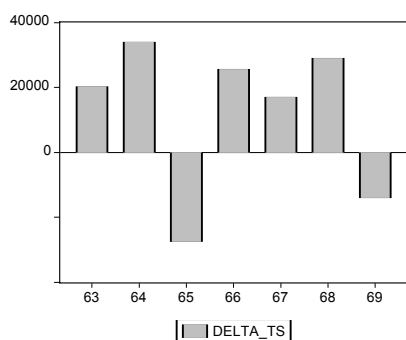
по рекурсивной *TS*-модели



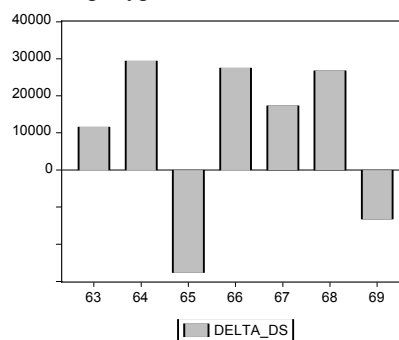
по рекурсивной *DS*-модели



Ошибки прогнозов:  
по рекурсивной *TS*-модели



по рекурсивной *DS*-модели



Характеристики точности прогнозов по рекурсивным моделям с инновационным выбросом:

	Рекурсивная DS	Рекурсивная TS
<b>Root Mean Squared Error</b>	23070.19	24831.23
<b>Mean Absolute Error</b>	21966.45	23946.98
<b>Mean Absolute Percent Error</b>	6.002693	6.570231

Сведем результаты в отношении ряда *M0*, полученные по моделям с инновационным выбросом, в одну таблицу.

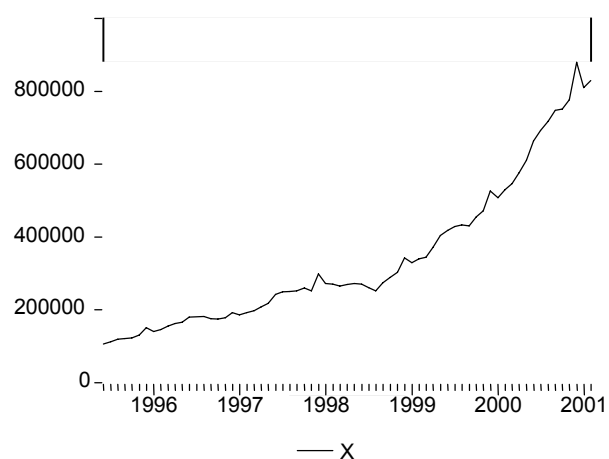
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	23070.19	24831.23	23660.00	35901.47
<b>Mean Absolute Error</b>	21966.45	23946.98	21726.06	33681.18
<b>Mean Absolute Percent Error</b>	6.002693	6.570231	5.966785	9.064165

И среди фиксированных, и среди рекурсивных моделей с инновационным выбросом лучшими по всем трем характеристикам точности прогнозов оказываются *DS*-модели. Среди *TS*-моделей рекурсивная модель оказалась лучше фиксированной. Рекурсивные *TS* и *DS*-модели дают близкие результаты. Наилучшей является рекурсивная *TS*-модель.

**Подведем итоги исследования прогнозов на один шаг вперед для ряда *M0*.** Среди фиксированных моделей в зависимости от типа выброса (инновационный или аддитивный) лучшей оказывается либо *TS*, либо *DS*-модель. В целом рекурсивные модели дают лучшие прогнозы по сравнению с фиксированными моделями. При этом среди рекурсивных моделей ни одна из двух (*TS* или *DS*) не является предпочтительной.

### 2.1.2. M1

За период 06.1995–02.2001 ряд M1 (в номинальных значениях) имеет вид:



При использовании критериев Перрона на наличие у этого ряда единичного корня с эндогенным выбором даты излома в упомянутой работе были получены следующие результаты (относящиеся к периоду с 06.1995 по 07.2000).

Для модели, допускающей сдвиг траектории и изменение наклона тренда в форме инновационного выброса (IO), в качестве даты излома указан 07.1999.

Для модели, допускающей только изменение наклона тренда (без сдвига траектории) в форме аддитивного выброса (AO), в качестве даты излома указан 02.1999.

В обоих случаях гипотеза единичного корня не отвергается, что, в согласии с результатами применения других критериев, приводит к решению о том, что поведение траектории ряда M1 на временном интервале с 06.1995 по 07.2000 соответствует поведению *DS*-ряда (т.е. поведению интегрированного ряда, приведение которого к стационарному ряду требует дифференцирования). Тем не менее, имея в виду задачу сравнения прогнозов, получаемых по моделям *TS*- и *DS*-рядов, мы построим для периода 06.1995–07.2000 подходящие модели обоих типов – *DS* и *TS*.

Здесь и далее среди альтернативных вариантов модели одного типа (*TS*

или  $DS$ ) мы будем отбирать модель авторегрессии, порядок которой определяется по наименьшему значению SIC информационного критерия Шварца (байесовского информационного критерия), проходящую стандартные диагностические тесты.

### DS-модели

Сначала построим авторегрессионные  $DS$ -модели с инновационным выбросом (дата излома 07.1999) и с аддитивным выбросом (дата излома 02.1999), имеющие порядок 12.

### Оцененная DS-модель с инновационным выбросом (дата излома 07.1999)

Dependent Variable: D(X)				
Method: Least Squares				
Date: 11/12/01 Time: 13:33				
Sample( adjusted): 1996:07 2000:07				
Included observations: 49 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6538.025	5679.194	1.151224	0.2577
DTB	6763.981	15653.45	0.432108	0.6684
DU	15474.11	12173.59	1.271121	0.2123
D( X(-1))	-0.111283	0.169099	-0.658094	0.5149
D( X(-2))	0.201149	0.162859	1.235112	0.2253
D( X(-3))	0.086458	0.165345	0.522895	0.6044
D( X(-4))	0.030484	0.158346	0.192518	0.8485
D( X(-5))	0.091538	0.153038	0.598140	0.5537
D( X(-6))	0.224769	0.151432	1.484287	0.1469
D( X(-7))	0.123352	0.162645	0.758411	0.4534
D( X(-8))	-0.254336	0.183272	-1.387754	0.1742
D( X(-9))	-0.409614	0.186994	-2.190522	0.0354
D( X(-10))	-0.264368	0.208246	-1.269497	0.2129
D( X(-11))	-0.328840	0.234154	-1.404377	0.1693
D( X(-12))	0.507417	0.227016	2.235154	0.0321
R-squared	0.608055	Mean dependent var		10454.73
Adjusted R-squared	0.446666	S.D. dependent var		17387.54
S.E. of regression	12933.97	Akaike info criterion		22.01989
Sum squared resid	5.69E+09	Schwarz criterion		22.59902
Log likelihood	-524.4872	F-statistic		3.767632
Durbin-Watson stat	1.832913	Prob( F-statistic)		0.000779

**DS-модель с аддитивным выбросом (дата излома 02.1999)**

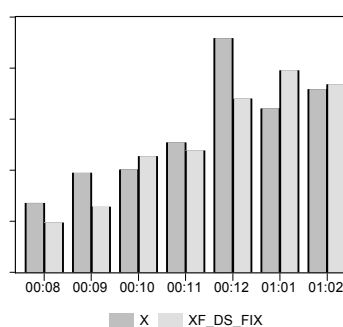
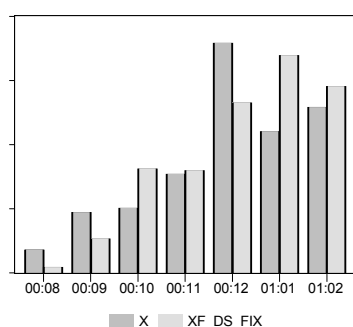
Dependent Variable: D(X)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4728.177	627.0684	7.540129	0.0000
DU	14890.38	1450.730	10.26406	0.0000
AR( 1)	-0.345214	0.169657	-2.034775	0.0495
AR( 2)	-0.079099	0.172553	-0.458402	0.6495
AR( 3)	-0.129310	0.162248	-0.796990	0.4308
AR( 4)	-0.137688	0.146297	-0.941151	0.3531
AR( 5)	-0.067956	0.141762	-0.479370	0.6347
AR( 6)	0.041004	0.138808	0.295399	0.7694
AR( 7)	-0.044951	0.144788	-0.310458	0.7581
AR( 8)	-0.322453	0.148128	-2.176852	0.0363
AR( 9)	-0.524660	0.160095	-3.277180	0.0024
AR( 10)	-0.426078	0.184944	-2.303825	0.0273
AR( 11)	-0.487277	0.195992	-2.486209	0.0178
AR( 12)	0.296755	0.196791	1.507967	0.1405
R-squared	0.667957	Mean dependent var		10454.73
Adjusted R-squared	0.544627	S.D. dependent var		17387.54
S.E. of regression	11733.35	Akaike info criterion		21.81321
Sum squared resid	4.82E+09	Schwarz criterion		22.35373
Log likelihood	-520.4237	F-statistic		5.416000
Durbin-Watson stat	1.837172	Prob( F-statistic)		0.000032

(Мы оставляем AR(12) в модели, несмотря на статистическую незначимость соответствующего коэффициента, поскольку при удалении этой переменной отвергается гипотеза некоррелированности ошибок.)

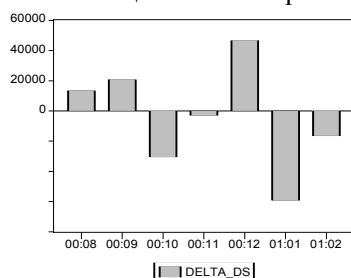
Если зафиксировать обе модели и производить прогнозы на один шаг вперед по этим двум моделям, то получаются следующие результаты:

Инновационный выброс:

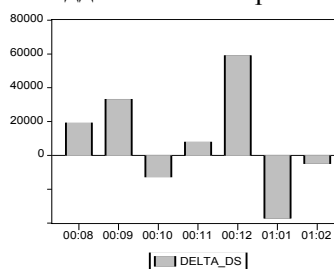
Аддитивный выброс:



Ошибки прогнозов:  
Инновационный выброс:



Аддитивный выброс:



Характеристики прогнозов по подобранным DS-моделям:

	DS инновац.	DS аддитивн.
Root Mean Squared Error	32677.59	30788.31
Mean Absolute Error	27062.91	25004.23
Mean Absolute Percent Error	3.374370	3.119278

Прогноз по второй модели оказывается предпочтительнее по всем трем показателям.

### TS-модели

**TS-модель, допускающая сдвиг траектории и изменение наклона тренда в форме инновационного выброса с датой излома 07.1999**

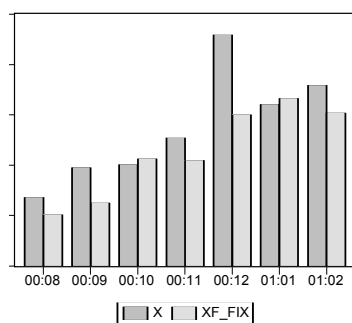
Dependent Variable: X				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	92605.11	38919.92	2.379376	0.0235
DU	-8972.727	12764.01	-0.702971	0.4872
T	7352.346	2947.812	2.494170	0.0180
DT	17196.43	6372.049	2.698729	0.0110
X(-1)	0.588699	0.170328	3.456275	0.0016
X(-2)	0.147850	0.176315	0.838555	0.4079
X(-3)	-0.116411	0.172833	-0.673549	0.5054
X(-4)	-0.098169	0.175323	-0.559930	0.5794
X(-5)	-0.030936	0.174361	-0.177426	0.8603
X(-6)	0.096587	0.164697	0.586453	0.5617
X(-7)	-0.044053	0.172331	-0.255630	0.7999
X(-8)	-0.299828	0.194321	-1.542947	0.1327
X(-9)	-0.338117	0.197411	-1.712761	0.0964
X(-10)	-0.078094	0.202998	-0.384705	0.7030
X(-11)	-0.219526	0.205182	-1.069909	0.2927
X(-12)	0.546394	0.213084	2.564221	0.0152
X(-13)	-0.578163	0.231096	-2.501836	0.0177
R-squared	0.995492	Mean dependent var		333059.7
Adjusted R-squared	0.993238	S.D. dependent var		138024.7
S.E. of regression	11350.15	Akaike info criterion		21.77964
Sum squared resid	4.12E+09	Schwarz criterion		22.43599
Log likelihood	-516.6012	F-statistic		441.6414
Durbin-Watson stat	1.958880	Prob( F-statistic)		0.000000

**TS-модель, допускающая изменение наклона тренда в форме аддитивного выброса с датой излома 02.1999**

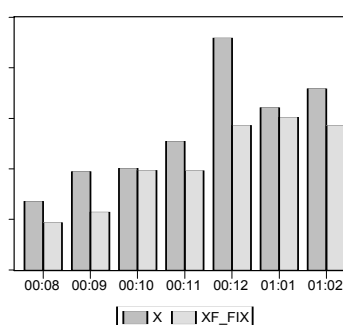
Dependent Variable: X				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	105239.9	2647.314	39.75347	0.0000
T	4823.404	92.89645	51.92238	0.0000
DT	14422.74	453.3884	31.81102	0.0000
AR( 1)	0.456479	0.161153	2.832571	0.0078
AR( 2)	0.064584	0.162178	0.398228	0.6930
AR( 3)	-0.160353	0.159935	-1.002611	0.3233
AR( 4)	-0.150698	0.161762	-0.931602	0.3583
AR( 5)	-0.058754	0.159910	-0.367422	0.7157
AR( 6)	0.058838	0.152973	0.384630	0.7030
AR( 7)	-0.072117	0.156758	-0.460052	0.6485
AR( 8)	-0.305773	0.173733	-1.760015	0.0877
AR( 9)	-0.324015	0.179801	-1.802079	0.0807
AR( 10)	-0.055112	0.183765	-0.299905	0.7661
AR( 11)	-0.201223	0.183716	-1.095297	0.2813
AR( 12)	0.544210	0.187449	2.903237	0.0065
AR( 13)	-0.537698	0.187655	-2.865350	0.0072
R-squared	0.996113	Mean dependent var		333059.7
Adjusted R-squared	0.994346	S.D. dependent var		138024.7
S.E. of regression	10378.13	Akaike info criterion		21.59054
Sum squared resid	3.55E+09	Schwarz criterion		22.20827
Log likelihood	-512.9682	F-statistic		563.8116

Если зафиксировать обе модели и производить последовательные прогнозы на один шаг вперед по этим двум моделям, то получаются следующие результаты:

Инновационный выброс:



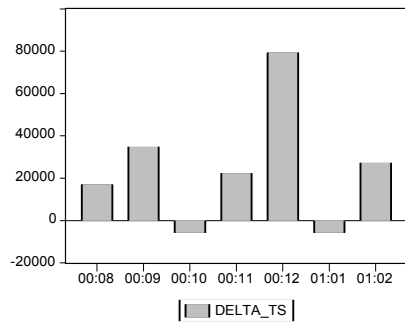
Аддитивный выброс:



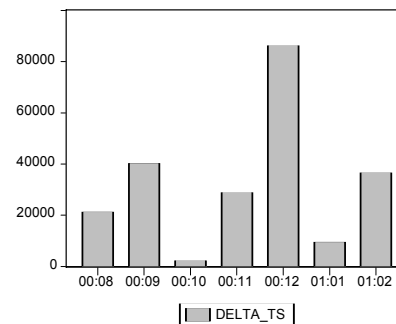


Соответственно, ошибки прогнозов имеют вид

Инновационный выброс:



Аддитивный выброс:



Характеристики прогнозов по подобранным *TS*-моделям:

	<b>TS инновац.</b>	<b>TS аддитивн.</b>
<b>Root Mean Squared Error</b>	36093.53	40981.79
<b>Mean Absolute Error</b>	27502.40	32088.90
<b>Mean Absolute Percent Error</b>	3.388481	3.958528

Прогноз по первой модели оказывается предпочтительнее по всем трем показателям.

Сведем полученные результаты в одну таблицу.

	<b>DS иннова.</b>	<b>TS иннова.</b>	<b>DS аддитив.</b>	<b>TS аддитив.</b>
<b>RMSE</b>	32677.59	36093.53	30788.31	40981.79
<b>MAE</b>	27062.91	27502.40	25004.23	32088.90
<b>MAPE</b>	3.374370	3.388481	3.119278	3.958528

Обе *DS*-модели лучше *TS*-моделей по всем трем характеристикам.

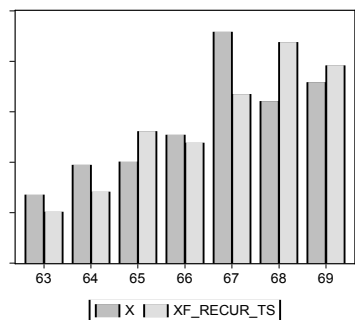
Приведем теперь результаты, получаемые при использовании рекурсивных моделей, использующих одни и те же переменные, но коэффициенты которых переоцениваются при поступлении новых наблюдений (т.е. количество запаздываний не изменяется, так же как не изменяются и даты изломов).

Для рекурсивных моделей на графиках по оси абсцисс откладываются номера последовательных наблюдений, начиная с 06.1995 (так что 63-е наблюдение соответствует 08.2000, а 69-е соответствует 02.2001).

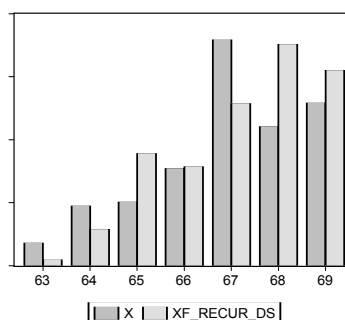
### Рекурсивные модели с инновационным выбросом

Прогнозы:

по рекурсивной *TS*-модели

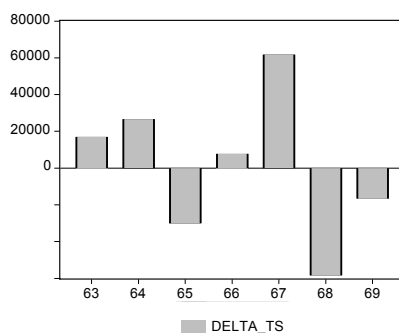


по рекурсивной *DS*-модели

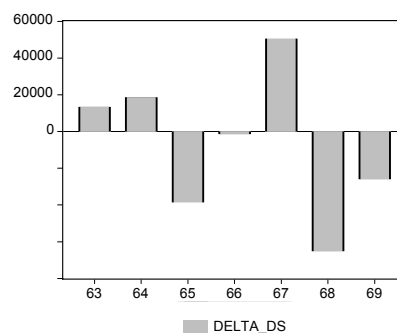


Ошибки прогнозов:

по рекурсивной *TS*-модели



по рекурсивной *DS*-модели



Характеристики точности прогнозов по рекурсивным моделям с инновационным выбросом:

	Рекурсивная DS	Рекурсивная TS
<b>Root Mean Squared Error</b>	36796.05	36710.86
<b>Mean Absolute Error</b>	30477.43	31124.61
<b>Mean Absolute Percent Error</b>	3.793081	3.874630

Сведем полученные результаты в отношении ряда М1 для моделей с инновационным выбросом в одну таблицу.

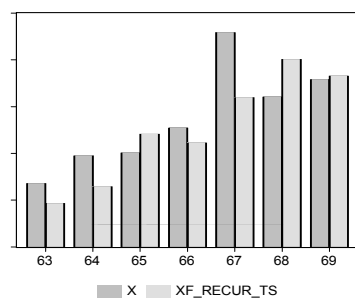
	DS рекурс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	36796.05	36710.86	32677.59	36093.53
<b>Mean Absolute Error</b>	30477.43	31124.61	27062.91	27502.40
<b>Mean Absolute Percent Error</b>	3.793081	3.874630	3.374370	3.388481

И среди фиксированных, и среди рекурсивных моделей с инновационным выбросом лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели (RMSE для рекурсивной *TS*-модели меньше, но незначительно). При этом и среди *TS*-моделей, и среди *DS*-моделей фиксированные модели оказались лучше рекурсивных. Наилучшей является фиксированная *DS*-модель.

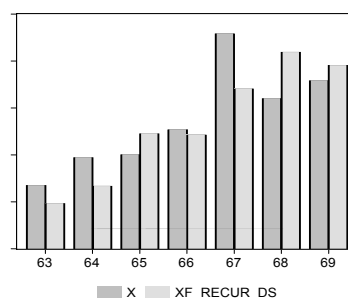
#### Рекурсивные модели с аддитивным выбросом

Прогнозы:

по рекурсивной *TS*-модели

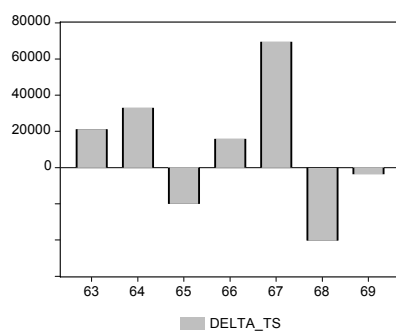


по рекурсивной *DS*-модели

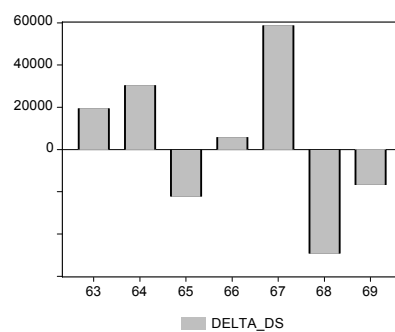


Ошибки прогнозов:

по рекурсивной *TS*-модели



по рекурсивной *DS*-модели



Характеристики точности прогнозов по рекурсивным моделям с аддитивным выбросом:

	<b>DS рекурс.</b>	<b>TS рекурс.</b>
<b>Root Mean Squared Error</b>	33750.29	35153.12
<b>Mean Absolute Error</b>	28915.77	20926.71
<b>Mean Absolute Percent Error</b>	3.604901	3.621122

Сведем полученные результаты в отношении ряда M1 для моделей с аддитивным выбросом в одну таблицу.

	<b>DS рекурс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	33750.29	35153.12	30788.31	40981.79
<b>Mean Absolute Error</b>	28915.77	29026.71	25004.23	32088.90
<b>Mean Absolute Percent Error</b>	3.604901	3.621122	3.119278	3.958528

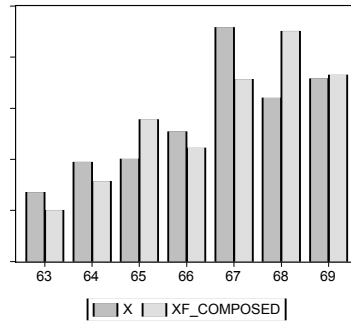
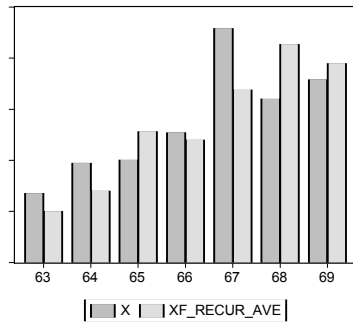
И среди фиксированных, и среди рекурсивных моделей с аддитивным выбросом лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели. При этом среди *TS*-моделей лучшей оказалась рекурсивная модель, а среди *DS*-моделей лучшей оказалась фиксированная модель. Наилучшей является фиксированная *DS*-модель.

**Подведем итоги исследования прогнозов на один шаг вперед для ряда M1.** И среди фиксированных, и среди рекурсивных моделей для каждого из двух типов выбросов (инновационный или аддитивный) лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели. При этом рекурсивные модели не обязательно дают лучшие прогнозы по сравнению с фиксированными моделями. Для обоих типов выбросов наилучшей оказалась фиксированная *DS*-модель.

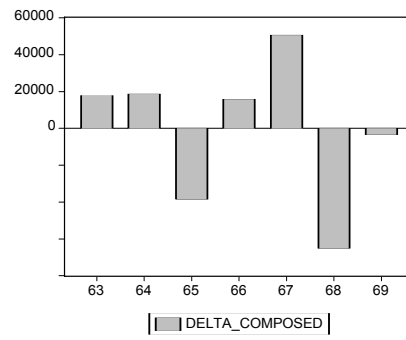
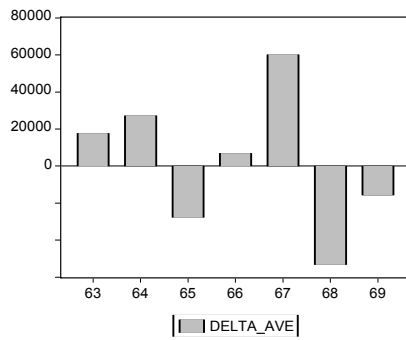
Приведем теперь результаты **комбинированных прогнозов**, получаемых следующими способами:

- (а) на каждом шаге построения рекурсивных прогнозов в качестве прогнозного берется среднее арифметическое всех четырех рассмотренных выше рекурсивных прогнозов;
- (б) на каждом шаге построения рекурсивных прогнозов в качестве прогнозного берется прогноз по модели, давшей наилучший прогноз значения ряда на предыдущем шаге.

Прогнозы:



Ошибки прогнозов:



Модели, давшие наилучшие прогнозы на 1, 2, ..., 7 шагах, соответственно:

*DS\_IO, DS\_IO, TS\_AO, DS\_IO, DS\_IO, TS\_IO, TS\_IO.*

Напомним значения характеристик прогнозов по каждой из четырех рекурсивных моделей:

	DS_IO	TS_IO	DS_AO	TS_AO
<b>RMSE</b>	36796.05	36710.86	33750.29	35153.12
<b>MAE</b>	30477.43	31124.61	28915.77	29026.71
<b>MAPE</b>	3.793081	3.874630	3.604901	3.621122

Хотя модель *DS\_AO* ни на одном из семи шагов не дала наилучшего прогноза, она, тем не менее, оказалась лучшей среди рассмотренных рекурсивных моделей по совокупности прогнозов.

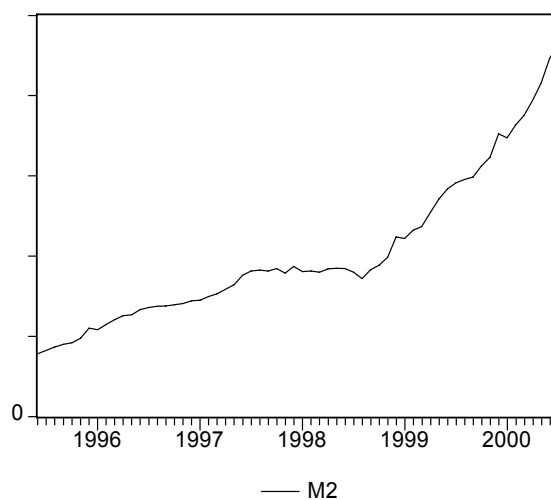
Характеристики прогнозов при применении стратегий (а) и (б):

	Способ (а) (усреднение)	Способ (б) (комбинирование)
RMSE	34966.44	36260.71
MAE	29789.99	29975.76
MAPE	3.711062	3.761013

Выбор модели, которая была наилучшей на предыдущем шаге, приводит к несколько худшим значениям трех характеристик полученной последовательности прогнозов по сравнению с простым усреднением альтернативных вариантов прогнозов. В то же время простое усреднение прогнозов служит лишь защитой от следования только одной модели, которая может в результате дать наихудшую последовательность прогнозов.

### 2.1.3. M2

За период 06.1995–02.2001 ряд M2 (номинальный показатель) имеет вид:



При использовании критериев Перрона на наличие у этого ряда единичного корня с эндогенным выбором даты излома в работе [Эконометрический анализ динамических рядов основных макроэкономических показателей (2001)] были получены следующие результаты (относящиеся к периоду с 06.1995 по 07.2000).

Для модели, допускающей сдвиг траектории и изменение наклона тренда в форме инновационного выброса (*IO*), в качестве даты излома указан 10.1999.

Для модели, допускающей только изменение наклона тренда (без сдвига траектории) в форме аддитивного выброса (*AO*), в качестве даты излома указан 05.1997.

В обоих случаях гипотеза единичного корня не отвергается, что, в согласии с результатами применения других критериев, приводит к решению о том, что поведение траектории ряда *M2* на временном интервале с 06.1995 по 07.2000 соответствует поведению *DS*-ряда. Имея в виду задачу сравнения прогнозов, получаемых по моделям *TS* и *DS*-рядов, построим для периода 06.1995–07.2000 подходящие модели обоих типов – *DS* и *TS*.

#### **DS-модели**

Сначала построим *DS*-модели с инновационным выбросом (дата излома 10.1999) и аддитивным выбросом (дата излома 05.1997). Оцененные модели имеют следующий вид.

#### **DS-модель с инновационным выбросом**

Dependent Variable: D(X)				
Method: Least Squares				
Sample: 1995:06 2000:07				
Included observations: 62				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3426.696	3623.040	0.945807	0.3491
DTB	-5425.164	14786.03	-0.366911	0.7153
DU	12888.66	8529.525	1.511064	0.1375
D(X(-1))	0.004036	0.137087	0.029442	0.9766
D(X(-2))	0.335098	0.134238	2.496290	0.0161
D(X(-3))	0.198757	0.142002	1.399685	0.1682
D(X(-4))	-0.000546	0.135054	-0.004044	0.9968
D(X(-5))	0.036323	0.136403	0.266290	0.7912
D(X(-6))	0.273612	0.135045	2.026077	0.0485
D(X(-7))	0.138695	0.160186	0.865839	0.3910
D(X(-8))	-0.164941	0.158634	-1.039758	0.3038
D(X(-9))	-0.426693	0.159670	-2.672350	0.0103
D(X(-10))	-0.058623	0.170621	-0.343588	0.7327
D(X(-11))	-0.145693	0.198169	-0.735195	0.4659
D(X(-12))	0.472155	0.187386	2.519684	0.0152
R-squared	0.625164	Mean dependent var	12790.32	
Adjusted R-squared	0.513510	S.D. dependent var	16295.92	
S.E. of regression	11366.21	Akaike info criterion	21.72156	
Sum squared resid	6.07E+09	Schwarz criterion	22.23619	
Log likelihood	-658.3684	F-statistic	5.599148	
Durbin-Watson stat	1.672096	Prob(F-statistic)	0.000004	

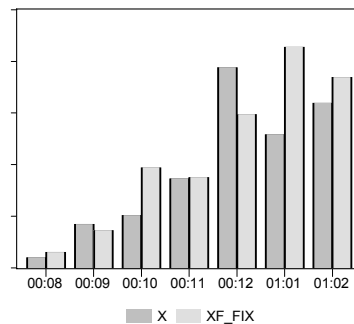
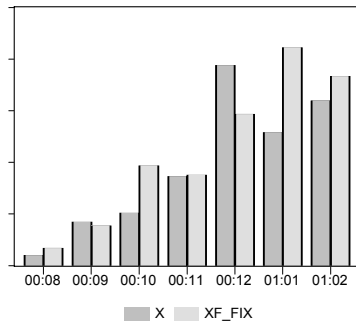
### DS-модель с аддитивным выбросом

Dependent Variable: D(X)				
Method: Least Squares				
Date: 11/22/01 Time: 10:29				
Sample: 1995:06 2000:07				
Included observations: 62				
Convergence achieved after 32 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2621.661	9620.741	0.272501	0.7864
DU	-4952.548	6111.386	-0.810381	0.4217
AR( 1)	0.080314	0.124527	0.644955	0.5220
AR( 2)	0.380210	0.127038	2.992891	0.0044
AR( 3)	0.233580	0.138363	1.688176	0.0979
AR( 4)	0.002231	0.134620	0.016570	0.9868
AR( 5)	0.017391	0.134480	0.129320	0.8976
AR( 6)	0.262875	0.134403	1.955865	0.0563
AR( 7)	0.170206	0.150058	1.134270	0.2623
AR( 8)	-0.104926	0.156309	-0.671272	0.5053
AR( 9)	-0.380568	0.157295	-2.419456	0.0194
AR( 10)	-0.014140	0.162464	-0.087036	0.9310
AR( 11)	-0.032809	0.155304	-0.211259	0.8336
AR( 12)	0.643909	0.154721	4.161736	0.0001
R-squared	0.612218	Mean dependent var		12790.32
Adjusted R-squared	0.507194	S.D. dependent var		16295.92
S.E. of regression	11439.76	Akaike info criterion		21.72326
Sum squared resid	6.28E+09	Schwarz criterion		22.20358
Log likelihood	-659.4209	F-statistic		5.829308
Durbin-Watson stat	1.708116	Prob( F-statistic)		0.000003

Если зафиксировать обе модели и производить прогнозы на один шаг вперед по этим двум моделям, то получаются следующие результаты:

Инновационный выброс:

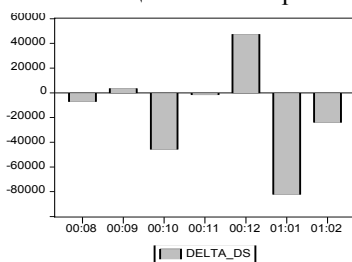
Аддитивный выброс:



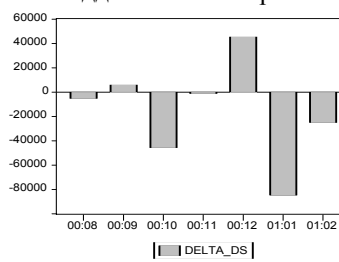


Ошибки прогнозов:

Инновационный выброс:



Аддитивный выброс:



Характеристики прогнозов по подобранным DS-моделям:

	DS (инновац.)	DS (аддитивн.)
Root Mean Squared Error	40870.17	41472.50
Mean Absolute Error	30168.47	30513.84
Mean Absolute Percent Error	2.814443	2.847573

По всем трем показателям качества обе модели достаточно близки, хотя формально лучшей оказывается модель с инновационным выбросом.

**TS-модели**

**TS-модель, допускающая сдвиг траектории и изменение наклона тренда в форме инновационного выброса с датой излома 10.1999**

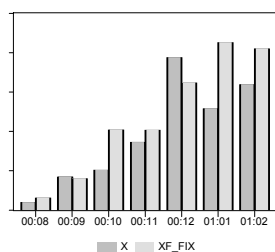
Dependent Variable: X				
Method: Least Squares				
Sample: 1995:06 2000:07				
Included observations: 62				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	36749.89	12099.40	3.037332	0.0040
DU	-4474.596	10582.46	-0.422831	0.6744
T	2886.380	835.1606	3.456077	0.0012
DT	8782.086	2681.435	3.275144	0.0020
X(-1)	0.846253	0.130008	6.509236	0.0000
X(-2)	0.282462	0.165456	1.707175	0.0947
X(-3)	-0.088658	0.171157	-0.517993	0.6070
X(-4)	-0.233731	0.170416	-1.371532	0.1770
X(-5)	0.000646	0.172620	0.003745	0.9970
X(-6)	0.279709	0.168430	1.660687	0.1037
X(-7)	-0.013597	0.186172	-0.073037	0.9421
X(-8)	-0.277851	0.201141	-1.381378	0.1740
X(-9)	-0.347935	0.205507	-1.693054	0.0974
X(-10)	0.290939	0.203619	1.428843	0.1600
X(-11)	-0.020465	0.224270	-0.091253	0.9277
X(-12)	0.567617	0.207125	2.740453	0.0088
X(-13)	-0.680590	0.179540	-3.790734	0.0004
R-squared	0.997889	Mean dependent var	408717.7	
Adjusted R-squared	0.997139	S.D. dependent var	189898.8	
S.E. of regression	10157.86	Akaike info criterion	21.51780	
Sum squared resid	4.64E+09	Schwarz criterion	22.10104	
Log likelihood	-650.0518	F-statistic	1329.635	
Durbin-Watson stat	1.686412	Prob(F-statistic)	0.000000	

**TS-модель, допускающая изменение наклона тренда в форме аддитивного выброса с датой излома 05.1997**

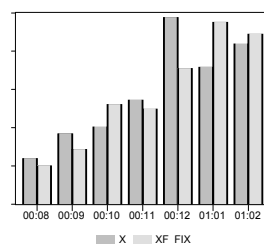
Dependent Variable: X				
Method: Least Squares				
Date: 11/22/01 Time: 11:24				
Sample( adjusted): 1996:07 2000:07				
Included observations: 49 after adjusting endpoints				
Convergence achieved after 26 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	216543.9	97470.58	2.221633	0.0333
T	665.5531	6387.900	0.104190	0.9176
DT	14928.06	9393.632	1.589168	0.1216
AR( 1)	0.981452	0.152779	6.424014	0.0000
AR( 2)	0.376224	0.210744	1.785219	0.0834
AR( 3)	-0.187398	0.222473	-0.842339	0.4057
AR( 4)	-0.321529	0.220437	-1.458599	0.1541
AR( 5)	0.007520	0.222916	0.033733	0.9733
AR( 6)	0.290384	0.218354	1.329876	0.1927
AR( 7)	0.004910	0.228411	0.021497	0.9830
AR( 8)	-0.362048	0.255537	-1.416814	0.1659
AR( 9)	-0.337684	0.256431	-1.316861	0.1970
AR( 10)	0.415446	0.253427	1.639316	0.1106
AR( 11)	0.101306	0.265421	0.381679	0.7051
AR( 12)	0.592220	0.257105	2.303415	0.0277
AR( 13)	-0.691692	0.190090	-3.638771	0.0009
R-squared	0.996845	Mean dependent var		461334.7
Adjusted R-squared	0.995410	S.D. dependent var		178906.8
S.E. of regression	12120.45	Akaike info criterion		21.90092
Sum squared resid	4.85E+09	Schwarz criterion		22.51866
Log likelihood	-520.5727	F-statistic		695.0137
Durbin-Watson stat	1.651115	Prob( F-statistic)		0.000000

Если зафиксировать обе модели и производить последовательные прогнозы на один шаг вперед по этим двум моделям, то получаются следующие результаты:

Инновационный выброс:

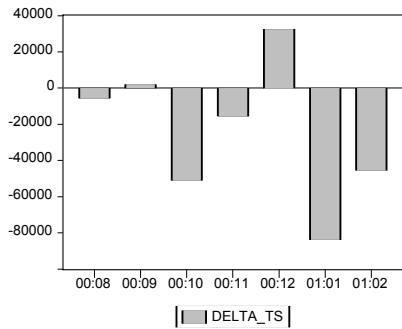


Аддитивный выброс:

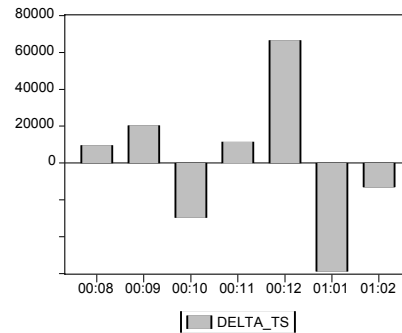


Соответственно, ошибки прогнозов имеют вид:

Инновационный выброс:



Аддитивный выброс:



Характеристики прогнозов по подобранным TS-моделям:

	TS (инновац.)	TS (аддитивн.)
<b>Root Mean Squared Error</b>	43134.08	36993.30
<b>Mean Absolute Error</b>	33700.94	29946.52
<b>Mean Absolute Percent Error</b>	3.154151	2.796299

Прогноз по второй модели оказывается предпочтительнее по всем трем показателям.

Сведем полученные результаты в одну таблицу.

	DS иннов.	TS иннов.	DS аддитив.	TS аддитив.
<b>RMSE</b>	40870.17	43134.08	41472.50	36993.30
<b>MAE</b>	30168.47	33700.94	30513.84	29946.52
<b>MAPE</b>	2.814443	3.154151	2.847573	2.796299

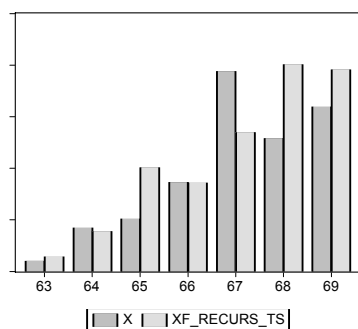
Здесь *DS*-модель оказалась лучшей только среди моделей с инновационным выбросом. Наилучшей, как и в случае ряда  $M_0$ , оказалась *TS*-модель с аддитивным выбросом.

Приведем теперь результаты, получаемые при использовании рекурсивных моделей (использующих те же переменные, те же количества запаздываний и даты изломов, что и в соответствующих фиксированных моделях).

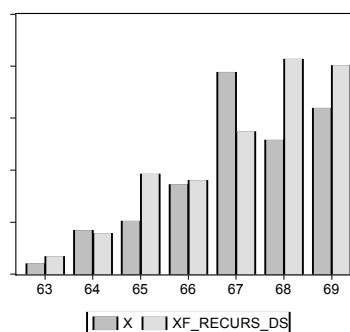
### Рекурсивные модели с инновационным выбросом

Прогнозы:

по рекурсивной *TS*-модели

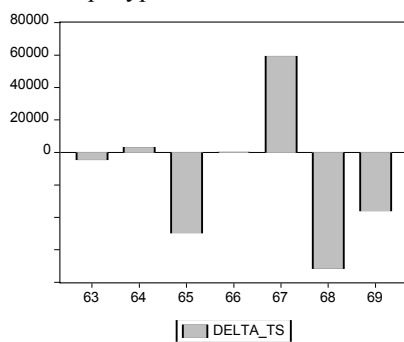


по рекурсивной *DS*-модели

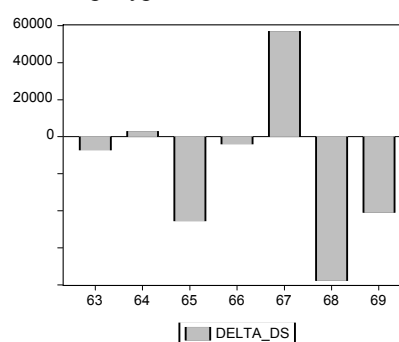


Ошибки прогнозов:

по рекурсивной *TS*-модели



по рекурсивной *DS*-модели



Характеристики точности прогнозов по рекурсивным моделям с инновационным выбросом:

	Рекурсивная DS	Рекурсивная TS
<b>Root Mean Squared Error</b>	43295.36	42186.68
<b>Mean Absolute Error</b>	33643.49	32129.98
<b>Mean Absolute Percent Error</b>	3.123998	2.982067

Рекурсивная *TS*-модель оказалась несколько лучше рекурсивной *DS*-модели.

Сведем полученные результаты в отношении ряда M2 для моделей с инновационным выбросом в одну таблицу.

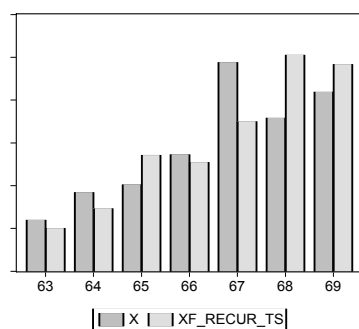
	DS рекурс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	43295.36	42186.68	40870.17	43134.08
<b>Mean Absolute Error</b>	33643.49	32129.98	30168.47	33700.94
<b>Mean Absolute Percent Error</b>	3.123998	2.982067	2.814443	3.154151

Рекурсивная *DS*-модель дает лучшие прогнозы по сравнению с фиксированной *DS*-моделью, тогда как прогнозы по рекурсивной *TS*-модели несколько хуже, чем по фиксированной *DS*-модели. Лучшей в классе моделей с инновационным выбросом оказалась фиксированная *DS*-модель.

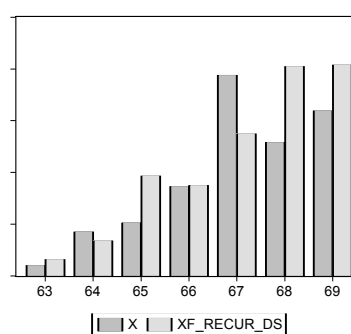
### Рекурсивные модели с аддитивным выбросом

Прогнозы:

по рекурсивной *TS*-модели

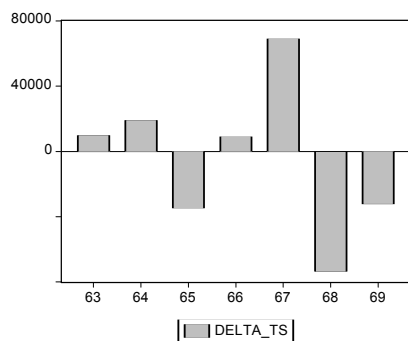


по рекурсивной *DS*-модели

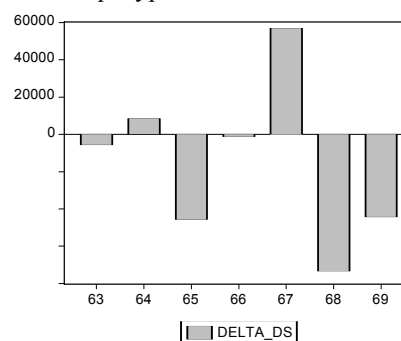


Ошибки прогнозов:

по рекурсивной *TS*-модели



по рекурсивной *DS*-модели



Характеристики точности прогнозов по рекурсивным моделям с аддитивным выбросом:

	Рекурсивная TS	Рекурсивная DS
<b>Root Mean Squared Error</b>	43020.02	42689.06
<b>Mean Absolute Error</b>	35293.53	33642.68
<b>Mean Absolute Percent Error</b>	3.284899	3.125200

Рекурсивная *DS*-модель имеет несколько лучшие характеристики прогнозов.

Сведем полученные результаты в отношении ряда M2 для моделей с аддитивным выбросом в одну таблицу.

	DS рекурс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	42689.06	43020.02	41472.50	36993.30
<b>Mean Absolute Error</b>	33642.68	35293.53	30513.84	29946.52
<b>Mean Absolute Percent Error</b>	3.125200	3.284899	2.847573	2.796299

Рекурсивные модели дают худшие прогнозы по сравнению с фиксированными моделями и в классе *DS*-моделей, и в классе *TS*-моделей. Лучшей в классе моделей с аддитивным выбросом оказалась фиксированная *DS*-модель.

**Подведем итоги исследования прогнозов на один шаг вперед для ряда M2.** В целом рекурсивные модели дают худшие прогнозы по сравнению с фиксированными моделями. Лучшей и в классе моделей с аддитивным выбросом, и в классе моделей с инновационным выбросом оказалась фиксированная *DS*-модель.

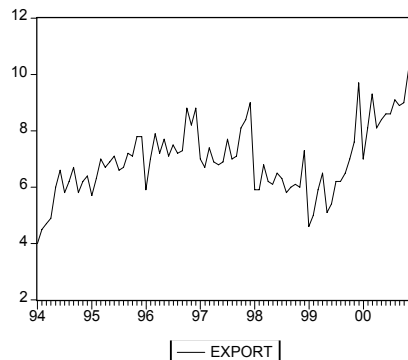
Проведенное эконометрическое моделирование и оценка различных методов прогнозирования денежных агрегатов M0, M1 и M2 показали, что для M1 и M2 наилучшей является фиксированная *DS*-модель, так что прогнозирование наиболее точно осуществляется для приростов денежных агрегатов без переоценки моделей. Содержательно это означает, что в то время как данные указывают на наличие стохастического тренда в динамике денежных агрегатов, приросты агрегатов вполне прогнозируемы. Это соответствует тому, что в относительно стабильных условиях Центральный банк придерживается стабильных ориентиров в проведении денежно-кредитной политики (как в отношении денежной базы, так и в отношении мультипликатора).

Полученный результат о том, что для прогнозирования временного ряда M0 предпочтительнее использовать рекурсивные модели, указывает на

то, что этот ряд подвержен более динамичным изменениям. Одно из объяснений этого может заключаться в том, что предложение наличных денег в отсутствие эффективных инструментов для стерилизации может зависеть от предложения валюты на валютном рынке, а значит, Центральный банк не может жестко контролировать  $M0$ .

## 2.2. Экспорт

В работе [Эконометрический анализ динамических рядов основных макроэкономических показателей (2001)] по результатам анализа данных за период с 01.1994 по 04.2000 этот ряд был отнесен к классу  $TS$ -рядов и квалифицирован как стационарный ряд с ненулевым средним. Мы рассмотрим поведение прогнозов для этого ряда на расширенном интервале до 12.2000. На этом интервале график ряда имеет вид:



Оцененные авторегрессионные модели поведения этого ряда на временном интервале 01.1994–04.2000

### TS-модель

Dependent Variable: X				
Method: Least Squares				
Sample( adjusted): 1995:02 2000:04				
Included observations: 63 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.239644	1.010418	2.216552	0.0313
X(-1)	0.651056	0.118845	5.478195	0.0000
X(-2)	0.015391	0.117147	0.131382	0.8960
X(-3)	0.146575	0.117572	1.246683	0.2184
X(-4)	-0.000202	0.121629	-0.001661	0.9987
X(-5)	0.022581	0.133819	0.168740	0.8667
X(-6)	-0.085018	0.133635	-0.636194	0.5276

Продолжение таблицы

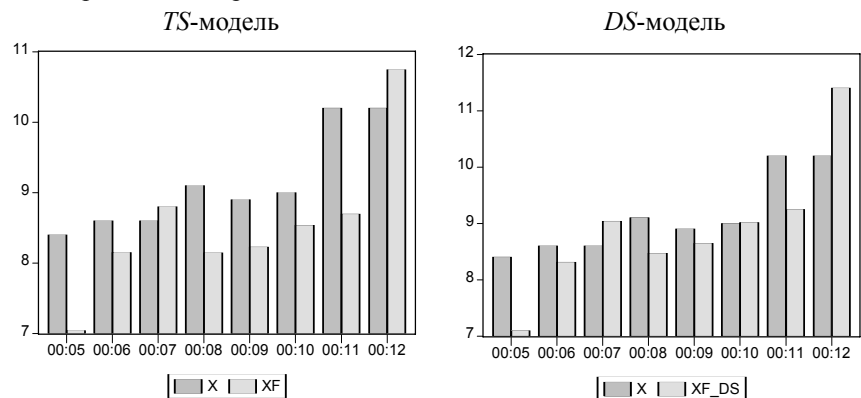
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X(-7)	0.024961	0.134457	0.185642	0.8535
X(-8)	0.028611	0.133500	0.214313	0.8312
X(-9)	-0.108398	0.132628	-0.817312	0.4177
X(-10)	-0.081318	0.133100	-0.610959	0.5441
X(-11)	-0.106808	0.131358	-0.813111	0.4201
X(-12)	0.780794	0.135034	5.782197	0.0000
X(-13)	-0.610254	0.130020	-4.693549	0.0000
R-squared	0.662984	Mean dependent var		6.982540
Adjusted R-squared	0.573572	S.D. dependent var		1.018148
S.E. of regression	0.664866	Akaike info criterion		2.214667
Sum squared resid	21.66028	Schwarz criterion		2.690919
Log likelihood	-55.76201	F-statistic		7.414900
Durbin-Watson stat	2.096796	Prob(F-statistic)		0.000000

### DS-модель

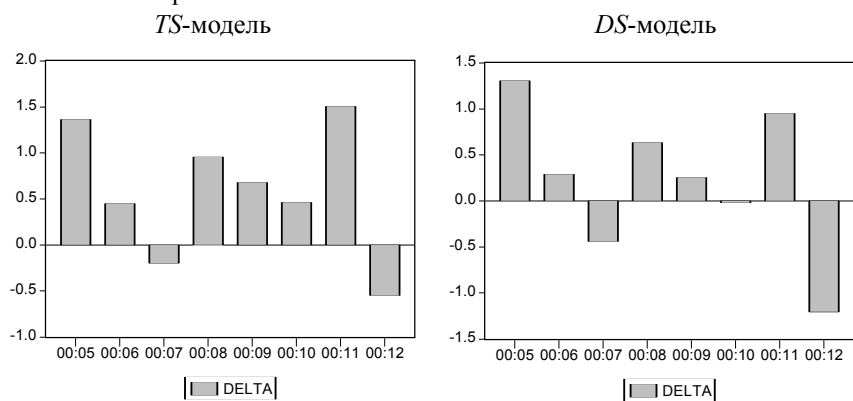
Dependent Variable: D(X)				
Method: Least Squares				
Sample( adjusted): 1995:02 2000:04				
Included observations: 63 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D( X(-1))	-0.277817	0.117524	-2.363925	0.0219
D( X(-2))	-0.229351	0.121724	-1.884185	0.0652
D( X(-3))	-0.055384	0.126158	-0.439003	0.6625
D( X(-4))	-0.052706	0.126601	-0.416318	0.6789
D( X(-5))	-0.016802	0.135283	-0.124196	0.9016
D( X(-6))	-0.090949	0.139185	-0.653441	0.5164
D( X(-7))	-0.048823	0.146741	-0.332715	0.7407
D( X(-8))	0.007220	0.146710	0.049216	0.9609
D( X(-9))	-0.069587	0.142771	-0.487403	0.6281
D( X(-10))	-0.120175	0.139854	-0.859290	0.3942
D( X(-11))	-0.190877	0.134192	-1.422409	0.1610
D( X(-12))	0.627777	0.132611	4.733969	0.0000
R-squared	0.597714	Mean dependent var		0.038095
Adjusted R-squared	0.510947	S.D. dependent var		0.978549
S.E. of regression	0.684322	Akaike info criterion		2.248867
Sum squared resid	23.88313	Schwarz criterion		2.657083
Log likelihood	-58.83931	Durbin-Watson stat		2.040815



Прогнозы по приведенным оцененным моделям:



Ошибки прогнозов:



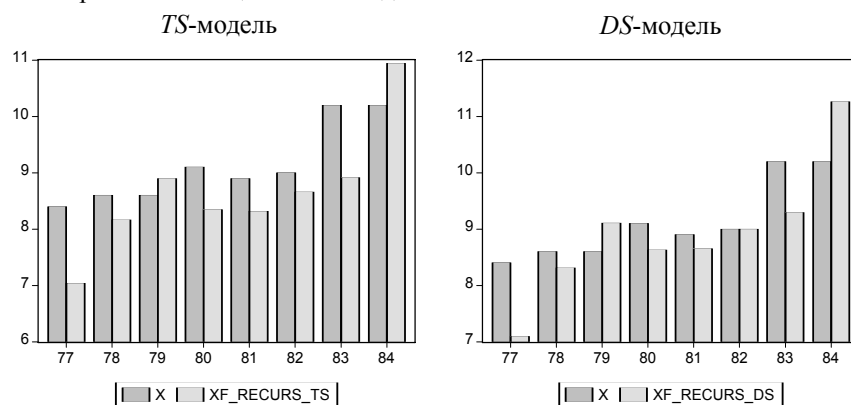
Характеристики прогнозов по подобранным моделям:

	DS	TS
<b>Root Mean Squared Error</b>	0.774216	0.883284
<b>Mean Absolute Error</b>	0.636348	0.769132
<b>Mean Absolute Percent Error</b>	6.889349	8.383397

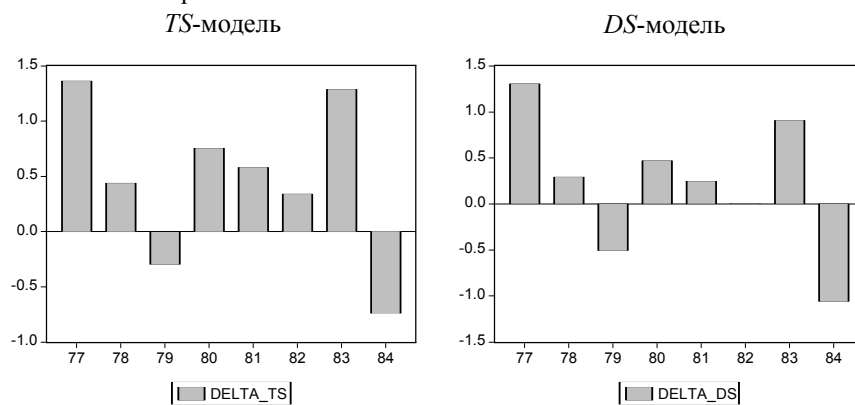
По всем трем показателям лучшее качество прогнозов дает DS-модель.

Рассмотрим теперь результаты прогнозирования по рекурсивным *TS*- и *DS*-моделям.

Прогнозы по оцененным моделям:



Ошибки прогнозов:



Характеристики прогнозов по рекурсивным моделям:

	<b>DS рекурс.</b>	<b>TS рекурс.</b>
<b>Root Mean Squared Error</b>	0.731282	0.818717
<b>Mean Absolute Error</b>	0.599139	0.725064
<b>Mean Absolute Percent Error</b>	6.513077	7.905067

Сведем полученные результаты в отношении ряда EXPORT в одну таблицу.

	DS рекурс.	TS рекурс.	DS фикс.	TS фикс.
Root Mean Squared Error	0.731282	0.818717	0.774216	0.883284
Mean Absolute Error	0.599139	0.725064	0.636348	0.769132
Mean Absolute Percent Error	6.513077	7.905067	6.889349	8.383397

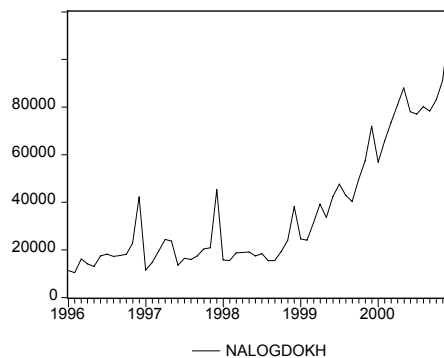
И среди фиксированных, и среди рекурсивных моделей лучшими оказываются *DS*-модели. И среди *DS*-моделей, и среди *TS*-моделей лучшими оказываются рекурсивные модели. Наилучшей является рекурсивная *DS*-модель.

Экспорт Российской Федерации более чем на 50% состоит из экспорта минерального сырья, черных и цветных металлов, цены на которые в значительной степени волатильны и определяются на мировых рынках ресурсов. Эти цены, в частности цены на нефть, являются нестационарными, что в значительной степени определяет наличие стохастического тренда во временном ряде экспорта. Дополнительной особенностью является то, что прогнозирование экспорта в приростах на рассматриваемом временном интервале лучше осуществляется на основе рекурсивных моделей. Это обстоятельство может быть частично объяснено изменением структуры экспорта, что привносит дополнительные инновационные выбросы в рассматриваемый временной ряд.

### 2.3. Налоговые доходы федерального бюджета

В работе [Эконометрический анализ динамических рядов основных макроэкономических показателей (2001)] этот ряд рассматривался на временном интервале с 01.1996 по 05.2000 и был классифицирован в этом периоде как *DS*-ряд.

На интервале 01.1996–12.2000 график ряда имеет вид:



При использовании критериев Перрона на наличие у этого ряда единичного корня с эндогенным выбором даты излома в упомянутой работе были получены следующие результаты (относящиеся к периоду с 01.1996 по 05.2000).

Для модели, допускающей сдвиг траектории и изменение наклона тренда в форме инновационного выброса (IO), в качестве даты излома указан 04.1999.

Для модели, допускающей только изменение наклона тренда (без сдвига траектории) в форме аддитивного выброса (AO), в качестве даты излома указан 02.1999.

В обоих случаях гипотеза единичного корня не отвергается, что, в согласии с результатами применения других критериев, приводит к решению о том, что поведение траектории ряда на временном интервале с 01.1996 по 05.2000 соответствует поведению *DS*-ряда. Тем не менее, имея в виду задачу сравнения прогнозов, получаемых по моделям *TS*- и *DS*-рядов, мы построим для периода 01.1996–05.2000 подходящие модели обоих типов – *DS* и *TS*.

#### **DS-модели**

Сначала построим авторегрессионные *DS*-модели с инновационным выбросом (дата излома 04.1999) и с аддитивным выбросом (дата излома 02.1999), имеющие порядок 11.

#### **Оцененная DS-модель с инновационным выбросом (дата излома 04.1999)**

Dependent Variable: D(X)				
Method: Least Squares				
Sample( adjusted): 1997:02 2000:05				
Included observations: 40 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1059.300	1049.066	1.009756	0.3223
DTB	-13770.96	6072.221	-2.267862	0.0322
DU	8327.450	4843.548	1.719287	0.0979
D(X(-1))	-0.325442	0.174342	-1.866695	0.0737
D(X(-2))	-0.242826	0.179405	-1.353504	0.1880
D(X(-3))	-0.171433	0.182824	-0.937694	0.3574
D(X(-4))	-0.049076	0.184115	-0.266554	0.7920
D(X(-5))	0.080119	0.186491	0.429614	0.6712
D(X(-6))	-0.024464	0.206353	-0.118556	0.9066
D(X(-7))	-0.024099	0.223645	-0.107754	0.9151
D(X(-8))	-0.144790	0.230824	-0.627276	0.5362
D(X(-9))	-0.356881	0.222879	-1.601234	0.1219
D(X(-10))	-0.341746	0.213861	-1.597985	0.1226

Продолжение таблицы

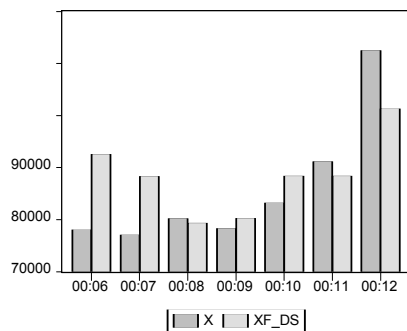
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D( X( -11))	-0.365560	0.193194	-1.892197	0.0701
D( X( -12))	0.392994	0.175103	2.244358	0.0339
R-squared	0.774788	Mean dependent var		1916.750
Adjusted R-squared	0.648670	S.D. dependent var		8880.354
S.E. of regression	5263.664	Akaike info criterion		20.25504
Sum squared resid	6.93E+08	Schwarz criterion		20.88837
Log likelihood	-390.1008	F-statistic		6.143329
Durbin-Watson stat	1.920896	Prob( F-statistic)		0.000045

**DS-модель с аддитивным выбросом (дата излома 02.1999)**

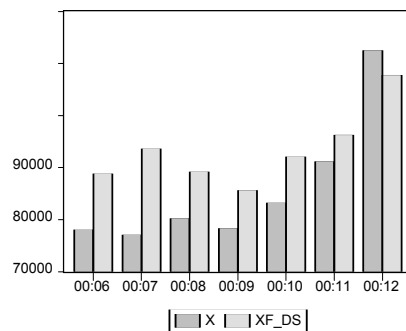
Dependent Variable: D(X)				
Method: Least Squares				
Sample(adjusted): 1996:05 2000:05				
Included observations: 49 after adjusting endpoints				
Convergence achieved after 5 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	238.1837	130.7869	1.821159	0.0769
DU	3592.132	333.1648	10.78185	0.0000
AR( 1)	-0.832716	0.117557	-7.083500	0.0000
AR( 2)	-0.748377	0.145532	-5.142359	0.0000
AR( 3)	-0.649621	0.165592	-3.923010	0.0004
AR( 4)	-0.534694	0.179638	-2.976508	0.0052
AR( 5)	-0.391297	0.183270	-2.135081	0.0396
AR( 6)	-0.411871	0.181896	-2.264319	0.0297
AR( 7)	-0.428597	0.183203	-2.339463	0.0250
AR( 8)	-0.473734	0.178118	-2.659665	0.0116
AR( 9)	-0.612492	0.166437	-3.680034	0.0008
AR( 10)	-0.708869	0.149669	-4.736254	0.0000
AR( 11)	-0.760419	0.124098	-6.127559	0.0000
R-squared	0.717094	Mean dependent var		1510.449
Adjusted R-squared	0.622792	S.D. dependent var		9668.323
S.E. of regression	5938.021	Akaike info criterion		20.43845
Sum squared resid	1.27E+09	Schwarz criterion		20.94036
Log likelihood	-487.7420	F-statistic		7.604223
Durbin-Watson stat	2.069674	Prob( F-statistic)		0.000001

Если зафиксировать обе модели и производить прогнозы на один шаг вперед по этим двум моделям, то получаются следующие результаты:

Инновационный выброс:

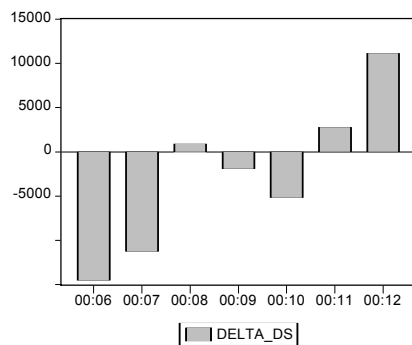


Аддитивный выброс:

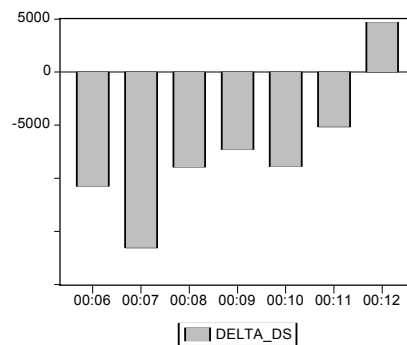


Ошибки прогнозов:

Инновационный выброс:



Аддитивный выброс:



Характеристики прогнозов по подобранным DS-моделям:

	DS инновац.	DS аддитивн.
<b>Root Mean Squared Error</b>	8459.231	9646.584
<b>Mean Absolute Error</b>	6809.506	8902.694
<b>Mean Absolute Percent Error</b>	7.994012	10.90060

### TS-модели

**TS-модель, допускающая сдвиг траектории и изменение наклона тренда в форме инновационного выброса с датой излома 04.1999**

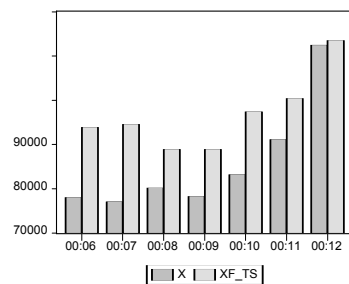
Dependent Variable: X				
Method: Least Squares				
Sample(adjusted): 1997:02 2000:05				
Included observations: 40 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8260.345	22877.30	0.361072	0.7213
DU	622.4802	5714.737	0.108925	0.9142
T	209.5077	161.5124	1.297162	0.2074
DT	2632.884	3030.193	0.868883	0.3939
X(-1)	0.385968	0.223207	1.729190	0.0972
X(-2)	0.021530	0.152492	0.141189	0.8890
X(-3)	0.035270	0.151872	0.232233	0.8184
X(-4)	0.090503	0.149418	0.605705	0.5506
X(-5)	0.056159	0.144855	0.387691	0.7018
X(-6)	-0.083065	0.160478	-0.517610	0.6097
X(-7)	-0.014157	0.163874	-0.086389	0.9319
X(-8)	-0.122748	0.169391	-0.724642	0.4760
X(-9)	-0.242626	0.160772	-1.509125	0.1449
X(-10)	-0.082384	0.157644	-0.522597	0.6063
X(-11)	-0.107512	0.156546	-0.686775	0.4991
X(-12)	0.680388	0.155324	4.380446	0.0002
X(-13)	-0.241768	0.217570	-1.111220	0.2780
R-squared	0.961063	Mean dependent var		33456.95
Adjusted R-squared	0.933976	S.D. dependent var		20596.03
S.E. of regression	5292.182	Akaike info criterion		20.28246
Sum squared resid	6.44E+08	Schwarz criterion		21.00024
Log likelihood	-388.6493	F-statistic		35.48084
Durbin-Watson stat	1.979157	Prob(F-statistic)		0.000000

**TS-модель, допускающая изменение наклона тренда  
в форме аддитивного выброса с датой излома 02.1999**

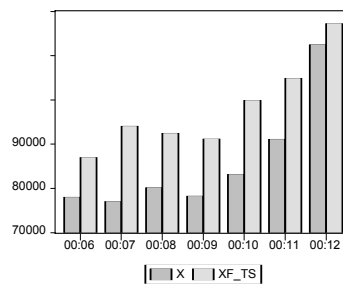
Dependent Variable: X				
Method: Least Squares				
Sample(adjusted): 1997:01 2000:05				
Included observations: 41 after adjusting endpoints				
Convergence achieved after 10 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	17176.29	2378.173	7.222474	0.0000
DU	2887.408	3785.067	0.762842	0.4527
T	141.7255	95.47941	1.484357	0.1502
DT	3519.175	309.2493	11.37973	0.0000
AR( 1)	-0.068358	0.170221	-0.401584	0.6914
AR( 2)	-0.092527	0.166782	-0.554775	0.5840
AR( 3)	-0.095203	0.163811	-0.581172	0.5663
AR( 4)	-0.025976	0.156108	-0.166395	0.8692
AR( 5)	0.014917	0.142789	0.104468	0.9176
AR( 6)	-0.142908	0.143724	-0.994324	0.3296
AR( 7)	-0.096370	0.142394	-0.676786	0.5048
AR( 8)	-0.158724	0.137935	-1.150716	0.2607
AR( 9)	-0.287779	0.139024	-2.069993	0.0489
AR( 10)	-0.199420	0.154979	-1.286753	0.2100
AR( 11)	-0.197585	0.157000	-1.258504	0.2198
AR( 12)	0.559493	0.165522	3.380172	0.0024
R-squared	0.964628	Mean dependent var		32920.44
Adjusted R-squared	0.943405	S.D. dependent var		20625.06
S.E. of regression	4906.662	Akaike info criterion		20.12037
Sum squared resid	6.02E+08	Schwarz criterion		20.78908
Log likelihood	-396.4675	F-statistic		45.45134
Durbin-Watson stat	1.437416	Prob(F-statistic)		0.000000

Если зафиксировать обе модели и производить последовательные прогнозы на один шаг вперед по этим двум моделям, то получаются следующие результаты:

Инновационный выброс:



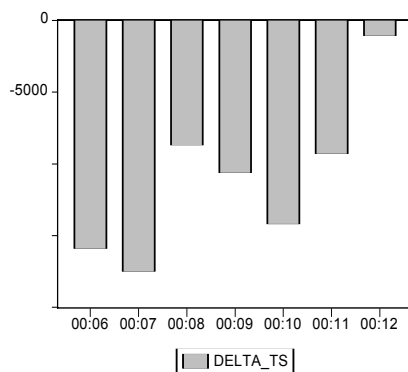
Аддитивный выброс:



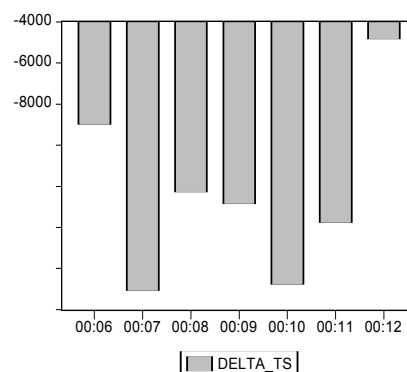


Соответственно, ошибки прогнозов имеют вид:

Инновационный выброс:



Аддитивный выброс:



Характеристики прогнозов по подобранным *TS*-моделям:

	<b>TS инновац.</b>	<b>TS аддитивн.</b>
<b>Root Mean Squared Error</b>	12180.72	12989.34
<b>Mean Absolute Error</b>	11051.46	12358.84
<b>Mean Absolute Percent Error</b>	13.68804	14.98790

Сведем полученные результаты в одну таблицу.

	<b>DS иннов.</b>	<b>TS иннов.</b>	<b>DS аддитивн.</b>	<b>TS аддитивн.</b>
<b>RMSE</b>	8459.231	12180.72	9646.584	12989.34
<b>MAE</b>	6809.506	11051.46	8902.694	12358.84
<b>MAPE</b>	7.994012	13.68804	10.90060	14.98790

Приведем теперь результаты, получаемые при использовании рекурсивных моделей, использующих одни и те же переменные, но коэффициенты которых переоцениваются при поступлении новых наблюдений (т.е. количество запаздываний не изменяется, так же как не изменяются и даты изломов).

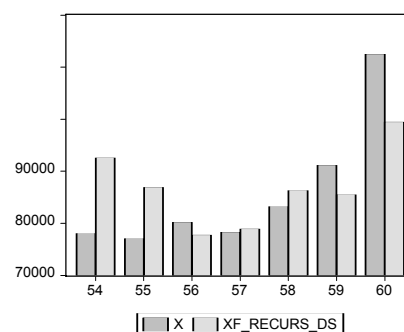
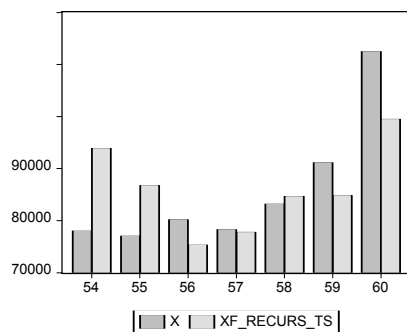
Для рекурсивных моделей на графиках по оси абсцисс откладываются номера последовательных наблюдений, начиная с 01.1996 (так что 53-е наблюдение соответствует 05.2000, а 60-е соответствует 12.2000).

### Рекурсивные модели с инновационным выбросом

Прогнозы:

по рекурсивной *TS*-модели

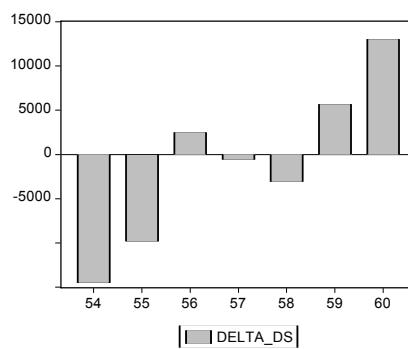
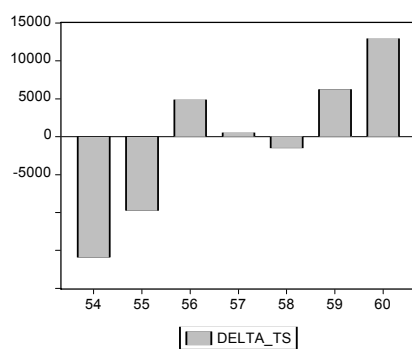
по рекурсивной *DS*-модели



Ошибки прогнозов:

по рекурсивной *TS*-модели

по рекурсивной *DS*-модели



Характеристики точности прогнозов по рекурсивным моделям с инновационным выбросом:

	Рекурсивная DS	Рекурсивная TS
<b>Root Mean Squared Error</b>	8656.144	9105.602
<b>Mean Absolute Error</b>	7019.826	7385.307
<b>Mean Absolute Percent Error</b>	8.095278	8.557180

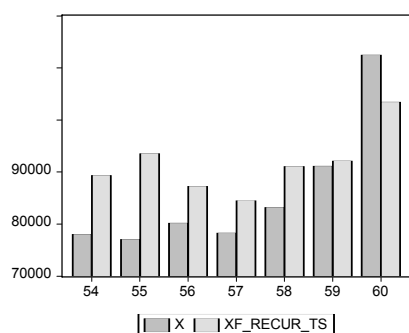
Сведем полученные результаты в отношении ряда налоговых доходов федерального бюджета для моделей с инновационным выбросом в одну таблицу.

	DS ре-курс.	TS ре-курс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	8656.144	9105.602	8459.231	12180.72
<b>Mean Absolute Error</b>	7019.826	7385.307	6809.506	11051.46
<b>Mean Absolute Percent Error</b>	8.095278	8.557180	7.994012	13.68804

И среди фиксированных, и среди рекурсивных моделей с инновационным выбросом лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели. Среди *DS*-моделей фиксированные модели оказались лучше рекурсивных. Однако среди *TS*-моделей рекурсивные модели обеспечивают значительное улучшение качества прогнозов. Наилучшей является фиксированная *DS*-модель.

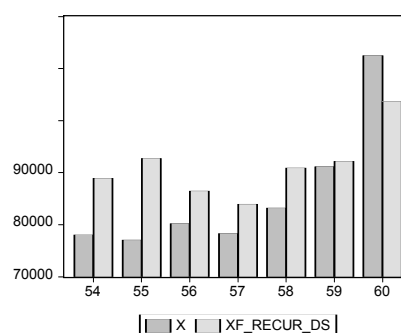
### Рекурсивные модели с аддитивным выбросом

по рекурсивной *TS*-модели



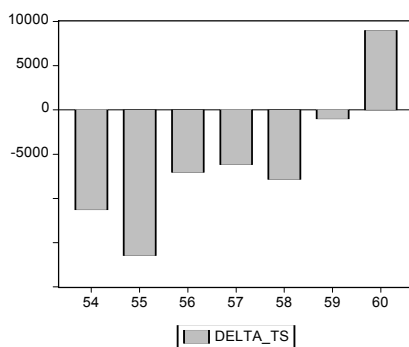
Прогнозы:

по рекурсивной *DS*-модели

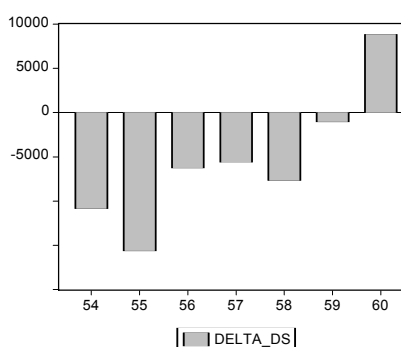


Ошибки прогнозов:

по рекурсивной *TS*-модели



по рекурсивной *DS*-модели



Характеристики точности прогнозов по рекурсивным моделям с аддитивным выбросом:

	<b>DS рекурс.</b>	<b>TS рекурс.</b>
<b>Root Mean Squared Error</b>	9044.562	9495.680
<b>Mean Absolute Error</b>	8003.194	8416.063
<b>Mean Absolute Percent Error</b>	9.648978	10.166833

Сведем полученные результаты в отношении ряда налоговых доходов федерального бюджета для моделей с аддитивным выбросом в одну таблицу.

	<b>DS рекурс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	9044.562	9495.680	9646.584	12989.34
<b>Mean Absolute Error</b>	8003.194	8416.063	8902.694	12358.84
<b>Mean Absolute Percent Error</b>	9.648978	10.166833	10.90060	14.98790

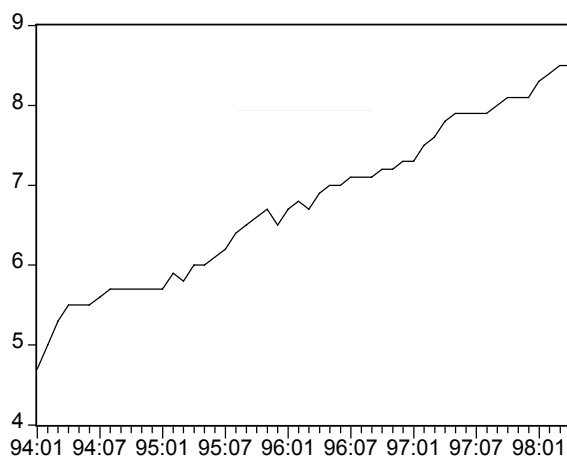
И среди фиксированных, и среди рекурсивных моделей с аддитивным выбросом лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели. При этом и среди *TS*-моделей, и среди *DS*-моделей лучшими оказались рекурсивные модели. Наилучшей является рекурсивная *DS*-модель.

**Подведем итоги исследования прогнозов на один шаг вперед для налоговых доходов федерального бюджета.** И среди фиксированных, и среди рекурсивных моделей для каждого из двух типов выбросов (инновационный или аддитивный) лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели. При этом рекурсивные модели не обязательно дают лучшие прогнозы по сравнению с фиксированными моделями. Наилучшей для инновационного выброса оказалась фиксированная *DS*-модель, а наилучшей для аддитивного выброса оказалась рекурсивная *DS*-модель.

При прогнозировании налоговых доходов федерального бюджета также оказалось, что данный временной ряд лучше прогнозировать в приростах. Как видно из динамики, ряд имеет значительную сезонную составляющую со значительными отклонениями в отдельные годы. Возможно поэтому рекурсивные модели не всегда лучше фиксированных, даже несмотря на налоговую реформу, проводимую в последние годы, которая не могла не сказаться на величине и динамике налоговых поступлений в федеральный бюджет РФ. Более подробно исследование качества прогнозирования налоговых доходов и поступлений отдельных налогов приведено в главе 5.

## 2.4. Безработица

Для анализа временного ряда безработицы будем использовать период с 01.1994 по 04.1998, на котором этот ряд в работе [Эконометрический анализ динамических рядов основных макроэкономических показателей (2001)] был определен как стационарный относительно линейного тренда. График ряда на этом интервале имеет вид:



— UNJOB

В качестве базового возьмем здесь период с 01.1994 по 09.1997, так что прогнозы будем производить, как и для предыдущих рядов, на последующие семь месяцев.

Построение *TS*-модели дает следующие результаты:

Dependent Variable: UNJOB				
Sample(adjusted): 1994:02 1997:09				
Included observations: 44 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.374932	0.537120	4.421607	0.0001
T	0.029559	0.007321	4.037459	0.0002
UNJOB(-1)	0.539134	0.108762	4.957033	0.0000
R-squared	0.989811	Mean dependent var		6.572727
Adjusted R-squared	0.989314	S.D. dependent var		0.848080
S.E. of regression	0.087670	Akaike info criterion		-1.964718
Sum squared resid	0.315130	Schwarz criterion		-1.843069
Log likelihood	46.22381	F-statistic		1991.396
Durbin-Watson stat	1.808843	Prob( F-statistic)		0.000000

(Здесь модель с включением в правую часть только запаздывания на один шаг уже проходит все стандартные тесты.)

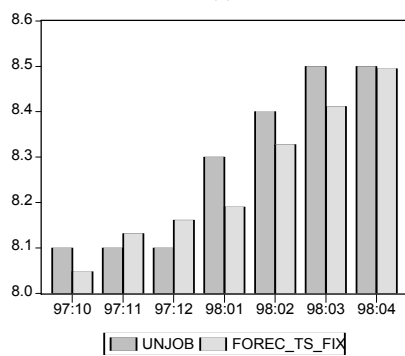
При построении *DS*-модели получаем:

Dependent Variable: D( UNJOB)				
Sample(adjusted): 1994:02 1997:09				
Included observations: 44 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.075000	0.015635	4.796958	0.0000
R-squared	0.000000	Mean dependent var		0.075000
Adjusted R-squared	0.000000	S.D. dependent var		0.103710
S.E. of regression	0.103710	Akaike info criterion		-1.671967
Sum squared resid	0.462500	Schwarz criterion		-1.631417
Log likelihood	37.78327	Durbin-Watson stat		1.989189

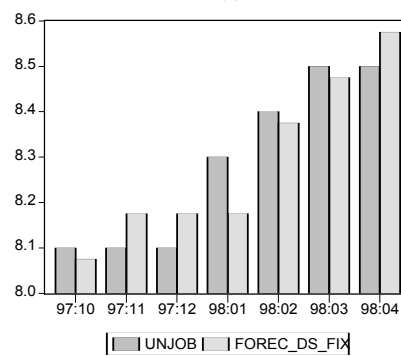
(Эта модель также проходит все стандартные тесты.)

Фиксируя эти две модели, получаем следующие результаты последовательного прогнозирования значений ряда на один шаг вперед.

*TS*-модель

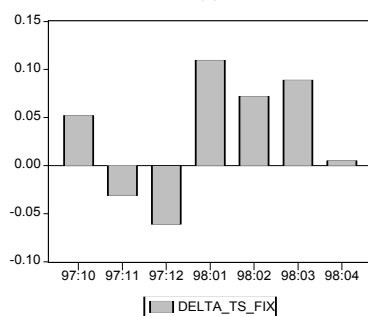


*DS*-модель

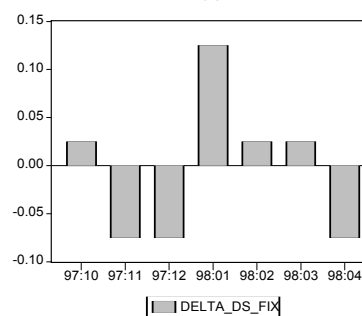


Ошибки прогнозов:

*TS*-модель



*DS*-модель



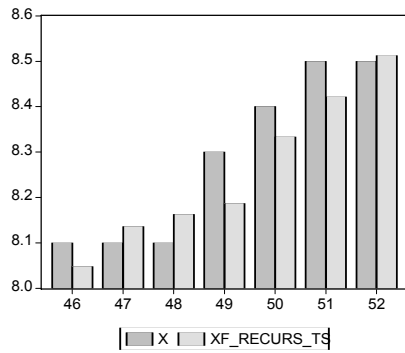
Характеристики прогнозов:

	TS	DS
<b>Root Mean Squared Error</b>	0.068207	0.070076
<b>Mean Absolute Error</b>	0.060048	0.060714
<b>Mean Absolute Percent Error</b>	0.724306	0.734373

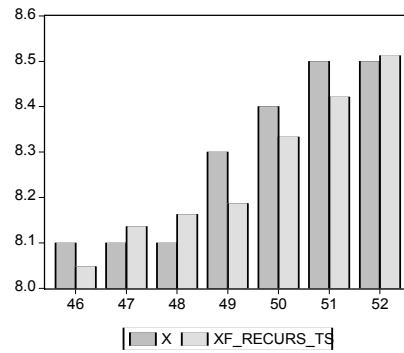
Значения указанных трех характеристик для *TS* и *DS*-моделей достаточно близки. Некоторое преимущество имеет *TS*-модель.

Приведем теперь результаты прогнозирования по рекурсивным *TS*- и *DS*-моделям.

*TS*-модель

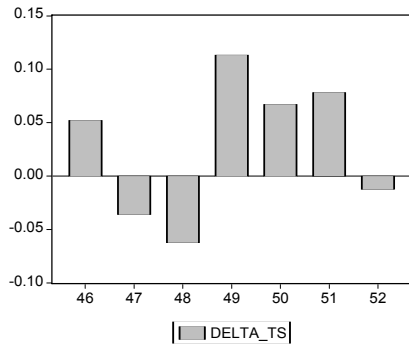


*DS*-модель

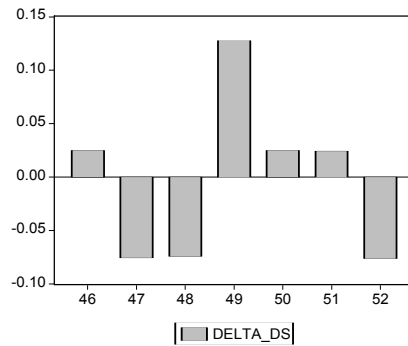


Ошибки прогнозов:

*TS*-модель



*DS*-модель



Характеристики одношаговых прогнозов по рекурсивным моделям:

	<b>DS рекурс.</b>	<b>TS рекурс.</b>
<b>Root Mean Squared Error</b>	0.070806	0.067143
<b>Mean Absolute Error</b>	0.061088	0.060246
<b>Mean Absolute Percent Error</b>	0.738836	0.727211

Для рекурсивных *TS*- и *DS*-моделей характеристики прогнозов опять достаточно близки, преимущество также имеет *TS*-модель.

Сведем полученные результаты в отношении ряда UNJOB в одну таблицу.

	<b>DS рекурс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	0.070806	0.067143	0.070076	0.068207
<b>Mean Absolute Error</b>	0.061088	0.060246	0.060714	0.060048
<b>Mean Absolute Percent Error</b>	0.738836	0.727211	0.734373	0.724306

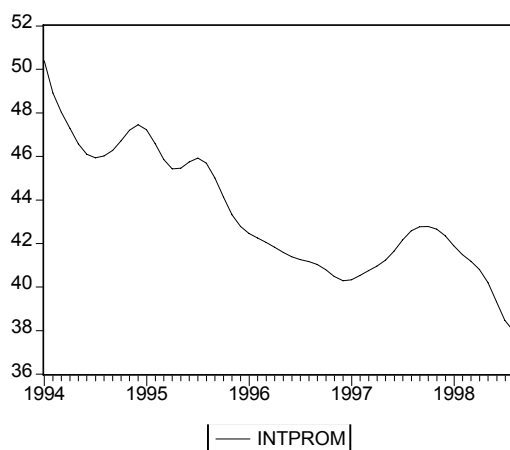
И среди фиксированных, и среди рекурсивных моделей несколько лучшими оказываются *TS*-модели. В целом все четыре модели имеют очень близкие показатели качества прогнозов. Как видно из графика динамики безработицы, этот показатель имеет достаточно стабильную динамику без резких колебаний с выраженным повышающимся трендом. Статистика по безработице в Российской Федерации очень слабо отражает занятость трудовых ресурсов, и это объясняется тем, что довольно большое количество работников формально числятся на предприятиях, работая неполную неделю или неполный рабочий день (некоторые предприятия формально функционируют, практически не работая, и т.п.). В целом официальная статистика по безработице является достаточно инертным, легко прогнозируемым показателем, что и объясняет полученные при анализе результаты.

## **2.5. Индекс интенсивности промышленного производства**

Следуя работе [Эконометрический анализ динамических рядов основных макроэкономических показателей (2001)], возьмем для рассмотрения докризисный период с 01.1994 по 08.1998, на котором ряд индекса интенсивности промышленного производства (сезонно сглаженный) был классифицирован как *DS*-ряд.



График ряда на данном интервале имеет вид:



Оцененная  $TS$ -модель авторегрессии относительно линейного тренда имеет вид:

Dependent Variable: X				
Sample: 1994:01 1998:01				
Included observations: 49				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.420358	0.248287	1.693031	0.0984
T	-0.000355	0.001167	-0.304595	0.7623
X(-1)	3.767049	0.140808	26.75316	0.0000
X(-2)	-6.570359	0.570864	-11.50950	0.0000
X(-3)	6.635075	1.172242	5.660157	0.0000
X(-4)	-3.519549	1.589186	-2.214687	0.0327
X(-5)	-0.182805	1.533573	-0.119202	0.9057
X(-6)	1.851247	1.054838	1.755005	0.0871
X(-7)	-1.373388	0.478584	-2.869689	0.0066
X(-8)	0.382818	0.108436	3.530373	0.0011
R-squared	0.999758	Mean dependent var		43.86466
Adjusted R-squared	0.999702	S.D. dependent var		2.682861
S.E. of regression	0.046324	Akaike info criterion		-3.126391
Sum squared resid	0.083692	Schwarz criterion		-2.740305
Log likelihood	86.59658	F-statistic		17884.22
Durbin-Watson stat	1.874209	Prob(F-statistic)		0.000000

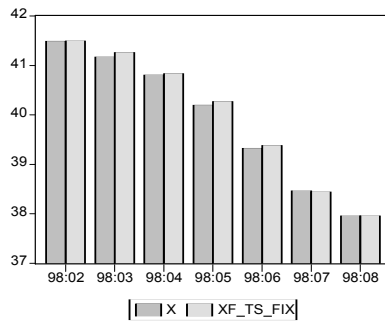
(Порядок модели определяется как и ранее).

Оцененная *DS*-модель имеет вид:

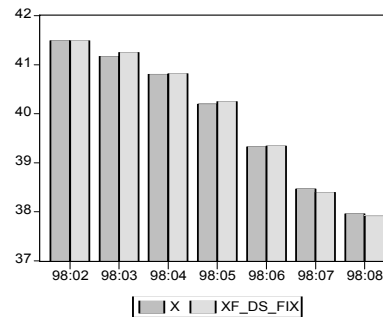
Dependent Variable: D(X)				
Method: Least Squares				
Sample: 1994:01 1998:01				
Included observations: 49				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.014830	0.008756	-1.693669	0.0979
D( X(-1))	2.878065	0.142754	20.16106	0.0000
D( X(-2))	-4.038304	0.454983	-8.875715	0.0000
D( X(-3))	3.221708	0.782165	4.118961	0.0002
D( X(-4))	-1.089103	0.900020	-1.210088	0.2332
D( X(-5))	-0.535873	0.723454	-0.740715	0.4631
D( X(-6))	0.817822	0.390195	2.095930	0.0423
D( X(-7))	-0.327252	0.111240	-2.941861	0.0053
R-squared	0.992350	Mean dependent var	-0.220853	
Adjusted R-squared	0.991044	S.D. dependent var	0.522865	
S.E. of regression	0.049481	Akaike info criterion	-3.026186	
Sum squared resid	0.100382	Schwarz criterion	-2.717317	
Log likelihood	82.14155	F-statistic	759.8284	
Durbin-Watson stat	1.787437	Prob(F-statistic)	0.000000	

Фиксируя эти две модели, получаем одношаговые прогнозы:

*TS*-модель

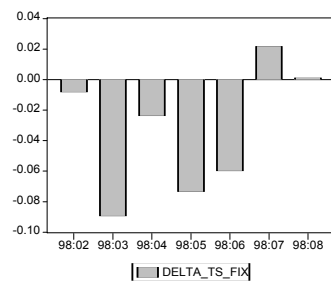


*DS*-модель

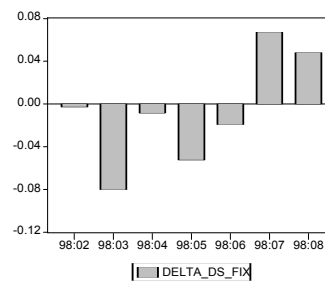


Ошибки прогнозов:

*TS*-модель



*DS*-модель

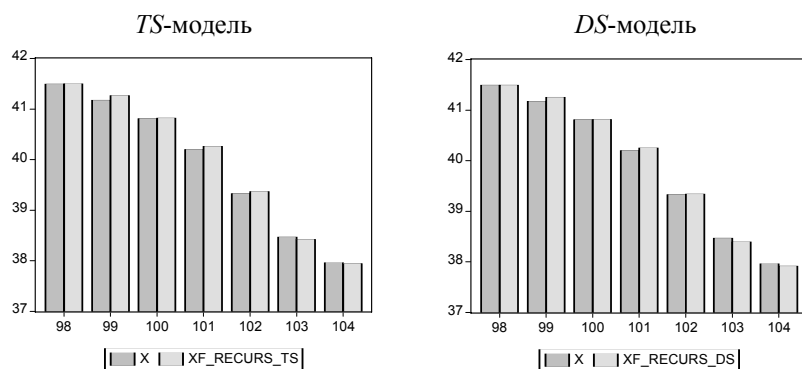


Характеристики прогнозов:

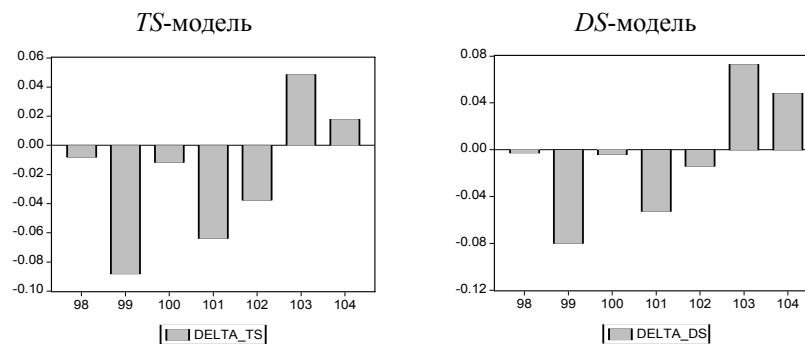
	TS	DS
<b>Root Mean Squared Error</b>	0.050764	0.048401
<b>Mean Absolute Error</b>	0.039615	0.039698
<b>Mean Absolute Percent Error</b>	0.098420	0.100229

Характеристики прогнозов по фиксированным моделям очень близки.

Приведем теперь результаты прогнозирования по рекурсивным *TS*- и *DS*-моделям.



Ошибки прогнозов:



Характеристики одношаговых прогнозов по рекурсивным моделям:

	DS рекурс.	TS рекурс.
<b>Root Mean Squared Error</b>	0.049396	0.048034
<b>Mean Absolute Error</b>	0.039283	0.039433
<b>Mean Absolute Percent Error</b>	0.099288	0.098643

Результаты оказываются весьма близкими.

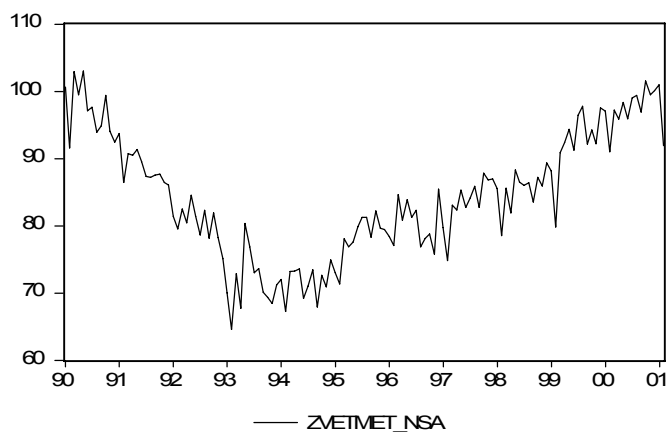
Сведем полученные результаты в отношении ряда интенсивности промышленного производства в одну таблицу.

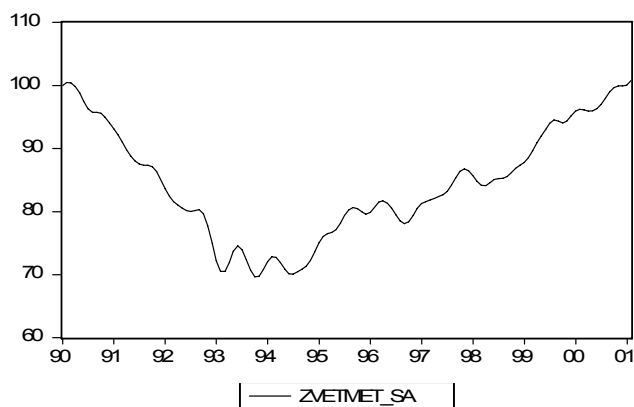
	DS рекурс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	0.049396	0.048034	0.048401	0.050764
<b>Mean Absolute Error</b>	0.039283	0.039433	0.039698	0.039615
<b>Mean Absolute Percent Error</b>	0.099288	0.098643	0.100229	0.098420

Ни одна из моделей не имеет определенного преимущества перед другими. Это согласуется с тем, что в краткосрочной перспективе (сглаженный) индекс промышленного производства имеет достаточно предсказуемую динамику и не содержит сезонных колебаний, которые можно было бы легко учитывать в эконометрических моделях временных рядов.

## 2.6. Индекс интенсивности производства цветных металлов

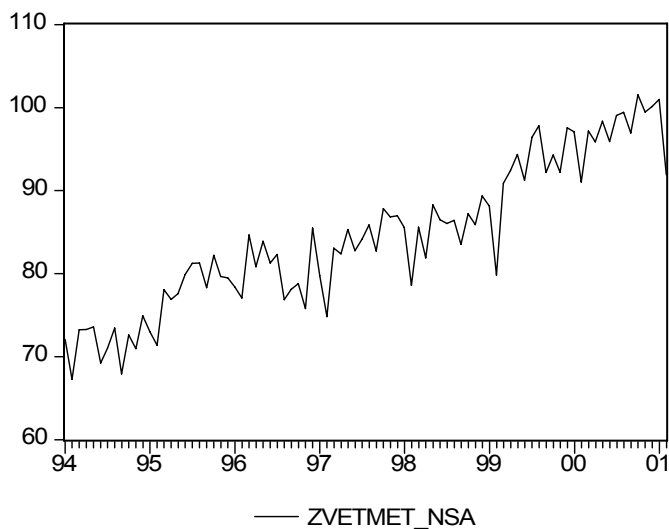
Индекс интенсивности производства цветных металлов был выбран в качестве примера для анализа индекса отраслевого промышленного производства. На всем временном интервале регистрации этого индекса графики поведения самого индекса и его сезонно скорректированного варианта имеют следующий вид:





Здесь мы выделим для рассмотрения период с 01.1994 до 02.2001. При этом мы избегаем проблем с идентификацией момента излома ряда и вместо этого проанализируем как сам ряд, так и его сезонно скорректированный вариант. Как и при исследовании предыдущих рядов, будем получать одношаговые прогнозы на один шаг вперед для семи будущих моментов времени. Поэтому первоначально построим авторегрессионные модели на временном интервале с 01.1994 до 07.2000.

Нескорректированный ряд



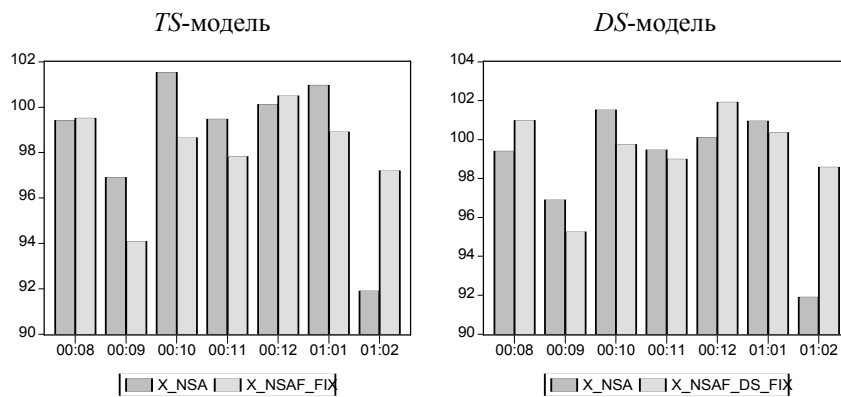
Оценивание модели с детерминированным линейным трендом дает следующий результат:

Dependent Variable: X NSA				
Method: Least Squares				
Sample: 1994:01 2000:07				
Included observations: 79				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	36.56804	15.62891	2.339770	0.0224
@TREND(1993:12)	0.161451	0.067919	2.377117	0.0204
X NSA(-1)	0.425699	0.117298	3.629225	0.0006
X NSA(-2)	0.003631	0.119202	0.030463	0.9758
X NSA(-3)	0.140500	0.115874	1.212525	0.2298
X NSA(-4)	0.021980	0.116278	0.189029	0.8507
X NSA(-5)	0.045544	0.115937	0.392835	0.6957
X NSA(-6)	-0.164483	0.117617	-1.398463	0.1668
X NSA(-7)	0.238075	0.116501	2.043543	0.0451
X NSA(-8)	-0.097152	0.113795	-0.853743	0.3964
X NSA(-9)	-0.059488	0.110920	-0.536313	0.5936
X NSA(-10)	-0.019746	0.110063	-0.179406	0.8582
X NSA(-11)	-0.105656	0.106267	-0.994247	0.3238
X NSA(-12)	0.370108	0.107821	3.432623	0.0011
X NSA(-13)	-0.317331	0.107602	-2.949130	0.0044
R-squared	0.911626	Mean dependent var		83.17016
Adjusted R-squared	0.892295	S.D. dependent var		8.083686
S.E. of regression	2.652942	Akaike info criterion		4.958398
Sum squared resid	450.4385	Schwarz criterion		5.408293
Log likelihood	-180.8567	F-statistic		47.15707
Durbin-Watson stat	2.053387	Prob(F-statistic)		0.000000

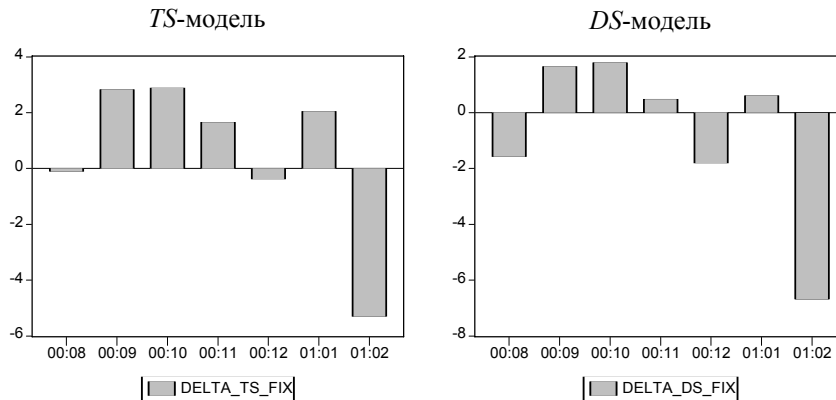
При оценивании модели в разностях получаем:

Dependent Variable: D(X NSA)				
Sample: 1994:01 2000:07				
Included observations: 79				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.970883	0.408373	2.377441	0.0203
X NSA DIF(-1)	-0.508794	0.116762	-4.357545	0.0000
X NSA DIF(-2)	-0.445117	0.130860	-3.401483	0.0011
X NSA DIF(-3)	-0.241818	0.136280	-1.774416	0.0806
X NSA DIF(-4)	-0.177286	0.136449	-1.299288	0.1984
X NSA DIF(-5)	-0.095401	0.134297	-0.710377	0.4800
X NSA DIF(-6)	-0.249181	0.132493	-1.880719	0.0644
X NSA DIF(-7)	0.031485	0.133050	0.236637	0.8137
X NSA DIF(-8)	-0.062970	0.128450	-0.490228	0.6256
X NSA DIF(-9)	-0.095828	0.125037	-0.766397	0.4462
X NSA DIF(-10)	-0.098018	0.122619	-0.799373	0.4269
X NSA DIF(-11)	-0.166679	0.116017	-1.436675	0.1555
X NSA DIF(-12)	0.259810	0.103544	2.509188	0.0146
R-squared	0.607039	Mean dependent var		0.352056
Adjusted R-squared	0.535592	S.D. dependent var		3.999653
S.E. of regression	2.725664	Akaike info criterion		4.992622
Sum squared resid	490.3302	Schwarz criterion		5.382531
Log likelihood	-184.2086	F-statistic		8.496303
Durbin-Watson stat	1.999610	Prob(F-statistic)		0.000000

Фиксируя эти две модели, получаем одношаговые прогнозы:



Ошибки прогнозов:

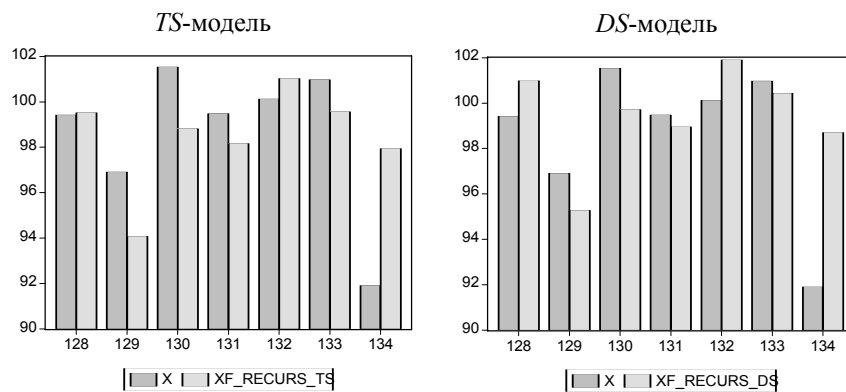


Характеристики прогнозов:

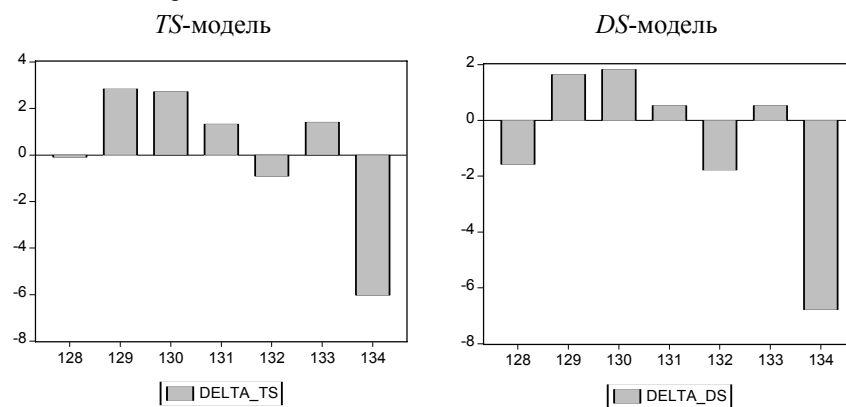
	TS	DS
<b>Root Mean Squared Error</b>	2.709401	2.850249
<b>Mean Absolute Error</b>	2.169326	2.082808
<b>Mean Absolute Percent Error</b>	2.241112	2.171102

Характеристики прогнозов по фиксированным моделям очень близки.

Приведем теперь результаты прогнозирования по рекурсивным *TS*- и *DS*-моделям.



Ошибки прогнозов:



Характеристики одношаговых прогнозов по рекурсивным моделям:

	DS рекурс.	TS рекурс.
<b>Root Mean Squared Error</b>	2.883328	2.835979
<b>Mean Absolute Error</b>	2.091729	2.187653
<b>Mean Absolute Percent Error</b>	2.181366	2.269576

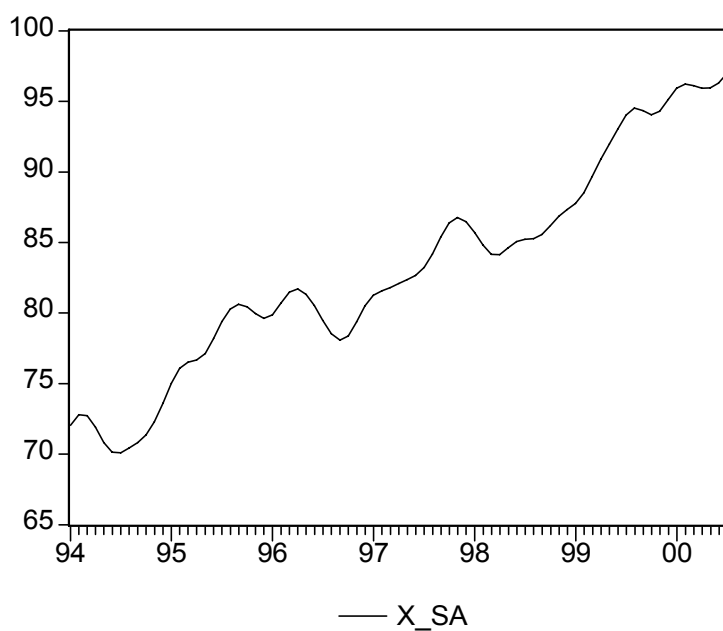


Сведем полученные результаты в отношении ряда ZVETMET\_NSA в одну таблицу.

	DS рекурс.	TS рекурс.	DS фикс.	TS фикс.
Root Mean Squared Error	2.883328	2.835979	2.850249	2.709401
Mean Absolute Error	2.091729	2.187653	2.082808	2.169326
Mean Absolute Percent Error	2.181366	2.269576	2.171102	2.241112

И среди *TS*-, и среди *DS*-моделей лучшие характеристики имеют прогнозы по фиксированным моделям. Прогноз по фиксированной *TS*-модели несколько лучше прогноза по фиксированной *DS*-модели с точки зрения RMSE, но несколько хуже последнего с точки зрения двух других характеристик. В целом все четыре прогноза имеют весьма близкие характеристики.

Сезонно скорректированный ряд



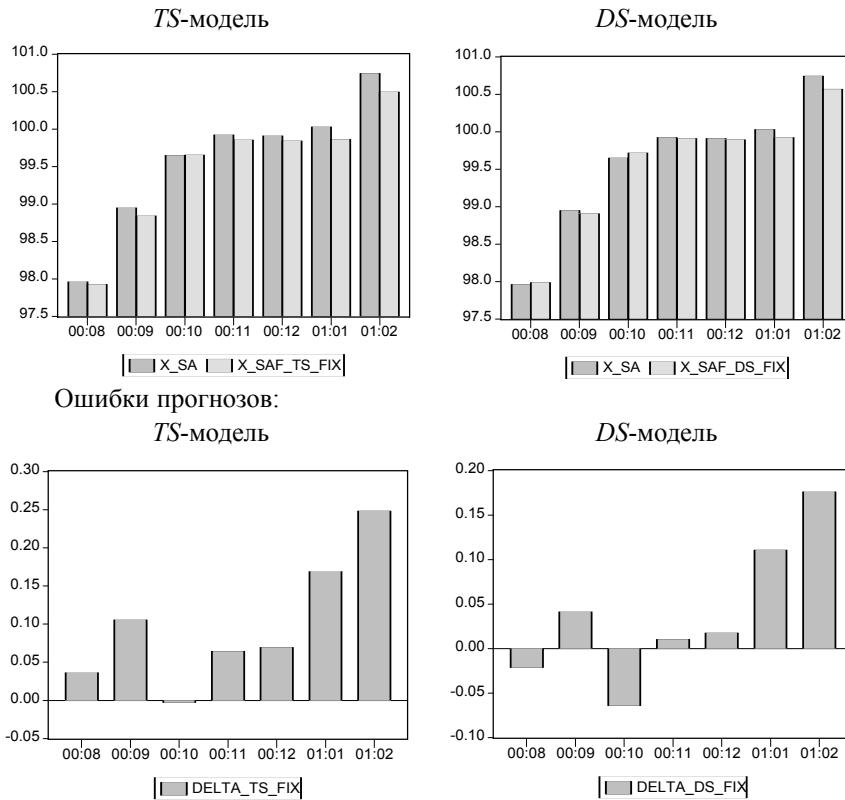
Оценивание модели с детерминированным линейным трендом дает следующий результат:

Dependent Variable: X SA				
Sample: 1994:01 2000:07				
Included observations: 79				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.889047	0.548959	3.441145	0.0010
@TREND( 1993:12)	0.008485	0.002484	3.415543	0.0011
X SA(-1)	3.631262	0.096718	37.54472	0.0000
X SA(-2)	-6.360191	0.332250	-19.14280	0.0000
X SA(-3)	6.989368	0.547534	12.76517	0.0000
X SA(-4)	-5.121163	0.541297	-9.460907	0.0000
X SA(-5)	2.396030	0.320370	7.478938	0.0000
X SA(-6)	-0.562043	0.090761	-6.192587	0.0000
R-squared	0.999833	Mean dependent var		83.16323
Adjusted R-squared	0.999817	S.D. dependent var		7.582406
S.E. of regression	0.102643	Akaike info criterion		-1.619359
Sum squared resid	0.748025	Schwarz criterion		-1.379414
Log likelihood	71.96466	F-statistic		60796.79
Durbin-Watson stat	1.901030	Prob(F-statistic)		0.000000

При оценивании модели в разностях получаем:

Dependent Variable: D( X SA)				
Sample: 1994:01 2000:07				
Included observations: 79				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.031516	0.016425	1.918845	0.0589
X_SA_DIF(-1)	2.714390	0.099284	27.33957	0.0000
X_SA_DIF(-2)	-3.824217	0.252220	-15.16224	0.0000
X_SA_DIF(-3)	3.322160	0.330006	10.06696	0.0000
X_SA_DIF(-4)	-1.812055	0.245999	-7.366110	0.0000
X_SA_DIF(-5)	0.513092	0.093295	5.499652	0.0000
R-squared	0.971546	Mean dependent var		0.331900
Adjusted R-squared	0.969598	S.D. dependent var		0.626785
S.E. of regression	0.109288	Akaike info criterion		-1.516751
Sum squared resid	0.871901	Schwarz criterion		-1.336793
Log likelihood	65.91168	F-statistic		498.5175
Durbin-Watson stat	1.791566	Prob(F-statistic)		0.000000

Фиксируя эти две модели, получаем одношаговые прогнозы:

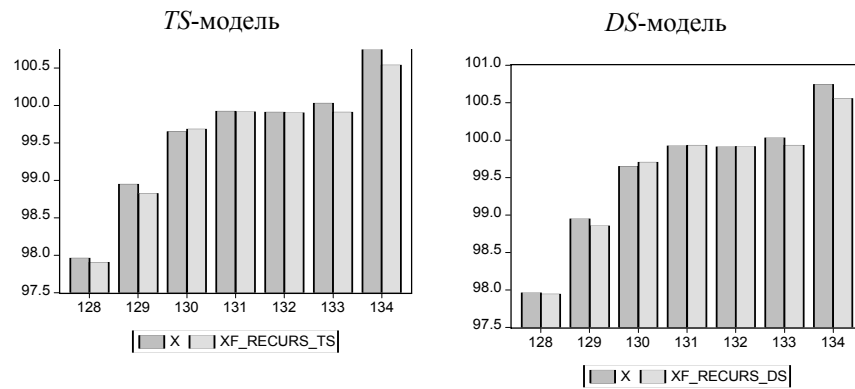


Характеристики прогнозов:

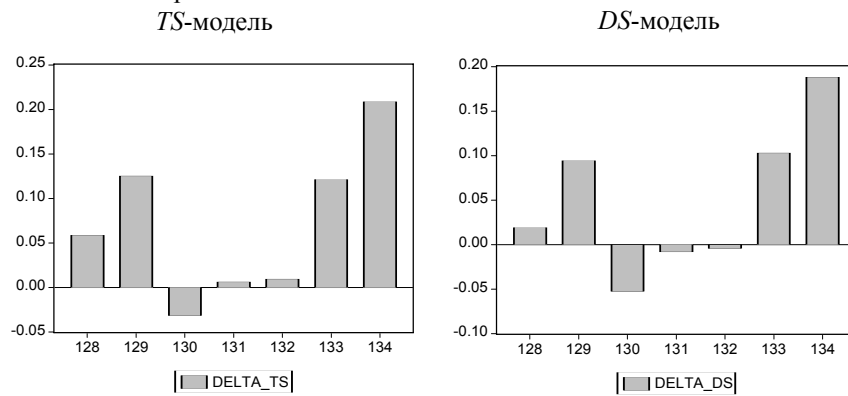
	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	0.084411	0.126271
<b>Mean Absolute Error</b>	0.063081	0.099413
<b>Mean Absolute Percent Error</b>	0.063052	0.099430

Здесь *DS*-модель оказалась лучшей по всем трем показателям. Отметим крайне малые по сравнению с моделью без сезонной коррекции значения MAPE, что, конечно, объясняется гладкостью траекторий сезонно скорректированного ряда и самой методикой построения сезонно скорректированного ряда, принимающей во внимание будущие значения «сырого» ряда.

Приведем теперь результаты прогнозирования по рекурсивным *TS*- и *DS*-моделям.



Ошибки прогнозов:



Характеристики одношаговых прогнозов по рекурсивным моделям:

	DS рекурс.	TS рекурс.
<b>Root Mean Squared Error</b>	0.090938	0.105791
<b>Mean Absolute Error</b>	0.066958	0.080089
<b>Mean Absolute Percent Error</b>	0.066982	0.080246

Сведем полученные результаты в отношении ряда ZVETMET\_SA в одну таблицу.

	DS ре-курс.	TS ре-курс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	0.090938	0.105791	0.084411	0.126271
<b>Mean Absolute Error</b>	0.066958	0.080089	0.063081	0.099413
<b>Mean Absolute Percent Error</b>	0.066982	0.080246	0.063052	0.099430

И среди фиксированных, и среди рекурсивных моделей лучшими являются *DS*-модели. Среди этих двух *DS*-моделей несколько лучшей оказалась фиксированная модель, хотя различие между характеристиками прогнозов очень мало.

## 2.7. Предпочтительность модели на временном интервале оценивания и качество прогнозов

Сделаем несколько замечаний относительно связи между предпочтительностью модели на временном интервале оценивания и качеством последовательности одношаговых прогнозов за пределами этого периода.

### Замечание 1

При сравнении *TS*- и *DS*-моделей мы, как и авторы работ, которые были упомянуты в главе 1, не занимались уточнением спецификации этих моделей в направлении исключения из них переменных со статистически не значимыми коэффициентами. И это имеет некоторую аргументацию. Поскольку мы анализируем качество прогнозов на один шаг вперед, полная модель, обладая наименьшей остаточной суммой квадратов по сравнению с редуцированными моделями, полученными на ее основе, имеет наилучшие показатели качества одношаговых прогнозов в пределах периода, на котором эта модель оценивалась, по крайней мере, в отношении RMSE.

Вместе с тем указанное преимущество полной модели никак не гарантирует того, что полная модель обязательно даст лучшее качество одношаговых прогнозов при выходе за пределы интервала, на котором модель оценивалась. Покажем это на примере.

При анализе одношаговых прогнозов для ряда M1 по *DS*-модели с аддитивным выбросом, датированным 02.1999, мы получили для использованной там фиксированной *DS*-модели следующие показатели качества одношаговых прогнозов, соответствующих периоду с 08.2000 по 02.2001.

RMSE=30788.31, MAE=25004.23, MAPE=3.119278.

В то же время, если исключить из этой модели составляющие  $AR(2), \dots, AR(7)$ , то показатели качества получаемой при этом редуцированной модели равны, соответственно,

$$RMSE=30337.52, MAE=24768.65, MAPE=3.090592.$$

Таким образом, показатели качества редуцированной модели оказываются лучшими по сравнению с аналогичными показателями для полной модели.

### Замечание 2

Указанная в предыдущем замечании редуцированная модель одновременно оказывается лучше полной модели и по информационному критерию Шварца: для нее значение этого критерия равно  $SIC=21.923$ , тогда как для полной модели  $SIC=22.354$ . Однако на основании этого все же нельзя делать выводы об обязательной предпочтительности для прогнозирования моделей с меньшими значениями  $SIC$ .

Если из той же полной модели исключить только составляющую  $AR(7)$ , то получим следующие показатели качества прогнозов на один шаг вперед по таким образом редуцированной модели:

$$RMSE=30897.48, MAE=25365.34, MAPE=3.168368.$$

Все три показателя у редуцированной модели хуже, чем у полной, хотя для редуцированной модели значение  $SIC=22.277$  меньше, чем для полной модели (22.354).

В качестве варианта  $DS$ -модели для ряда  $M1$ , содержащего в качестве объясняющих совершенно другие переменные, рассмотрим оцененную модель.

Dependent Variable: D( X)				
Method: Least Squares				
Sample(adjusted): 1995:07 2000:07				
Included observations: 61 after adjusting endpoints				
Convergence achieved after 5 iterations				
Backcast: 1994:10 1995:06				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4048.204	1086.735	3.725107	0.0005
@SEAS( 12)	26585.19	4198.554	6.331988	0.0000
@SEAS( 1)	-21519.15	4231.340	-5.085658	0.0000
T3	1037.048	130.7445	7.931870	0.0000
MA( 9)	-0.417871	0.124729	-3.350225	0.0015
R-squared	0.713867	Mean dependent var		9607.820
Adjusted R-squared	0.693429	S.D. dependent var		15969.54
S.E. of regression	8842.148	Akaike info criterion		21.09086
Sum squared resid	4.38E+09	Schwarz criterion		21.26388
Log likelihood	-638.2712	F-statistic		34.92834
Durbin-Watson stat	2.070271	Prob (F-statistic)		0.000000

Здесь @SEAS(1), @SEAS(12) – сезонные DUMMY, а T3 – переменная, равная 0 на временном интервале до 07.1998 включительно и равная t-42, начиная с 08.1998 (43-е наблюдение).

Сравним приведенные результаты для этой модели с результатами, полученными при оценивании фиксированной *DS*-модели для M1 с аддитивным выбором, рассмотренной в разделе 2.1.1, а именно:

R-squared	0.667957	Mean dependent var	10454.73
Adjusted R-squared	0.544627	S.D. dependent var	17387.54
S.E. of regression	11733.35	Akaike info criterion	21.81321
Sum squared resid	4.82E+09	Schwarz criterion	22.35373
Log likelihood	-520.4237	F-statistic	5.416000
Durbin-Watson stat	1.837172	Prob(F-statistic)	0.000032

Оцененная альтернативная модель имеет лучшие показатели ( $R^2$ , AIC, SIC). Сравним характеристики качества последовательности одношаговых прогнозов на временном интервале 08.2000–02.2001, получаемые по этим двум моделям.

	<b>DS альтернат.</b>	<b>DS аддитивн.</b>
<b>Root Mean Squared Error</b>	37314.6	30788.31
<b>Mean Absolute Error</b>	26437.34	25004.23
<b>Mean Absolute Percent Error</b>	3.270095	3.119278

Альтернативная модель дает худшие прогнозы по всем трем характеристикам.

### Замечание 3

Из предыдущих двух замечаний следует, что ни выбор по информационному критерию Шварца, ни выбор по коэффициенту детерминации  $R^2$  не гарантируют того, что в рамках *DS*- или *TS*-моделей выбранная по этим критериям из нескольких вариантов модель обязательно даст лучшие результаты прогнозов за пределами интервала, по которому эти модели оценивались.

С этой точки зрения выбор для сравнения качества прогнозов именно «полных» *TS*- и *DS*-моделей оправдан лишь с точки зрения определенности выбора модели.

## 2.8. Сравнение прогнозов, полученных по выбранным моделям, с «наивными» прогнозами

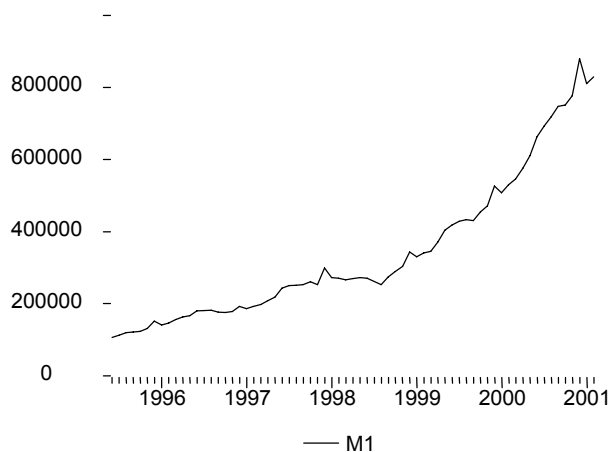
В разделах 2.1–2.6 мы провели сравнение прогнозов для соответствующих временных рядов по фиксированным и рекурсивным моделям в уровнях и в разностях.

При этом для наилучших среди рассмотренных по каждому ряду альтернативных вариантов были получены следующие средние процентные ошибки (MAPE):

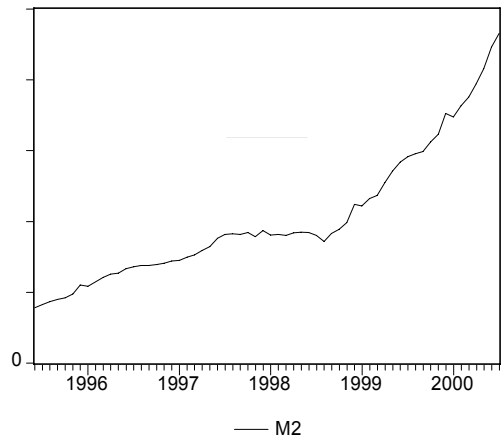
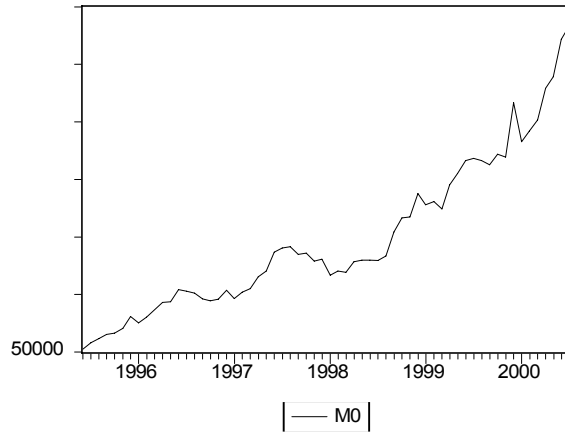
Ряд	MAPE
M1	3.119
M0	3.748
M2	2.814
Экспорт	7.363
Безработица	0.724
Индекс интенсивности промышленного производства	0.098
Индекс интенсивности производства цветных металлов (NSA)	2.171
Индекс интенсивности производства цветных металлов (SA)	0.063

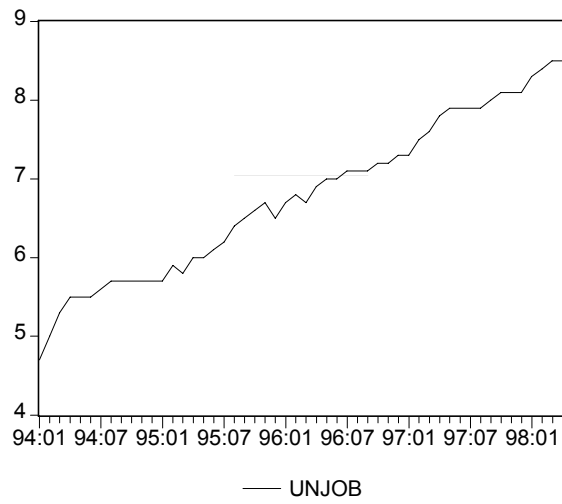
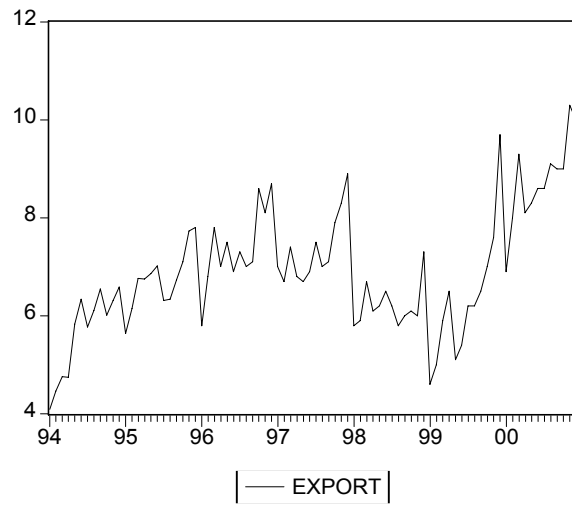
Если формально упорядочить список этих рядов по величине MAPE, то наихудший прогноз получен для ряда «Экспорт», а наилучший – для ряда «Индекс интенсивности производства цветных металлов (SA)».

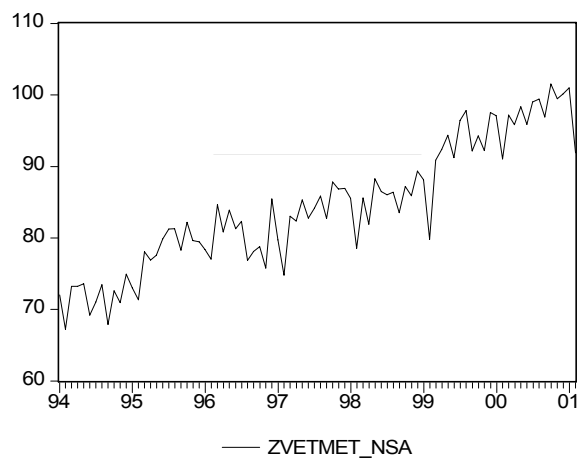
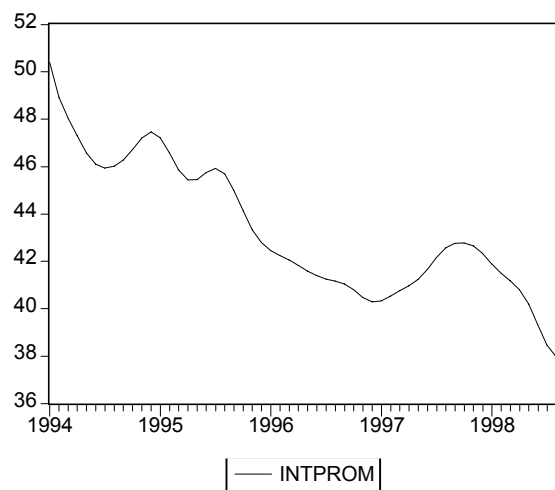
Вспомним, однако, общий характер поведения данных рядов на рассмотренных временных интервалах:

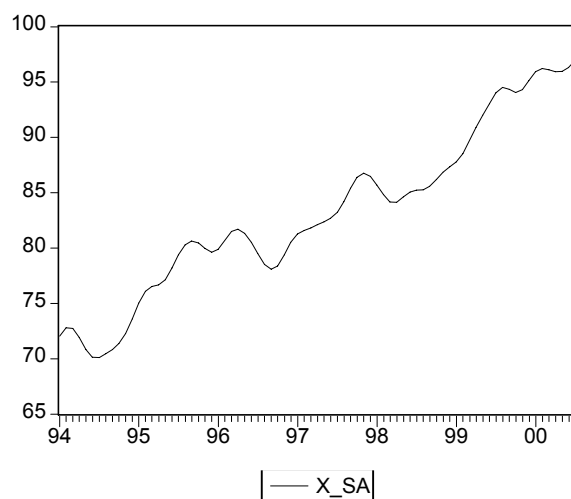












Все эти ряды, за исключением ряда «Экспорт», имеют на рассмотренных временных интервалах выраженный линейный тренд. Поэтому представляется естественным сравнить результаты прогнозирования по выбранным выше моделям, для которых приведены соответствующие значения MAPE, с результатами «наивных» прогнозов, строящихся по линейному тренду, выделенному из данных на временном периоде оценивания или на части этого периода (если за весь период оценивания происходит заметное изменение наклона тренда – см. денежные ряды).

Для рядов с изломом детерминированного тренда используем для выделения линейного тренда интервалы времени, следующие за моментом излома, указанным процедурой Перрона в модели с аддитивным выбросом.

Приведем результаты выделения линейного тренда по каждому из перечисленных рядов, получаемые экстраполяцией по моделям линейного тренда прогнозы и характеристики соответствующих последовательностей одношаговых прогнозов.

Ряд	Период оценивания	Период прогнозирования	MAPE
M1	03.1993-07.2000	08.2000-02.2001	7.770
M0	03.1993-07.2000	08.2000-02.2001	8.550
M2	03.1993-07.2000	08.2000-02.2001	5.912
Экспорт	01.1994-04.2000	05.2000-12.2000	20.166
Безработица	01.1994-09.1997	10.1997-04.1998	0.704
Интпром	01.1994-01.1998	02.1998-08.1998	2.471
Цветмет (NSA)	01.1994-07.2000	08.2000-02.2001	3.364
Цветмет (SA)	01.1994-07.2000	08.2000-02.2001	2.694

Для удобства значения MAPE для каждого ряда, полученные по модели линейного тренда и по ранее подобранной модели, сведем в одну таблицу.

Ряд	MAPE Selected model	MAPE Linear model
M1	3.119	7.770
M0	3.748	8.550
M2	2.814	5.912
Экспорт	7.363	20.166
Безработица	0.724	0.704
Индекс интенсивности Промышленного производства	0.098	2.471
Индекс интенсивности Производства цветных Металлов (NSA)	2.171	3.364
Индекс интенсивности Производства цветных Металлов (SA)	0.063	2.694

Из этой таблицы видно, что для денежных рядов средняя процентная ошибка прогноза по ранее подобранной модели в два с лишним раза меньше ошибки прогноза по оцененному линейному тренду. Для экспорта первая ошибка меньше в 2,7 раза, а по индексу производства цветных металлов (NSA) – примерно в 1,5 раза.

Для сезонно скорректированных рядов – индекса интенсивности промышленного производства и индекса интенсивности производства цветных металлов – наблюдается наиболее существенное снижение ошибки прогнозов по сравнению с прогнозом по линейному тренду: в 25,2 и 42,8 раза соответственно.

Наконец, для ряда, описывающего динамику безработицы, прогноз по оцененному линейному тренду оказался даже лучше, чем по подобранной ранее *TS*-модели:

$$X_t = 2.375 + 0.0296 * t + 0.539 * X_{t-1}.$$

И это несмотря на то, что на временном интервале оценивания для последней модели  $R^2_{\text{adj}} = 0.989$ ,  $AIC = -1.96$  и  $SIC = -1.84$ , тогда как для оцененной модели линейного тренда  $R^2_{\text{adj}} = 0.981$ ,  $AIC = -1.31$  и  $SIC = -1.23$ , т.е. на временном интервале оценивания модель линейного тренда хуже по всем показателям. Это еще раз подтверждает положение, которое было сформулировано ранее в замечании 3 раздела 2.7.

### Глава 3. Влияние на сравнительное качество последовательностей прогнозов длины интервала, на котором это сравнение производится

Данная глава по своей структуре в значительной степени повторяет предыдущую; существенным отличием является то, что те же самые методы, которые были использованы для прогнозирования выше, применяются для проверки качества прогнозирования на более длительный интервал (12–15 месяцев). Это необходимо для того, чтобы понять, сохраняются ли отмеченные выше тенденции (в отношении сравнительного качества прогнозов по альтернативным моделям) при расширении интервала, для которого строится последовательность одношаговых прогнозов.

#### 3.1. Денежные ряды

##### 3.1.1. MO

В качестве расширенного интервала берется интервал **08.2000–11.2001**.  
 Модели с аддитивным выбросом:  
 прогнозы за период 08.2000–02.2001 (7 точек):

	DS ре-курс.	TS ре-курс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	18103.33	16733.83	20762.60	18035.41
<b>Mean Absolute Error</b>	13663.46	13698.90	19084.87	16031.46
<b>Mean Absolute Percent Error</b>	3.737642	3.747511	5.247909	4.371771

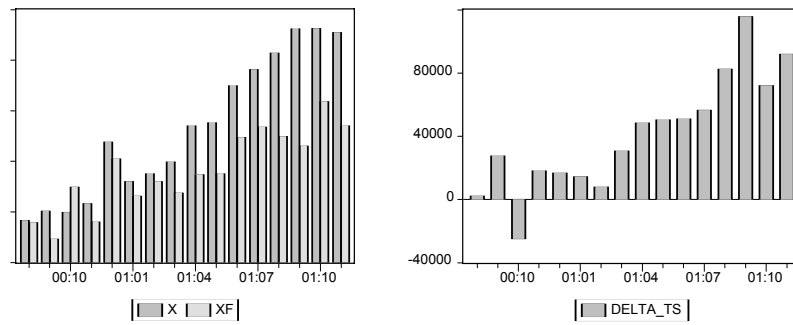
прогнозы за период 08.2000–11.2001 (16 точек):

	DS ре-курс.	TS ре-курс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	17422.22	21420.81	24404.03	54664.95
<b>Mean Absolute Error</b>	13679.97	18606.90	21787.85	44497.54
<b>Mean Absolute Percent Error</b>	3.309110	4.320607	5.128878	9.520612

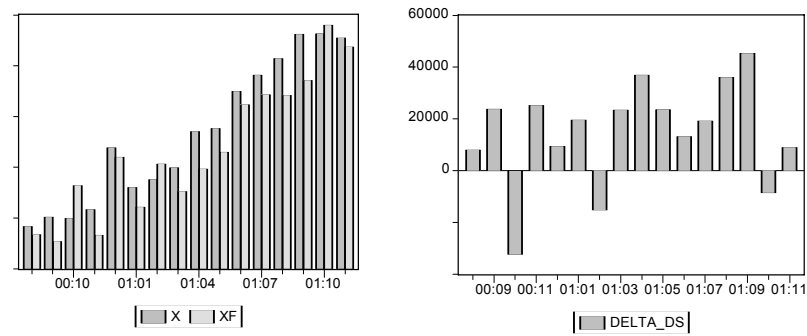
Обращает на себя внимание существенное ухудшение качества прогнозов по фиксированной *TS*-модели при незначительном ухудшении (и даже улучшении – по MAPE) качества прогнозов по фиксированной *DS*-модели.

Графически это выражается следующим образом.

Для фиксированной *TS*-модели:

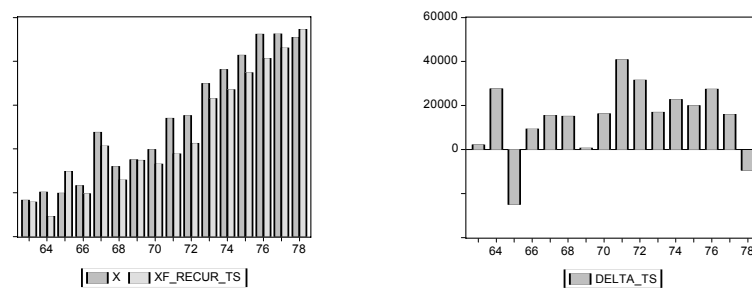


Для фиксированной *DS*-модели:



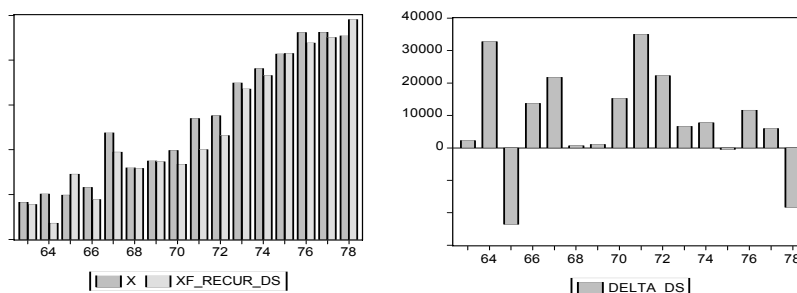
Что касается рекурсивных моделей, то для *TS*-моделей качество прогнозов ухудшается по всем трем характеристикам, а для *DS*-моделей оно даже улучшается по MAPE.

Рекурсивная *TS*-модель:





Рекурсивная *DS*-модель:



Инновационный выброс:

прогнозы за период 08.2000–02.2001 (7 точек):

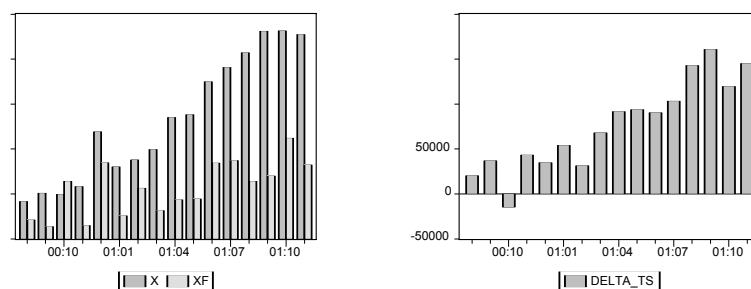
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	23070.19	24831.23	23660.00	35901.47
<b>Mean Absolute Error</b>	21966.45	23946.98	21726.06	33681.18
<b>Mean Absolute Percent Error</b>	6.002693	6.570231	5.966785	9.064165

прогнозы за период 08.2000–11.2001 (16 точек):

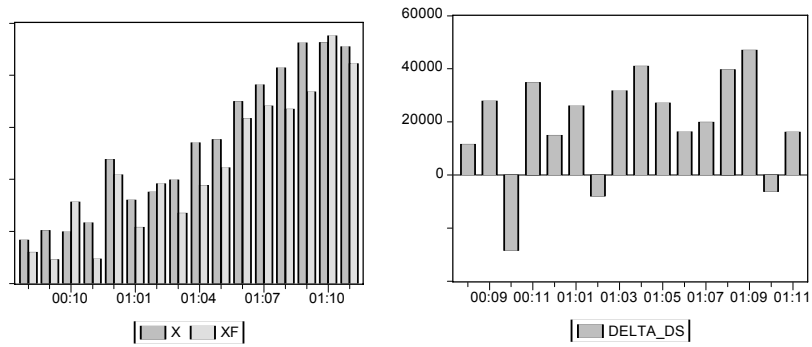
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	19346.11	20645.95	27501.13	90596.73
<b>Mean Absolute Error</b>	16948.18	17302.25	24844.06	78204.61
<b>Mean Absolute Percent Error</b>	4.204860	4.362469	5.861262	16.96909

Здесь также наблюдается существенное ухудшение качества прогнозов по фиксированной *TS*-модели при незначительном ухудшении (и даже улучшении – по MAPE) качества прогнозов по фиксированной *DS*-модели. Графически это выражается следующим образом.

Для фиксированной *TS*-модели:

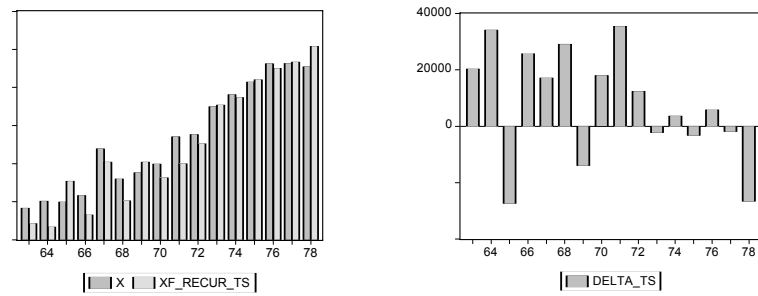


Для фиксированной *DS*-модели:

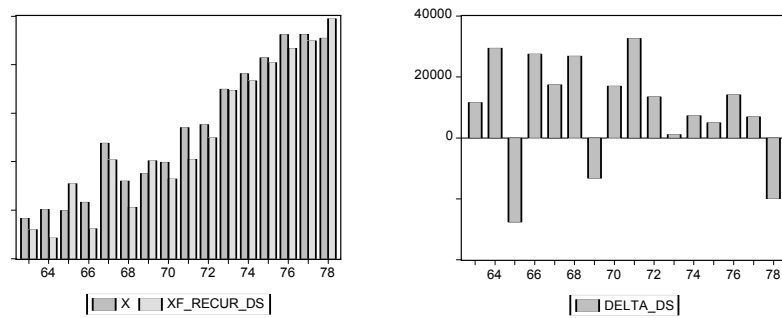


Что касается рекурсивных моделей, то и для *TS*-, и для *DS*-моделей все три характеристики качества прогнозов улучшаются. Несколько лучшее качество остается у прогнозов, построенных по рекурсивной *DS*-модели.

Рекурсивная *TS*-модель:



Рекурсивная *DS*-модель:



Сведем вместе результаты, полученные для расширенного интервала прогнозирования.

Аддитивный выброс:

	DS ре-курс.	TS ре-курс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	17422.22	21420.81	24404.03	54664.95
<b>Mean Absolute Error</b>	13679.97	18606.90	21787.85	44497.54
<b>Mean Absolute Percent Error</b>	3.309110	4.320607	5.128878	9.520612

Инновационный выброс:

	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	19346.11	20645.95	27501.13	90596.73
<b>Mean Absolute Error</b>	16948.18	17302.25	24844.06	78204.61
<b>Mean Absolute Percent Error</b>	4.204860	4.362469	5.861262	16.96909

Подведем итоги исследования прогнозов на один шаг вперед для ряда M0.

Интервал прогнозирования 08.2000–02.2001.

Среди фиксированных моделей с аддитивным выбросом лучшей по всем трем характеристикам точности прогнозов оказывается *TS*-модель. Среди рекурсивных моделей ни одна из двух не является более предпочтительной. При этом и среди *TS*-моделей, и среди *DS*-моделей рекурсивные модели оказались лучше фиксированных.

И среди фиксированных, и среди рекурсивных моделей с инновационным выбросом лучшими по всем трем характеристикам точности прогнозов оказываются *DS*-модели. Среди *TS*-моделей рекурсивная модель оказалась лучше фиксированной. Рекурсивные *TS*- и *DS*-модели дают близкие результаты.

Интервал прогнозирования 08.2000–11.2001.

И среди фиксированных и среди рекурсивных моделей для каждого из двух типов выбросов (инновационный или аддитивный) лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели, причем для фиксированных моделей такое преимущество трехкратное. Рекурсивные модели не дают определенного преимущества в классе *DS*-моделей; однако они обеспечивают значительно лучшее качество прогнозов в классе *TS*-моделей. Для обоих типов выбросов наилучшей оказалась рекурсивная *DS*-модель.

### 3.1.2. M1

В качестве расширенного интервала берется интервал **08.2000–11.2001**.

Модели с аддитивным выбросом:

прогнозы за период 08.2000–02.2001 (7 точек):

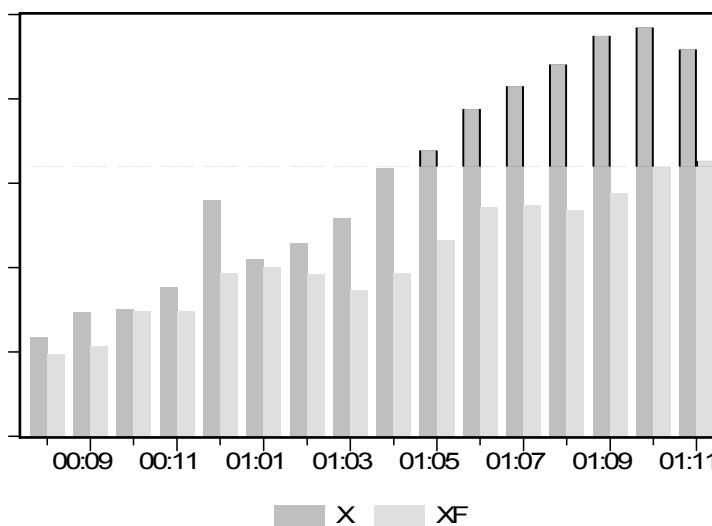
	DS рекурс.	TS ре-курс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	33750.29	35153.12	30788.31	40981.79
<b>Mean Absolute Error</b>	28915.77	29026.71	25004.23	32088.90
<b>Mean Absolute Percent Error</b>	3.604901	3.621122	3.119278	3.958528

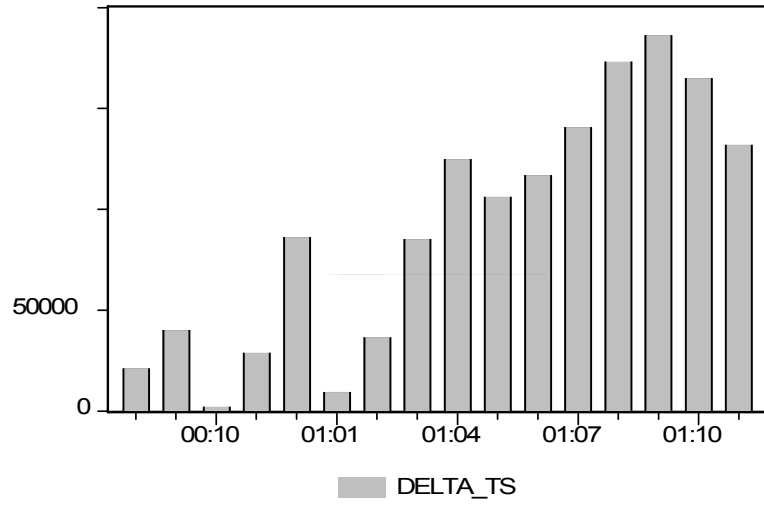
прогнозы за период 08.2000–11.2001 (16 точек):

	DS ре-курс.	TS ре-курс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	37646.23	41033.29	35359.73	108578.5
<b>Mean Absolute Error</b>	30672.38	33871.66	30334.93	90923.73
<b>Mean Absolute Percent Error</b>	3.440899	3.775183	3.362247	9.368012

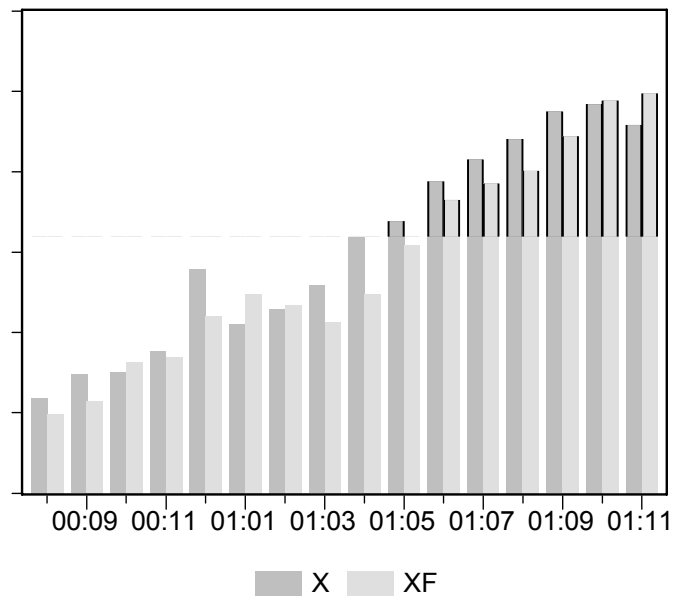
Обращает на себя внимание существенное ухудшение качества прогнозов по фиксированной *TS*-модели при весьма незначительном ухудшении качества прогнозов по фиксированной *DS*-модели. Графически это выражается следующим образом.

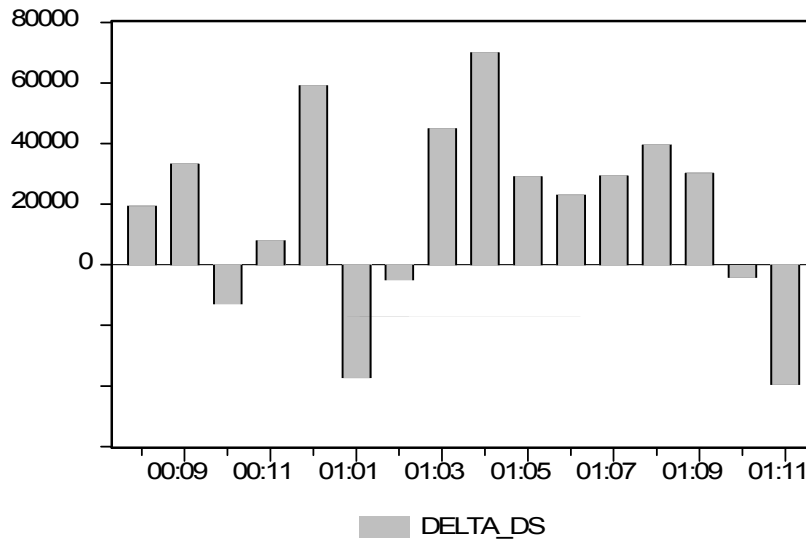
Для фиксированной *TS*-модели:



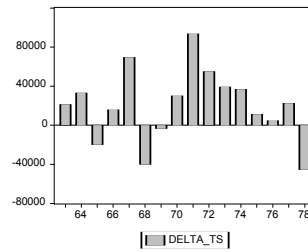
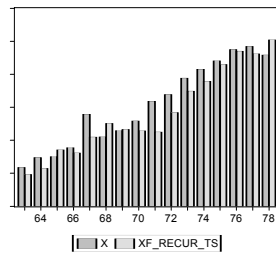


Для фиксированной DS-модели:

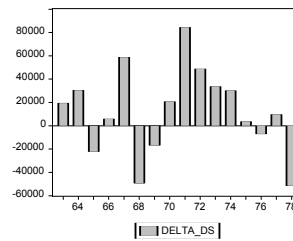
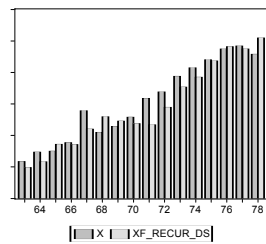




Среди рекурсивных моделей по MAPE несколько улучшаются результаты для *DS*-модели и несколько ухудшаются результаты по *TS*-модели.  
Для *TS*-модели:



Для *DS*-модели:



Инновационный выброс:  
 прогнозы за период 08.2000–02.2001 (7 точек):

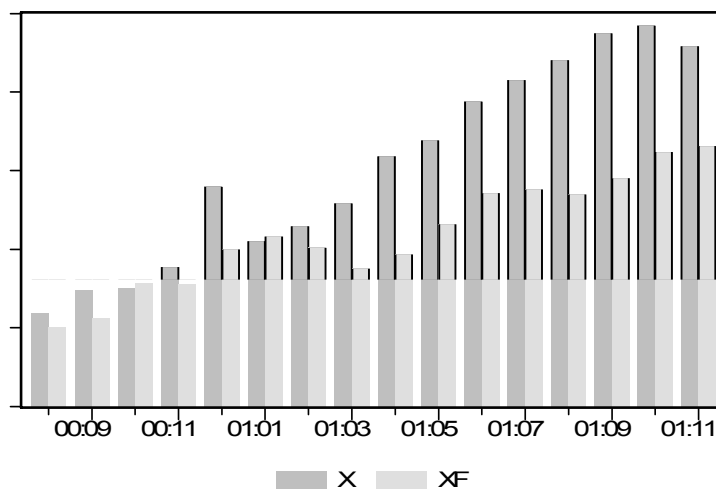
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	36796.05	36710.86	32677.59	36093.53
<b>Mean Absolute Error</b>	30477.43	31124.61	27062.91	27502.40
<b>Mean Absolute Percent Error</b>	3.793081	3.874630	3.374370	3.388481

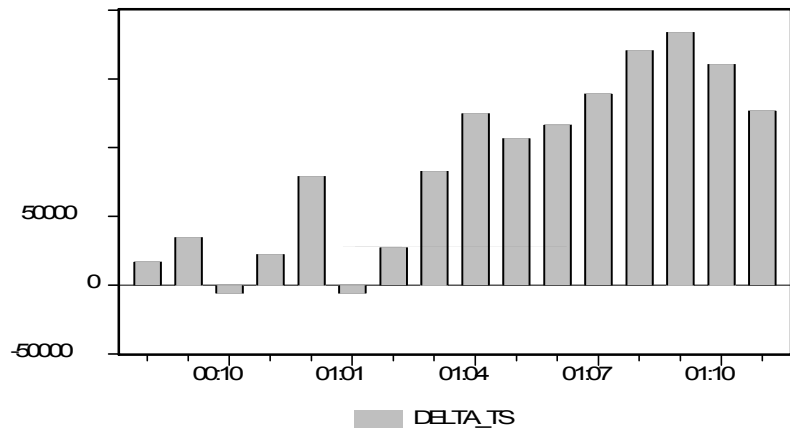
прогнозы за период 08.2000–11.2001 (16 точек):

	DS рекурс	TS рекурс	DS фикс	TS фикс
<b>Root Mean Squared Error</b>	38323.81	41215.39	33651.74	106400.4
<b>Mean Absolute Error</b>	28954.54	33153.51	27219.82	87877.65
<b>Mean Absolute Percent Error</b>	3.279413	3.721022	3.052948	9.018130

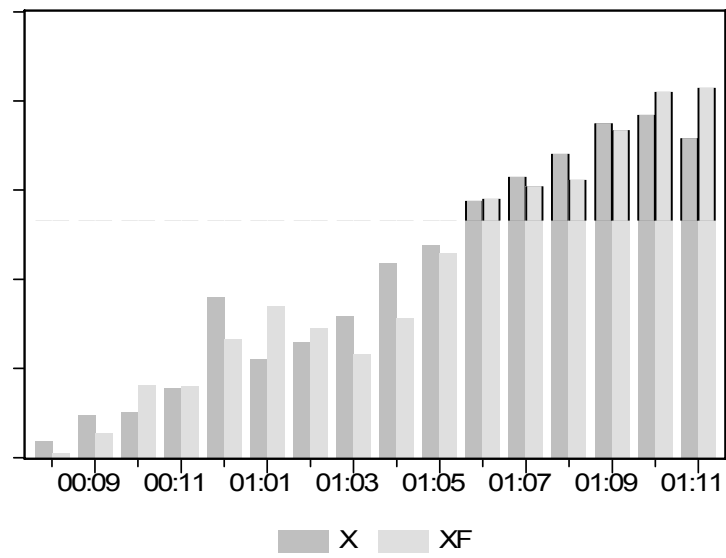
Как и в случае моделей с аддитивным выбросом, наблюдается существенное ухудшение качества прогнозов по фиксированной *TS*-модели, тогда как с точки зрения MAPE прогнозы по фиксированной *DS*-модели даже улучшаются. Графически это выражается следующим образом.

Для фиксированной *TS*-модели:

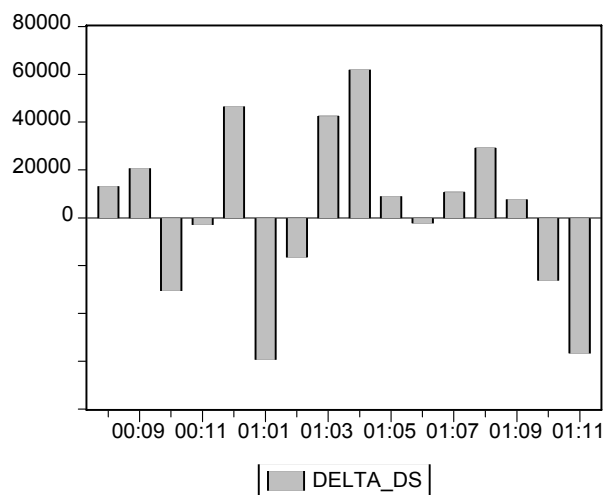




Для фиксированной DS-модели:

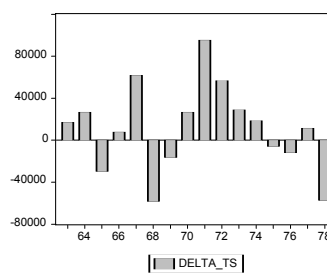
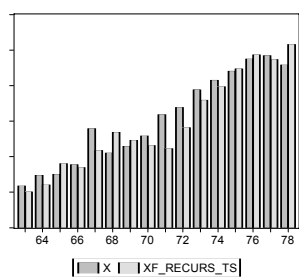




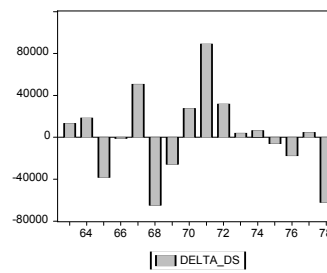
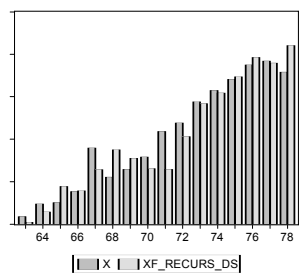


Что касается рекурсивных моделей, то и для *TS*-, и для *DS*-моделей значений MAPE уменьшается.

Для рекурсивной *TS*-модели:



Для рекурсивной *DS*-модели:



Сведем вместе результаты, полученные для расширенного интервала прогнозирования.

Аддитивный выброс:

	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	37646.23	41033.29	35359.73	108578.5
<b>Mean Absolute Error</b>	30672.38	33871.66	30334.93	90923.73
<b>Mean Absolute Percent Error</b>	3.440899	3.775183	3.362247	9.368012

Инновационный выброс:

	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	38323.81	41215.39	33651.74	106400.4
<b>Mean Absolute Error</b>	28954.54	33153.51	27219.82	87877.65
<b>Mean Absolute Percent Error</b>	3.279413	3.721022	3.052948	9.018130

Подведем итоги исследования прогнозов на один шаг вперед для ряда M1.

Интервал прогнозирования 08.2000–02.2001.

И среди фиксированных, и среди рекурсивных моделей для каждого из двух типов выбросов (инновационный или аддитивный) лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели. При этом рекурсивные модели не обязательно дают лучшие прогнозы по сравнению с фиксированными моделями. Для обоих типов выбросов наилучшей оказалась фиксированная *DS*-модель.

Интервал прогнозирования 08.2000–11.2001.

И среди фиксированных, и среди рекурсивных моделей для каждого из двух типов выбросов (инновационный или аддитивный) лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели, причем для фиксированных моделей такое преимущество трехкратное. Рекурсивные модели не дают определенного преимущества в классе *DS*-моделей; однако они обеспечивают значительно лучшее качество прогнозов в классе *TS*-моделей. Для обоих типов выбросов наилучшей оказалась фиксированная *DS*-модель.

### 3.1.3. M2

В качестве расширенного интервала опять берется интервал **08.2000–11.2001**.

Модели с аддитивным выбросом:  
прогнозы за период 08.2000–02.2001 (7 точек):

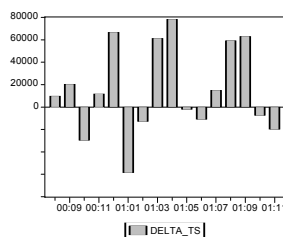
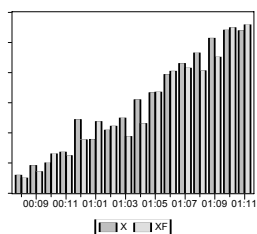
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	42689.06	43020.02	41472.50	36993.30
<b>Mean Absolute Error</b>	33642.68	35293.53	30513.84	29946.52
<b>Mean Absolute Percent Error</b>	3.125200	3.284899	2.847573	2.796299

прогнозы за период 08.2000–11.2001 (16 точек):

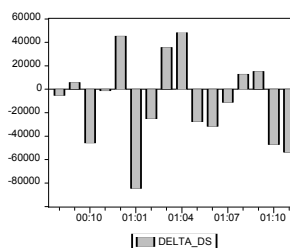
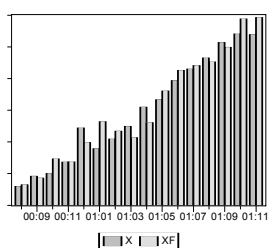
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	39133.66	41913.71	37951.93	41623.77
<b>Mean Absolute Error</b>	28768.55	31083.92	31062.16	32920.79
<b>Mean Absolute Percent Error</b>	2.457265	2.656437	2.598554	2.762772

В отличие от ситуации с прогнозированием рядов M1 и M0 здесь качество прогнозов по фиксированной *TS*-модели даже улучшается (по MAPE), как и качество (по MAPE) прогнозов по фиксированной *DS*-модели. Графически это выражается следующим образом.

Для фиксированной *TS*-модели:

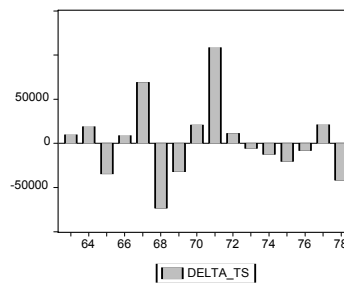
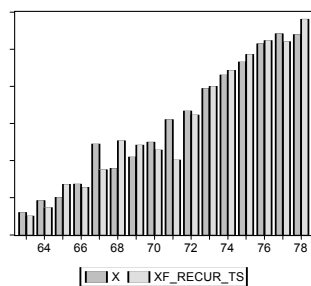


Для фиксированной *DS*-модели:

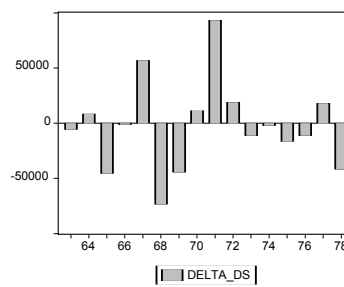
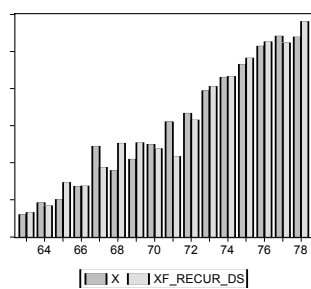


Что касается рекурсивных моделей, то и для *TS*-, и для *DS*-моделей происходит улучшение прогнозов по всем трем характеристикам.

Для *TS*-модели:



Для *DS*-модели:



Инновационный выброс:

прогнозы за период 08.2000–02.2001 (7 точек):

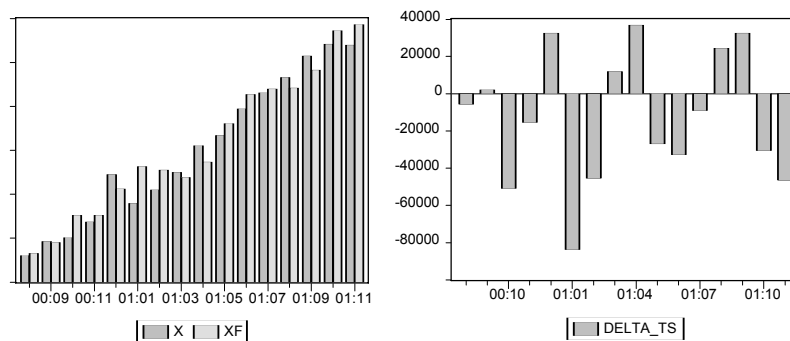
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	43295.36	42186.68	40870.17	43134.08
<b>Mean Absolute Error</b>	33643.49	32129.98	30168.47	33700.94
<b>Mean Absolute Percent Error</b>	3.123998	2.982067	2.814443	3.154151

прогнозы за период 08.2000–11.2001 (16 точек):

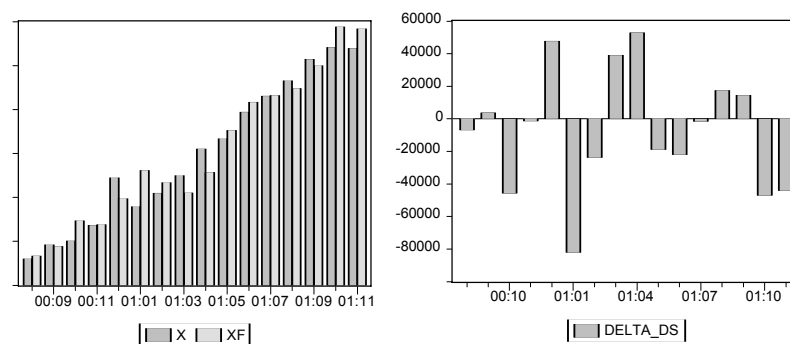
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	39331.72	43931.79	36671.29	36353.01
<b>Mean Absolute Error</b>	28245.74	35045.52	29290.59	30451.56
<b>Mean Absolute Percent Error</b>	2.428232	2.925264	2.465415	2.561185

Здесь наблюдается улучшение качества прогнозов (по всем трем характеристикам) и по фиксированной *TS*-модели, и по фиксированной *DS*-модели. Графически это выражается следующим образом.

Для фиксированной *TS*-модели:



Для фиксированной *DS*-модели:



Сведем вместе результаты, полученные для расширенного интервала прогнозирования.

Аддитивный выброс:

	DS рекурс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	39133.66	41913.71	37951.93	41623.77
<b>Mean Absolute Error</b>	28768.55	31083.92	31062.16	32920.79
<b>Mean Absolute Percent Error</b>	2.457265	2.656437	2.598554	2.762772

Инновационный выброс:

	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	39331.72	43931.79	36671.29	36353.01
<b>Mean Absolute Error</b>	28245.74	35045.52	29290.59	30451.56
<b>Mean Absolute Percent Error</b>	2.428232	2.925264	2.465415	2.561185

Подведем итоги исследования прогнозов на один шаг вперед для ряда M2.

Интервал прогнозирования 08.2000–02.2001.

Инновационный выброс:

Среди фиксированных моделей лучшей является *DS*-модель, а среди рекурсивных моделей – *TS*-модель. Среди *TS*-моделей лучше рекурсивная, а среди *DS*-моделей – фиксированная. Лучшей по всем трем показателям является фиксированная *DS*-модель.

Аддитивный выброс:

Среди фиксированных моделей лучшей является *TS*-модель, а среди рекурсивных моделей – *DS*-модель. И среди *TS*-, и среди *DS*-моделей лучшими являются фиксированные модели. Лучшей по всем трем показателям является фиксированная *TS*-модель.

Интервал прогнозирования 08.2000–11.2001.

И среди фиксированных, и среди рекурсивных моделей для каждого из двух типов выбросов (инновационный или аддитивный) несколько лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели; однако преимущество это незначительное. Рекурсивные модели имеют некоторое преимущество в классе *DS*-моделей; однако в классе *TS*-моделей этого не наблюдается. Для обоих типов выбросов наилучшей по двум последним показателям оказалась рекурсивная *DS*-модель.

### 3.2. Экспорт

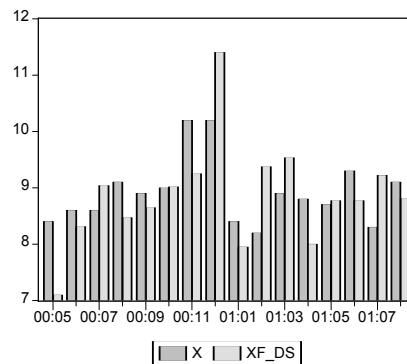
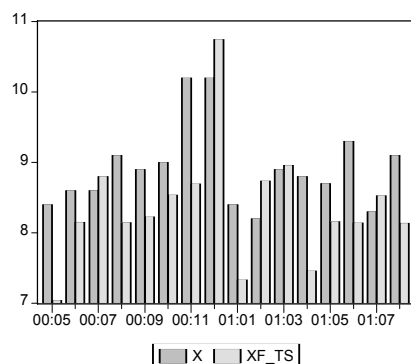
В качестве расширенного интервала берется интервал **05.2000–08.2001**.

Фиксированные модели

Прогнозы по приведенным в главе 2 оцененным моделям:

*TS*-модель

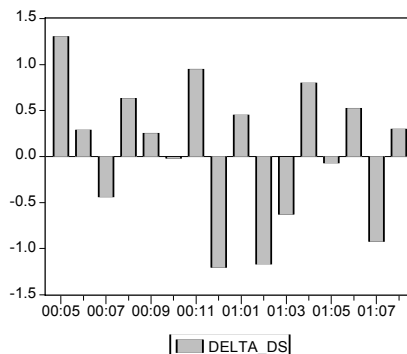
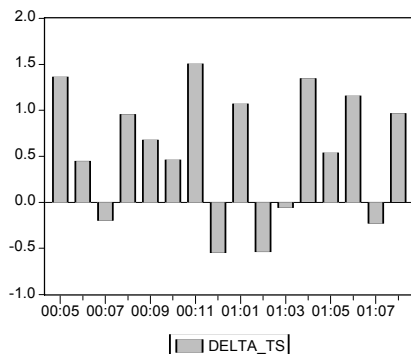
*DS*-модель



Ошибки прогнозов:

*TS*-модель

*DS*-модель



Характеристики прогнозов по подобранным фиксированным моделям (расширенный интервал):

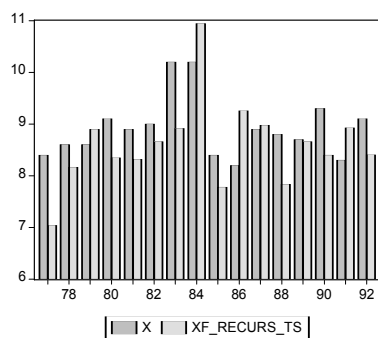
	<b>DS</b>	<b>TS</b>
<b>Root Mean Squared Error</b>	0.734464	0.869509
<b>Mean Absolute Error</b>	0.622717	0.753316
<b>Mean Absolute Percent Error</b>	6.989588	8.393028

На расширенном интервале *DS*-модель также оказывается предпочтительнее по всем трем показателям.

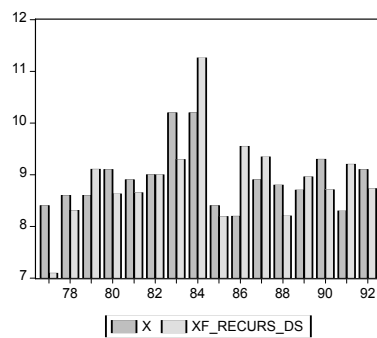
Рекурсивные модели

Прогнозы по рекурсивным моделям:

*TS*-модель

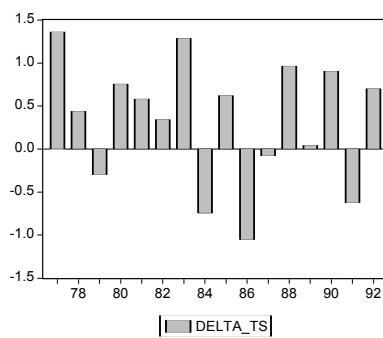


*DS*-модель

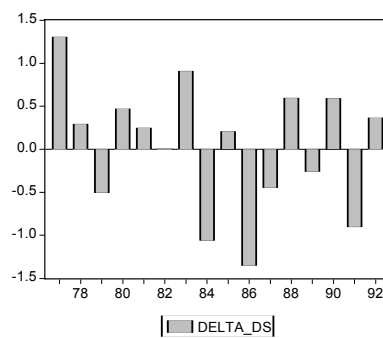


Ошибки прогнозов:

*TS*-модель



*DS*-модель



Характеристики прогнозов по подобранным рекурсивным моделям (расширенный интервал):

	<b>DS</b>	<b>TS</b>
<b>Root Mean Squared Error</b>	0.710022	0.770122
<b>Mean Absolute Error</b>	0.594670	0.673983
<b>Mean Absolute Percent Error</b>	6.693006	7.542889



На расширенном интервале *DS*-модель также оказывается предпочтительнее по всем трем показателям.

Сведем полученные результаты вместе.

Прогнозы на временном интервале 05.2000–12.2000 (8 точек):

	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	0.815693	0.901445	0.774216	0.883284
<b>Mean Absolute Error</b>	0.676787	0.794594	0.636348	0.769132
<b>Mean Absolute Percent Error</b>	7.362792	8.649441	6.889349	8.383397

прогнозы на временном интервале 05.2000–08.2001 (16 точек):

	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	0.710022	0.770122	0.734464	0.869509
<b>Mean Absolute Error</b>	0.594670	0.673983	0.622717	0.753316
<b>Mean Absolute Percent Error</b>	6.693006	7.542889	6.989588	8.393028

При переходе к расширенному интервалу характеристики прогнозов по рекурсивным моделям улучшаются, а характеристики прогнозов по фиксированным моделям практически не изменяются.

### 3.3. Налоговые доходы федерального бюджета

В качестве расширенного интервала берется интервал **06.2000–09.2001**.

Модели с аддитивным выбросом:

прогнозы за период 06.2000–12.2000 (7 точек):

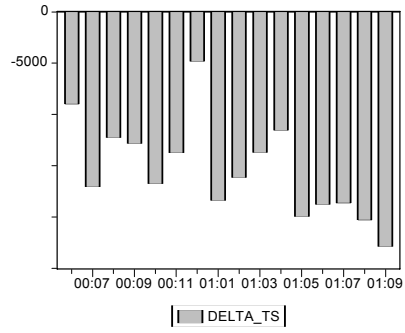
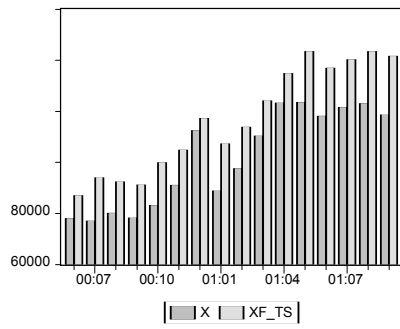
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	9044.562	9495.680	9646.584	12989.34
<b>Mean Absolute Error</b>	8003.194	8416.063	8902.694	12358.84
<b>Mean Absolute Percent Error</b>	9.648978	10.166833	10.90060	14.98790

прогнозы за период 06.2000–09.2001 (16 точек):

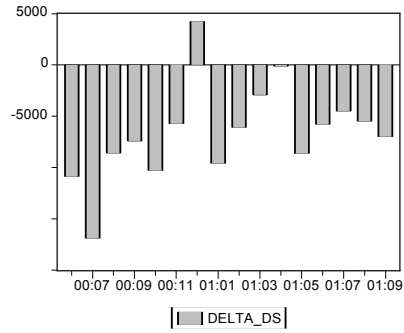
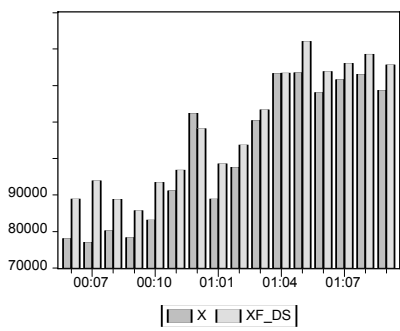
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	6676.117	6983.570	8038.587	16075.68
<b>Mean Absolute Error</b>	5097.416	5367.743	7134.696	15428.55
<b>Mean Absolute Percent Error</b>	5.740951	6.044299	7.761660	15.44266

Имеет место ухудшение качества прогнозов по фиксированной *TS*-модели при улучшении качества прогнозов по фиксированной *DS*-модели. Графически это выражается следующим образом.

Для фиксированной  $TS$ -модели:

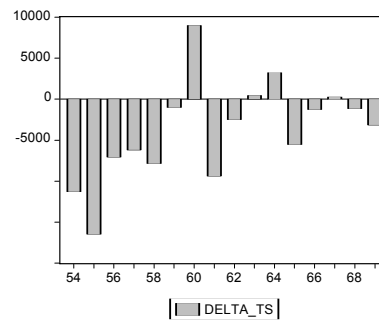
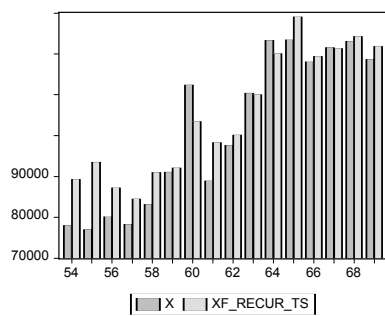


Для фиксированной  $DS$ -модели:

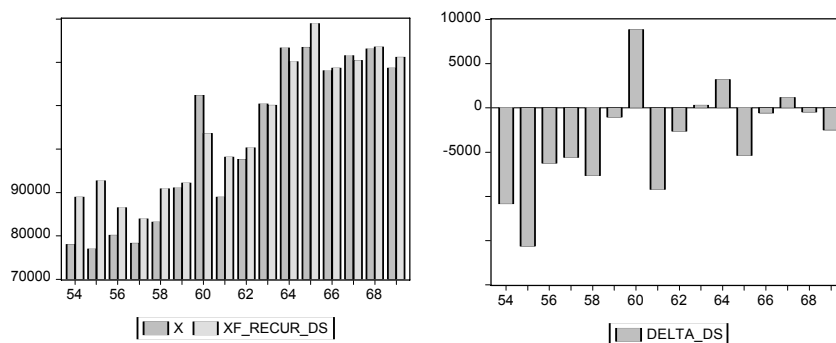


Для обеих рекурсивных моделей ( $TS$  и  $DS$ ) качество прогнозов существенно улучшается.

Для  $TS$ -модели:



Для *DS*-модели:



**Инновационный выброс:**

прогнозы на временном интервале 06.2000–12.2000 (7 точек):

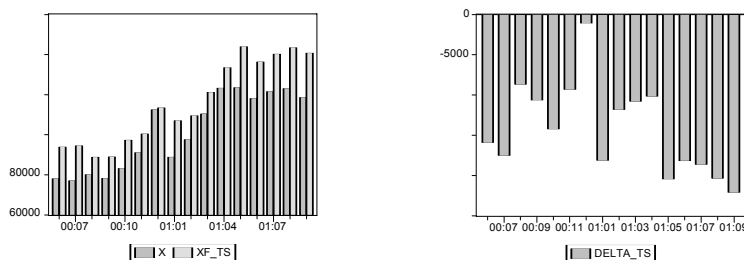
	<b>DS рекурс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	8656.144	9105.602	8459.231	12180.72
<b>Mean Absolute Error</b>	7019.826	7385.307	6809.506	11051.46
<b>Mean Absolute Percent Error</b>	8.095278	8.557180	7.994012	13.68804

прогнозы на временном интервале 06.2000–09.2001 (16 точек):

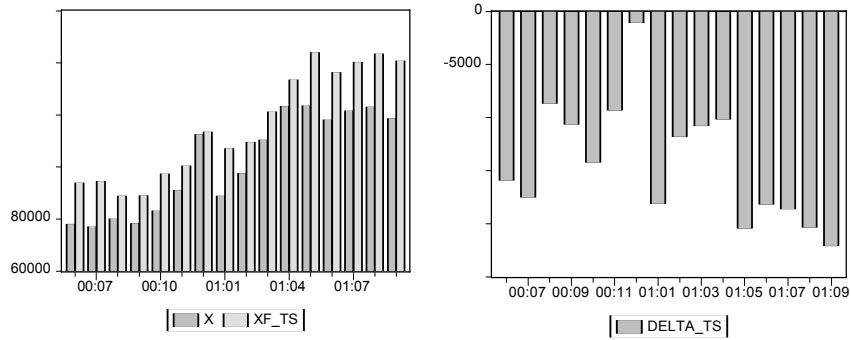
	<b>DS рекурс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	6554.468	6656.385	6620.289	15255.10
<b>Mean Absolute Error</b>	5033.279	4708.978	5168.779	14245.22
<b>Mean Absolute Percent Error</b>	5.365670	5.147208	5.485214	14.29462

Как и в случае моделей с аддитивным выбросом, наблюдается ухудшение качества прогнозов по фиксированной *TS*-модели, тогда как прогнозы по фиксированной *DS*-модели улучшаются. Графически это выражается следующим образом.

Для фиксированной *TS*-модели:

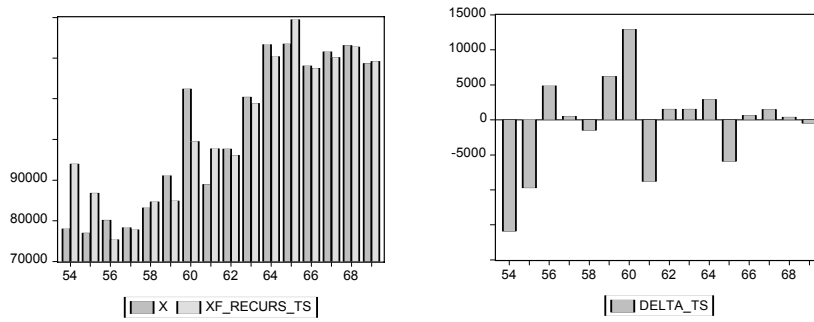


Для фиксированной *DS*-модели:

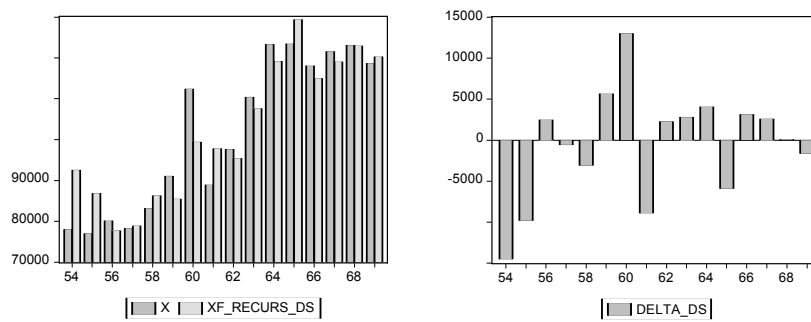


Что касается рекурсивных моделей, то и для *DS*-, и для *TS*-моделей качество прогнозов заметно улучшается.

Для рекурсивной *TS*-модели:



Для рекурсивной *DS*-модели:



Сведем вместе результаты, полученные для расширенного интервала прогнозирования.

Аддитивный выброс:

	<b>DS рекурс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	6676.117	6983.570	8038.587	16075.68
<b>Mean Absolute Error</b>	5097.416	5367.743	7134.696	15428.55
<b>Mean Absolute Percent Error</b>	5.740951	6.044299	7.761660	15.44266

Инновационный выброс:

	<b>DS ре-курс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	6554.468	6656.385	6620.289	15255.10
<b>Mean Absolute Error</b>	5033.279	4708.978	5168.779	14245.22
<b>Mean Absolute Percent Error</b>	5.365670	5.147208	5.485214	14.29462

Подведем итоги исследования прогнозов на один шаг вперед для налоговых доходов федерального бюджета.

Интервал прогнозирования 06.2000–12.2000.

И среди фиксированных, и среди рекурсивных моделей для каждого из двух типов выбросов (инновационный или аддитивный) лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели. При этом рекурсивные модели не обязательно дают лучшие прогнозы по сравнению с фиксированными моделями. Для обоих типов выбросов наилучшей оказалась фиксированная *DS*-модель.

Интервал прогнозирования 06.2000–09.2001.

Среди фиксированных моделей для каждого из двух типов выбросов (инновационный или аддитивный) лучшими по указанным характеристикам точности прогнозов оказываются *DS*-модели, причем такое преимущество более чем двукратное. Среди рекурсивных моделей в случае инновационного выброса некоторое преимущество имеет *TS*-модель, а в случае аддитивного выброса – *DS*-модель.

Рекурсивные модели имеют преимущество и в классе *DS*-моделей и в классе *TS*-моделей; при этом они обеспечивают более чем двукратное улучшение качества прогнозов в классе *TS*-моделей. Для аддитивного выброса наилучшей оказалась рекурсивная *DS*-модель, а для инновационного выброса чуть предпочтительнее выглядит рекурсивная *TS*-модель.

## **Общий вывод**

При расчете характеристик качества прогнозов по 7 (или 8 для временного ряда экспорта) последовательным прогнозам рекурсивные модели оказались лучшими только в 41,7% случаев, а при расчете характеристик качества прогнозов по 16 последовательным прогнозам рекурсивные модели оказались лучшими в 90,9% случаев. Лучшее качество прогнозирования рекурсивных моделей на более длительный период можно объяснить тем, что они дополнительно учитывают инновации и возможные изменения в структуре временного ряда и используют эту информацию для прогнозирования. Наличие такого рода изменений и структурных сдвигов может быть обусловлено многими причинами, главными среди которых являются трансформационные сдвиги в переходной экономике РФ, научно-технический и технологический прогресс, экзогенное влияние внешнеэкономических факторов и различных шоков и др.

## **Глава 4. Источники ошибок прогнозов и некоторые методы их коррекции**

В главе 1 уже указывалось на то, что экономические прогнозы, получаемые стандартными эконометрическими методами, часто оказываются довольно неточными, и это достаточно хорошо иллюстрируется результатами анализа нескольких российских макроэкономических рядов, проведенного в главах 2 и 3.

Анализ ошибок макроэкономических прогнозов и выявление их причин является достаточно сложной задачей, детальный подход к решению которой часто позволяет усовершенствовать методику прогнозирования и понять характеристики происходящих процессов. Этому вопросу посвящено довольно много работ различных авторов. В этой связи укажем, например, на анализ прогнозов по моделям для экономики Великобритании, проведенный в работах [WallisK.F., M.J.Andrews, D.N.F.Bell, P.G.Fisher, J.D.Whitley (1984)], [WallisK.F., M.J.Andrews, P.G.Fisher, J.Longbottom, J.D.Whitley (1986)], [WallisK.F. (1989)], [WallisK.F., J.D.Whitley (1991)], и на подобный анализ индустрии прогнозов в США, проведенный в работах [KleinL.R. (1991)] и [KleinL.R., E.Burmeister *eds* (1976)]. Детальному рассмотрению причин ошибок прогнозирования экономических временных рядов и теоретическому обоснованию некоторых эмпирических процедур, часто приводящих на практике к улучшению прогнозов, посвящены две недавно изданные книги [Clements M.P., D.F.Hendry (1998b)], [Clements M.P., D.F.Hendry (2001)].

### **4.1. Источники ошибок прогнозов**

В исторической перспективе, первоначально в центре внимания эконометристов были вопросы спецификации и оценивания моделей. Ожидалось, что «хорошие» модели, хорошо оцененные и хорошо протестированные, дают и хорошие прогнозы, тогда как «плохие» модели (например, не проходящие диагностические тесты) будут давать плохие прогнозы. Однако в действительности это оказалось не так: степень конгруэнтности или неконгруэнтности модели (т.е. степень соответствия модели реальным данным, по которым она построена) не является ни необходимым, ни достаточным условием успеха или неудачи прогноза, обращенного в будущее. Более того, прогнозы ряда по одному только его прошлому могут оказываться лучшими

по сравнению с прогнозами, построенными на основании структурных моделей.

При анализе прогностических процедур следует различать понятия «предсказуемости» временного ряда и «прогнозируемости» ряда.

Значения ряда  $X_t$  *не предсказуемы*, если знание его значений до момента  $T-1$  включительно ничего не прибавляет к знанию о распределении вероятностей значения  $X_T$ . Предсказуемость (непредсказуемость) – это свойство самого ряда, точнее, это свойство стохастического процесса, порождающего ряд, относительно информационного множества  $\Omega_{T-1}$ , включающего прошлые значения этого ряда.

Понятие прогнозируемости ряда более сложное: оно может быть связано с некоторой мерой потерь, и тогда прогнозируемость понимается как получение меньших потерь при использовании соответствующей процедуры прогнозирования по сравнению, скажем, с прогнозом по историческому среднему значений ряда.

Предсказуемость необходима, но не достаточна для прогнозируемости: например, может быть просто не известной форма зависимости условного среднего для  $X_t$  относительно  $\Omega_{T-1}$ .

Вместе с тем успех прогнозирования зависит от:

- наличия регулярностей в поведении ряда;
- информативности этих регулярностей относительно будущего;
- методов, использующих эти регулярности;
- исключения нерегулярностей, сводящих на нет пользу от наличия регулярностей.

Можно говорить о несостоятельности прогноза как такового, если качество прогнозов значительно ухудшается при переходе от периода, на котором модель оценивается, к будущим значениям ряда, и можно говорить о том, что прогноз плох по сравнению с некоторыми стандартными или конкурирующими моделями.

Можно также по-разному говорить о точности прогнозов. Можно понимать под точностью определенность прогноза (точный прогноз имеет малую неопределенность), но можно понимать под точностью прогноза и то, что он оказывается близким «в среднем» к истинным значениям.

Заметим, что, используя для прогнозирования стохастические модели, мы просто по этой причине практически никогда не можем абсолютно точно предсказать будущие значения ряда, даже если нам известна истинная модель порождения данных (процесс порождения данных, *data generating*



*process – DGP*), так что ошибка предсказания в таких моделях существует всегда, и эти ошибки накапливаются в процессе динамического прогнозирования на несколько периодов времени вперед. К этим ошибкам, проистекающим из самой природы процесса порождения данных, добавляются также ошибки, связанные с незнанием процесса порождения данных и, как следствие, вынужденным использованием для целей прогнозирования статистической модели (*statistical model – SM*), которая строится на основании той или иной экономической теории, более ранних аналогичных исследований и характера имеющихся в распоряжении статистических данных и, как правило, не совпадает с (истинным) процессом порождения данных.

Используемая статистическая модель может быть неправильно специфицирована в отношении входящих в нее переменных и в отношении динамических связей между переменными. Но даже если статистическая модель специфицирована правильно с этой точки зрения, так, что по форме она совпадает с процессом порождения данных, неопределенность в процессе прогнозирования вносит необходимость статистического оценивания коэффициентов модели по имеющимся статистическим данным: оцененные коэффициенты не совпадают с истинными.

Еще одним источником ошибок прогнозов является неправильное измерение значений, непосредственно используемых при прогнозировании. Так, последние (наиболее свежие) значения макроэкономических переменных часто впоследствии подвергаются пересмотру (ревизии), что легко заметить, сравнивая исторические данные о значениях макроэкономических рядов, опубликованные в одном и том же издании в различные годы.

Clements и Hendry (в цитируемых выше работах) анализируют относительное влияние на качество прогнозов каждого из перечисленных возможных источников ошибок прогнозов и приходят к выводу о том, что наиболее драматическое ухудшение качества прогнозов связано со скачкообразным изменением параметров DGP при переходе от периода, на котором модель подбирается и оценивается, к периоду, для которого прогноз собственно и составляется. Разумеется, при этом имеется в виду, что построенная статистическая модель не учитывает (не предполагает) такого изменения DGP.

## 4.2. Календарный эффект

Для переменных интервального типа некоторую роль в возникновении ошибок прогноза может играть неучет календарного фактора. Рассмотрим в этой связи влияние календарного фактора, учитывающего количество дней в месяце, на характеристики прогнозов денежных рядов.

Для учета календарного фактора дополним правые части использованных ранее фиксированных моделей с аддитивным выбросом переменной, значения которой равны количеству дней в месяце (31, 30, 28 или 29).

Для ряда M0:

	08.2000–02.2001		08.2000–11.2001	
	Без учета	С учетом	Без учета	С учетом
<b>RMSE</b>	20762.60	19773.54	24404.03	23987.22
<b>MAE</b>	19084.87	17246.75	21787.85	21443.12
<b>MAPE</b>	5.247909	4.757094	5.128878	5.00533

Хотя переменная, отражающая календарный эффект, имеет статистически незначимый коэффициент (Р-значение=0.2248), ее учет привел к некоторому, но не очень существенному улучшению характеристик прогнозов, вычисляемых и для короткого и для расширенного интервала.

Для ряда M1:

	08.2000–02.2001		08.2000–11.2001	
	Без учета	С учетом	Без учета	С учетом
<b>RMSE</b>	30788.31	30803.10	35259.73	35139.85
<b>MAE</b>	25004.23	25015.30	30334.93	30190.61
<b>MAPE</b>	3.119278	3.119104	3.362247	3.347867

Здесь коэффициент при календарной переменной также статистически незначим, но с существенно большим Р-значением (0.8588). При учете этой переменной характеристики прогнозов практически не изменяются.

Для ряда M2:

	08.2000–02.2001		08.2000–11.2001	
	Без учета	С учетом	Без учета	С учетом
<b>RMSE</b>	41472.50	41062.65	37951.93	38869.82
<b>MAE</b>	30513.84	30165.17	31062.16	31184.74
<b>MAPE</b>	2.847573	2.806430	2.598554	2.619281

Коэффициент при календарной переменной статистически незначим с Р-значением 0.83788. При учете этой переменной характеристики прогнозов незначительно улучшаются на коротком периоде и незначительно ухудшаются на расширенном периоде.

### 4.3. Коррекция прогнозов методами «back-on-track» и «back-on-average»

В этом разделе мы исследуем возможность улучшения качества прогнозов путем коррекции получаемых на каждом шаге прогнозов с использованием описываемого ниже (в 4.3.2) метода «back-on-track» и его модификации – метода «back-on-average».

#### 4.3.1. M0

Здесь мы ограничимся только моделями с аддитивным выбросом.

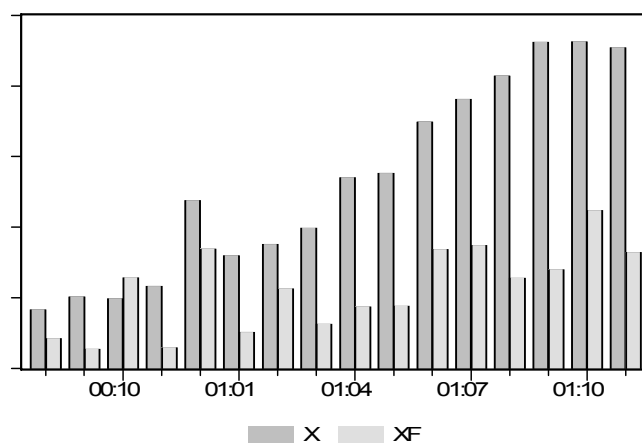
Прогнозы на временном интервале 08.2000–11.2001 (16 точек):

	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	17422.22	21420.81	24404.03	54664.95
<b>Mean Absolute Error</b>	13679.97	18606.90	21787.85	44497.54
<b>Mean Absolute Percent Error</b>	3.309110	4.320607	5.128878	9.520612

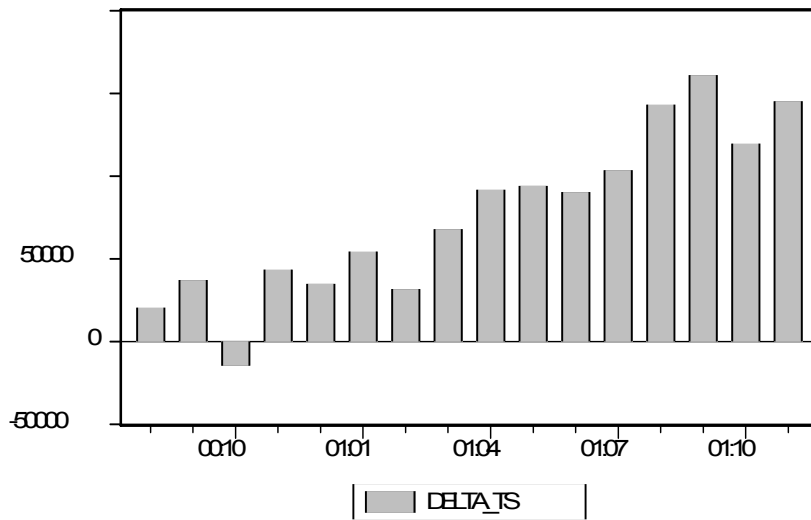
Наихудшие результаты дает прогнозирование по фиксированной *TS*-модели, приводящее к систематическому смещению прогнозов.

Для фиксированной *TS*-модели:

Наблюдения и прогнозы:



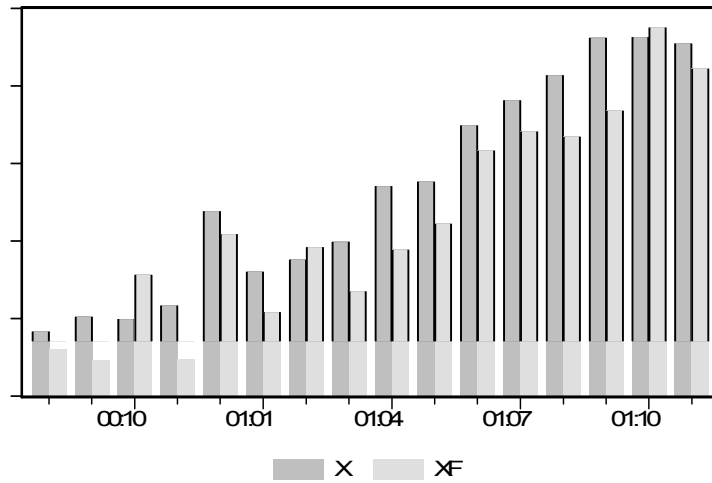
Ошибки прогнозов:



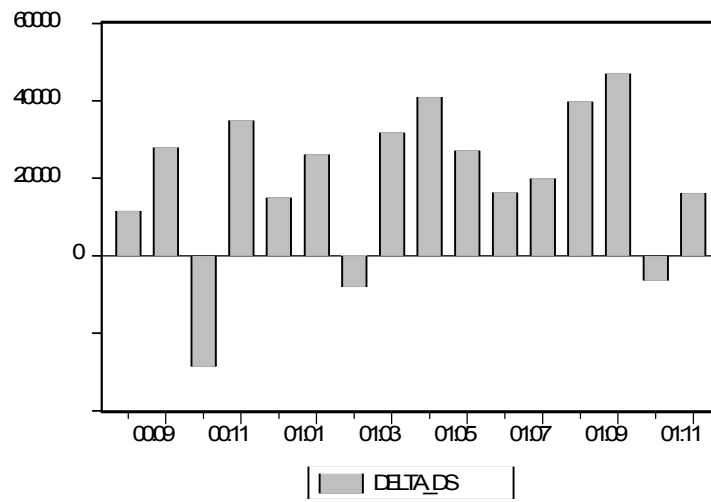
Смещение, хотя и не столь значительное, имеют и прогнозы по фиксированной *DS*-модели:

Для фиксированной *DS*-модели:

Наблюдения и прогнозы:



Ошибки прогнозов

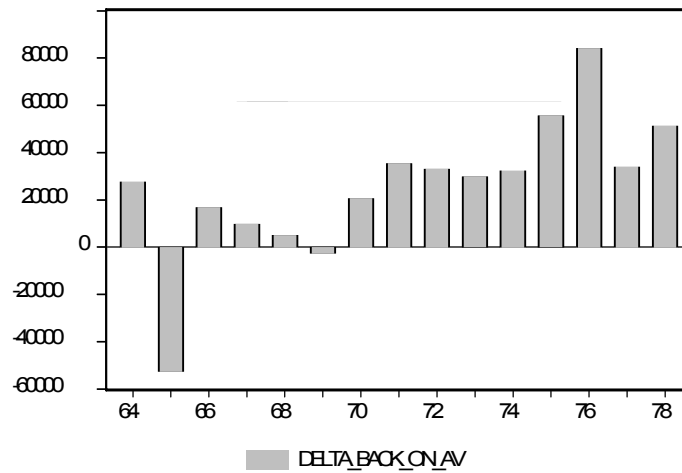


Проанализируем, что дает коррекция прогнозов по этим моделям методами «back-on-track» и «back-on-ave».

**TS-модель**

«Back-on-ave»:

Ошибки прогнозов

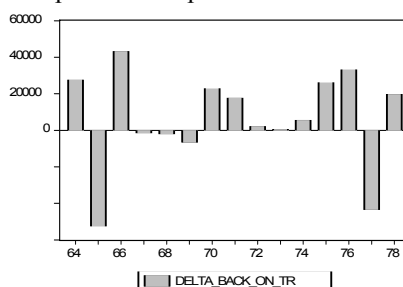


Характеристики прогнозов:

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	38846.15
<b>Mean Absolute Error</b>	32756.52
<b>Mean Absolute Percent Error</b>	7.247948

Скорректированные прогнозы улучшаются по сравнению с фиксированной *TS*-моделью, но оказываются хуже прогнозов по рекурсивной *TS*-модели. «Back-on-track»:

Ошибки скорректированных прогнозов



Характеристики скорректированных прогнозов

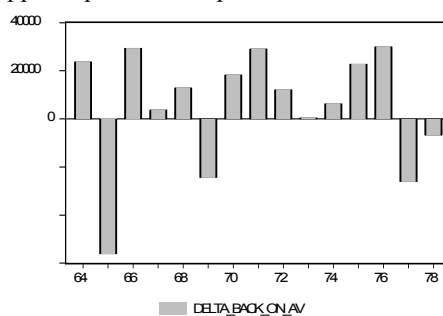
	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	26312.80
<b>Mean Absolute Error</b>	20334.04
<b>Mean Absolute Percent Error</b>	4.829495

Скорректированные прогнозы улучшаются по сравнению с фиксированной *TS*-моделью, лучше прогнозов, скорректированных методом «back-on-ave», но хуже прогнозов, полученных по рекурсивной *TS*-модели.

**DS-модель**

«Back-on-ave»:

Ошибки скорректированных прогнозов



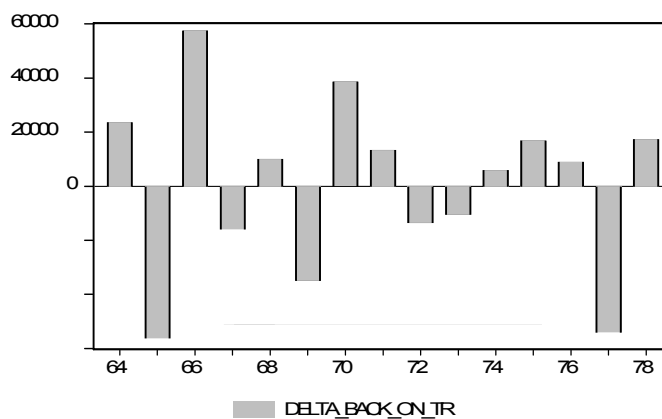
Характеристики скорректированных прогнозов

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	24368.25
<b>Mean Absolute Error</b>	20210.75
<b>Mean Absolute Percent Error</b>	4.904167

Характеристики скорректированных прогнозов несколько лучше, чем у прогнозов по фиксированной *DS*-модели, но хуже, чем у прогнозов по рекурсивной *DS*-модели.

«Back-on-track»:

Ошибки скорректированных прогнозов



Характеристики скорректированных прогнозов

	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	30762.63
<b>Mean Absolute Error</b>	25194.41
<b>Mean Absolute Percent Error</b>	6.142304

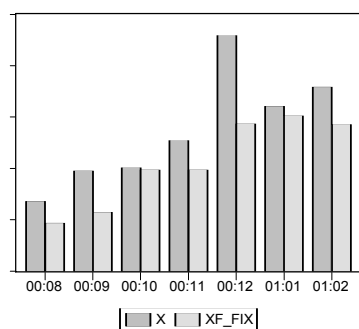
Характеристики прогнозов, скорректированных методом «back-on-track», хуже прогнозов, полученных по фиксированной и рекурсивной *DS*-моделям, а также прогнозов, скорректированных методом «back-on-ave».

### 4.3.2. M1

Модель в уровнях (*TS*-модель) с аддитивным выбросом

Напомним полученные в главе 2 результаты прогнозирования по оцененной фиксированной *TS*-модели:

Наблюдения и прогнозы



Ошибки прогнозов

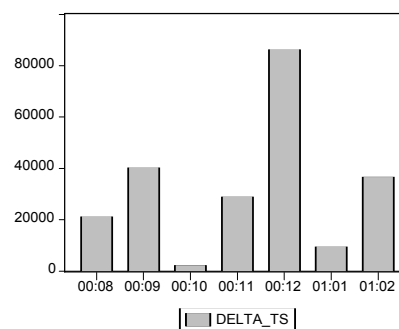
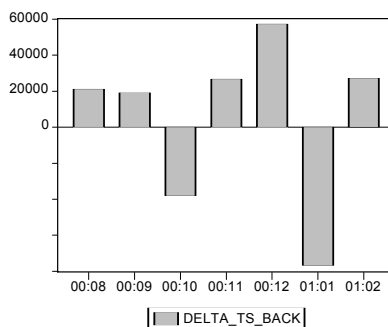


График ошибок прогноза указывает на явную смещенность прогнозов, полученных по использованной модели: прогнозные значения систематически недооценивают истинных значений ряда. В связи с этим проанализируем, что дало бы использование в процессе прогнозирования по указанной модели методики «back-on-track», исследовавшейся в работах [Clements P., D.F.Hendry (1998a)] и [Clements P., D.F.Hendry (1998b)]. Методика «back-on-track» состоит в коррекции каждого очередного одношагового прогноза, построенного по фиксированной модели, прибавлением ему наблюдаемой ошибки предыдущего прогноза, построенного по той же модели.

Следуя этой методике, получаем следующие результаты.

Ошибки скорректированных прогнозов:



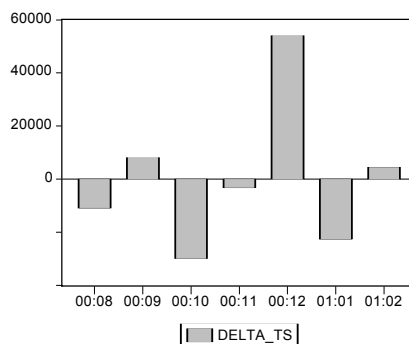


### Характеристики прогнозов

	Фиксир. модель	Back-on-track
<b>Root Mean Squared Error</b>	40981.79	42890.17
<b>Mean Absolute Error</b>	32088.90	38012.71
<b>Mean Absolute Percent Error</b>	3.958528	4.750199

Систематическое смещение прогнозов устранено, однако все три характеристики в таблице ухудшились.

Рассмотрим теперь гипотетическую ситуацию, когда нам в силу каких-то причин известно, что при переходе к интервалу прогнозирования произойдет скачкообразное повышение уровня ряда на 32088.9 (это есть среднее значение delta\_ts на временном интервале прогнозирования). При наличии такой информации для прогнозирования следует использовать модифицированную фиксированную модель с постоянной составляющей, увеличенной на 32088.9. Ошибки прогнозов по модифицированной модели таковы:



Здесь так же, как и при применении метода «back-on-track», систематическое смещение одношаговых прогнозов устраняется, однако на этот раз это происходит без существенного возрастания вариабельности, что отражается на характеристиках точности прогнозов:

	С однократной коррекцией константы в фиксированной модели
<b>Root Mean Squared Error</b>	25491.36
<b>Mean Absolute Error</b>	19057.50
<b>Mean Absolute Percent Error</b>	2.356272

Все три характеристики весьма заметно улучшились.

Таким образом, однократная коррекция константы в фиксированной модели дала значительно лучшие прогнозы по сравнению с исходной фикс-

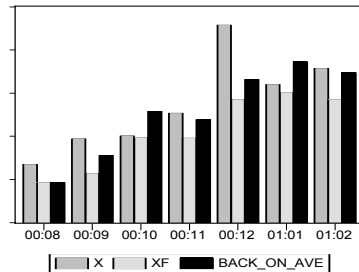
сированной моделью и по сравнению с методикой последовательной коррекции этой константы методом «back-on-track». Проблема, однако, в том, что информация о предстоящем сдвиге уровня ряда на прогнозируемом периоде может быть недоступной, даже если этот сдвиг в действительности планируется. Но и при наличии такой информации нельзя ожидать гарантированного осуществления планируемого сдвига, так что и в этом случае значения, приведенные в последней таблице, являются лишь ориентирами потенциально возможного улучшения характеристик прогноза посредством однократной коррекции константы в фиксированной модели.

#### Модификация метода «back-on-track» – метод «back-on-ave»

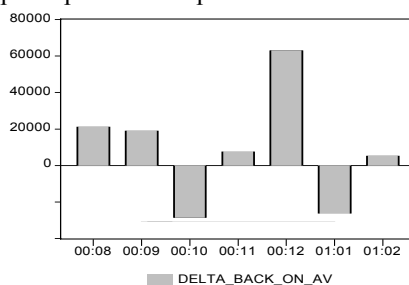
Здесь мы рассмотрим возможную модификацию метода «back-on-track» отличающуюся от последнего тем, что к каждому очередному одношаговому прогнозу, построенному по фиксированной модели, прибавляется среднее арифметическое наблюдаемых ошибок предыдущих одношаговых прогнозов, построенных по той же модели. Мы называем эту модификацию методом «back-on-average» или, для краткости, «back-on-ave».

Следуя методике «back-on-ave», получаем следующие результаты.

Наблюдаемые значения (X), прогнозы по фиксированной TS-модели (XF) и прогнозы, получаемые коррекцией по методу «back-on-ave» (BACK\_ON\_AVE):



Ошибки скорректированных прогнозов



#### Характеристики скорректированных прогнозов

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	30187.74
<b>Mean Absolute Error</b>	24423.78
<b>Mean Absolute Percent Error</b>	3.046623

Применение процедуры «back-on-ave» для коррекции прогнозов по фиксированной модели устранило систематическое смещение прогнозов и привело к значительному улучшению точностных характеристик прогнозов. Последние хотя и не достигают минимальных значений, соответствующих модифицированной фиксированной модели с увеличенной (на 32088.9) постоянной составляющей, но существенно приближаются к этим минимальным значениям. Более того, характеристики прогнозов, полученные с применением метода «back-on-ave», оказались лучше характеристик прогнозов, полученных по всем другим вариантам моделей с аддитивным выбросом, что видно из уже приводившейся ранее таблицы.

	<b>DS ре-курс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	33750.29	35153.12	30788.31	40981.79
<b>Mean Absolute Error</b>	28915.77	29026.71	25004.23	32088.90
<b>Mean Absolute Percent Error</b>	3.604901	3.621122	3.119278	3.958528

Обратимся теперь к расширенному интервалу **08.2000–11.2001**, использованному в главе 3.

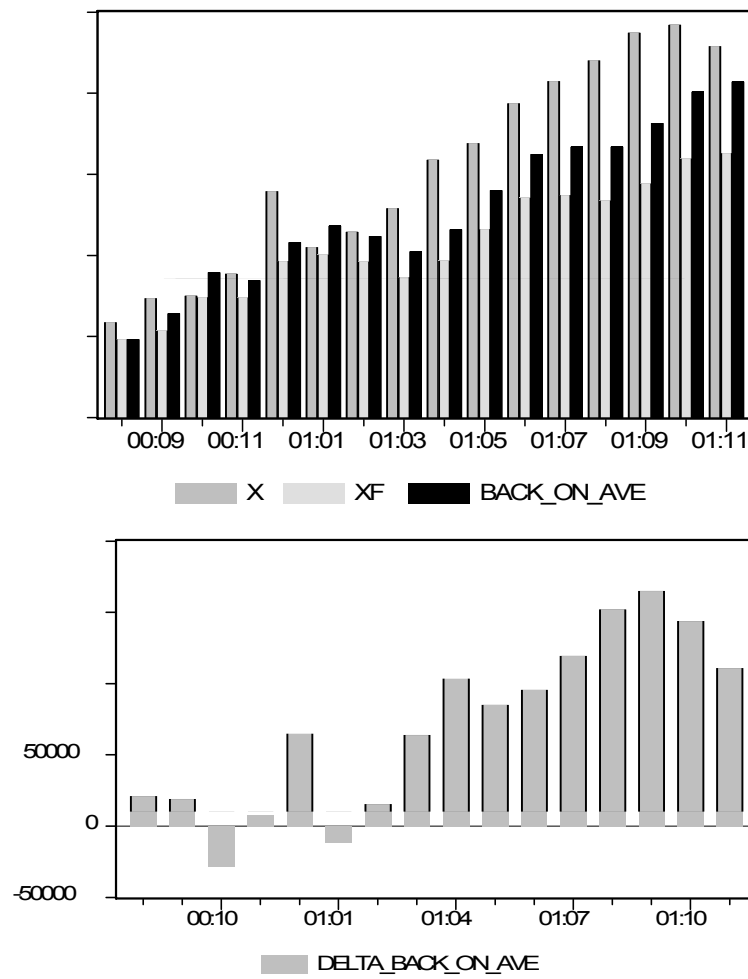
Напомним полученные там результаты последовательного одношагового прогнозирования значений ряда M1 для этого интервала (модели с аддитивным выбросом).

	<b>DS ре-курс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	37646.23	41033.29	35359.73	108578.5
<b>Mean Absolute Error</b>	30672.38	33871.66	30334.93	90923.73
<b>Mean Absolute Percent Error</b>	3.440899	3.775183	3.362247	9.368012

Качество прогнозов резко отличается в худшую сторону для фиксированной *TS*-модели.

Фактически данная таблица уже указывает на два возможных метода предупреждения получения столь неудовлетворительных результатов: использование для прогнозирования рекурсивных моделей и моделей в разностях. Проверим возможности двух других методов, уже использованных в данном разделе: «back-on-track» и «back-on-average» («back-on-ave»).

«Back-on- ave»:

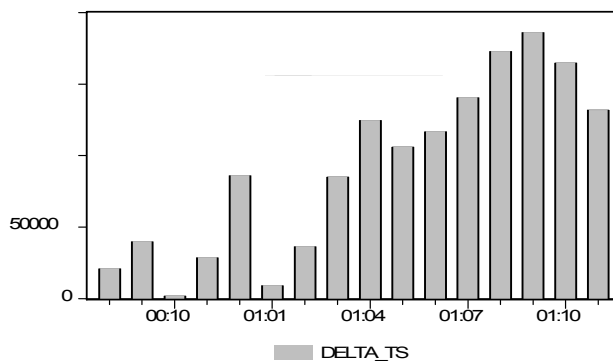


Характеристики скорректированной последовательности прогнозов на расширенном интервале:

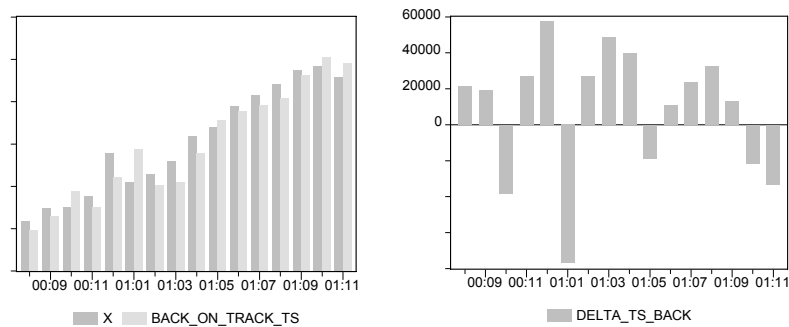
	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	62716.33
<b>Mean Absolute Error</b>	53497.27
<b>Mean Absolute Percent Error</b>	5.608093

На расширенном интервале характеристики прогнозов, скорректированных методом «back-on-ave», по-прежнему значительно лучше характеристик прогнозов по фиксированной *TS*-модели. Однако на этот раз они заметно хуже прогнозов, получаемых по рекурсивным *DS*-моделям и по обеим *DS*-моделям.

Последнее объясняется поведением ошибок прогнозов по фиксированной *TS*-модели на расширенном интервале: на второй половине расширенного интервала эти ошибки весьма высоки, тогда как на первой половине этого интервала они относительно малы. Соответственно, среднеарифметические значения ошибок, используемые для коррекции, оказываются слишком малыми для того, чтобы компенсировать их возрастание на второй половине интервала.



В этой ситуации представляется вероятным более успешное применение метода «back-on-track» с коррекцией только на предыдущую ошибку прогноза. Применение этого метода приводит здесь к следующим результатам.



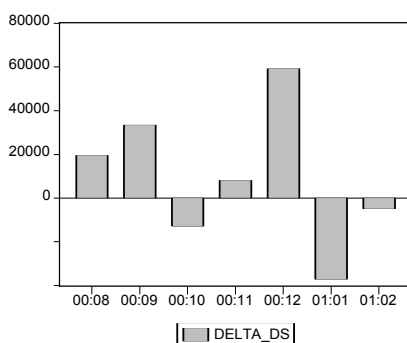
	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	35886.93
<b>Mean Absolute Error</b>	31706.01
<b>Mean Absolute Percent Error</b>	3.628869

Характеристики последовательности прогнозов, скорректированных методом «back-on-track», лучше характеристик прогнозов, полученных по рекурсивной *TS*-модели, но все же несколько хуже характеристик, получаемых по обеим *DS*-моделям.

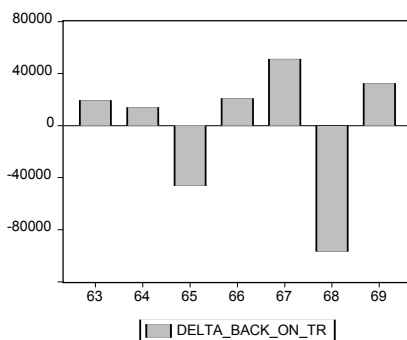
### Модель в разностях (DS-модель) с аддитивным выбросом

Ошибки прогнозов

Фиксированная модель:



Back-on-track:



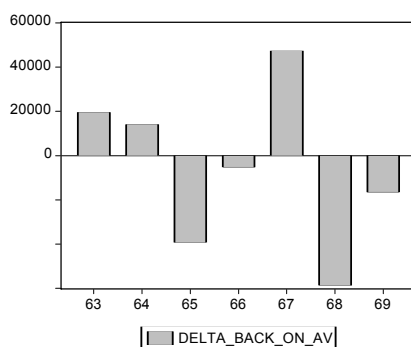
Характеристики прогнозов

	<b>Фиксир. модель</b>	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	30788.31	47972.41
<b>Mean Absolute Error</b>	25004.23	40060.90
<b>Mean Absolute Percent Error</b>	3.119278	5.004921

Здесь, в отличие от *TS*-модели, выраженное систематическое смещение прогнозов по фиксированной модели отсутствует. Характеристики последовательности скорректированных (методом «back-on-track») прогнозов оказываются хуже характеристик нескорректированных прогнозов.

Результаты применения метода «back-on-ave»:

Ошибки прогнозов



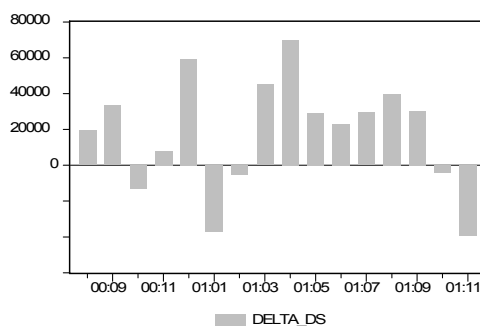
Характеристики прогнозов

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	33987.66
<b>Mean Absolute Error</b>	28623.01
<b>Mean Absolute Percent Error</b>	3.583805

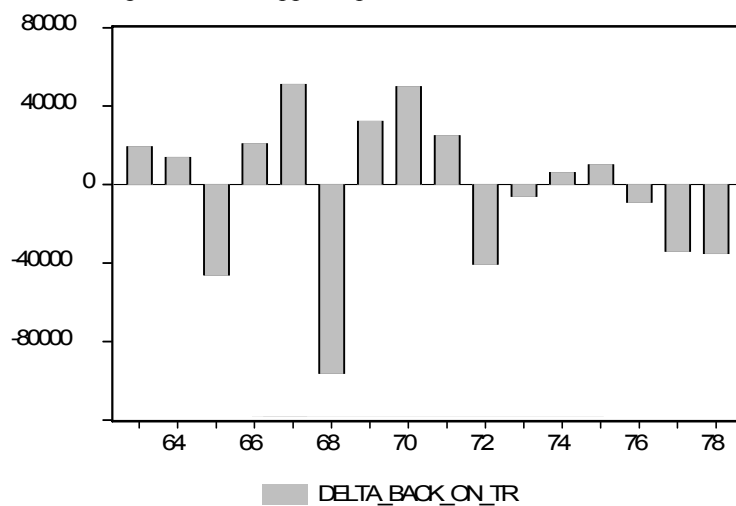
Характеристики последовательности прогнозов, скорректированных методом «back-on-ave», оказались лучше характеристик прогнозов, полученных по рекурсивной *DS*-модели, но хуже характеристик, получаемых по фиксированной *DS*-модели, что видно из приводившейся выше таблицы.

Обратимся теперь к расширенному интервалу **08.2000–11.2001**.

Ошибки прогнозов по фиксированной *DS*-модели



Ошибки прогнозов, скорректированных методом «back-on-track»

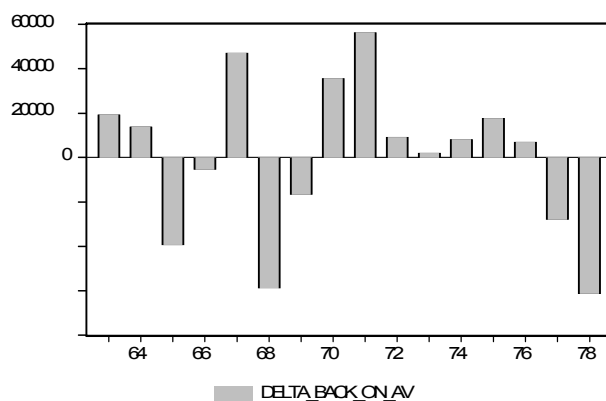


Характеристики прогнозов

	Фиксир. модель	Back-on-track
<b>Root Mean Squared Error</b>	35359.73	38409.81
<b>Mean Absolute Error</b>	30334.93	31134.30
<b>Mean Absolute Percent Error</b>	3.362247	3.597017

Характеристики прогнозов в результате коррекции ухудшились.  
«Back-on-ave»:

Ошибки скорректированных прогнозов





Характеристики скорректированной последовательности прогнозов на расширенном интервале:

	Back-on-ave
Root Mean Squared Error	33164.02
Mean Absolute Error	26333.67
Mean Absolute Percent Error	3.008401

Коррекция методом «back-on-ave» привела к наилучшей последовательности одношаговых прогнозов.

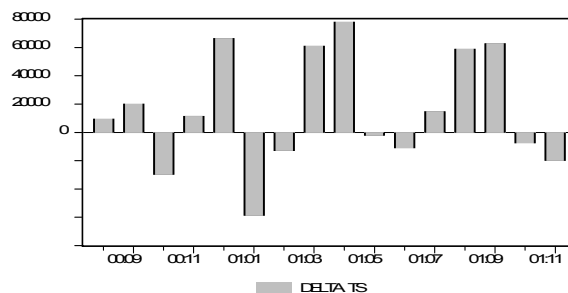
#### 4.3.3. M2

И здесь мы ограничимся только моделями с аддитивным выбросом.

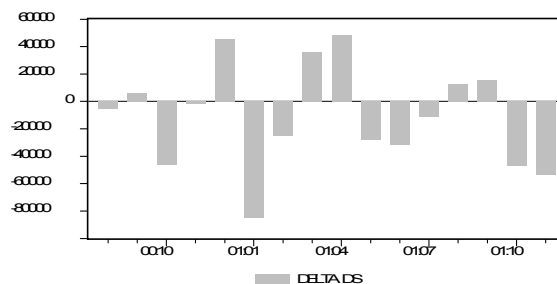
Прогнозы на временном интервале 08.2000–11.2001 (16 точек):

	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
Root Mean Squared Error	39133.66	41913.71	37951.93	41623.77
Mean Absolute Error	28768.55	31083.92	31062.16	32920.79
Mean Absolute Percent Error	2.457265	2.656437	2.598554	2.762772

Ошибки прогнозов по фиксированной TS-модели



по фиксированной DS-модели



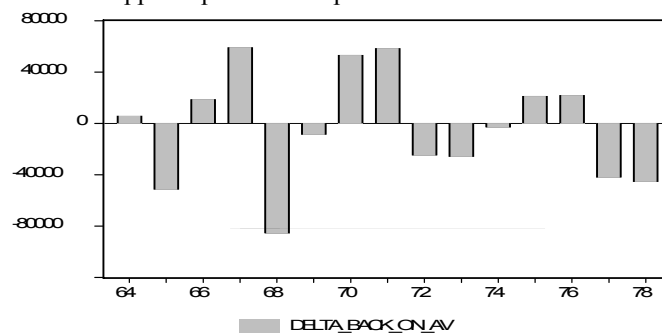
Систематическое смещение отсутствует и у прогнозов по *DS*-моделям, и у прогнозов по *TS*-моделям.

Проанализируем, что при этом дает коррекция прогнозов методами «back-on-track» и «back-on-ave».

**TS-модель:**

«Back-on-ave»:

Ошибки скорректированных прогнозов



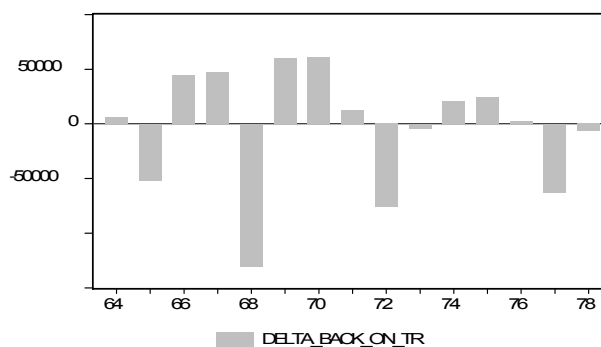
Характеристики скорректированных прогнозов

	Back-on-ave
Root Mean Squared Error	41913.43
Mean Absolute Error	35127.43
Mean Absolute Percent Error	2.960978

Характеристики скорректированных прогнозов хуже характеристик прогнозов по фиксированной и по рекурсивной *TS*-моделям.

«Back-on-track»:

Ошибки скорректированных прогнозов



Характеристики скорректированных прогнозов

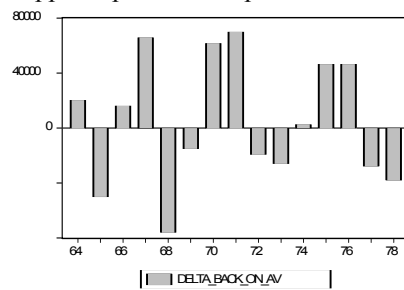
	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	52884.03
<b>Mean Absolute Error</b>	40505.50
<b>Mean Absolute Percent Error</b>	3.506794

Характеристики прогнозов, скорректированных методом «back-on-track», хуже характеристик всех остальных прогнозов, полученных в классе *TS*-моделей.

**DS-модель**

«Back-on-ave»:

Ошибки скорректированных прогнозов



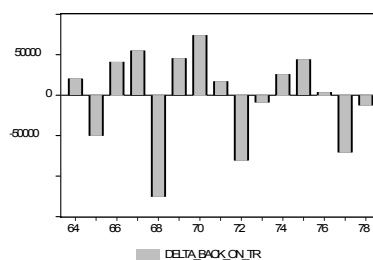
Характеристики скорректированных прогнозов

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	44569.30
<b>Mean Absolute Error</b>	38767.36
<b>Mean Absolute Percent Error</b>	3.261524

Характеристики прогнозов, скорректированных методом «back-on-ave», хуже характеристик прогнозов, полученных по фиксированной и рекурсивной *DS*-моделям.

«Back-on-track»:

Ошибки скорректированных прогнозов



#### Характеристики скорректированных прогнозов

	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	55119.29
<b>Mean Absolute Error</b>	45006.80
<b>Mean Absolute Percent Error</b>	3.856839

Характеристики прогнозов, скорректированных методом «back-on-track», хуже характеристик всех остальных прогнозов, полученных в классе *DS*-моделей.

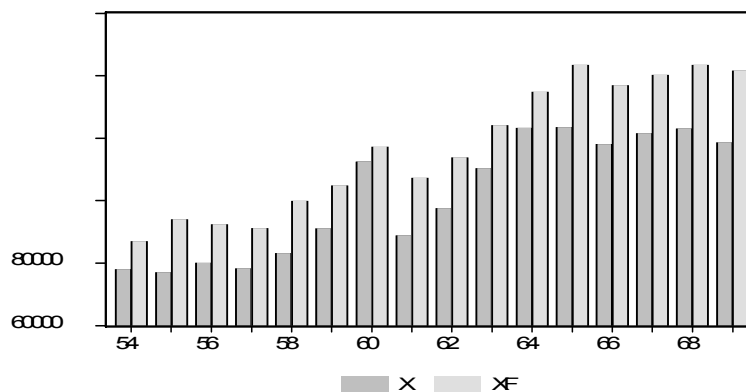
#### 4.3.4. Налоговые доходы федерального бюджета

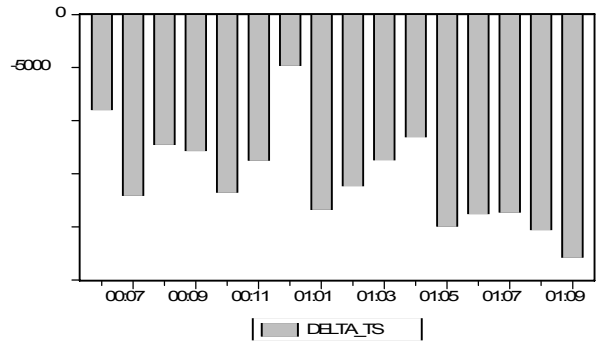
Ограничимся моделями с аддитивным выбросом.

Прогнозы на временном интервале **06.2000–09.2001** (16 точек):

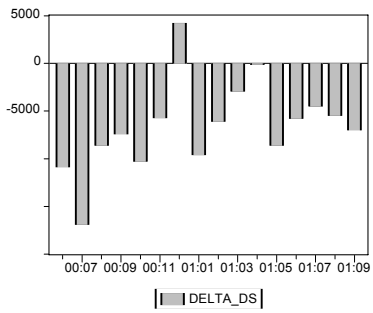
	<b>DS ре-курс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	6676.117	6983.570	8038.587	16075.68
<b>Mean Absolute Error</b>	5097.416	5367.743	7134.696	15428.55
<b>Mean Absolute Percent Error</b>	5.740951	6.044299	7.761660	15.44266

Наихудшие результаты дает прогнозирование по фиксированной *TS*-модели (заметим, что сам ряд на временном интервале оценивания фиксированной модели был классифицирован как *DS*-ряд), приводящее к систематическому смещению прогнозов – переоцениванию истинных значений ряда:





Впрочем, смещение, хотя и не столь значительное, имеют и прогнозы по фиксированной *DS*-модели:

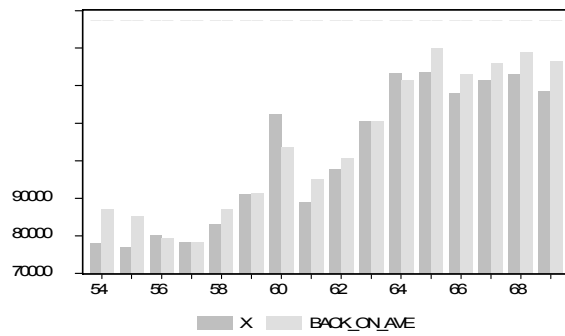


Проанализируем, что дает коррекция прогнозов по этим моделям методами «back-on-track» и «back-on-ave».

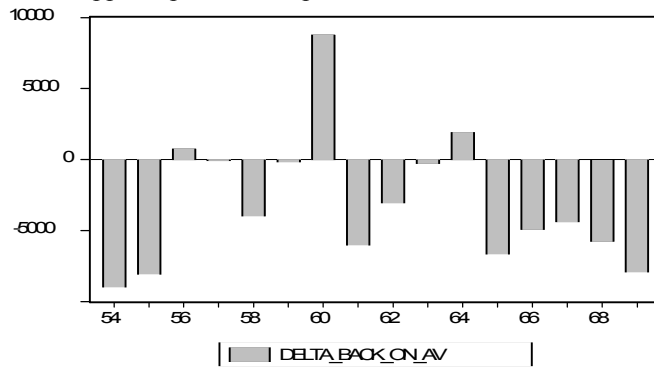
**TS-модель:**

«Back-on-ave»:

Наблюдения и скорректированные прогнозы



Ошибки скорректированных прогнозов



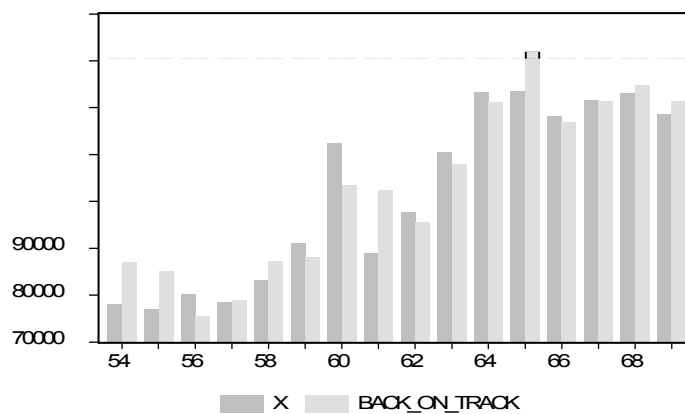
Характеристики скорректированных прогнозов

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	5437.116
<b>Mean Absolute Error</b>	4485.736
<b>Mean Absolute Percent Error</b>	4.504539

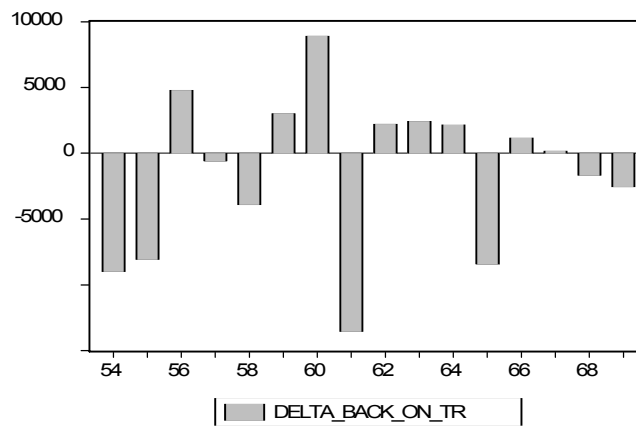
Характеристики прогнозов, скорректированных методом «back-on-ave», лучше характеристик прогнозов, полученных по фиксированной и рекурсивной *DS*-моделям.

«Back-on-track»:

Наблюдения и скорректированные прогнозы



Ошибки скорректированных прогнозов



Характеристики скорректированных прогнозов

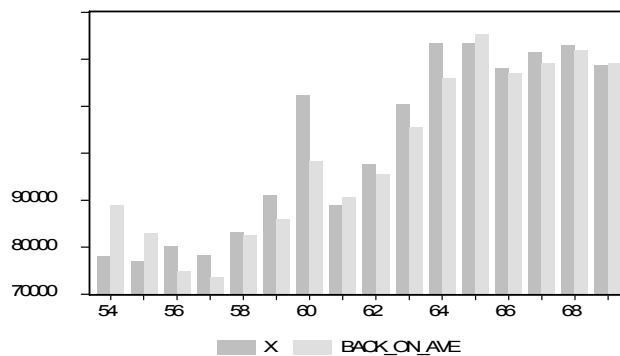
	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	5887.558
<b>Mean Absolute Error</b>	4543.872
<b>Mean Absolute Percent Error</b>	4.853807

Характеристики прогнозов, скорректированных методом «back-on-track», лучше характеристик прогнозов, полученных по фиксированной и рекурсивной *TS*-моделям, но несколько хуже характеристик прогнозов, скорректированных методом «back-on-ave».

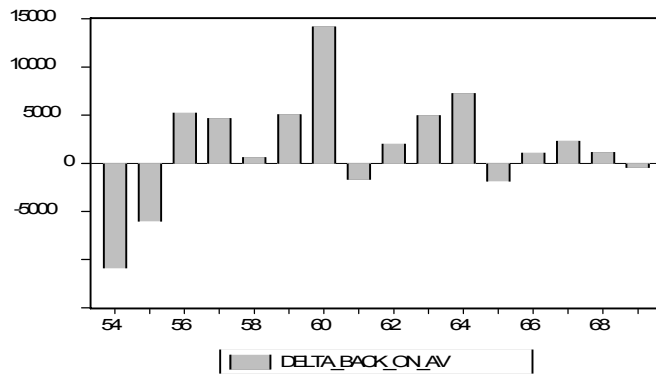
**DS-модель**

«Back-on-ave»:

Наблюдения и прогнозы



Ошибки скорректированных прогнозов



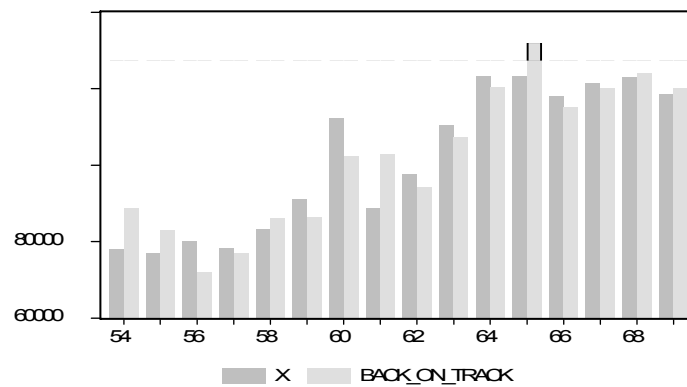
Характеристики скорректированных прогнозов

	Back-on-ave
Root Mean Squared Error	5750.589
Mean Absolute Error	4359.457
Mean Absolute Percent Error	4.591145

Характеристики прогнозов, скорректированных методом «back-on-ave», лучше характеристик прогнозов, полученных по фиксированной и рекурсивной *DS*-моделям.

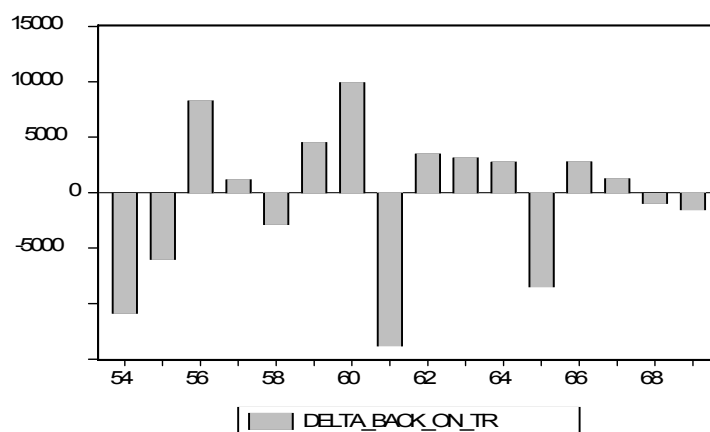
«Back-on-track»:

Наблюдения и прогнозы





### Ошибки скорректированных прогнозов



### Характеристики скорректированных прогнозов

	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	6419.555
<b>Mean Absolute Error</b>	5136.422
<b>Mean Absolute Percent Error</b>	5.476168

Характеристики прогнозов, скорректированных методом «back-on-track», лучше характеристик прогнозов, полученных по фиксированной и рекурсивной *DS*-моделям, но несколько хуже характеристик прогнозов, скорректированных методом «back-on-ave».

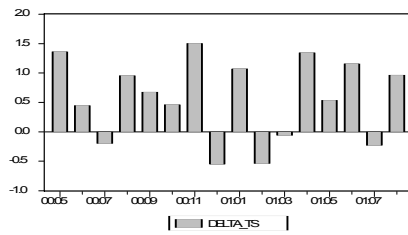
#### 4.3.5. Экспорт

Напомним результаты прогнозирования на расширенном интервале 05.2000-08.2001 (16 точек):

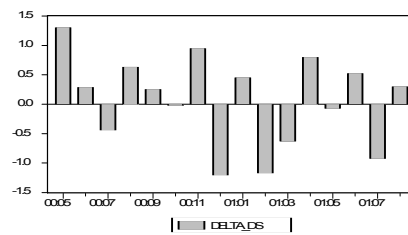
	<b>DS ре-курс.</b>	<b>TS рекурс.</b>	<b>DS фикс.</b>	<b>TS фикс.</b>
<b>Root Mean Squared Error</b>	0.710022	0.770122	0.734464	0.869509
<b>Mean Absolute Error</b>	0.594670	0.673983	0.622717	0.753316
<b>Mean Absolute Percent Error</b>	6.693006	7.542889	6.989588	8.393028

Ошибки прогнозов по фиксированным моделям:

*TS*-модель



*DS*-модель

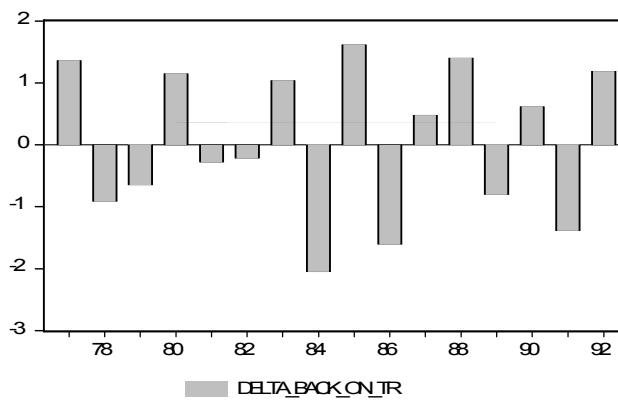


Выраженного систематического смещения прогнозов по обеим фиксированным моделям не наблюдается. Тем не менее попробуем произвести коррекцию этих прогнозов в режимах «back-on-track» и «back-on-ave».

***TS*-модель**

«Back-on-track»:

Ошибки прогнозов



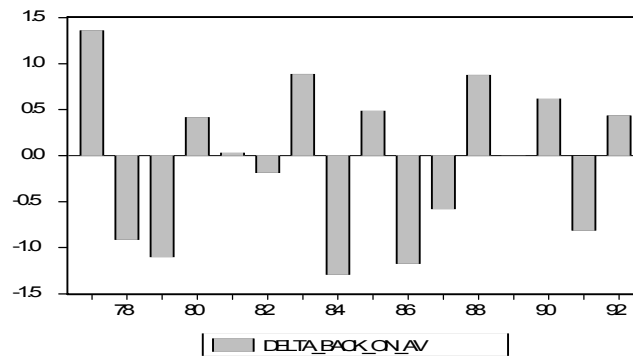
Характеристики прогнозов

	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	1.162296
<b>Mean Absolute Error</b>	1.047766
<b>Mean Absolute Percent Error</b>	11.79271

Характеристики прогнозов, скорректированных методом «back-on-track», значительно хуже характеристик прогнозов, полученных по фиксированной и рекурсивной *TS*-моделям.

«Back-on-ave»:

Ошибки прогнозов



Характеристики прогнозов

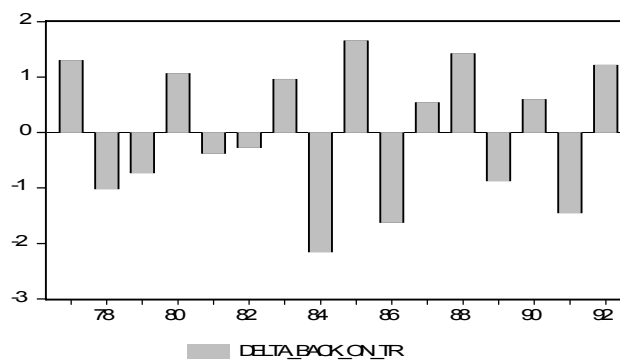
	Back-on-ave
Root Mean Squared Error	0.810727
Mean Absolute Error	0.698312
Mean Absolute Percent Error	7.863206

Характеристики прогнозов, скорректированных методом «back-on-ave», несколько **хуже** характеристик прогнозов, полученных по фиксированной и рекурсивной *TS*-моделям, и значительно лучше характеристик прогнозов, скорректированных методом «back-on-track».

**DS-модель**

«Back-on-track»:

Ошибки прогнозов



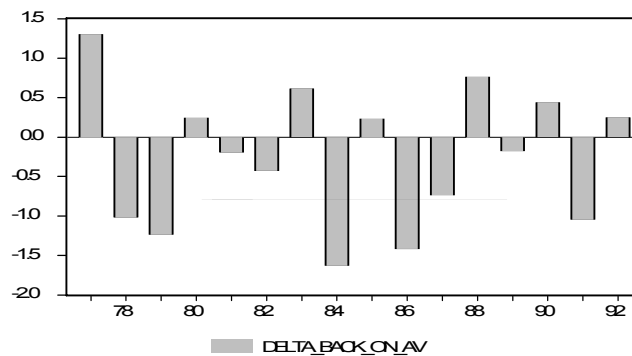
### Характеристики прогнозов

	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	1.189322
<b>Mean Absolute Error</b>	1.080290
<b>Mean Absolute Percent Error</b>	12.165979

Характеристики прогнозов, скорректированных методом «back-on-track», значительно хуже характеристик прогнозов, полученных по фиксированной и рекурсивной *DS*-моделям.

«Back-on-ave»:

Ошибки прогнозов



### Характеристики прогнозов

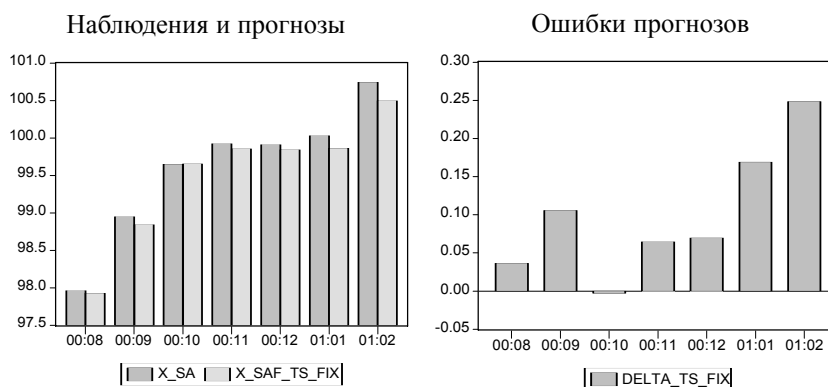
	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	0.870558
<b>Mean Absolute Error</b>	0.732049
<b>Mean Absolute Percent Error</b>	8.266621

Характеристики прогнозов, скорректированных методом «back-on-ave», хуже характеристик прогнозов, полученных по фиксированной и рекурсивной *DS*-моделям, но значительно лучше характеристик прогнозов, скорректированных методом «back-on-track».

#### 4.3.6. Индекс интенсивности производства цветных металлов

Сезонно скорректированный ряд

Вспомним итоги прогнозирования для этого ряда по подобранной фиксированной *TS*-модели:

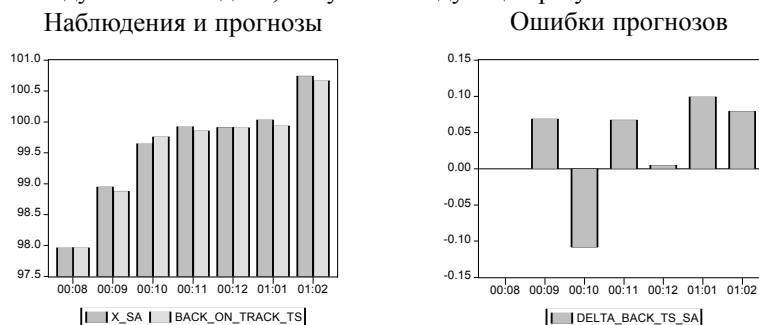


Характеристики прогнозов

	TS
Root Mean Squared Error	0.126271
Mean Absolute Error	0.099413
Mean Absolute Percent Error	0.099430

График ошибок прогноза указывает на явную смещенность прогнозов, полученных по использованной модели: прогнозные значения систематически недооценивают истинных значений ряда. В связи с этим проанализируем, что дало бы использование в процессе прогнозирования по указанной модели методики «back-on-track».

Следуя этой методике, получаем следующие результаты:



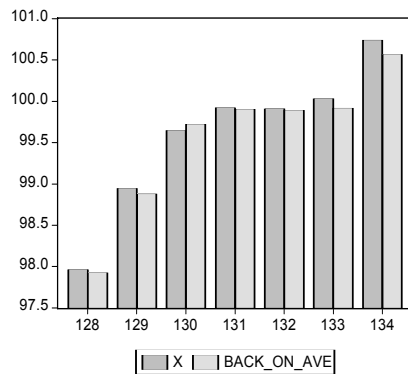
### Характеристики прогнозов

	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	0.074136
<b>Mean Absolute Error</b>	0.066350
<b>Mean Absolute Percent Error</b>	0.066537

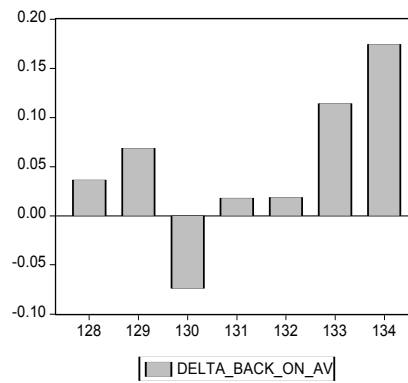
Качество прогноза с использованием методики «back-on-track» значительно улучшилось. По RMSE прогноз оказался даже лучше не только прогноза по фиксированной *TS*-модели, но и остальных трех прогнозов (по фиксированной *DS*-модели, по рекурсивной *TS*-модели и по рекурсивной *DS*-модели).

«Back-on-ave»:

Наблюдения и прогнозы



Ошибки прогнозов



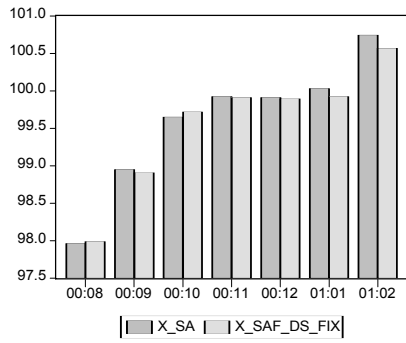
### Характеристики прогнозов

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	0.089205
<b>Mean Absolute Error</b>	0.072079
<b>Mean Absolute Percent Error</b>	0.072144

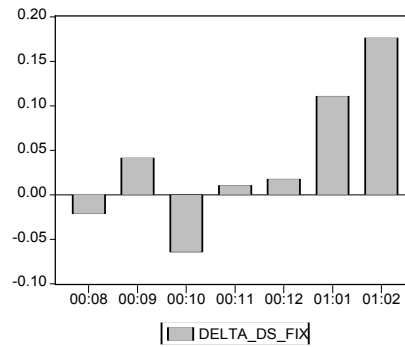
Качество скорректированных прогнозов лучше, чем у прогнозов по фиксированной и рекурсивной *TS*-моделям, но несколько хуже, чем у прогнозов, скорректированных методом «back-on-track».

Применим те же две методики коррекции прогноза к фиксированной *DS*-модели.

Ранее мы получили следующее:  
Наблюдения и прогнозы



Ошибки прогнозов

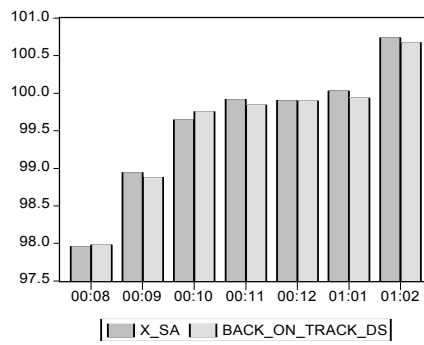


Характеристики прогнозов

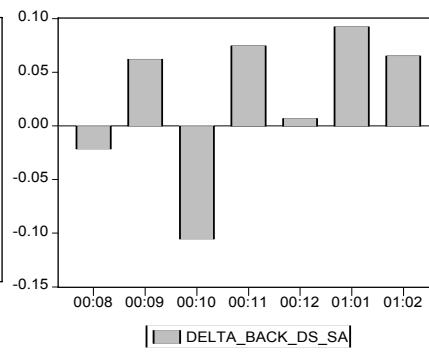
	<b>DS</b>
<b>Root Mean Squared Error</b>	0.084411
<b>Mean Absolute Error</b>	0.063081
<b>Mean Absolute Percent Error</b>	0.063052

При использовании методики «back-on-track» получаем следующие результаты:

Наблюдения и прогнозы



Ошибки прогнозов

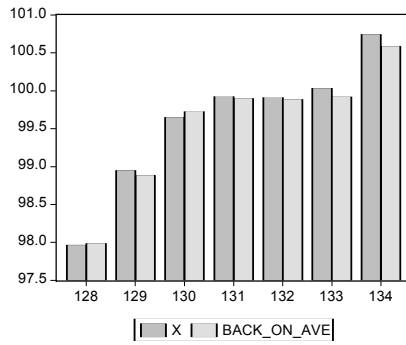


	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	0.069753
<b>Mean Absolute Error</b>	0.061431
<b>Mean Absolute Percent Error</b>	0.061579

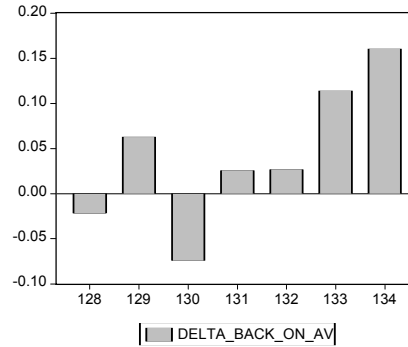
Качество прогнозов улучшилось, но незначительно.

«Back-on-ave»:

Наблюдения и прогнозы



Ошибки прогнозов



Характеристики прогнозов

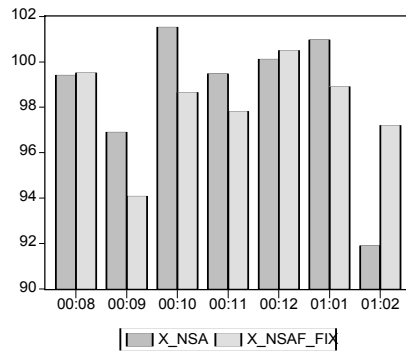
	Back-on-ave
<b>Root Mean Squared Error</b>	0.084417
<b>Mean Absolute Error</b>	0.099174
<b>Mean Absolute Percent Error</b>	0.069202

Характеристики скорректированных прогнозов лучше, чем у прогнозов, полученных по рекурсивной *DS*-модели, но несколько хуже, чем у прогнозов, полученных по фиксированной *DS*-модели, и у прогнозов, скорректированных методом «back-on-track».

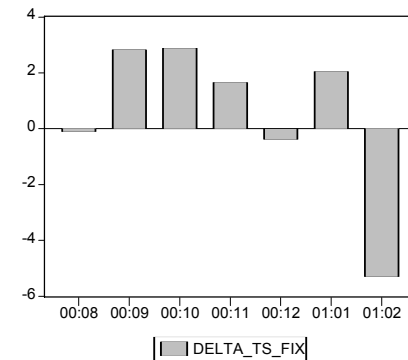
Нескорректированный ряд

Итоги прогнозирования для этого ряда по оцененной фиксированной *TS*-модели:

Наблюдения и прогнозы



Ошибки прогнозов



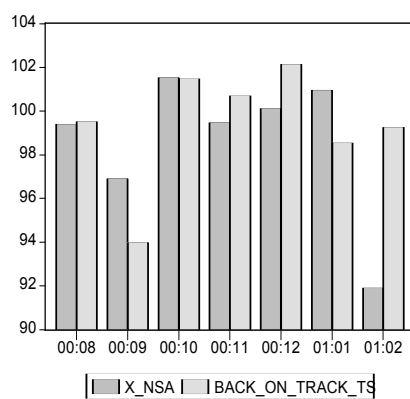


Характеристики прогнозов

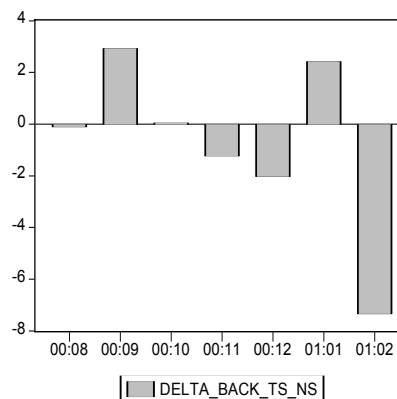
	TS
<b>Root Mean Squared Error</b>	2.709401
<b>Mean Absolute Error</b>	2.169326
<b>Mean Absolute Percent Error</b>	2.241112

Следуя методике «back-on-track», получаем следующие результаты:

Наблюдения и прогнозы



Ошибки прогнозов



Характеристики прогнозов

	TS
<b>Root Mean Squared Error</b>	3.252290
<b>Mean Absolute Error</b>	2.302860
<b>Mean Absolute Percent Error</b>	2.405861

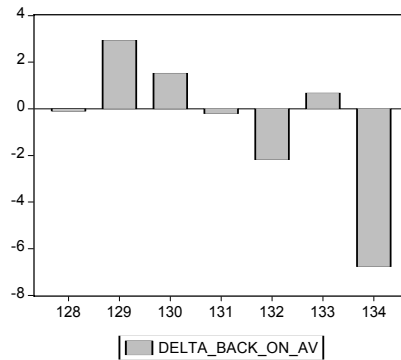
В этом случае методика «back-on-track» не дала улучшения прогнозов, и это связано с существенным ухудшением прогноза на февраль 2001 г.

Вместе с тем, если сравнивать альтернативные прогнозы, не включая в рассмотрение прогноз на этот месяц, то тогда получаем следующие результаты:

	TS-фикс.	«Back-on-track»
<b>Root Mean Squared Error</b>	1.973102	1.832792
<b>Mean Absolute Error</b>	1.648535	1.463207
<b>Mean Absolute Percent Error</b>	1.654562	1.475603

На этом периоде методика «back-on-track» улучшает прогнозы.

«Back-on-ave»:  
Ошибки прогнозов



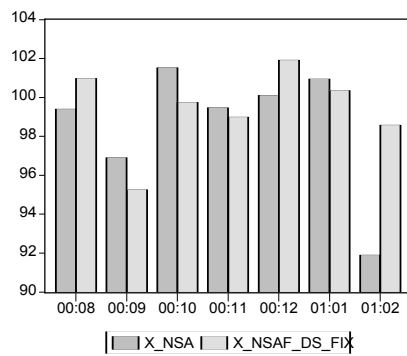
Характеристики скорректированных прогнозов

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	2.980953
<b>Mean Absolute Error</b>	2.059774
<b>Mean Absolute Percent Error</b>	2.154163

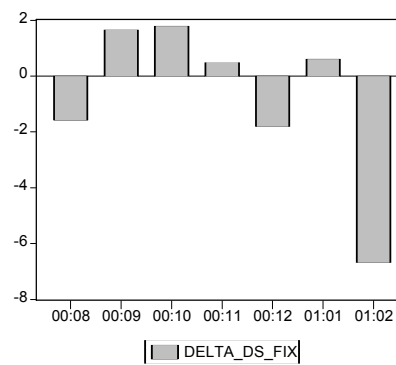
Методика «back-on-ave» несколько улучшает прогнозы по фиксированной и рекурсивной *TS*-моделям.

Рассмотрим теперь итоги прогнозирования ряда по подобранной фиксированной *DS*-модели:

Наблюдения и прогнозы



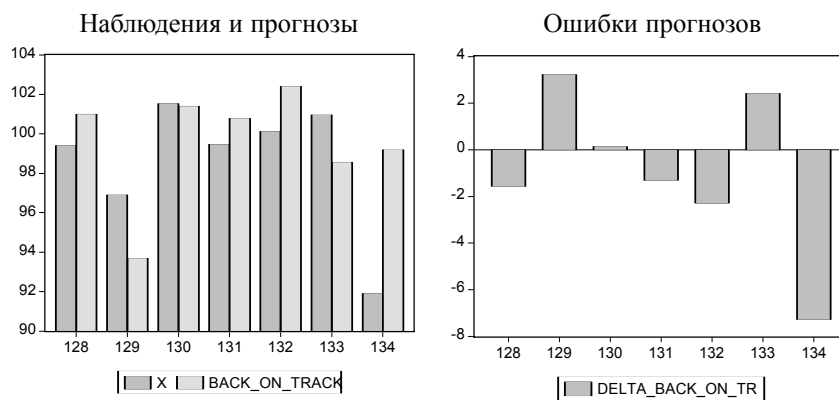
Ошибки прогнозов



### Характеристики прогнозов

	<b>DS</b>
<b>Root Mean Squared Error</b>	2.850249
<b>Mean Absolute Error</b>	2.082808
<b>Mean Absolute Percent Error</b>	2.171102

Результаты применения корректирующих процедур.  
«Back-on-track»:

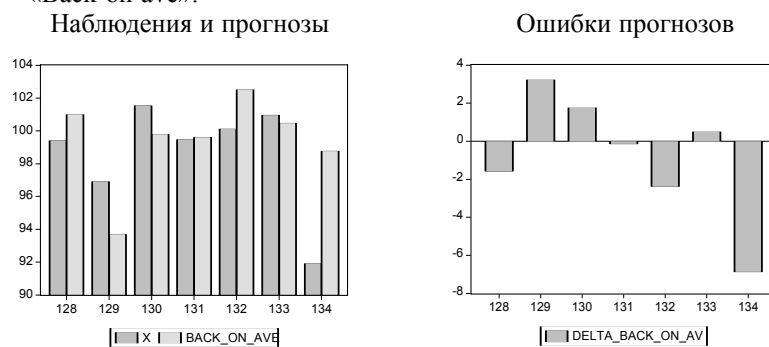


### Характеристики прогнозов

	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	3.352380
<b>Mean Absolute Error</b>	2.603443
<b>Mean Absolute Percent Error</b>	2.708143

Применение методики «back-on-track» привело к ухудшению характеристик прогнозов.

«Back-on-ave»:



#### Характеристики скорректированных прогнозов

	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	3.141902
<b>Mean Absolute Error</b>	2.349225
<b>Mean Absolute Percent Error</b>	2.446948

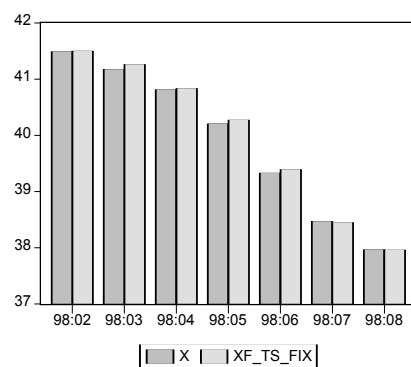
Применение методики «back-on-ave» также привело к ухудшению характеристик прогнозов, хотя и не столь значительному, как при применении методики «back-on-track».

#### 4.3.7. Индекс интенсивности промышленного производства

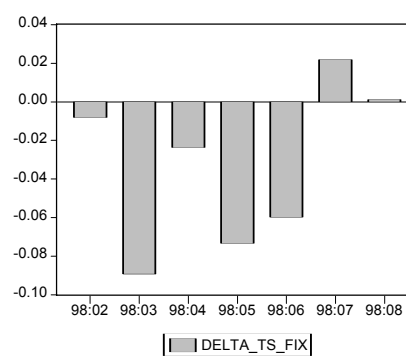
Ранее мы рассмотрели прогнозы для этого (сезонно скорректированного) ряда по фиксированным *TS*- и *DS*-моделям, оцененным на временном интервале с 01.1994 по 01.1998.

Прогнозы по фиксированной *TS*-модели:

Наблюдения и прогнозы



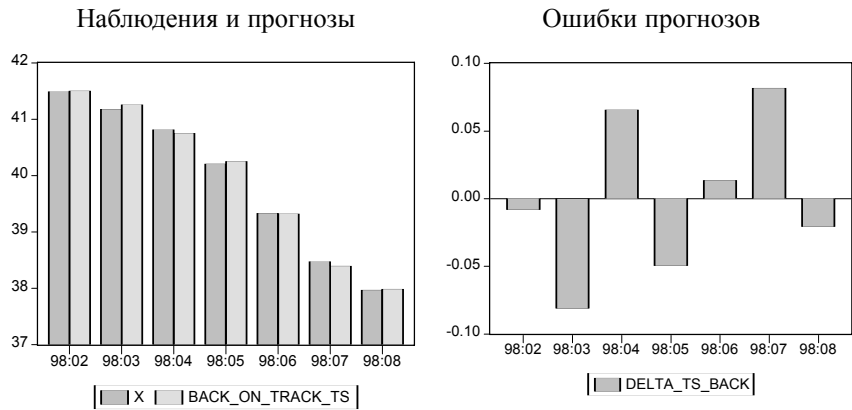
Ошибки прогнозов



#### Характеристики прогнозов

	<b>TS</b>
<b>Root Mean Squared Error</b>	0.050764
<b>Mean Absolute Error</b>	0.039615
<b>Mean Absolute Percent Error</b>	0.098420

При использовании методики «back-on-track» получаем следующие результаты:

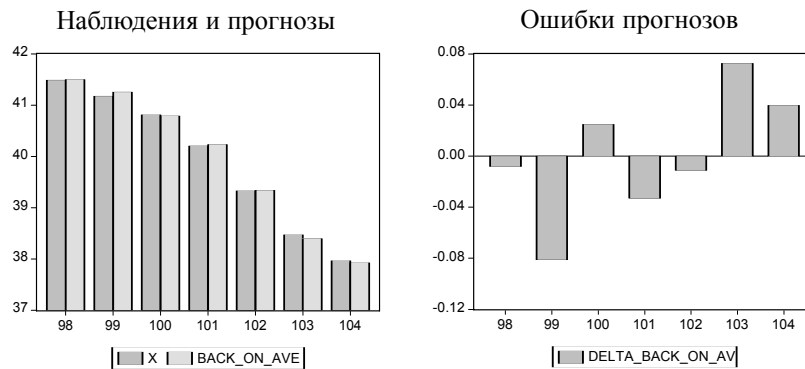


Характеристики прогнозов

	TS
<b>Root Mean Squared Error</b>	0.054357
<b>Mean Absolute Error</b>	0.045757
<b>Mean Absolute Percent Error</b>	0.114544

Прогнозы в целом ухудшились: систематическое смещение прогнозов устраняется за счет их большей вариабельности, что, в конечном счете, привело к ухудшению характеристик прогнозов.

При использовании методики «back-on-ave» результаты таковы:

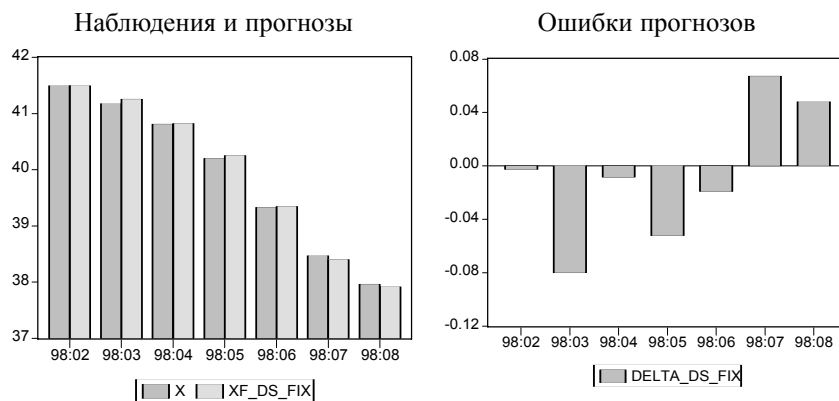


Характеристики прогнозов

	Back-on-ave
<b>Root Mean Squared Error</b>	0.046849
<b>Mean Absolute Error</b>	0.038710
<b>Mean Absolute Percent Error</b>	0.097472

Характеристики прогнозов по фиксированной *TS*-модели практически не изменяются при коррекции прогнозов методом «back-on-ave».

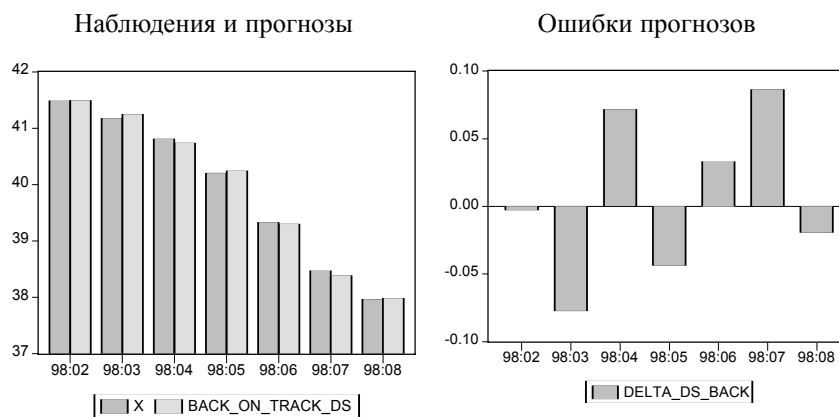
Прогнозы по фиксированной *DS*-модели:



Характеристики прогнозов

	<b>DS</b>
<b>Root Mean Squared Error</b>	0.048401
<b>Mean Absolute Error</b>	0.039698
<b>Mean Absolute Percent Error</b>	0.100229

При использовании методики «back-on-track» получаем следующие результаты:



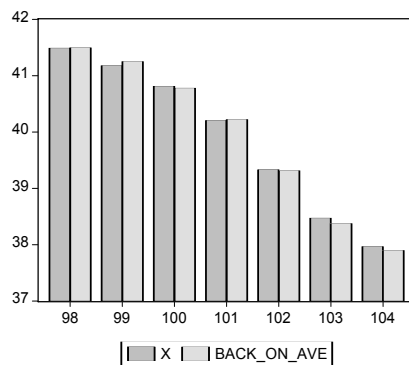
### Характеристики прогнозов

	<b>Back-on-track</b>
<b>Root Mean Squared Error</b>	0.055997
<b>Mean Absolute Error</b>	0.047726
<b>Mean Absolute Percent Error</b>	0.119700

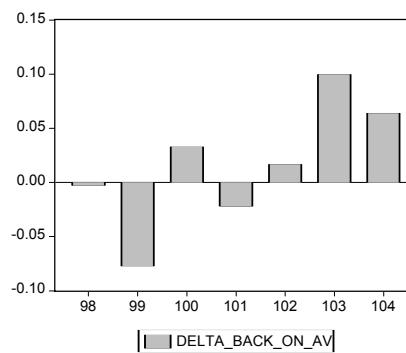
И в этом случае уменьшение смещения прогнозов, достигнутое ценой возрастания их вариабельности, в конечном счете, привело к ухудшению характеристик прогнозов.

При использовании методики «back-on-ave» результаты таковы:

Наблюдения и прогнозы



Ошибки прогнозов



### Характеристики прогнозов

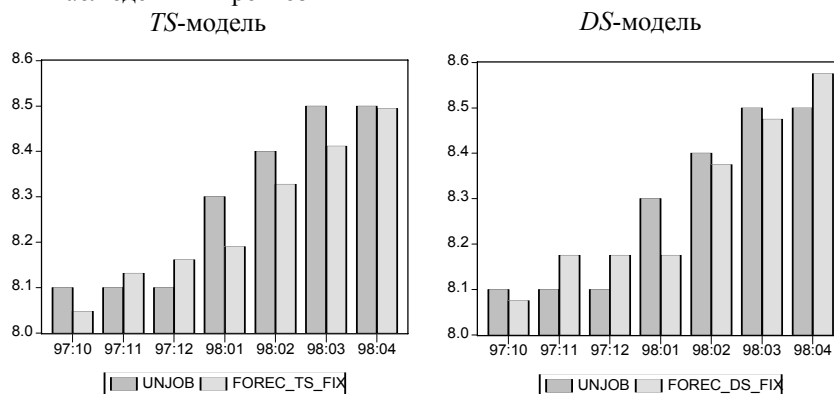
	<b>Back-on-ave</b>
<b>Root Mean Squared Error</b>	0.055864
<b>Mean Absolute Error</b>	0.045015
<b>Mean Absolute Percent Error</b>	0.114184

Уменьшение смещения прогнозов достигается ценой возрастания их вариабельности, что в конечном счете привело к ухудшению характеристик прогнозов.

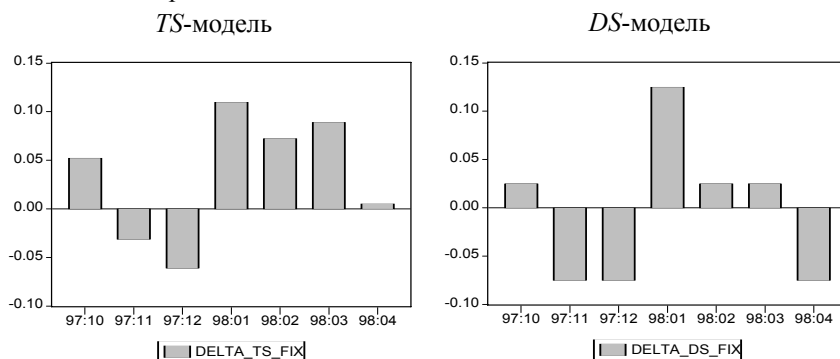
#### 4.3.8. Безработица

В главе 1 были получены следующие результаты для прогнозов по оцененным фиксированным *TS*- и *DS*-моделям (оценивание на временном интервале 01.1994-09.1997, одношаговые прогнозы на временном интервале 10.1997-04.1998).

Наблюдения и прогнозы



Ошибки прогнозов



Характеристики прогнозов по фиксированным и рекурсивным моделям

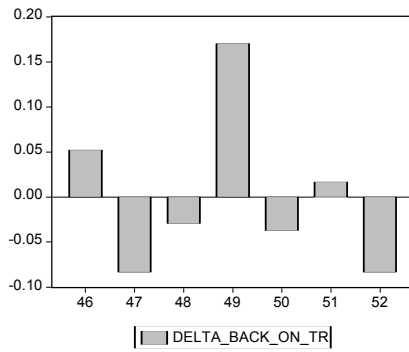
	DS ре-курс.	TS рекурс.	DS фикс.	TS фикс.
<b>Root Mean Squared Error</b>	0.070806	0.067143	0.070076	0.068207
<b>Mean Absolute Error</b>	0.061088	0.060246	0.060714	0.060048
<b>Mean Absolute Percent Error</b>	0.738836	0.727211	0.734373	0.724306



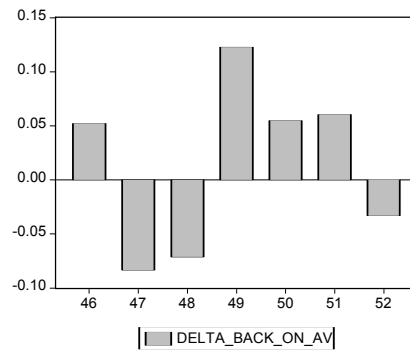
Скорректированные прогнозы:

**TS-модель:**

«Back-on-track»



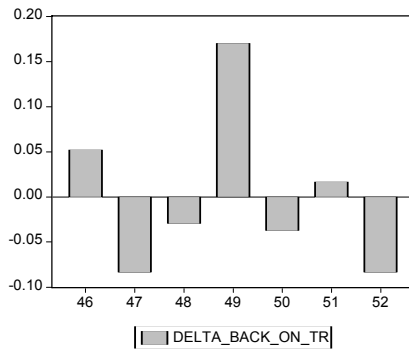
«Back-on-ave»



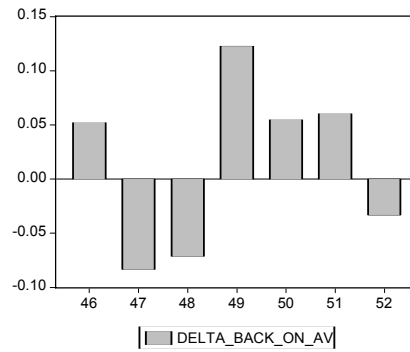
	Back-on-track	Back-on-ave
<b>Root Mean Squared Error</b>	0.083033	0.073324
<b>Mean Absolute Error</b>	0.067588	0.068324
<b>Mean Absolute Percent Error</b>	0.816525	0.827066

**DS-модель:**

«Back-on-track»



«Back-on-ave»



	Back-on-track	Back-on-ave
<b>Root Mean Squared Error</b>	0.083033	0.733244
<b>Mean Absolute Error</b>	0.067588	0.068324
<b>Mean Absolute Percent Error</b>	0.816525	0.827066

Для ряда данных о безработице обе корректирующие процедуры приводят фактически к одним и тем же характеристикам, которые оказались хуже, чем для прогнозов и по фиксированным, и по рекурсивным моделям.

## Глава 5. Моделирование и прогноз налоговых поступлений в консолидированный и федеральный бюджеты РФ

Целью данной главы является выбор наиболее подходящих моделей и методов для регулярного построения краткосрочных прогнозов поступлений основных налогов в консолидированный и федеральный бюджеты РФ. В этой главе будут рассмотрены эконометрические модели, основанные на авторегрессионных процессах и процессах скользящего среднего (ARIMA), а также модель оценки поступлений (Revenue Estimating Model, REM)<sup>2</sup>. Для каждого из исследуемых временных рядов построены два типа эконометрических моделей – модели для процессов с детерминированным трендом (*TS*-модели) и модели для процессов со стохастическим трендом (*DS*-модели); при этом даже при ложной классификации, мы будем рассматривать обе модели, и нас будет интересовать разница в качестве прогнозов рассматриваемых *TS* и *DS*-моделей. Для исследования выбраны шесть временных рядов: суммарные налоговые поступления и поступления налога на прибыль в консолидированный и федеральный бюджеты, поступления НДС в консолидированный бюджет и поступления подоходного налога в консолидированный бюджет. В конце главы представлены результаты сравнения многошаговых прогнозов эконометрических моделей с соответствующими прогнозами, которые получены в модели оценки поступлений.

Общие замечания:

- При расчетах используются ежемесячные данные по налоговым поступлениям (источник – Министерство финансов РФ) в постоянных ценах декабря 1993 г. В качестве дефлятора использовался индекс потребительских цен по данным Госкомстата РФ, как наиболее корректный и достоверный регулярно рассчитываемый ценовой индекс.

---

<sup>2</sup> Модель оценки поступлений была предложена в качестве примера для сравнения результатов прогнозирования при помощи простой модели – калькулятора для расчета прогноза с использованием простых корректировок предыдущих значений налоговых поступлений и эконометрических ARIMA-моделей. Авторы выражают благодарность профессору Р. Конраду (Public Finance Group, Duke Center for International Development) за предоставление каркаса данной модели и полезные замечания и комментарии при подготовке прогнозов.

- Для проверки существования единичного корня и наличия тренда во временных рядах использовался ADF-тест и процедура Dolado, Jenkinson'a и Sosvilla-Rivero. Исследование на автокорреляцию остатков проводилось на основе Q-статистики и теста множителей Лагранжа (LM-тест). Выбор наилучшей модели осуществляется на основании информационных критериев Akaikie (AIC) и Schwarz'a (BIC).

При построении одношаговых прогнозов применялись три метода:

1. На каждый последующий месяц прогноз рассчитывается с помощью уравнения регрессии вновь оцененного по фактическим данным за все предшествующие месяцы (ниже – метод 1).
2. Прогноз с добавлением фактических данных без переоценки коэффициентов модели (ниже – метод 2).
3. Прогноз на каждый последующий месяц равен сумме прогноза по второму методу и ошибки прогноза в предыдущем периоде (ниже – метод 3).

## 5.1. Поступления подоходного налога

Поступления подоходного налога в рассматриваемый период характеризуются значительными колебаниями; при этом в некотором смысле в отдельные периоды 1995–1998 и 1998–2002 гг. поступления характеризуются относительно стабильной динамикой с повторяющейся сезонностью (см. *рис. 5.1.1*). Поскольку одной из главных целей работы является построение прогнозов на будущие периоды, для моделирования был выбран последний интервал времени с относительно стабильной динамикой: с октября 1998 г. по март 2002 г.

При проверке с помощью ADF-теста, добавляя при этом от 1 до 3 лагов согласно правилу  $T^{1/3}$  ( $T$  – количество наблюдений), гипотеза о наличии единичного корня в рассматриваемом временном ряде на данном интервале отвергается. При этом проверки указывают на то, что данный ряд является стационарным относительно тренда. Для того, чтобы учесть регулярные пики в декабре (выплата вознаграждений и премий по результатам года), июле (выплата отпускных) и снижение поступлений в январе (новогодние каникулы, невысокие поступления после декабря в условиях относительно высокой инфляции), были использованы фиктивные переменные, равные единице в соответствующие месяцы и нулю – в остальные месяцы года ( $DUM12$ ,  $DUM07$ ,  $DUM01$ ).

Используемые для расчетов переменные:

*INC* – реальные поступления подоходного налога.

*T* – линейный тренд.

*DUM12* – фиктивная переменная, равная единице в декабре и нулю во все остальные месяцы.

*DUM01* – фиктивная переменная, равная единице в январе и нулю во все остальные месяцы.

*DUM07* – фиктивная переменная, равная единице в июле и нулю во все остальные месяцы.



Рис. 5. 1.1

Таблица 5.1.1

**TS-модель поступлений по подоходному налогу (1998.10–2002.03)**

Dependent Variable: INC				
Method: Least Squares				
Sample: 1998:10 2002:03				
Included observations: 42				
Convergence achieved after 98 iterations				
Backcast: 1997:07 1998:09				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.368642	0.028271	-13.03962	0.0000
@TREND	0.008539	0.000266	32.12224	0.0000
DUM12	0.382276	0.015178	25.18601	0.0000
DUM01	-0.177080	0.016995	-10.41934	0.0000
DUM07	0.157104	0.018094	8.682851	0.0000
AR( 1)	-0.734847	0.112648	-6.523371	0.0000
AR( 3)	0.477936	0.085368	5.598534	0.0000
MA( 3)	-0.181721	0.081469	-2.230549	0.0326
SMA( 12)	0.885662	0.000118	7481.354	0.0000
R-squared	0.985717	Mean dependent var		0.532409
Adjusted R-squared	0.982254	S.D. dependent var		0.147497
S.E. of regression	0.019649	Akaike info criterion		-4.834204
Sum squared resid	0.012740	Schwarz criterion		-4.461846
Log likelihood	110.5183	F-statistic		284.6740
Durbin-Watson stat	1.906247	Prob(F-statistic)		0.000000

На основании значений выборочной автокорреляционной и частной автокорреляционной функций, а также критерия Шварца, в уравнение были добавлены авторегрессионные члены первого и третьего порядков и члены скользящего среднего с лагами в три и двенадцать периодов. Результаты LM-теста для первых 16 лагов не выявляют автокорреляции остатков.

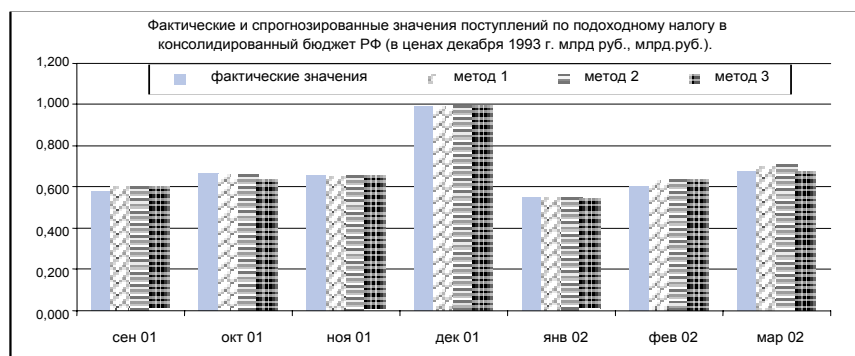


Рис. 5.1.2. *TS*-модель, фактические данные и одношаговые прогнозы

Как уже упоминалось выше, мы будем оценивать для временных рядов и *TS*, и *DS*-модели, и обращать внимание на качество прогнозов вне зависимости от результатов теста на единичный корень. Ниже приведены результаты оценок и прогнозирования поступлений подоходного налога по *DS*-модели.

Таблица 5.1.2

#### DS-модель поступлений подоходного налога (1998.10-2002.03)

Dependent Variable: D( INC)				
Method: Least Squares				
Sample( adjusted): 1998:10 2002:03				
Included observations: 42 after adjusting endpoints				
Convergence achieved after 22 iterations				
Backcast: 1998:06 1998:09				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.011317	0.004652	2.432575	0.0204
DUM01	-0.340414	0.027533	-12.36391	0.0000
DUM12	0.275983	0.024858	11.10259	0.0000
DUM07	0.069992	0.021531	3.250717	0.0026
D( INC( -1))	-0.349235	0.039997	-8.731509	0.0000
D( INC( -2))	-0.236336	0.038229	-6.182137	0.0000
D( INC( -8))	0.107250	0.033499	3.201599	0.0030
MA( 4)	-0.931761	0.046427	-20.06926	0.0000
R-squared	0.977238	Mean dependent var	0.008588	
Adjusted R-squared	0.972552	S.D. dependent var	0.167083	
S.E. of regression	0.027681	Akaike info criterion	-4.166466	
Sum squared resid	0.026053	Schwarz criterion	-3.835481	
Log likelihood	95.49579	F-statistic	208.5321	
Durbin-Watson stat	2.285610	Prob(F-statistic)	0.000000	

По результатам тестов для устранения коррелированности остатков в уравнение был добавлен член скользящего среднего четвертого порядка.

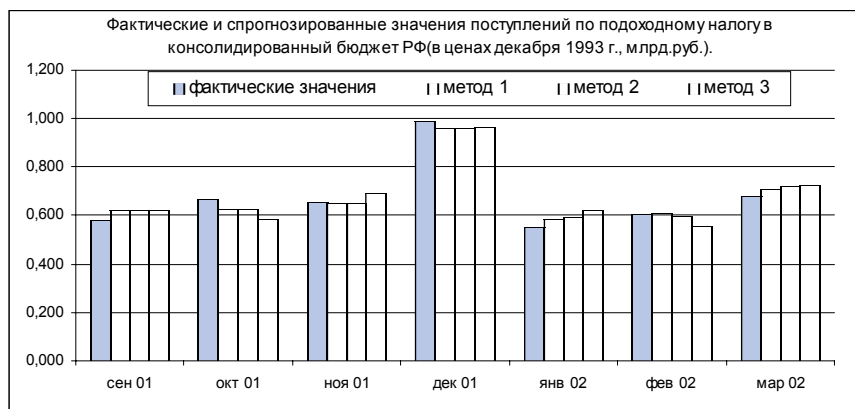


Рис. 5.1.3. DS-модель, фактические данные и одношаговые прогнозы

Судя по диаграммам и таблице, представленной ниже, качество прогноза зависит в основном от типа модели (*TS* или *DS*), а не от выбранного метода. Можно отметить, что *TS*-модели, в целом, дают более качественный одношаговый прогноз для поступлений по подоходному налогу. Наименьшая относительная ошибка отклонения получается при использовании первого метода, т.е. метода с переоценкой коэффициентов регрессии при добавлении новых данных.

Таблица 5.1.3

**Характеристики одношаговых прогнозов налоговых поступлений по подоходному налогу (в консолидированный бюджет) *TS*- и *DS*-моделей**

	Метод 1		Метод 2		Метод 3	
	<i>TS</i>	<i>DS</i>	<i>TS</i>	<i>DS</i>	<i>TS</i>	<i>DS</i>
Среднее квадратичное отклонение	0.018	0.030	0.020	0.033	0.019	0.053
Среднее абсолютное отклонение	0.013	0.026	0.014	0.029	0.014	0.050
Среднее абсолютное отклонение в %	2.004%	4.016%	2.260%	4.511%	2.227%	7.910%

## 5.2. Поступления налога на прибыль

Для моделирования налога на прибыль был выбран период с января 1999 г. до марта 2002 г. Динамика поступлений приведена на *рис. 5.2.1.*



*Рис. 5.2.1.*

Проверка гипотезы о наличии единичного корня указывает на то, что временной ряд поступлений налога на прибыль является нестационарным в уровнях и стационарным в первых разностях. Однако, как и для подоходного налога, для налога на прибыль мы оценивали и *TS*, и *DS*-модель, чтобы сравнить качество прогнозов рассматриваемыми тремя методами по обоим моделям. Моделирование поступлений налога на прибыль в уровнях (*TS*-модель) дает следующие результаты:

Используемые для расчетов переменные:

*PRF* – реальные поступления налога на прибыль в консолидированный бюджет.

*T* – тренд.

Анализ остатков показывает, что тренд нельзя исключить из уравнения даже несмотря на то, что коэффициент оказывается незначимым. Для учета сезонности в уравнение был добавлен авторегрессионный член *AR(12)*. Результаты *LM*-теста указывают на отсутствие автокорреляции остатков.



Таблица 5.2.1

**TS-модель**

Dependent Variable: PRF				
Method: Least Squares				
Sample: 1999:01 2002:03				
Included observations: 39				
Convergence not achieved after 100 iterations				
Backcast: 1997:10 1998:12				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.664341	4.472225	1.490162	0.1457
T	-0.039065	0.029266	-1.334835	0.1911
AR( 12)	0.646295	0.131007	4.933304	0.0000
MA( 2)	-0.004078	0.108964	-0.037428	0.9704
MA( 3)	0.140874	0.104210	1.351820	0.1856
SMA( 12)	0.885771	1.97E-05	44974.73	0.0000
R-squared	0.857029	Mean dependent var		1.073164
Adjusted R-squared	0.835367	S.D. dependent var		0.429261
S.E. of regression	0.174173	Akaike info criterion		-0.516902
Sum squared resid	1.001091	Schwarz criterion		-0.260969
Log likelihood	16.07958	F-statistic		39.56319
Durbin-Watson stat	1.600468	Prob(F-statistic)		0.000000

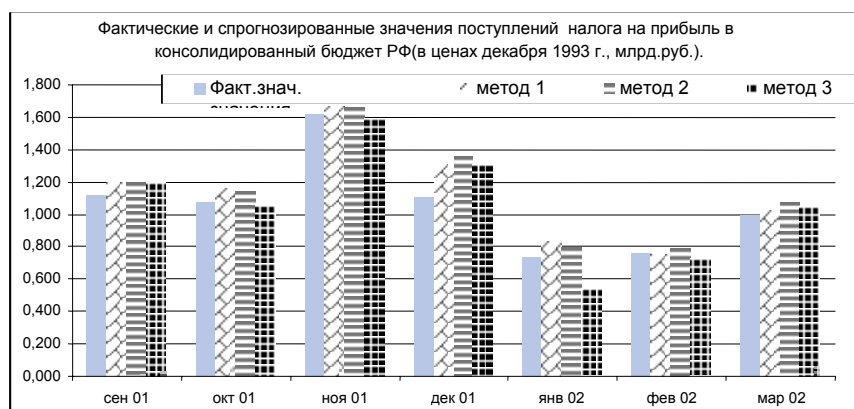


Рис. 5.2.2. TS-модель, фактические данные и одношаговые прогнозы

В соответствующей модели со стохастическим трендом мы также учитываем сезонность добавлением авторегрессионных членов и членов скользящего среднего двенадцатого порядка (как показали предварительные результаты, это позволяет существенно улучшить прогнозы, несмотря на то, что значительно сокращает количество степеней свободы в регрессии).

Таблица 5.2.2

**DS-модель**

Dependent Variable: D( PRF)				
Method: Least Squares				
Sample: 1999:01 2002:03				
Included observations: 39				
Convergence achieved after 99 iterations				
Backcast: 1997:11 1998:12				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.011705	0.041977	0.278847	0.7821
D( PRF( -12))	0.754360	0.125916	5.990986	0.0000
D( PRF( -1))	-0.364116	0.111278	-3.272130	0.0025
MA( 2)	-0.197575	0.100874	-1.958638	0.0584
SMA( 12)	0.885650	0.000230	3848.928	0.0000
R-squared	0.858105	Mean dependent var		0.003020
Adjusted R-squared	0.841411	S.D. dependent var		0.461479
S.E. of regression	0.183776	Akaike info criterion		-0.430991
Sum squared resid	1.148301	Schwarz criterion		-0.217713
Log likelihood	13.40432	F-statistic		51.40325
Durbin-Watson stat	2.095453	Prob(F-statistic)		0.000000

Для устранения автокорреляции остатков был также добавлен член MA(2).

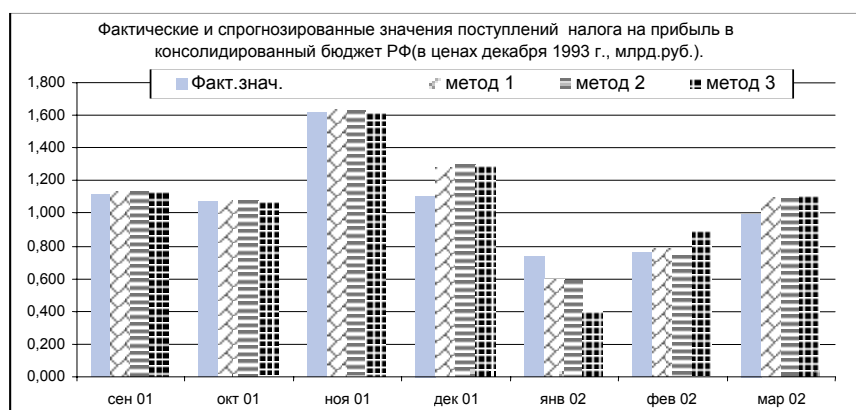


Рис. 5.2.3. DS-модель, фактические данные и одношаговые прогнозы

В случае налога на прибыль, так же как для подоходного налога, наименьшее относительное отклонение одношаговых прогнозов от фактических значений получается при использовании первого метода. В качестве прогнозов *TS* - и *DS*-моделей для поступлений налога на прибыль существенной разницы не наблюдается.

Таблица 5.2.3

**Характеристики одношаговых прогнозов поступлений налога на прибыль (в консолидированный бюджет) TS- и DS-моделей**

	Метод 1		Метод 2		Метод 3	
	TS	DS	TS	DS	TS	DS
Среднее квадратичное отклонение	0.101	0.093	0.112	0.098	0.114	0.158
Среднее абсолютное отклонение	0.080	0.069	0.087	0.069	0.085	0.111
Среднее абсолютное отклонение в %	7.736%	7.354%	8.361%	7.413%	9.134%	13.067%

### 5.3. Поступления налога на прибыль (в федеральный бюджет РФ)

Предварительные оценки показали, что удовлетворительную модель для временного ряда поступлений налога на прибыль в федеральный бюджет РФ удастся построить только на сокращенном интервале с декабря 1999 г. до марта 2002 г.



Рис. 5.3.1

Результаты ADF-теста указывают на то, что на рассматриваемом интервале процесс классифицируется как *TS*. На основании критерия Шварца и значений выборочной автокорреляционной и частной автокорреляционной функций, в уравнение с детерминированным трендом были добавлены члены *AR(1)* и *AR(2)*.

Используемые для расчетов переменные:

*FPRF* – реальные поступления налога на прибыль в федеральный бюджет.

Таблица 5.3.1

**TS-модель**

Dependent Variable: FPRF				
Method: Least Squares				
Sample: 1999:12 2002:03				
Included observations: 28				
Convergence achieved after 37 iterations				
Backcast: 1998:12 1999:11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.542277	0.028108	19.29241	0.0000
AR( 2)	-0.470293	0.180270	-2.608827	0.0154
AR( 1)	0.363175	0.180367	2.013534	0.0554
MA( 12)	0.885815	0.000130	6806.169	0.0000
R-squared	0.723397	Mean dependent var		0.510739
Adjusted R-squared	0.688822	S.D. dependent var		0.175689
S.E. of regression	0.098005	Akaike info criterion		-1.676026
Sum squared resid	0.230521	Schwarz criterion		-1.485711
Log likelihood	27.46437	F-statistic		20.92233
Durbin-Watson stat	1.730642	Prob(F-statistic)		0.000001

Тест множителей Лагранжа не выявил коррелированности остатков уравнения регрессии.

Ниже на рисунке приведены результаты прогнозирования поступлений налога на прибыль в федеральный бюджет на основе построенной модели тремя методами (с переоценкой модели при добавлении новых данных, без переоценки и с учетом ошибки предыдущего периода).

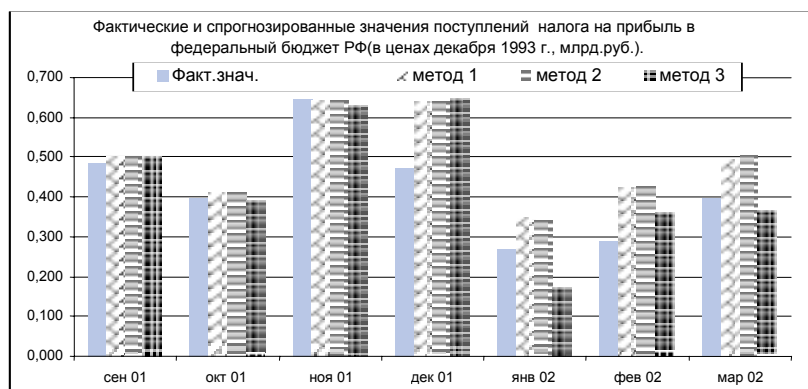


Рис. 5.3.2. TS-модель, фактические данные и одношаговые прогнозы

Аналогичным образом, не учитывая результаты ADF-теста, для сравнения прогнозов построим *DS*-модель, предполагая, что ряд поступлений имеет стохастический тренд. Результаты оценки соответствующей модели приведены ниже. В модели со стохастическим трендом автокорреляция остатков также не наблюдается.

Таблица 5.3.2

**DS-модель**

Dependent Variable: D( FPRF)				
Method: Least Squares				
Sample: 1998:12 2002:03				
Included observations: 40				
Convergence achieved after 6 iterations				
Backcast: 1997:12 1998:11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.004927	0.026075	0.188952	0.8512
D( FPRF( -2))	-0.455761	0.121944	-3.737456	0.0007
D( FPRF( -1))	-0.247316	0.120996	-2.043994	0.0485
D( FPRF( -10))	-0.426217	0.122114	-3.490320	0.0013
MA( 12)	0.871342	0.046472	18.74966	0.0000
R-squared	0.792363	Mean dependent var		0.003313
Adjusted R-squared	0.768633	S.D. dependent var		0.185712
S.E. of regression	0.089329	Akaike info criterion		-1.876520
Sum squared resid	0.279286	Schwarz criterion		-1.665410
Log likelihood	42.53040	F-statistic		33.39083
Durbin-Watson stat	2.094007	Prob(F-statistic)		0.000000

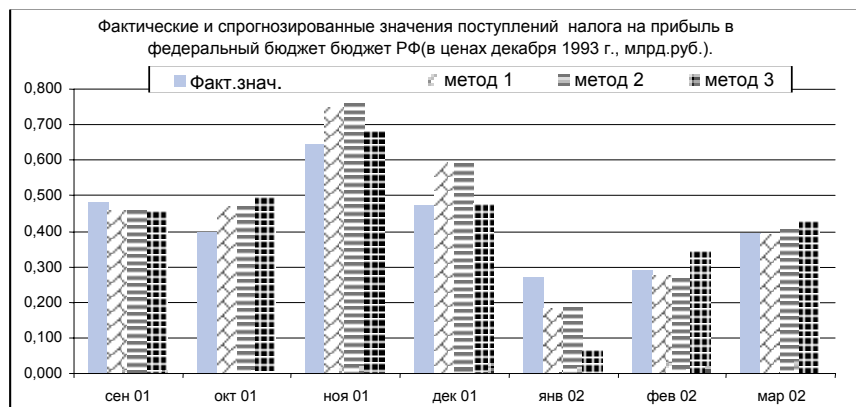


Рис. 5.3.3. *DS*-модель, фактические данные и одношаговые прогнозы

Среднее относительное отклонение спрогнозированных значений от фактических достаточно велико. Для моделей с детерминированным трендом (*TS*) оно достигает 20% при использовании первых двух методов прогнозирования и около 16% для третьего. Это можно объяснить тем, что при ограниченном количестве данных и нестабильной динамике поступлений третий метод позволяет компенсировать структурные изменения за счет коррекции текущих прогнозов на величину отклонения предыдущих. Но в случае стохастического тренда (*DS*-модель) это не так, более точный прогноз дают первый и второй методы. В целом можно отметить, что *DS*-модели, в данном случае, имеют относительно более качественные прогнозы.

Таблица 5.3.3

**Характеристики одношаговых прогнозов поступлений налога на прибыль (в федеральный бюджет) *TS*- и *DS*-моделей**

	Метод 1		Метод 2		Метод 3	
	<i>TS</i>	<i>DS</i>	<i>TS</i>	<i>DS</i>	<i>TS</i>	<i>DS</i>
Среднее квадратичное отклонение	0.095	0.073	0.096	0.076	0.081	0.090
Среднее абсолютное отклонение	0.074	0.060	0.074	0.064	0.059	0.066
Среднее абсолютное отклонение в %	20.600%	14.567%	20.602%	15.373%	16.028%	20.105%

#### 5.4. Поступления налога на добавленную стоимость

На выбранном интервале исследования (с декабря 1998 г. до января 2002 г.) данные поступлений налога на добавленную стоимость характеризуются относительно стабильной динамикой. Для корректировки пика поступлений в декабре 2001 г., обусловленного сокращением недоимки на сумму около 23 млрд руб., в уравнение добавлена соответствующая фиктивная переменная. Динамика поступлений НДС приведена на *рис. 5.4.1*.

Используемые для расчетов переменные:

*VAT* – реальные поступления НДС (в консолидированный бюджет РФ).

*DUM1201* – фиктивная переменная, равная единице в декабре 2001 г. и нулю во все остальные месяцы.

На основании результатов ADF-теста (с включением первых четырех лагов), мы отвергаем гипотезу о наличии единичного корня. Для учета сезонности в уравнение были добавлены авторегрессионный член и член скользящего среднего двенадцатого порядка. В результате моделирования поступлений НДС как стационарного ряда были получены следующие результаты – *табл. 5.4.1* :



Рис. 5.4.1

Таблица 5.4.1

**TS-модель**

Dependent Variable: VAT				
Method: Least Squares				
Sample: 1998:12 2002:01				
Included observations: 38				
Convergence achieved after 11 iterations				
Backcast: 1997:10 1998:11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.512976	0.060432	25.03621	0.0000
DUM1201	1.054938	0.118482	8.903755	0.0000
AR( 12)	0.280280	0.070583	3.970939	0.0004
MA( 2)	0.162753	0.052755	3.085045	0.0041
SMA( 12)	0.846364	0.040215	21.04593	0.0000
R-squared	0.916580	Mean dependent var		1.331455
Adjusted R-squared	0.906469	S.D. dependent var		0.349917
S.E. of regression	0.107015	Akaike info criterion		-1.509620
Sum squared resid	0.377921	Schwarz criterion		-1.294148
Log likelihood	33.68278	F-statistic		90.64726
Durbin-Watson stat	2.665606	Prob(F-statistic)		0.000000

Соответственно, результаты прогнозирования по *TS*-модели с использованием трех рассматриваемых в данной главе методов дают следующие значения:

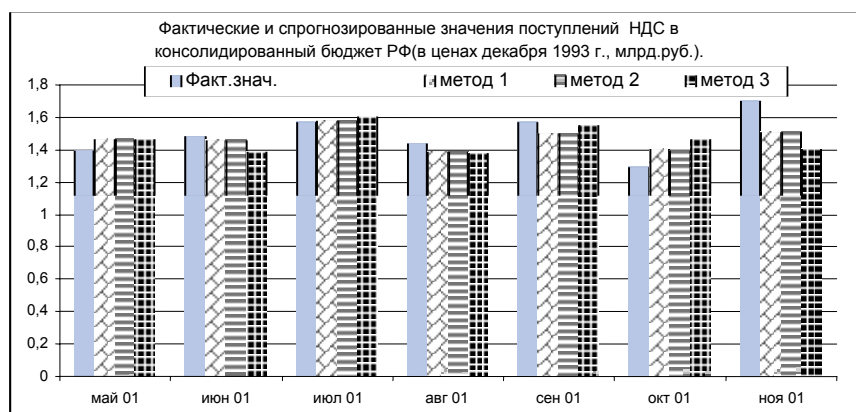


Рис. 5.4.2. *TS*-модель, фактические данные и одношаговые прогнозы

Аналогичным образом, для сравнения качества прогнозов по *DS*- и *TS*-моделям, несмотря на результаты теста на стационарность, проведем также оценивание, предполагая, что поступления НДС имеют стохастический тренд. В соответствующей модели со стохастическим трендом, для учета сезонности также был добавлен авторегрессионный член с лагом в двенадцать месяцев. В результате моделирования были получены следующие оценки:

Таблица 5.4.2

***DS*-модель**

Dependent Variable: D(VAT)				
Method: Least Squares				
Sample: 1998:12 2002:01				
Included observations: 38				
Convergence achieved after 11 iterations				
Backcast: 1997:09 1998:11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.007055	0.048276	-0.146148	0.8848
D(VAT(-1))	-0.767654	0.131004	-5.859796	0.0000
D(VAT(-12))	0.446102	0.108687	4.104456	0.0003
DUM1201	1.019370	0.172790	5.899465	0.0000
MA(1)	-0.073347	0.071738	-1.022434	0.3145
MA(3)	0.113705	0.056322	2.018840	0.0522
SMA(12)	0.880468	0.035290	24.94964	0.0000
R-squared	0.891457	Mean dependent var		0.012365
Adjusted R-squared	0.870448	S.D. dependent var		0.397558
S.E. of regression	0.143094	Akaike info criterion		-0.885803
Sum squared resid	0.634756	Schwarz criterion		-0.584142
Log likelihood	23.83025	F-statistic		42.43339
Durbin-Watson stat	2.462904	Prob(F-statistic)		0.000000



Несмотря на то, что коэффициент при  $MA(1)$  является незначимым, мы оставляем его в уравнении, так как в случае его исключения остатки оказываются автокоррелированными. Результаты прогнозирования по  $DS$ -модели приведены на рис. 5.4.3.

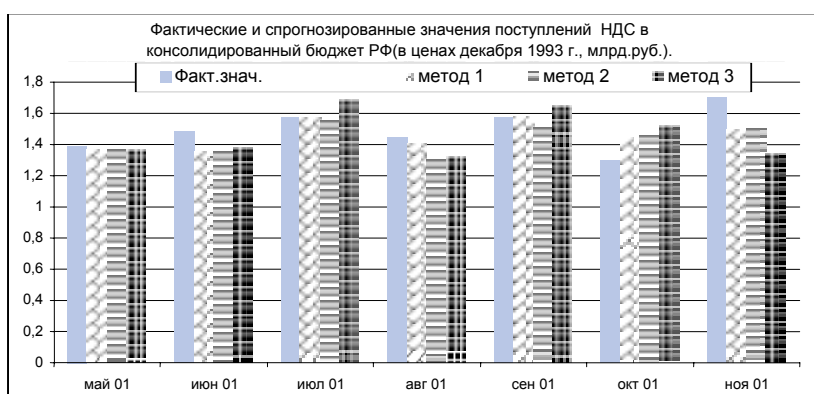


Рис. 5.4.3.  $DS$ -модель, фактические данные и одношаговые прогнозы

Так же, как и в случае подоходного налога, при относительно стабильной динамике поступлений наблюдаются небольшие отклонения спрогнозированных значений от фактических. Качество прогноза  $TS$ -моделей невысоко, но превосходит соответствующие прогнозы  $DS$ -моделей для всех методов прогнозирования. Как и в предыдущих рассмотренных случаях, наиболее точные прогнозы дают первые два метода.

Таблица 5.4.3

**Характеристики одношаговых прогнозов поступлений НДС (в консолидированный бюджет) для  $TS$ - и  $DS$ -моделей**

	Метод 1		Метод 2		Метод 3	
	TS	DS	TS	DS	TS	DS
Среднее квадратичное отклонение	0.097	0.113	0.094	0.126	0.141	0.184
Среднее абсолютное отклонение	0.078	0.081	0.076	0.107	0.108	0.151
Среднее абсолютное отклонение в %	5.204%	5.454%	5.046%	7.233%	7.191%	9.954%

## 5.5. Суммарные налоговые поступления

В качестве базового периода для моделирования суммарных налоговых поступлений в консолидированный бюджет РФ был выбран период с декабря 1998 г. по январь 2002 г. Динамика поступлений приведена на *рис. 5.5.1*.

Согласно результатам ADF-теста мы отвергаем гипотезу о наличии единичного корня, однако, как и ранее, будем оценивать обе модели (*TS* и *DS*) для сравнения качества прогнозов. Результаты оценок *TS*-модели приведены ниже. Для того, чтобы учесть наличие сезонности, в уравнение был добавлен авторегрессионный член двенадцатого порядка. Результаты LM-теста для построенной модели для первых шестнадцати лагов не выявили автокорреляции остатков.



*Рис. 5.5.1*

Используемые для расчетов переменные:

*TAX* – реальные суммарные налоговые поступления в консолидированный бюджет РФ.

*Таблица 5.5.1*

### *TS*-модель

Dependent Variable: TAX				
Method: Least Squares				
Sample: 1998:12 2002:01				
Included observations: 38				
Convergence achieved after 47 iterations				
Backcast: 1998:02 1998:11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.352150	0.923057	9.048361	0.0000
AR( 1)	0.669244	0.051148	13.08443	0.0000
SAR( 12)	0.728283	0.048402	15.04649	0.0000

Таблица 5.5.1. (продолжение)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA( 10)	-0.903911	0.000153	-5907.261	0.0000
R-squared	0.923898	Mean dependent var		4.839853
Adjusted R-squared	0.917183	S.D. dependent var		1.280565
S.E. of regression	0.368520	Akaike info criterion		0.940658
Sum squared resid	4.617439	Schwarz criterion		1.113035
Log likelihood	-13.87249	F-statistic		137.5895
Durbin-Watson stat	2.231670	Prob(F-statistic)		0.000000

Прогнозирование суммарных налоговых поступлений по построенной модели с использованием трех рассматриваемых методов дает следующие результаты:

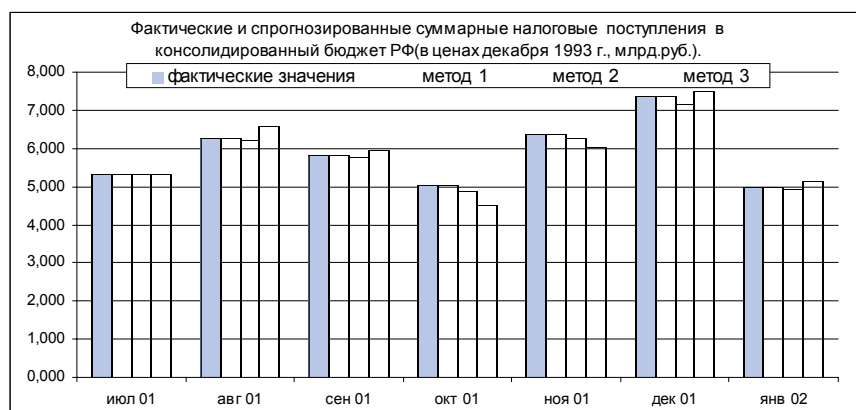


Рис. 5.5.2. *TS*-модель, фактические данные и одношаговые прогнозы

Оценки и построение прогнозов по *DS*-модели дают следующие результаты:

Таблица 5.5.2

***DS*-модель**

Dependent Variable: D( TAX)				
Method: Least Squares				
Sample: 1998:12 2002:01				
Included observations: 38				
Convergence achieved after 40 iterations				
Backcast: 1997:12 1998:11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.048226	0.118756	0.406092	0.6873
D( TAX(-2))	-0.271057	0.119323	-2.271633	0.0298

Таблица 5.5.2 (продолжение)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D( TAX( -12))	0.589665	0.096749	6.094792	0.0000
D( TAX( -1))	-0.187112	0.097694	-1.915290	0.0642
MA( 12)	0.885703	0.000101	8776.951	0.0000
R-squared	0.914548	Mean dependent var		0.034381
Adjusted R-squared	0.904190	S.D. dependent var		1.197624
S.E. of regression	0.370703	Akaike info criterion		0.975249
Sum squared resid	4.534889	Schwarz criterion		1.190721
Log likelihood	-13.52974	F-statistic		88.29506
Durbin-Watson stat	2.382760	Prob(F-statistic)		0.000000

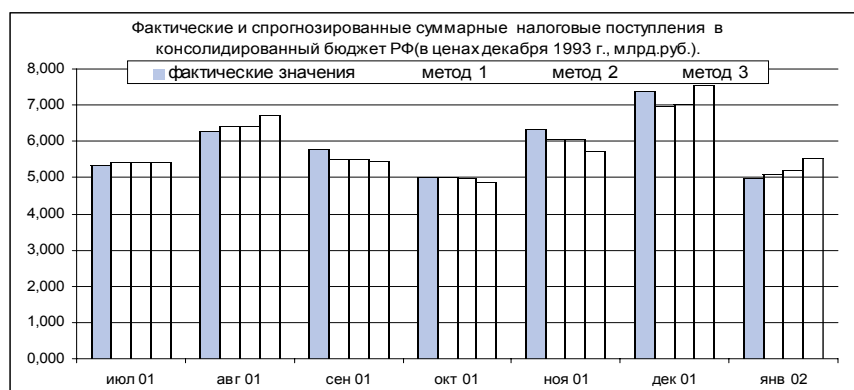


Рис. 5.5.3. *DS*-модель, фактические данные и одношаговые прогнозы

Использование намеренно ложной спецификации для моделирования временного ряда суммарных налоговых поступлений, в данном случае как процессов типа *DS*, не улучшает качество прогнозирования для всех рассмотренных методов. Как и для большинства рассмотренных выше временных рядов, наиболее точные одношаговые прогнозы суммарных налоговых поступлений дают первые два метода.

Таблица 5.5.3

**Характеристики одношаговых прогнозов суммарных налоговых поступлений (в консолидированный бюджет) для *TS*- и *DS*-моделей**

	Метод 1		Метод 2		Метод 3	
	<i>TS</i>	<i>DS</i>	<i>TS</i>	<i>DS</i>	<i>TS</i>	<i>DS</i>
Среднее квадратичное отклонение	0.301	0.343	0.288	0.351	0.386	0.499
Среднее абсолютное отклонение	0.270	0.301	0.276	0.305	0.345	0.406
Среднее абсолютное отклонение в %	5.014%	5.318%	4.881%	5.440%	6.167%	7.210%

## 5.6. Суммарные налоговые поступления (в федеральный бюджет РФ)

На основе анализа динамики суммарных налоговых поступлений в федеральный бюджет в качестве базового периода для моделирования был выбран период с декабря 1998 г. по январь 2002 г. Динамика поступлений приведена на *рис. 5.6.1*.

На рассматриваемом интервале согласно ADF-тесту мы отвергаем гипотезу о наличии единичного корня. Как показывают результаты оценок, временной ряд налоговых поступлений в федеральный бюджет является стационарным относительно тренда. Как и ранее, для моделирования будем использовать и *TS*-, и *DS*-спецификацию, сравнивая качество прогнозов.

Результаты оценки *TS*-модели приведены ниже. Для учета регулярных пиков поступлений, приходящихся на декабрь каждого года, в уравнение была добавлена соответствующая фиктивная переменная. Результаты LM-теста не выявляют автокорреляции остатков в построенной модели.



*Рис. 5.6.1*

Используемые для расчетов переменные:

*FTAX* – реальные суммарные налоговые поступления в федеральный бюджет РФ.

*T* – тренд.

*DUM12* – фиктивная переменная, равная единице в декабре и нулю во все остальные месяцы.

Таблица 5.6.1

**TS-модель**

Dependent Variable: FTAX				
Method: Least Squares				
Sample: 1998:12 2002:01				
Included observations: 38				
Convergence achieved after 41 iterations				
Backcast: 1997:11 1998:11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.765316	0.822514	-4.577812	0.0001
T	0.062886	0.008164	7.703082	0.0000
DUM12	1.011838	0.056569	17.88694	0.0000
MA( 1)	0.961698	0.040999	23.45688	0.0000
SMA( 12)	0.467951	0.131552	3.557158	0.0012
R-squared	0.926175	Mean dependent var		2.672254
Adjusted R-squared	0.917227	S.D. dependent var		0.906002
S.E. of regression	0.260660	Akaike info criterion		0.270881
Sum squared resid	2.242143	Schwarz criterion		0.486353
Log likelihood	-0.146738	F-statistic		103.5009
Durbin-Watson stat	1.579306	Prob( F-statistic)		0.000000

Оценки прогнозов суммарных налоговых поступлений в федеральный бюджет РФ по построенной модели дают следующие результаты:

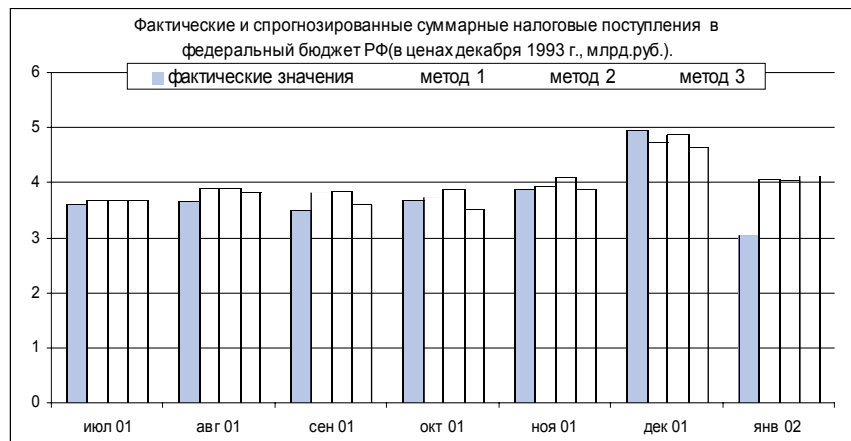


Рис. 5.6.2. TS-модель, фактические данные и одношаговые прогнозы

Оценки DS-модели дают следующие результаты (добавлены члены, учитывающие сезонность):

Таблица 5.6.2

**DS-модель**

Dependent Variable: D( FTAX)				
Method: Least Squares				
Sample( adjusted): 1999:02 2002:01				
Included observations: 36 after adjusting endpoints				
Convergence achieved after 9 iterations				
Backcast: 1998:02 1999:01				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.024307	0.051838	-0.468895	0.6422
D( FTAX(-12))	1.032298	0.129821	7.951703	0.0000
MA( 12)	-0.816168	0.059737	-13.66272	0.0000
R-squared	0.715528	Mean dependent var		0.050093
Adjusted R-squared	0.698287	S.D. dependent var		0.515622
S.E. of regression	0.283223	Akaike info criterion		0.394489
Sum squared resid	2.647099	Schwarz criterion		0.526449
Log likelihood	-4.100807	F-statistic		41.50216
Durbin-Watson stat	1.889166	Prob(F-statistic)		0.000000

На основании результатов LM-теста для первых шестнадцати лагов мы отвергаем гипотезу о наличии автокорреляции остатков.

Оценки прогнозов по *DS*-модели дают следующие результаты:

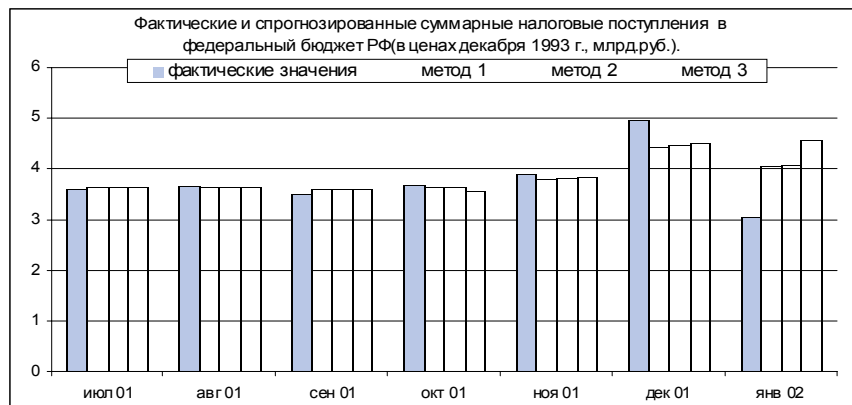


Рис. 5.6.3. *DS*-модель, фактические данные и одношаговые прогнозы

В данном случае, несмотря на определение исследуемого временного ряда как процесса со стационарным трендом, относительно более качест-

венные прогнозы для первых двух методов дают *DS*-модели. Для *TS*-модели третий метод дает более точные одношаговые прогнозы, чем в случае стохастического тренда.

Таблица 5.6.3

**Характеристики одношаговых прогнозов суммарных налоговых поступлений (в федеральный бюджет) для *TS*- и *DS*-моделей**

	Метод 1		Метод 2		Метод 3	
	TS	DS	TS	DS	TS	DS
Среднее квадратичное отклонение	0.426	0.428	0.425	0.429	0.427	0.600
Среднее абсолютное отклонение	0.289	0.254	0.315	0.246	0.267	0.323
Среднее абсолютное отклонение в %	8.526%	7.183%	9.315%	7.013%	7.852%	9.559%

**5.7. Сравнение результатов многошаговых прогнозов эконометрических моделей и модели оценки поступлений (Revenue Estimating Model)**

Для оценки качества многошаговых прогнозов произведем их расчет на основании построенных эконометрических моделей, при этом прогнозы на несколько шагов вперед рассчитываются с использованием спрогнозированных значений для прогноза на следующий период без переоценки коэффициентов уравнения.

Для оценки качества и сравнения результатов многошаговых прогнозов, полученных при помощи эконометрических методов, был также рассчитан многошаговый прогноз по модели оценки поступлений.

Модель оценки поступлений (Revenue Estimating Model, REM) – это модель-калькулятор для прогнозирования налоговых поступлений на основе информации о налоговых поступлениях за предыдущие месяцы. Расчет прогнозных значений в *REM*-модели проводится в постоянных ценах и основывается на значениях поступлений за соответствующий период базового года с учетом изменения возможных изменений ставок и базы налогов (если какие-либо изменения произошли, они учитываются простой корректировкой на соответствующий множитель). Кроме того, в модели осуществляется дополнительная корректировка прогнозных значений на относительное изменение поступлений текущего года по сравнению с предыду-



щим годом, причем, чем ближе к концу текущего года мы прогнозируем поступления, тем в большей степени (пропорционально больше значение весового коэффициента) прогноз опирается на информации о поступлениях текущего года по сравнению с поступлениями того же месяца прошлого года.

Для определения прогностических способностей рассматриваемых моделей также проведено сопоставление абсолютных и относительных отклонений многошаговых прогнозов ARIMA и REM-моделей от фактических данных.

### 5.7.1. Поступления подоходного налога

На рис. 5.7.1 приведены результаты прогнозирования (многошаговый прогноз) поступлений подоходного налога в консолидированный бюджет РФ на последние 9 месяцев 2001 г.

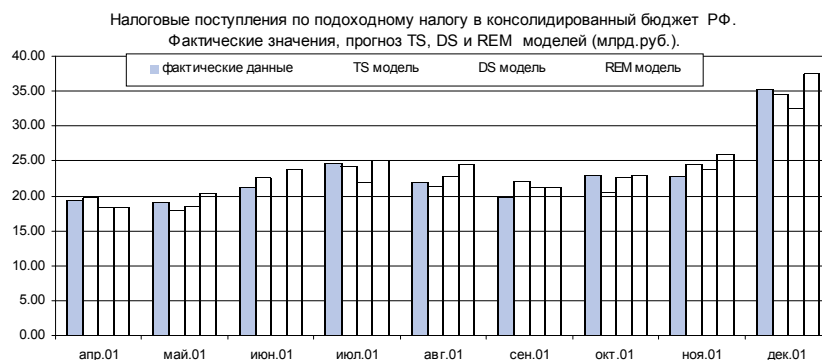


Рис. 5.7.1. Ретропрогноз TS-, DS- и REM- моделей на девять месяцев 2001 г.

Таблица 5.7.1

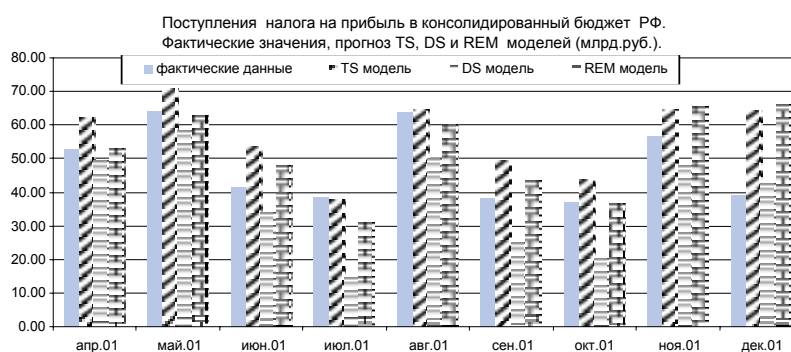
#### Характеристики многошаговых прогнозов поступлений подоходного налога для TS-, DS- и REM- моделей

	TS	DS	REM
Среднее квадратичное отклонение	1.468	1.552	1.976
Среднее абсолютное отклонение	1.239	1.325	1.684
Среднее абсолютное отклонение в %	5.696%	5.577%	7.447%

Для подоходного налога прогнозы, полученные с помощью эконометрических моделей, по своим качественным характеристикам близки между собой (несколько лучше *DS*-модель) и лучше модели оценки поступлений *REM*.

### 5.7.2. Поступления налога на прибыль

На *рис. 5.7.2* приведены результаты расчета прогнозных значений поступлений налога на прибыль в консолидированный бюджет РФ в последние 9 месяцев 2001 г. с использованием рассматриваемых *TS*-, *DS*- и *REM*-моделей.



*Рис. 5.7.2.* Ретропрогноз *TS*-, *DS*- и *REM*-моделей на девять месяцев 2001 г.

*Таблица 5.7.2*

#### Характеристики многошаговых прогнозов поступлений налога на прибыль (в консолидированный бюджет РФ) для *TS*-, *DS*- и *REM*-моделей

	<b>TS</b>	<b>DS</b>	<b>REM</b>
<b>Среднее квадратичное отклонение</b>	11.390	12.342	10.257
<b>Среднее абсолютное отклонение</b>	9.146	10.546	6.716
<b>Среднее абсолютное отклонение в %</b>	21.103%	24.339%	15.862%

Для поступлений налога на прибыль многошаговые прогнозы *TS*- и *DS*-моделей так же, как и рассмотренные выше одношаговые прогнозы, имеют достаточно большое отклонение от фактических данных. Из представленных моделей относительно наилучшими характеристиками прогнозов обладает *REM*-модель.

### 5.7.3. Поступления налога на прибыль (в федеральный бюджет РФ)

На рис. 5.7.3 приведены результаты расчета прогнозных значений поступлений налога на прибыль в федеральный бюджет РФ в последние 9 месяцев 2001 г. с использованием рассматриваемых *TS*-, *DS*- и *REM*-моделей.

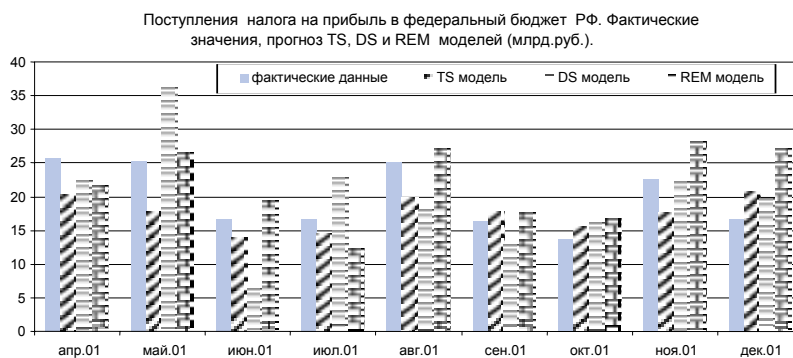


Рис. 5.7.3. Ретропрогноз *TS*-, *DS*- и *REM*-моделей на девять месяцев 2001 г.

Таблица 5.7.3

#### Характеристики многошаговых прогнозов поступлений налога на прибыль (в федеральный бюджет РФ) для *TS*-, *DS*- и *REM*-моделей

	TS	DS	REM
Среднее квадратичное отклонение	4.380	6.285	4.771
Среднее абсолютное отклонение	3.937	5.268	3.896
Среднее абсолютное отклонение в %	18.894%	27.052%	21.087%

Результаты многошагового прогнозирования поступлений налога на прибыль в федеральный бюджет РФ показывают, что, как и для поступлений в консолидированный бюджет, все представленные модели обладают достаточно большими отклонениями спрогнозированных значений. Однако в данном случае относительно более точный прогноз дает *TS*-модель.

#### 5.7.4. Поступления налога на добавленную стоимость

Ниже приведены результаты расчета прогнозных значений поступлений налога на добавленную стоимость в консолидированный бюджет РФ с использованием рассматриваемых *TS*-, *DS*- и *REM*-моделей.

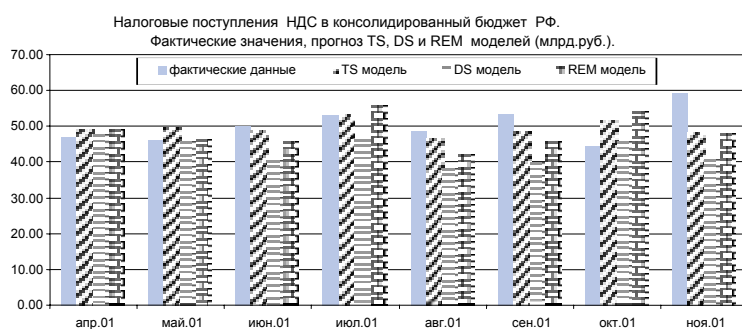


Рис. 5.7.4. Ретропрогноз *TS*-, *DS*- и *REM*-моделей на восемь месяцев 2001 г.

Таблица 5.7.4

#### Характеристики многошаговых прогнозов поступлений НДС (в консолидированный бюджет РФ) для *TS*-, *DS*- и *REM*-моделей

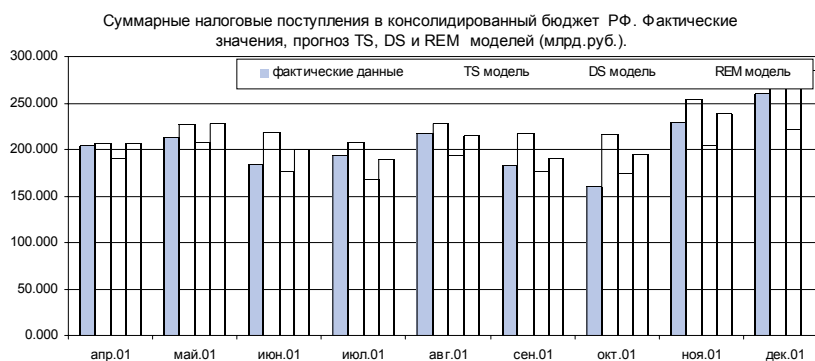
	<b>TS</b>	<b>DS</b>	<b>REM</b>
<b>Среднее квадратичное отклонение</b>	5.292	9.923	6.743
<b>Среднее абсолютное отклонение</b>	4.053	7.712	5.645
<b>Среднее абсолютное отклонение в %</b>	7.932%	14.484%	11.063%

В отличие от поступлений других налогов, для НДС многошаговое прогнозирование осуществлялось на временной интервал без декабря 2001 г. Это связано с тем, что резкий выброс поступлений в этот месяц, обусловленный значительным сокращением недоимки, приводит к значительной ошибке прогноза. По этой причине, а также чтобы значительная ошибка прогнозов в декабре 2001 г. не искажала сравнения моделей по остальным прогнозам, для ретропрогноза использовались 8 месяцев 2001 г.

Сравнение прогнозов показывает, что для поступлений налога на добавленную стоимость качество прогнозов *TS*-модели лучше остальных.

### 5.7.5. Суммарные налоговые поступления

На *рис. 5.7.5* приведены результаты расчета прогнозных значений суммарных налоговых поступлений в консолидированный бюджет РФ в последние 9 месяцев 2001 г. с использованием рассматриваемых *TS*-, *DS*- и *REM*-моделей.



*Рис. 5.7.5.* Ретропрогноз *TS*-, *DS*- и *REM*-моделей на девять месяцев 2001 г.

*Таблица 5.7.5*

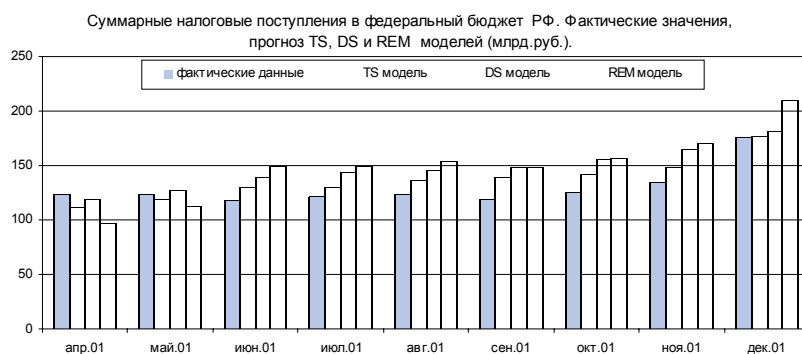
### Характеристики многошаговых прогнозов суммарных налоговых поступлений (в консолидированный бюджет) для *TS*-, *DS*- и *REM*-моделей

	<i>TS</i>	<i>DS</i>	<i>REM</i>
Среднее квадратичное отклонение	27.752	20.927	16.632
Среднее абсолютное отклонение	23.184	18.017	12.912
Среднее абсолютное отклонение в %	12.243%	8.508%	6.623%

Как следует из сводных результатов многошагового прогнозирования, для суммарных налоговых поступлений в консолидированный бюджет *REM*-модель дает существенно более качественный прогноз, чем эконометрические модели.

### 5.7.6 Суммарные налоговые поступления (в федеральный бюджет РФ)

На *рис. 5.7.6* приведены результаты расчета прогнозных значений суммарных налоговых поступлений в федеральный бюджет РФ в последние 9 месяцев 2001 г. с использованием рассматриваемых *TS*-, *DS*- и *REM*-моделей.



*Рис. 5.7.6.* Ретропрогноз *TS*-, *DS*- и *REM*-моделей на девять месяцев 2001 г.

*Таблица 5.7.6*

**Характеристики многошаговых прогнозов суммарных налоговых поступлений (в федеральный бюджет РФ) для *TS*-, *DS*- и *REM*-моделей**

	<b>TS</b>	<b>DS</b>	<b>REM</b>
<b>Среднее квадратичное отклонение</b>	12.513	21.198	29.258
<b>Среднее абсолютное отклонение</b>	11.242	18.445	28.439
<b>Среднее абсолютное отклонение в %</b>	9.080%	14.736%	22.074%

Для суммарных налоговых поступлений в федеральный бюджет прогноз эконометрической модели с детерминированным трендом обладает наименьшим средним относительным отклонением 9%. При этом, в отличие от поступлений в консолидированный бюджет РФ, *REM*-модель обладает наихудшими характеристиками прогноза среди рассматриваемых моделей.

## 5.8. Основные результаты и выводы

- Для временных рядов налоговых поступлений с относительно стабильной динамикой (таких, как: поступления подоходного налога, НДС и суммарные налоговые поступления в консолидированный бюджет) наиболее качественными характеристиками одношаговых прогнозов обладают *TS*-модели.
- Для временных рядов поступлений налога на прибыль и суммарных налоговых поступлений в федеральный бюджет более точными оказываются одношаговые прогнозы *DS*-моделей.
- Использование намеренно «ложной» классификации процесса по результатам теста на стационарность, в целом, не улучшает качество одношаговых прогнозов.
- Для временных рядов с относительно стабильной динамикой поступлений (поступления подоходного налога, НДС и суммарные налоговые поступления в консолидированный бюджет) качество одношаговых прогнозов, полученных с помощью рекурсивного метода (первый метод), существенно лучше третьего метода (корректирующего прогноз на ошибку предыдущего прогноза) и незначительно превосходит второй метод (не требующий переоценки коэффициентов уравнения).
- Для временных рядов поступлений подоходного налога, НДС и суммарных налоговых поступлений в федеральный бюджет наиболее предпочтительными моделями при многошаговом прогнозировании являются *TS*-модели. Среднее относительное отклонение спрогнозированных значений от фактических для них не превышает 10%.
- Для поступлений налога на прибыль относительное расхождение многошаговых прогнозов с фактическими данными, для всех рассматриваемых моделей, составляет около 20%.
- Для суммарных налоговых поступлений в консолидированный бюджет *REM*- модель дает лучший прогноз (средняя относительная ошибка составляет около 6%), чем эконометрические

модели (значения 14% и 8% для *TS*- и *DS*-моделей соответственно).

- Ошибки многошаговых прогнозов, полученных на основании *REM*- модели, являются более стабильным, и в целом равны средним ошибкам эконометрических моделей.



## ЧАСТЬ II. Структуры случайных векторов

### Глава 6. Прогнозирование с использованием структур случайных векторов

Во второй части работы рассматривается подход к построению прогнозов временных рядов, основанный на использовании структур случайных векторов. Данный подход, в отличие от традиционного эконометрического подхода, используемого в первой части, не предполагает каких-либо гипотез о конкретном функциональном виде случайной составляющей временного ряда, при этом расчет прогноза производится на основе имеющейся фактической информации о совместном распределении рассматриваемых случайных величин.

Кратко формулируются основные положения теории структур случайных векторов. Ставится задача прогнозирования временных рядов с использованием их лаговой структуры как структуры соответствующих случайных векторов и излагаются алгоритмы построения трендовых и вероятностных моделей и построения прогнозов с их использованием. Представлены результаты расчетов по десяти временным рядам: инфляция, денежные агрегаты  $M0$ ,  $M1$ ,  $M2$ , экспорт, импорт, безработица, налоговые доходы федерального бюджета, доходы федерального бюджета и ВВП. В Заключении приводятся краткие выводы и излагаются пути совершенствования предложенных моделей. В Приложениях приведены макросы (программы) построения моделей и прогнозов на их основе, а также исходные данные для расчетов.

#### 6.1. Предварительные замечания

##### *6.1.1. Непосредственные связи между элементами статистической системы*

Для изучения закономерностей механизма, определяющего состояние и поведение многих сложных систем, обычно имеются лишь результаты наблюдений за реализациями состояния и поведения системы. В связи с этим часто бывает целесообразно моделировать изучаемую систему некоторым

случайным вектором, предполагая, что наблюдаемые значения характеристик системы являются его реализациями.

Полное и в то же время достаточно сжатое описание статистической системы (системы, моделируемой случайным вектором) полностью определяется совместной функцией распределения вероятностей компонент индуцируемой ею случайной многомерной величины. Априорные гипотезы, выделяющие класс статистических систем, могут отражать как содержательные свойства систем, входящих в этот класс, так и формальные особенности случайных векторов, моделирующих системы класса.

Количественному исследованию сложной системы обычно предшествует качественный содержательный анализ, позволяющий сформулировать некоторую априорную гипотезу о специфике изучаемой системы. Первым этапом такого анализа является этап идентификации системы. В хорошо идентифицированной системе не должно быть элементов, не зависящих друг от друга, хотя бы через посредство других элементов системы. В противном случае система может быть разбита на независимые подсистемы, которые исследуются по отдельности. В то же время, если все элементы системы взаимосвязаны, то система оказывается недоидентифицирована. В связи с этим возникает важная проблема определения непосредственных и опосредованных связей между элементами системы.

Естественно считать, что элементы сложной системы взаимосвязаны, если изменение состояния одного элемента с необходимостью влечет за собой изменение состояния другого. Однако этому качественному представлению о взаимосвязи можно поставить в соответствие множество различных формальных определений.

На практике часто в качестве характеристики связи между элементами статистической системы используют корреляционную матрицу соответствующего случайного вектора. Наличие или отсутствие корреляции между переменными интерпретируется как наличие или отсутствие связи между ними. На самом деле коррелированность компонент случайного вектора может означать лишь то, что соответствующие элементы статистической системы связаны одновременно с одним или несколькими другими ее элементами.

Если каждому элементу системы поставить в соответствие вершину помеченного графа, ребра которого отвечают наличию непосредственных связей между элементами, отвечающими вершинам, инцидентным этим ребрам, то получится очень наглядное представление непосредственных связей в системе. Этот граф называется *графом непосредственных связей* или *структурой* изучаемой системы.

Пусть, например, рассматривается система  $X = \{x_1, x_2, x_3, x_4, x_5\}$ , состоящая из пяти элементов. Пусть эта система имеет структуру  $\Gamma$ , изображенную на рис. 6.1.1. В этом случае переменная  $x_1$  непосредственно связана с переменными  $x_2, x_3, x_4$ ;  $x_2$  с переменными  $x_1, x_4$ ;  $x_3$  с переменными  $x_1, x_4$ ;  $x_4$  с переменными  $x_1, x_2, x_3, x_5$ ;  $x_5$  с переменной  $x_4$ . Таким образом, если интерес представляет только поведение второго и пятого элементов системы в зависимости от остальных, то можно вместо одной 5-мерной системы  $\{x_1, x_2, x_3, x_4, x_5\}$  рассмотреть две подсистемы  $\{x_1, x_2, x_3, x_4\}$  и  $\{x_4, x_5\}$  меньшей размерности.

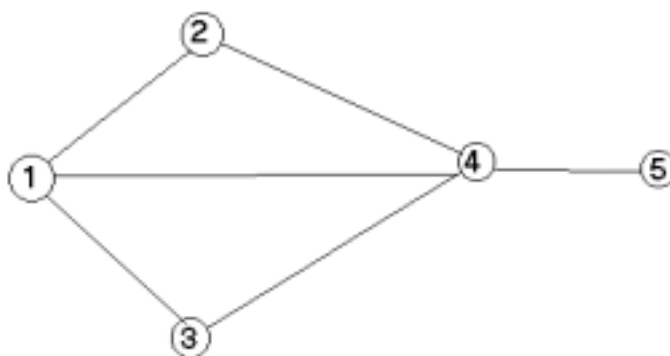


Рис. 6.1.1. Структура системы  $X = \{x_1, x_2, x_3, x_4, x_5\}$

### 6.1.2. Структуры случайных векторов

Структура статистической системы (случайного вектора, моделирующего статистическую систему) – достаточно широкое понятие. Отличаясь высокой наглядностью, структуры очень полезны при качественном анализе сложных статистических систем<sup>3</sup>. Выделяя для каждого элемента системы группу элементов, непосредственно с ним связанных, структуры позволяют существенно упростить и количественный анализ, снижая размерность рассматриваемой в каждый момент системы. В связи с этим возникает актуальная задача – задача построения структуры по статистической информации (обычно весьма ограниченной) о системе. К сожалению, в общем случае

<sup>3</sup> См.: Гаврилец Ю.Н. Социально-экономическое планирование (системы и модели). М.: Экономика, 1974.

построение структуры статистической системы представляет собой весьма трудоемкую работу (вычислительная сложность данной задачи – башня экспонент).

Не будем приводить строгого формального определения структуры случайного вектора, а приведем лишь критерий для проверки того, что заданный граф является структурой данного случайного вектора. Обозначим через  $\Gamma(i)$  множество вершин графа  $\Gamma$ , смежных с вершиной  $i$  (т.е. множество вершин, соединенных ребром с вершиной  $i$ ). Справедливо следующее утверждение<sup>4</sup>.

**Теорема.** Помеченный граф  $\Gamma$  является структурой  $n$ -мерной случайной величины  $\xi$  с совместным распределением вероятностей  $p(x) = p(x_1, \dots, x_n)$  тогда и только тогда, когда для любых подмножеств  $A$  и  $B$  множества индексов  $I = \{1, \dots, n\}$  таких, что  $\Gamma(A) \subset B \subset I \setminus A$ , выполняется соотношение

$$p(x_A | x_{I \setminus A}) = p(x_A | x_B) = p(x_A | x_{\Gamma(A)})$$

где  $\Gamma(A) = \bigcup_{i \in A} \Gamma(i) \setminus A$ , т.е.  $\Gamma(A)$  – объединение всех индексов связанных с индексами из множества  $A$  и не входящих в него;  $x_A$  – подвектор вектора  $x$  с индексами из  $A$ .

В рассмотренном выше примере (рис. 6.1.1)  $\Gamma(1) = \{2, 3, 4\}$ ;  $\Gamma(2) = \{1, 4\}$ ;  $\Gamma(3) = \{1, 4\}$ ;  $\Gamma(4) = \{1, 2, 3, 5\}$ ;  $\Gamma(5) = \{4\}$ ; а, например,  $\Gamma(1, 2) = \{2, 3, 4\} \cup \{1, 4\} \setminus \{1, 2\} = \{3, 4\}$ .

Для практических расчетов и оценок взаимосвязей между элементами статистических систем использование многомерных распределений вероятностей не конструктивно. В связи с этим целесообразно ввести эквивалентное определение структуры случайного вектора в терминах теории информации<sup>5</sup>. Понятия «энтропия» и «количество информации» представляют собой естественные меры неопределенности и зависимости компонент многомерной случайной величины.

Энтропия дискретной случайной величины  $\xi$  (т.е. случайной величины, принимающей не более, чем счетное число значений), принимаю-

<sup>4</sup> Юдин А.Д. Структуры наборов псевдонезависимых случайных величин // Модели и методы исследования социально-экономических процессов. М.: ЦЭМИ АН СССР, 1976.

<sup>5</sup> См., например: Яглом А.М., Яглом И.М. Вероятность и информация. М.: Наука, 1973.

щей значения  $x \in X$  с вероятностями  $p(x)$ , равна

$$H(\xi) = - \sum_{x \in X} p(x) \log_2 p(x)$$

Количество информации о случайной величине  $\xi_A$ , содержащееся в случайном векторе  $\xi_B$ , определяется как

$$I(\xi_A, \xi_B) = H(\xi_A) + H(\xi_B) - H(\xi_A, \xi_B)$$

При этом энтропия и количество информации измеряются в битах.

В силу свойств количества информации, приведенную теорему можно сформулировать в более конструктивном виде.

**Теорема.** *Помеченный граф  $\Gamma$  является структурой  $n$ -мерной случайной величины  $\xi$  с совместным распределением вероятностей  $p(x) = p(x_1, \dots, x_n)$  тогда и только тогда, когда для любых подмножеств  $A$  и  $B$  множества индексов  $I = \{1, \dots, n\}$  таких, что  $\Gamma(A) \subset B \subset I \setminus A$ , выполняется соотношение*

$$I(\xi_A, \xi_{I \setminus A}) = I(\xi_A, \xi_B) = I(\xi_A, \xi_{\Gamma(A)}),$$

где

$$\Gamma(A) = \bigcup_{i \in A} \Gamma(i) \setminus A.$$

Знание структуры случайного вектора, индуцированного изучаемой системой, позволяет существенно упростить ее качественный анализ. В частности, по известной структуре можно построить рекурсивную систему<sup>6</sup>, что дает возможность упростить также и количественный анализ исследуемой системы. Однако, как уже указывалось, построение структур случайных векторов в общем случае является весьма трудоемкой задачей.

### 6.1.3. Существенная размерность и информативные структуры

При известной структуре случайного вектора, можно построить большое количество различных рекурсивных систем, определяемых перестановкой индексов компонент исследуемой системы. Для любой перестановки

---

<sup>6</sup> Рекурсивная система-система, в которой переменные можно упорядочить таким образом, чтобы переменная с большим индексом зависела только от переменных с меньшими индексами.

$\{\alpha_i\}_{i=1}^n$  множества индексов  $I = \{1, \dots, n\}$  справедливо

$$p(x) = \prod_{i=1}^n p(x_{\alpha_i} | x_{\Delta(\alpha_i)}),$$

где  $\Delta(\alpha_i) = \Gamma(\alpha_i) \cap \{\alpha_1, \dots, \alpha_{i-1}\}$ .

При этом возникают две задачи:

- определение перестановки индексов, при которой размерность сомножителей композиции, задающей совместное распределение вероятностей случайного вектора, минимальна;
- оценка соответствующих частных распределений вероятностей.

Существенной размерностью  $n$ -мерного случайного вектора с совместным распределением вероятностей  $p(x)$  называется минимальное число  $k$ , для которого существует такая перестановка индексов  $\{\alpha_i\}_{i=1}^n$ , что соответствующая декомпозиция распределения вероятностей  $p(x)$  имеет размерность сомножителей не выше  $k$ .

В силу того, что в практических исследованиях обычно имеется весьма ограниченный статистический материал, с достаточной степенью достоверности можно строить совместные распределения вероятностей лишь очень невысокой размерности. Таким образом, становится актуальной проблема восстановления по имеющимся данным совместного распределения вероятностей случайного вектора невысокой существенной размерности наилучшим образом (в некотором заранее заданном смысле) отражающем имеющуюся информацию.

Ниже в качестве характеристики непосредственных связей статистических систем будут рассматриваться информативные структуры этих систем. Грубо говоря<sup>7</sup>, информативные структуры  $k$ -го порядка — это структуры случайных векторов  $k$ -существенной размерности, ближайших в смысле минимума различающей информации к модели исследуемой статистической системы. Другими словами, информативная структура  $k$ -го порядка  $n$ -мерного случайного вектора — это структура другого  $n$ -мерного случайного вектора, совместное распределение вероятностей которого удовлетворяет следующим свойствам. Оно является, во-первых, композицией  $k$ -мерных

<sup>7</sup> Подробнее о существенной размерности случайных векторов см.: Юдин А.Д. О выделении существенных связей в многомерной случайной величине // Модели социально-экономических процессов и социальное планирование. М.: Наука, 1979.

частных распределений вероятностей исходного случайного вектора, и, во-вторых, ближайшим в смысле минимума различающей информации к неизвестному распределению вероятностей случайного вектора, моделирующего изучаемую статистическую систему. Соответствующее распределение вероятностей называется информативной аппроксимацией  $k$ -го порядка истинного распределения вероятностей. Все  $k$ -мерные маргинальные распределения предполагаются известными или же их можно построить по имеющемуся статистическому материалу с достаточной степенью достоверности и надежности<sup>8</sup>.

#### **6.1.4. Прогнозирование с использованием структур случайных векторов**

Обычно при качественном анализе сложной системы всю рассматриваемую систему показателей можно разбить на две принципиально различные группы показателей: зависимые переменные, поведение которых необходимо исследовать, и факторные показатели, определяющих в значительной мере поведение первой группы показателей. В свою очередь, факторные показатели также подразделяются на две подгруппы: управляемые переменные, на которые возможно непосредственное управленческое воздействие, и наблюдаемые характеристики среды, непосредственно воздействовать на которые нельзя.

Знание структуры системы (информативной структуры) позволяет построить рекурсивную систему с известным распределением вероятностей:

$$x_i = f_i(x_{\Delta(i)}) \quad i = 1, \dots, n,$$

где  $\Delta(i)$  – множество индексов, определяемых для каждого показателя выявленными информативными связями. В качестве функций  $f_i(x_{\Delta(i)})$ , в соответствии с которыми осуществляется прогноз, можно взять обычные линейные регрессии или, что, вообще говоря, должно приводить к лучшему результату, условные математические ожидания по полученным при построении информативных структур частным распределениям вероятностей. Линейные регрессии дают наилучший результат в случае, когда индуцируе-

---

<sup>8</sup> Понятие информативной структуры было введено в работе: Юдин А.Д. Об информативных структурах многомерных случайных величин // Известия АН СССР. Техническая кибернетика, 1977, № 6. Подробнее см., например, Юдин Д.Б., Юдин А.Д. Математики измеряют сложность. М.: Знание, 1985 (сер. «Число и мысль» Вып. 8).

мый рассматриваемой системой случайный вектор является нормально распределенным<sup>9</sup>.

При использовании второго подхода для построения прогноза фактически применяется исходное определение регрессии как условного математического ожидания. В этом случае функции прогноза выглядят следующим образом:

$$\tilde{x}_i = f_i(x_{j(i)}, x_{k(i)}) = \sum_{x_i} x_i P\{X_i = x_i | X_{\Delta(i)} = x_{\Delta(i)}\} \quad i = 1, \dots, n,$$

где множество  $X_i$  – множество возможных значений  $i$ -го показателя системы,  $P\{X_i = x_i | X_{\Delta(i)} = x_{\Delta(i)}\}$  – условное распределение вероятностей того, что  $i$ -й показатель примет значение равное  $x_i$  при том, что  $X_{\Delta(i)} = x_{\Delta(i)}$ . Поскольку ниже мы рассматриваем только информативные структуры третьего порядка, то все множества  $\Delta(i)$  содержат по два элемента:  $\Delta(i) = \{j(i), k(i)\}$ .

## 6.2. Постановка проблемы и алгоритмы

### 6.2.1. Постановка проблемы

#### 6.2.1.1. Трендовый прогноз

Будем рассматривать произвольный временной ряд  $x_t$  (в общем случае векторный, размерности  $n$ ) как реализацию суммы двух процессов: детерминированного тренда, отражающего внутренние закономерности исследуемого процесса, и стохастического тренда некоторого случайного процесса, отражающего случайные отклонения от детерминированного тренда, а также неучтенные в детерминированном тренде факторы. Другими словами, будем считать, что

$$x_t = f(t) + \xi(y_{t-1}, y_{t-2}, \dots, y_{t-p} \dots),$$

где  $f$  – детерминированная функция (в общем случае вектор-функция), зависящая от времени;  $y_t = x_t - f(t)$  – отклонение значения временного ряда в момент времени  $t$  от тренда;  $\xi$  – случайная функция (в общем случае случайная вектор-функция), зависящая от предыстории отклонений временного ряда от тренда.

<sup>9</sup> См.: Заруцкий В.И. О выделении некоторых графов связей для нормальных векторов большой размерности // Алгоритмическое и программное обеспечение прикладного статистического анализа. М.: Наука, 1980.



Для построения функции прогноза значения ряда в момент времени  $t + 1$  по известным значениям ряда  $x_t, x_{t-1}, \dots, x_{t-\tau}, \dots$  достаточно знать функцию тренда  $f(t)$  и условное распределение вероятностей  $p(y_{t+1} = y | y_t, y_{t-1}, \dots, y_{t-\tau}, \dots)$ . В этом случае наилучший прогноз представляет собой сумму значения тренда в момент времени  $t + 1$  и условного математического ожидания реализации случайной величины при известных значениях  $x_t, x_{t-1}, \dots, x_{t-\tau}, \dots$ , т.е.

$$\tilde{x}_{t+1} = f(t+1) + \int_{y \in Y} yp(y_{t+1} = y | y_t, y_{t-1}, \dots, y_{t-\tau}, \dots) dy.$$

Таким образом, если известна структура случайного вектора  $\xi$ , то оптимальный прогноз величины  $x_{t+1}$  задается выражением

$$\tilde{x}_{i,t+1} = f(t+1) + \int_{y \in Y_i} yp(y_{i,t+1} = y | y_{\Gamma(t,i)}) dy, i = 1, \dots, n,$$

где  $\Gamma$  – структура случайного вектора  $\xi$ . (Напомним, что структурой случайного вектора называется граф, ребра которого отвечают наличию непосредственных связей между соответствующими компонентами случайного вектора. Таким образом, множество  $\Gamma(t, i)$  – множество переменных с соответствующими лагами, от которых непосредственно зависит переменная  $x_i$ .

В настоящей работе мы рассматриваем только скалярные временные ряды. Поэтому

$$\tilde{x}_{t+1} = \varphi(t+1 | \Gamma(t)) = f(t+1) + \int_{x \in X} xp(x_{t+1} = x | x_{\Gamma(t)}) dx. \quad (6.2.1)$$

Следовательно, для построения функции прогнозирования указанного вида необходимо произвести следующие действия:

1. Определить функциональный вид тренда.
2. Построить тренд.
3. Построить ряд отклонений от тренда.
4. Построить структуру (информативную структуру) лаговых показателей отклонений от тренда.
5. Построить условное распределение вероятностей отклонений от тренда в зависимости от выявленных лаговых отклонений.

Прогноз, построенный с использованием предлагаемого здесь подхода, основанного на выявлении структуры случайных отклонений от детерминированного тренда временного ряда, будем называть для краткости *трендовым прогнозом*.

### 6.2.1.2. Вероятностный прогноз

Рассмотрим еще один подход к построению прогноза, также основанный на выявлении структуры случайного вектора. Однако в этом случае предполагается, что в исходном временном ряде  $x_t$  отсутствует детерминированная составляющая (детерминированный тренд). Для краткости будем называть прогноз, построенный на основе этого подхода, *вероятностным прогнозом*.

Таким образом, в данном случае будем рассматривать произвольный временной ряд  $x_t$  (в общем случае векторный) как реализацию стохастического тренда некоторого случайного процесса, отражающего случайные отклонения от некоторого среднего значения. Другими словами, будем считать, что

$$x_t = \xi(x_{t-1}, x_{t-2}, \dots, x_{t-p} \dots),$$

где  $\xi$  случайная функция (в общем случае случайная вектор функция), зависящая от предыстории.

В рассматриваемом случае для построения функции прогноза значения ряда в момент времени  $t+1$  по известным значениям ряда  $x_p, x_{t-1}, \dots, x_{t-p} \dots$  достаточно знать условное распределение вероятностей  $p(x_{t+1} = x | x_p, x_{t-1}, \dots, x_{t-p} \dots)$ . Наилучший прогноз представляет собой в этой ситуации условное математическое ожидание реализации случайной величины  $\zeta$  при известных значениях  $x_p, x_{t-1}, \dots, x_{t-p} \dots$ , т.е.

$$\hat{x}_{t+1} = \int_{x \in X} xp(x_{t+1} = x | x_t, x_{t-1}, \dots, x_{t-p}, \dots) dx.$$

Однако достаточно часто ряд  $x_t$  имеет ярко выраженную тенденцию. В таких ситуациях данные следует детрендить. Будем для этого переходить к темповым характеристикам ряда<sup>10</sup>. Введем новый ряд:

$$y_t = \frac{x_{t+1}}{x_t}, t = 1, \dots, T-1.$$

Будем считать значения ряда  $y_t$  реализациями случайного вектора  $\zeta$ . Таким образом, если известна структура случайного вектора  $\zeta$ , то оптимальный прогноз величины  $y_{t+1}$  задается выражением

$$\tilde{y}_{i,t+1} = \int_{y \in Y_i} yp(y_{i,t+1} = y | y_{\Delta(i,t)}) dy, i = 1, \dots, n,$$

где  $\Delta$  структура случайного вектора  $\zeta$ . Таким образом, прогноз  $x_{t+1}$  определяется соотношением

<sup>10</sup> Такой способ детрендинга эквивалентен детрендингованию ряда логарифмов.

$$\tilde{x}_{i,t+1} = x_t \int_{y \in Y_i} yp(y_{i,t+1} = y | y_{\Delta(t,i)}) dy, i = 1, \dots, n.$$

Поскольку мы рассматриваем только скалярные временные ряды, то

$$\mathfrak{E}_{t+1} = \psi(\xi + 1 | \Delta(t)) = x_t \int_{y \in Y} yp(y_{t+1} = y | y_{\Delta(t)}) dy. \quad (6.2.2)$$

Следовательно, для построения функции прогнозирования указанного вида необходимо произвести следующие действия:

1. Рассчитать темповые характеристики (если нужно) ряда.
2. Построить структуру (информативную структуру) лаговых показателей темповых характеристик ряда.
3. Построить условное распределение вероятностей темповых характеристик ряда в зависимости от выявленных лаговых характеристик.

### 6.2.1.3. Использование прогнозирующих функций

Полученная в результате проведенных действий функция прогнозирования  $\varphi(t | \Gamma(t))$  вида (6.2.1) или  $\psi(t | \Delta(t))$  (6.2.2) и является прогнозирующей функцией исходного временного ряда  $x_t$ . Для проверки построенной функции прогнозирования производится имитация исходного временного ряда в соответствии с соотношением

$$\mathfrak{E}_{t+1} = \begin{cases} \mathfrak{E}_{t+1}^{(1)} = \varphi(\xi + 1 | \Gamma(t)) \\ \mathfrak{E}_{t+1}^{(2)} = \psi(\xi + 1 | \Delta(t)) \end{cases} t = \tau, \dots, T-1,$$

где  $T$  – длина ряда  $x_t$ , использованного для построения функции прогнозирования, а  $\tau$  – максимальный лаг, входящий в множества  $\Gamma$  и  $\Delta$ . Качество имитации будем оценивать величиной средней относительной ошибки, точнее, величиной средней абсолютной процентной ошибки (MAPE – Mean Absolute Percent Error):

$$\delta = \frac{100}{T - \tau + 1} \sum_{t=\tau}^T \frac{|x_t - \mathfrak{E}_t|}{x_t}.$$

После проверки качества имитации строятся одношаговые прогнозы развития временного ряда  $x_t$ : три с использованием функции  $\varphi(t | \Gamma(t))$  и два с использованием функции  $\psi(t | \Delta(t))$ . Качество прогнозов оценивается, так же как и качество имитации, величиной средней относительной ошибки (MAPE)

$$\mathcal{E} = \frac{100}{\mathcal{F}} \sum_{t=1}^{\mathcal{F}} \frac{|x_t - \mathcal{E}_t|}{x_t} .$$

Здесь  $\mathcal{F}$  – горизонт прогнозирования.

Рассмотрим пять методов прогнозирования.

1. Трендовый прогноз без обучения. В этом случае и функция тренда  $f(t)$ , и условное распределение вероятностей  $p(y_{t+1} = y | y_{\Gamma(t)})$  случайного вектора  $\xi$  строятся на основе отклонений от тренда исходного временного ряда  $\{y_t\}_{t=1}^T$ . После этого с использованием полученных функций строится последовательность одношаговых прогнозов значений временного ряда  $\{\mathcal{E}_{T+t}\}_{t=1}^{\mathcal{F}}$ , зависящих от истинных значений  $\{x_{T+t}\}_{t=0}^{\mathcal{F}-1}$ .
2. Трендовый прогноз с обучением без коррекции тренда. Этот случай отличается от предыдущего тем, что на каждом шаге построения последовательности одношаговых прогнозов пересчитывается условное распределение вероятностей  $p(y_{t+1} = y | y_{\Gamma(t)})$  случайного вектора  $\xi$  с использованием всех значений ряда отклонений от тренда до момента прогнозирования, т.е.  $\{y_t\}_{t=1}^T$ .
3. Трендовый прогноз с обучением и коррекцией тренда. В этой ситуации на каждом шаге прогнозирования уточняется не только условное распределение вероятностей, но и функция тренда. Другими словами, данный метод прогнозирования аналогичен «рекуррентному прогнозу», рассмотренному в первой части.
4. Вероятностный прогноз без обучения. В этом случае условное распределение вероятностей  $p(y_{t+1} = y | y_{\Delta(t)})$  случайного вектора  $\zeta$  строится на основе темповых характеристик  $\{y_t\}_{t=1}^{T-1}$  исходного временного ряда  $\{x_t\}_{t=1}^T$ . После этого с использованием полученных функций строится последовательность одношаговых прогнозов значений временного ряда  $\{\mathcal{E}_{T+t}\}_{t=1}^{\mathcal{F}}$  за-

висящих от истинных значений  $\{x_{T+t} \}_{t=0}^{t=T-1}$ .

5. Вероятностный прогноз с обучением. Этот случай отличается от предыдущего тем, что на каждом шаге построения прогноза пересчитывается условное распределение вероятностей  $p(y_{t+1} = y | y_{\Delta(t)})$  случайного вектора  $\zeta$  с использованием всех значений ряда до момента прогнозирования, т.е.

$\{x_{\tau} \}_{\tau=1}^t$ .

## 6.2.2. Алгоритмы моделирования и прогнозирования

### 6.2.2.1. Построение тренда

Здесь можно рассмотреть два подхода. Первый основан на качественном анализе исходного динамического ряда.

Сначала промежуток времени, на котором задан анализируемый временной ряд, разбивается на интервалы относительно схожего поведения. В данном случае в значительной мере используются результаты работы «Эконометрический анализ динамических рядов основных макроэкономических показателей» (М.: ИЭПП, 2001).

В результате такого анализа определяется функциональный вид тренда анализируемого временного ряда, т.е. определяется функция тренда  $f(t, a)$  с точностью до некоторого векторного параметра  $a$ . В данной работе в большинстве случаев используется кусочно-линейный тренд с изломом, приходящимся на вторую половину (август-сентябрь) 1998 г. На каждом из выделенных временных интервалов решается отвечающая соответствующему функциональному виду тренда система нормальных линейных уравнений, которая и определяет неизвестные параметры функции тренда. Для практической реализации нахождения параметров функции тренда используется построение регрессии по данным  $\{x_t \}_{t=1}^T$  на  $f(t, (1, \dots, 1))$ . Это позволяет получить не только значения векторного параметра  $a$ , но и статистические характеристики как значений параметров, так и самой функции тренда.

Второй подход основан на рассмотрении функции тренда временного ряда в виде некоторого полинома, т.е. приближенного разложения функции тренда в степенной ряд. Задав некоторый порядок полинома  $p$ , можно построить регрессию данных  $\{x_t \}_{t=1}^T$  на значения  $1, t, t^2, \dots, t^p$ . После удаления

статистически незначимых регрессоров получаем полиномиальную функцию тренда.

В данной работе использован только первый подход к построению функции тренда.

### 6.2.2.2. Выявление лаговой структуры

Для случайных величин  $\xi$  и  $\zeta$  строится условная информативная структура третьего порядка. Другими словами, определяются два лаговых показателя, знание значений которых в совокупности несет наибольшее количество информации о значении исследуемой переменной в текущий момент времени. Поскольку в настоящей работе рассматриваются только одномерные временные ряды, то задача построения информативных структур существенно упрощается.

В этом случае достаточно получить количество информации о значении показателя в момент времени  $t$ , содержащееся в любой паре лаговых показателей. После этого выбирается пара, отвечающая максимальному значению количества информации. Лаги, задаваемые этой парой, являются наиболее информативными с точки зрения определения значения анализируемого показателя в момент времени  $t$ . По определению отношение полученного количества информации к энтропии исследуемого показателя (коэффициент информативности) определяет долю неопределенности этого показателя, снижаемую за счет знания значений соответствующих лаговых показателей.

Рассмотрим схему практической реализации выявления пары наиболее информативных лагов. Соответствующий макрос (Информации3()) приведен в Приложении П1.1.

По определению количество информации  $I(\eta_1, (\eta_2, \eta_3))$  о случайной величине  $\eta_1$ , содержащееся в паре случайных величин  $\eta_2$  и  $\eta_3$ , равно:  $I(\eta_1, (\eta_2, \eta_3)) = H(\eta_1) + H(\eta_2, \eta_3) - H(\eta_1, \eta_2, \eta_3)$ , где  $H(\eta)$  – энтропия случайной величины  $\eta$  с совместным распределением вероятностей  $p(x) = P\{\eta = x\}$ , определяемая как  $H(\eta) = -\sum_{x \in X} p(x) \log_2 p(x)$ . Количество информации и энтропия измеряются в битах.

На первом этапе выявления информативных лагов задается максимально возможное запаздывание (величина  $\tau$ ). В данной работе рассматриваются лаги не превышающие одного года, т.е. для месячных данных  $\tau = 12$ , а для квартальных  $\tau = 4$ . Таким образом, рассматриваются 13- и 5-мерные случайные векторы, для которых определяется наиболее информативная зависимость первой компоненты от пары остальных.

На следующем этапе производится нормировка исходных данных ( $id(1), \dots, id(nn)$ ,  $nn$  – длина ряда по которому строится модель) так, чтобы они изменялись в диапазоне от 0 до 1:

$$idn(i) = \frac{id(i) - \min_{1 \leq j \leq nn} id(j)}{\max_{1 \leq j \leq nn} id(j) - \min_{1 \leq j \leq nn} id(j)}.$$

Кроме того, на этом шаге задается уровень табуляции данных ( $eps$ ). Дело в том, что чем ниже уровень табуляции, тем более точным получается результат, однако, в силу ограниченности данных, одновременно снижается надежность получаемого результата. Поэтому уровень табуляции выбирается таким образом, чтобы в каждую градацию попадало в среднем 4–5 значений ряда, т.е.

$$\frac{4}{nn} \leq eps \leq \frac{5}{nn}.$$

На третьем этапе рассчитываются приближенные значения вероятностей (частоты) того, что соответствующие случайные величины принимают значения из каждой из градаций. Для этого определяются центры градаций каждого из табулированных распределений вероятностей

$$\left( \frac{(k-1) \times eps}{2}, k = 1, \dots, n_i \right), \text{ где } n_i = \left\lceil \frac{\sqrt{i}}{eps} \right\rceil, i = 1, 2, 3 - \text{ минимальное}$$

целое число не меньшее  $\frac{\sqrt{i}}{eps}$ ) и рассчитываются расстояния в одно- двух- и трехмерном пространствах от данных за каждую единицу времени до соответствующих центров. После этого подсчитывается количество попавших в каждую градацию элементов для анализируемого показателя  $N_0(k)$ ,  $k = 1, \dots, n_1$ , всех пар лаговых показателей  $N_{ij}(k)$ ,  $i = 1, \dots, 11, j = i+1, \dots, 12, k = 1, \dots, n_2$  и троек, включающих анализируемый показатель и всевозможные пары лаговых показателей  $N_{0ij}(k)$ ,  $i = 1, \dots, 11, j = i+1, \dots, 12, k = 1, \dots, n_3$ , как точек с расстоянием от соответствующего центра не превышающем  $\frac{eps}{2}$ . Энтропии показателей определяются как

$$h1 = \sum_{k=1}^{n_1} N_0(k) \log_2 N_0(k) - \log_2 nn,$$

$$h2(i, j) = \sum_{k=1}^{n_2} N_{ij}(k) \log_2 N_{ij}(k) - \log_2 nn, i = 1, \dots, 11, j = i+1, \dots, 12,$$

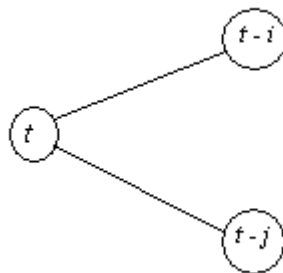
$$h3(0, i, j) = \sum_{k=1}^{n_3} N_{0ij}(k) \log_2 N_{0ij}(k) - \log_2 nn, i = 1, \dots, 11, j = i+1, \dots, 12$$

Количества информации рассчитываются по формуле:

$$I(t, t-i, t-j) = h1 + h2(i, j) - h3(0, i, j), i = 1, \dots, 11, j = i + 1, \dots, 12.$$

Максимальное из полученных  $I(t, t-i, t-j)$  и определяет лаговую структуру модели прогнозирования. Таким образом, если  $(i_0, j_0) = \operatorname{argmax}\{I(t, t-i, t-j) \mid i = 1, \dots, 11, j = i + 1, \dots, 12\}$ , то значения исследуемого показателя с лагами  $(i_0, j_0)$  несут в совокупности наибольшую информацию о его значении в момент времени  $t$ , т.е. лаговая структура будет иметь вид

### 6.2.2.3. Построение условного математического ожидания



Основой моделей прогнозирования (6.2.1) и (6.2.2) являются условные распределения вероятностей того, что случайная величина  $\xi$  (или  $\zeta$ ) принимает некоторое значение в момент времени  $t$  при известных ее значениях в моменты времени  $t-i$  и  $t-j$ , где запаздывания  $i$  и  $j$  определены так, как это указано в 2.2.2. Опишем алгоритм построения соответствующих условных распределений вероятностей. В Приложении П1.2 приведен макрос (Прогноз3()), реализующий этот алгоритм.

На первом этапе, так же как и для выявления наиболее информативных лагов, производится нормировка исходных данных так, чтобы они изменялись в диапазоне от 0 до 1:

$$idn(i) = \frac{id(i) - \min_{1 \leq j \leq nn} id(j)}{\max_{1 \leq j \leq nn} id(j) - \min_{1 \leq j \leq nn} id(j)},$$

где  $id(1), \dots, id(nn)$ ,  $nn$  – длина ряда по которому строится модель. На этом же шаге задается уровень табуляции данных ( $eps$ ) и задаются лаги, определенные ранее в результате работы программы Информации3(), от значений которых строится условное распределение вероятностей.



На следующем этапе для заданных значений лаговых переменных  $z_{t-i} = z^{(1)}$  и  $z_{t-j} = z^{(2)}$  определяется набор значений прогнозируемой переменной  $z_t = z^{(0)}$ , таких что расстояние в трехмерном пространстве от тройки значений  $\{z_t, z_{t-i}, z_{t-j}\}$  до  $\{z^{(0)}, z^{(1)}, z^{(2)}\}$  не превосходит  $eps$ , т.е.

$$ro(\{z_t, z_{t-i}, z_{t-j}\} \{z^{(0)}, z^{(1)}, z^{(2)}\}) = \sqrt{(z_t - z^{(0)})^2 + (z_{t-i} - z^{(1)})^2 + (z_{t-j} - z^{(2)})^2} \leq eps$$

Если таких наборов  $N$ , то условное математическое ожидание значения  $z_{t+1}$  при заданных  $z_{t-i+1} = z^{(1)}$  и  $z_{t-j+1} = z^{(2)}$  определяется как  $\frac{1}{N} \sum z^{(0)}$ .

#### 6.2.2.4. Построение прогноза

Реализуем два метода имитации и пять методов прогнозирования следующим образом.

**1. Трендовая имитация.** Служит для проверки качества трендовой модели прогнозирования за период  $13, \dots, T$  при месячных данных или  $5, \dots, T$  – при квартальных, где  $T$  – длина ряда  $x_t$ , использованного для построения функции прогнозирования. Полученная функция прогнозирования  $\varphi(t|\Gamma(t))$  вида (6.2.1) используется в данном случае следующим образом.

Строится ряд  $\{x_t\}_{t=\tau}^T$  ( $\tau = 5$  или  $13$  в зависимости от того, квартальные или месячные данные рассматриваются):

$$x_t = \varphi(t|\Gamma(t)) = f(t) + \sum_{k=1}^N x^{(k)} p \{x_t = x^{(k)} | x_{t-i}, x_{t-j}\} \Gamma(t) = \{i, j\} t = \tau, \dots, T$$

Здесь  $N$  – количество градаций, полученное при табулировании случайной величины  $\xi$ . Качество имитации оценивается величиной средней относительной ошибки (MAPE):

$$\delta = \frac{100}{T - \tau + 1} \sum_{t=\tau}^T \frac{|x_t - \hat{x}_t|}{x_t}.$$

Реализуется трендовая имитация использованием макроса Прогноз3() с параметрами: tau1 = i, tau2 = j, t = T, tt =  $\tau$ , ttt = T, eps = 0,005.

**2. Трендовый прогноз без обучения.** Служит для проверки качества трендовой модели прогнозирования за период  $T \div T + T'$ , где  $T$  – длина ряда  $x_t$ , использованного для построения функции прогнозирования, а  $T'$  – горизонт прогнозирования. Функция прогнозирования  $\varphi(t|\Gamma(t))$  используется в данном случае следующим образом. Строится ряд :  $\{x_t\}_{t=T+1}^{T+T'}$

$$\xi_t = \varphi(t|\Gamma(t)) = f(t) + \sum_{k=1}^N x^{(k)} p \left\{ \xi_t = x^{(k)} \mid x_{t-i}, x_{t-j} \right\} \Gamma(t) = \{i, j\} t = T+1, \dots, T+T'$$

Здесь  $N$  – количество градаций, полученное при табулировании случайной величины  $\xi$ . Качество прогноза также оценивается величиной средней абсолютной процентной ошибки (MAPE)

$$\delta = \frac{100}{T'} \sum_{t=\tau}^T \frac{|x_t - \xi_t|}{x_t}.$$

Реализуется трендовый прогноз без обучения использованием макроса Прогноз3() с параметрами: tau1 = i, tau2 = j, t = T, tt =  $\tau$ , ttt = T + T', eps = 0,005.

**3. Трендовый прогноз с обучением без коррекции тренда.** Служит также для проверки качества трендовой модели прогнозирования за период  $T \div T + T'$ , где  $T$  – длина ряда  $x_t$ , использованного для построения функции прогнозирования, а  $T'$  – горизонт прогнозирования. Функция прогнозирования  $\varphi(t|\Gamma(t))$  используется в данном случае следующим образом. Строится ряд  $\{x_t\}_{t=T+1}^{T+T'}$ :

$$\xi_t = \varphi(t|\Gamma(t)) = f(t) + \sum_{k=1}^N x^{(k)} p \left\{ \xi_t = x^{(k)} \mid x_{t-i}, x_{t-j} \right\} \Gamma(t) = \{i, j\} t = T+1, \dots, T+T'$$

Здесь  $N$  – количество градаций, полученное при табулировании случайной величины  $\xi$ . Однако в отличие от предыдущего случая в данном случае на каждом шаге уточняется условное распределение вероятностей  $p \left\{ \xi_t = x^{(k)} \mid x_{t-i}, x_{t-j} \right\}$ . Качество прогноза оценивается величиной средней абсолютной процентной ошибки (MAPE)

$$\delta = \frac{100}{T'} \sum_{t=\tau}^T \frac{|x_t - \xi_t|}{x_t}.$$

Реализуется трендовый прогноз с обучением без коррекции тренда  $T'$ -кратным использованием макроса Прогноз3(). При  $k$ -м ( $k = 1, \dots, T'$ ) применении макроса применяются параметры: tau1 = i, tau2 = j, t = T + k – 1, tt =  $\tau$ , ttt = T + k, eps = 0,005.

**4. Трендовый прогноз с обучением и коррекцией тренда.** Служит также для проверки качества трендовой модели прогнозирования на временном интервале  $T \div T + T'$ , где  $T$  – длина ряда  $x_t$ , использованного для построения функции прогнозирования, а  $T'$  – горизонт прогнозирования. Функция прогнозирования  $\varphi(t|\Gamma(t))$  используется в данном случае следую-

щим образом. Строится ряд  $\{x_t\}_{t=T+1}^{T+T'}$ :

$$\xi_t = \varphi(\Gamma(t)) = f(t) + \sum_{k=1}^N x^{(k)} p \left\{ \xi_t = x^{(k)} \mid x_{t-i}, x_{t-j} \right\} \Gamma(t) = \{i, j\} t = T+1, \dots, T+T'$$

Здесь  $N$  – количество градаций, полученное при табулировании случайной величины  $\xi$ . Однако в отличие от предыдущего случая в данном случае на каждом шаге уточняется не только условное распределение вероятностей  $p \left\{ \xi_t = x^{(k)} \mid x_{t-i}, x_{t-j} \right\}$ , но и функция тренда  $f(t)$ . Качество прогноза оценивается величиной средней абсолютной процентной ошибки (MAPE)

$$\delta = \frac{100}{T'} \sum_{t=\tau}^T \frac{|x_t - \xi_t|}{x_t}.$$

Реализуется трендовый прогноз с обучением и коррекцией тренда  $T'$ -кратным использованием макроса Прогноз3(). Перед  $k$ -м ( $k = 1, \dots, T'$ ) применением макроса производится пересчет тренда по ряду  $\{x_t\}_{t=1}^{T+k-1}$ , после чего к макросу применяются параметры: tau1 =  $i$ , tau2 =  $j$ , t =  $T + k - 1$ , tt =  $\tau$ , ttt =  $T + k$ , eps = 0,005.

**5. Вероятностная имитация.** Служит для проверки качества вероятностной модели прогнозирования на временном интервале  $13 \div T$  при месячных данных или  $5 \div T$  – при квартальных, где  $T$  – длина ряда  $x_t$ , использованного для построения функции прогнозирования. Полученная функция прогнозирования  $\psi(t|\Delta(t))$  вида (6.2.2) используется в данном случае следующим образом. Строится ряд  $\{x_t\}_{t=\tau}^T$  ( $\tau = 5$  или  $13$  в зависимости от того, квартальные или месячные данные рассматриваются):

$$\xi_t = \psi(\zeta|\Delta(t)) = x_{t-1} \sum_{k=1}^N y^{(k)} p \left\{ \xi_t = y^{(k)} \mid y_{t-i}, y_{t-j} \right\} \Delta(t) = \{i, j\} t = \tau, \dots, T$$

Здесь  $N$  – количество градаций, полученное при табулировании случайной величины  $\zeta$ . Качество имитации оценивается величиной средней относительной ошибки (MAPE):

$$\delta = \frac{100}{T - \tau + 1} \sum_{t=\tau}^T \frac{|x_t - \xi_t|}{x_t}.$$

Реализуется вероятностная имитация использованием макроса Прогноз3() с параметрами: tau1 =  $i$ , tau2 =  $j$ , t =  $T$ , tt =  $\tau$ , ttt =  $T$ , eps = 0,005.

**6. Вероятностный прогноз без обучения.** Служит для проверки качества вероятностной модели прогнозирования на временном интервале  $T \div T + T'$ , где  $T$  – длина ряда  $x_t$ , использованного для построения функции прогнозирования, а  $T'$  – горизонт прогнозирования. Функция прогнозирования  $\psi(t|\Delta(t))$  используется в данном случае следующим образом. Строится ряд  $\{x_t\}_{t=T+1}^{T+T'}$ :

$$\varepsilon_t = \psi(t|\Delta(t)) = x_{t-1} \sum_{k=1}^N y^{(k)} p \{y_t = y^{(k)} | y_{t-i}, y_{t-j}\} \Delta(t) = \{\xi, j\} t = \tau, \dots, T$$

Здесь  $N$  – количество градаций, полученное при табулировании случайной величины  $\zeta$ . Качество прогноза оценивается величиной средней абсолютной процентной ошибки (MAPE)

$$\delta = \frac{100}{T'} \sum_{t=\tau}^T \frac{|x_t - \varepsilon_t|}{x_t} .$$

Реализуется вероятностный прогноз без обучения с использованием макроса Прогноз3() с параметрами: tau1 = i, tau2 = j, t = T, tt =  $\tau$ , ttt =  $T + T'$ , eps = 0,005.

**7. Вероятностный прогноз с обучением.** Служит также для проверки качества вероятностной модели прогнозирования на временном интервале  $T \div T + T'$ , где  $T$  – длина ряда  $x_t$ , использованного для построения функции прогнозирования, а  $T'$  – горизонт прогнозирования. Функция прогнозирования  $\psi(t|\Delta(t))$  используется в данном случае следующим образом. Строится ряд

$$\varepsilon_t = \psi(t|\Delta(t)) = x_{t-1} \sum_{k=1}^N y^{(k)} p \{y_t = y^{(k)} | y_{t-i}, y_{t-j}\} \Delta(t) = \{\xi, j\} t = \tau, \dots, T$$

Здесь  $N$  – количество градаций, полученное при табулировании случайной величины  $\zeta$ . Однако в отличие от предыдущего случая, в данном случае на каждом шаге уточняется условное распределение вероятностей  $p \{y_t = y^{(k)} | y_{t-i}, y_{t-j}\}$ . Качество прогноза оценивается величиной средней абсолютной процентной ошибки (MAPE)

$$\delta = \frac{100}{T'} \sum_{t=\tau}^T \frac{|x_t - \varepsilon_t|}{x_t} .$$

Реализуется вероятностный прогноз с обучением  $T'$ -кратным использованием макроса Прогноз3(). При  $k$ -м ( $k = 1, \dots, T'$ ) применении макроса используются параметры: tau1 = i, tau2 = j, t =  $T + k - 1$ , tt = , ttt =  $T + k$ , eps = 0,005.

## 6.3. Результаты расчетов

### 6.3.1. Инфляция

В качестве исходных данных для построения трендовой модели прогнозирования инфляции используются месячные данные за период 01.1991 – 08.2000. Данные за период 09.2000 – 12.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на *рис. 6.3.1.*

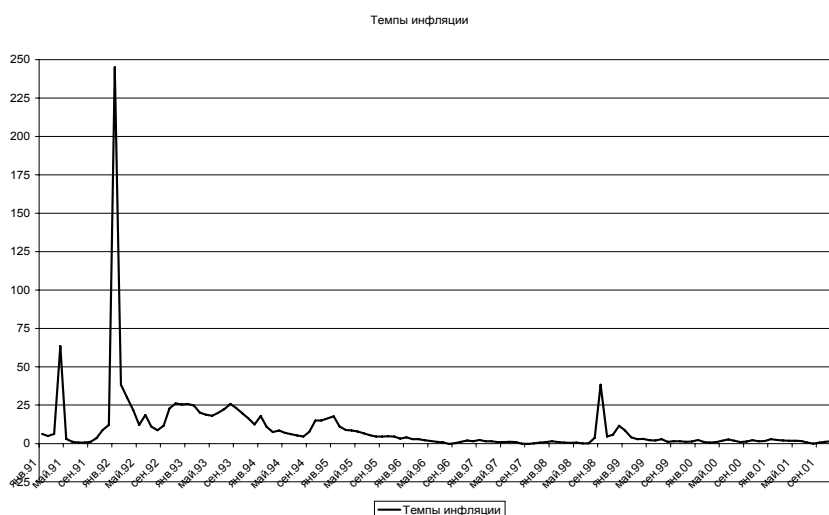


Рис. 6.3.1. Исходные данные

В соответствии с проведенным ранее исследованием ряда инфляции<sup>11</sup> разобьем весь интервал наблюдений на 3 подинтервала: 01.1991 – 12.1991; 01.1992 – 08.1998; 09.1998 – 08.2000. Для каждого из этих интервалов построим детерминированный логарифмический тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = 68,767 - 28,826 \ln t, & t = 1, \dots, 12; \\ f_2(t) = 119,028 - 27,854 \ln t, & t = 13, \dots, 92; \\ f_3(t) = 293,892 - 62,311 \ln t, & t = 93, \dots, 116. \end{cases}$$

<sup>11</sup> Здесь и далее при выборе функционального вида тренда мы ориентируемся на проведенное ранее исследование «*Эконометрический анализ динамических рядов основных макроэкономических показателей*». М.: ИЭПП, 2001.

Характеристики построенных трендов приведены в *табл. 6.3.1*.

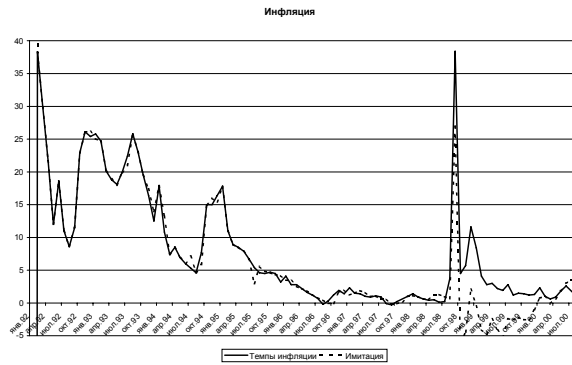
Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 1 и 5 месяцев с коэффициентом информативности, равным 0,872. Это означает, что значения инфляции 1 и 5 месяцев назад определяют сегодняшнюю инфляцию на 87,2%. Таким образом, модель прогнозирования инфляции имеет вид

$$\tilde{x}_t = f_3(t) + \sum_{y \in Y} \alpha y(y | y_{t-1}, y_{t-5}).$$

*Таблица 6.3.1*

Тренд	R	Нормированный R <sup>2</sup>	F-статистика	t-статистика c <sub>0</sub>	t-статистика c <sub>1</sub>
$f_1(t)$	0,527	0,175	2,695	1,908	-1,542
$f_2(t)$	0,528	0,269	30,132	6,052	-5,489
$f_3(t)$	0,550	0,270	9,520	3,131	-3,085

На *рис.6.3.2* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации близки между собой. Средняя относительная ошибка имитации (MAPE) за период 01.1992 – 08.2000 составляет 0,40%<sup>12</sup>.



*Рис. 6.3.2. Имитация*

<sup>12</sup> В связи с тем, что значение инфляции может равняться 0, для расчета относительной ошибки прогноза используется не темп прироста, а темп роста.

Для проверки качества прогнозирования использованы данные за период 09.2000 – 12.2001. Построено три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-5})$ , построенное на первоначальных исходных данных) и две с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-5})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В табл. 6.3.2 и на рис. 6.3.3 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_3$  меняются от шага к шагу. Соответствующие коэффициенты приведены в табл. 6.3.3. Средняя ошибка прогноза без обучения составляет 6,21%, а с обучением – 1,00% без коррекции тренда и 0,76% – с коррекцией.

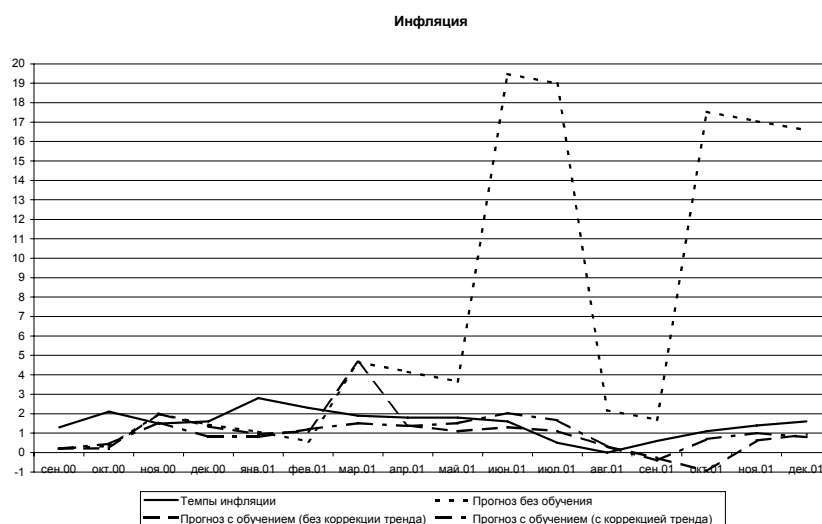


Рис. 6.3.3. Прогноз

Таблица 6.3.2

## Трендовый прогноз инфляции

Дата	ИПЦ	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
сен.00	1,3	0,20	1,08	0,20	1,08	0,20	1,08
окт.00	2,1	0,35	1,72	0,22	1,84	0,45	1,62
ноя.00	1,5	2,00	0,49	2,00	0,49	1,55	0,05
дек.00	1,6	1,43	0,17	1,34	0,26	0,80	0,78
январь.01	2,8	1,07	1,68	0,95	1,80	0,80	1,95
фев.01	2,3	0,57	1,69	1,10	1,18	1,20	1,07
мар.01	1,9	4,66	2,71	4,66	2,71	1,51	0,38
апр.01	1,8	4,16	2,31	1,40	0,40	1,36	0,43
май.01	1,8	3,66	1,82	1,10	0,69	1,51	0,28
Июнь.01	1,6	19,48	17,60	1,30	0,29	2,04	0,43
Июль.01	0,5	18,98	18,39	1,11	0,60	1,66	1,15
авг.01	0,0	2,18	2,18	0,31	0,31	0,31	0,31
сен.01	0,6	1,69	1,09	-0,27	0,87	-0,41	1,00
окт.01	1,1	17,53	16,25	-0,95	2,03	0,70	0,39
ноя.01	1,4	17,05	15,44	0,62	0,77	1,01	0,39
дек.01	1,6	16,58	14,74	0,93	0,66	0,79	0,79
Средняя ошибка прогноза		6,21			1,00		0,76

Таблица 6.3.3

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
117	293,892	_-62,311	121	227,708	_-47,963	125	177,273	_-37,058	129	152,812	_-31,781
118	276,099	_-58,450	122	211,052	_-44,359	126	168,425	_-35,148	130	148,167	_-30,781
119	256,748	_-54,254	123	197,754	_-41,483	127	160,763	_-33,495	131	142,712	_-29,607
120	241,722	_-50,998	124	186,909	_-39,139	128	156,183	_-32,508	132	137,006	_-28,380

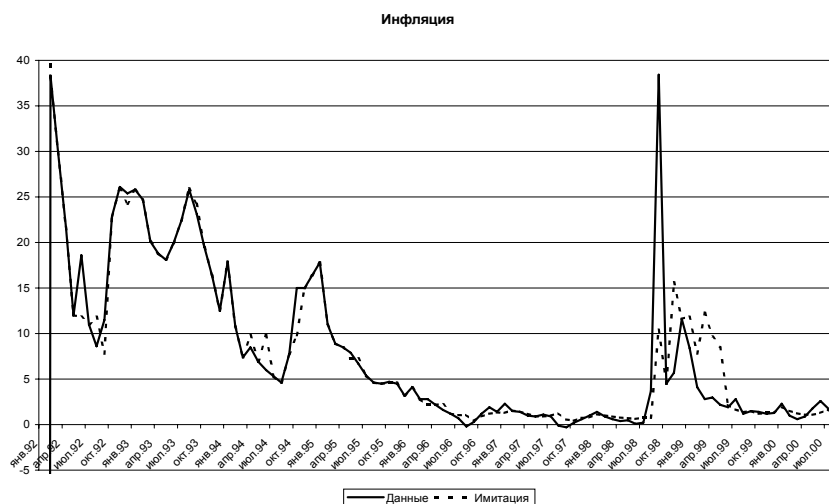
Рассмотрим теперь вероятностную модель прогнозирования. В этом случае будем считать сами исходные данные реализациями некоторой случайной величины. Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель инфляции и лаговые значения до 12 месяцев, т.е. для системы  $\{x_t, x_{t-1}, \dots, x_{t-12}\}$ . Получаем зависимость показателя инфляции от лага в 1 и 12 месяцев с коэффициентом ин-



формативности, равным 0,832. Это означает, что значения инфляции 1 и 12 месяцев назад определяют сегодняшнюю инфляцию на 83,2%. Таким образом, модель прогнозирования инфляции в данном случае имеет вид

$$\tilde{x}_t = \sum_{x \in X} xp(x|x_{t-1}, x_{t-12}).$$

На *рис. 6.3.4* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации и в этом случае близки между собой. Средняя относительная ошибка имитации (МАРЕ) за период 01.1992–08.2000 составляет 0,94% (против 0,40% в предыдущем случае).



*Рис. 6.3.4.* Имитация

Для проверки качества прогнозирования, как и выше, использованы данные за период 09.2000 – 12.2001. Построены две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(x_t|x_{t-1}, x_{t-12})$ , построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(x_t|x_{t-1}, x_{t-12})$  уточняется за счет добавления новой точки). В *табл. 6.3.4* и на *рис. 6.3.5* приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (МАРЕ) без обучения составляет 0,45%, а с обучением – 0,41%. Таким образом, в данном случае прогнозные свойства модели оказались выше.

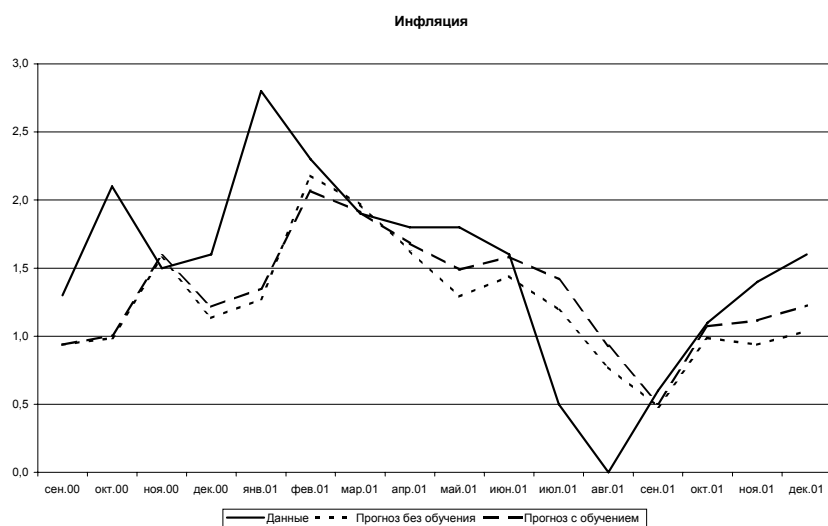


Рис. 6.3.5. Прогноз

Таблица 6.3.4

**Вероятностный прогноз инфляции**

Дата	ИПЦ	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	Ошибка
сен.00	1,3	0,94	0,36	0,94	0,36
окт.00	2,1	0,99	1,09	1,01	1,07
ноя.00	1,5	1,58	0,07	1,59	0,09
дек.00	1,6	1,13	0,46	1,22	0,38
январь.01	2,8	1,27	1,49	1,35	1,41
фев.01	2,3	2,18	0,12	2,07	0,23
мар.01	1,9	1,97	0,07	1,91	0,01
апр.01	1,8	1,63	0,17	1,68	0,12
май.01	1,8	1,29	0,50	1,49	0,31
июнь.01	1,6	1,44	0,16	1,58	0,02
июль.01	0,5	1,19	0,69	1,42	0,91
авг.01	0,0	0,77	0,77	0,93	0,93
сен.01	0,6	0,48	0,12	0,51	0,09
окт.01	1,1	0,99	0,11	1,07	0,03
ноя.01	1,4	0,94	0,46	1,11	0,28
дек.01	1,6	1,04	0,55	1,23	0,37
Средняя ошибка прогноза			0,45		0,41

### 6.3.2. Денежные агрегаты

#### 6.3.2.1. Денежный агрегат $M0$

В качестве исходных данных для построения модели прогнозирования денежного агрегата  $M0$  используются месячные данные за период 12.1990–07.2000. Данные за период 08.2000–11.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на рис. 6.3.6.

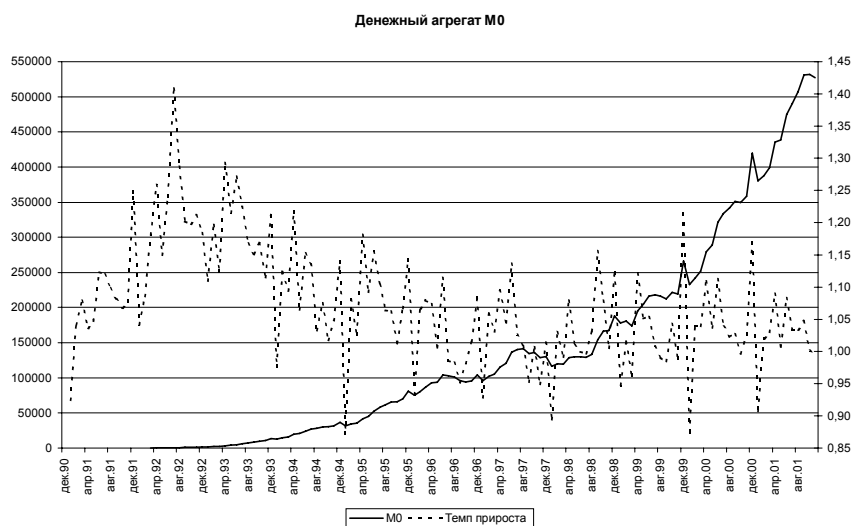


Рис. 6.3.6. Исходные данные

В соответствии с проведенным ранее исследованием ряда  $M0$  на временном интервале 06.1995–07.2000 данный ряд является  $DS$ -рядом. В то же время визуальный анализ графика ряда позволяет предположить изменение характера его поведения начиная со второй половины 1998 г. Разобьем весь интервал наблюдений на 2 подинтервала: 12.1990–08.1998 и 09.1998–07.2000. Для каждого из этих интервалов построим детерминированный линейный тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = -33648,821 + 1777,917t, & t = 1, \dots, 92; \\ f_2(t) = -498483,752 + 6869,279t, & t = 93, \dots, 116. \end{cases}$$

Характеристики построенных трендов приведены в *табл. 6.3.5*.

*Таблица 6.3.5*

Тренд	R	Нормированный R <sup>2</sup>	F-статистика	t-статистика c <sub>0</sub>	t-статистика c <sub>1</sub>
$f_1(t)$	0,948	0,898	803,356	-10,018	28,344
$f_2(t)$	0,949	0,896	198,324	-9,758	14,083

Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 2 и 12 месяцев с коэффициентом информативности, равным 0,875. Это означает, что значения  $M0$  2 и 12 месяцев назад определяют сегодняшнее значение  $M0$  на 87,5%. Таким образом, модель прогнозирования инфляции имеет вид

$$\tilde{x}_t = f_2(t) + \sum_{y \in Y} yp(y | y_{t-2}, y_{t-12}).$$

На *рис. 6.3.7* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации полностью совпадают между собой.



*Рис. 6.3.7. Имитация*

Для проверки качества прогнозирования использованы данные за период 09.2000 – 11.2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-2}, y_{t-12})$ , построенное на первоначальных исходных данных) и два

с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-2}, y_{t-1})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В табл. 6.3.6. и на рис. 6.3.8. приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_2$  меняются от шага к шагу. Соответствующие коэффициенты приведены в табл. 6.3.7. Средняя ошибка прогноза (MAPE) без обучения составляет 23,99%, а с обучением 4,18% без коррекции тренда и 5,14% с коррекцией.

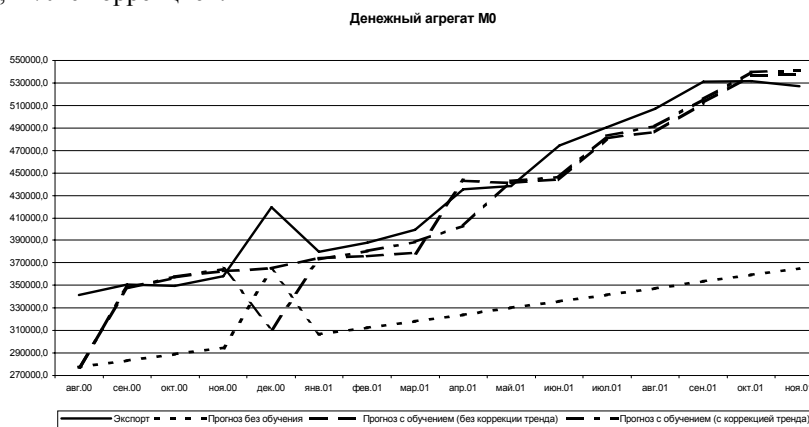


Рис. 6.3.8. Прогноз

Таблица 6.3.6

**Трендовый прогноз М0**

Дата	М0	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
авг 00	341627	277192,1	18,86	277192,1	18,86	277192,1	18,86
сен 00	350935	283049,8	19,34	347484,4	0,98	347820,8	0,89
окт 00	349669	288907,5	17,38	356793,0	2,04	357432,8	2,22
ноя 00	358351	294765,2	17,74	362650,7	1,20	364353,5	1,67
дек 00	419262	363325,4	13,34	365057,4	12,93	310995,4	25,82
январь 01	380127	306480,5	19,37	374366,0	1,52	372863,4	1,91
фев 01	387959	312338,2	19,49	375924,3	3,10	380543,5	1,91

Таблица 6.3.6 (продолжение)

Дата	M0	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
мар 01	399395	318195,9	20,33	378957,1	5,12	388418,4	2,75
апр 01	435350	324053,6	25,56	442692,3	1,69	402716,0	7,50
май 01	438312	329911,3	24,73	441207,2	0,66	443203,0	1,12
июн 01	474692	335769,0	29,27	444169,2	6,43	446336,0	5,97
июл 01	490598	341626,6	30,37	480549,4	2,05	482953,8	1,56
авг 01	507107	347484,3	31,48	486407,1	4,08	491682,2	3,04
сен 01	530972	353342,0	33,45	512964,8	3,39	515831,7	2,85
окт 01	531481	359199,7	32,42	536829,4	1,01	539938,6	1,59
ноя 01	527287	365057,4	30,77	537338,9	1,91	540645,2	2,53
Средняя ошибка прогноза			23,99		4,18		5,14

Таблица 6.3.7

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
117	-498483,752	6869,279	121	-589828,211	7773,412
118	-532312,207	7205,323	122	-638019,335	8245,874
119	-560816,089	7487,540	123	-646040,710	8324,259
120	-576472,511	7642,044	124	-652322,303	8385,443

Таблица 6.3.7 (продолжение)

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
125	-659299,727	8453,185	129	-745936,517	9287,212
126	-680298,565	8656,399	130	-770499,627	9521,892
127	-694883,545	8797,091	131	-797546,854	9779,485
128	-720806,658	9046,351	132	-816958,807	9963,775

Рассмотрим теперь вероятностную модель прогнозирования. Поскольку в исходных данных явно содержится детерминированный тренд, то преобразуем их в темповые характеристики, т.е. введем новый ряд:

$$y_t = \frac{x_t}{x_{t-1}}, t = 1, \dots, T.$$

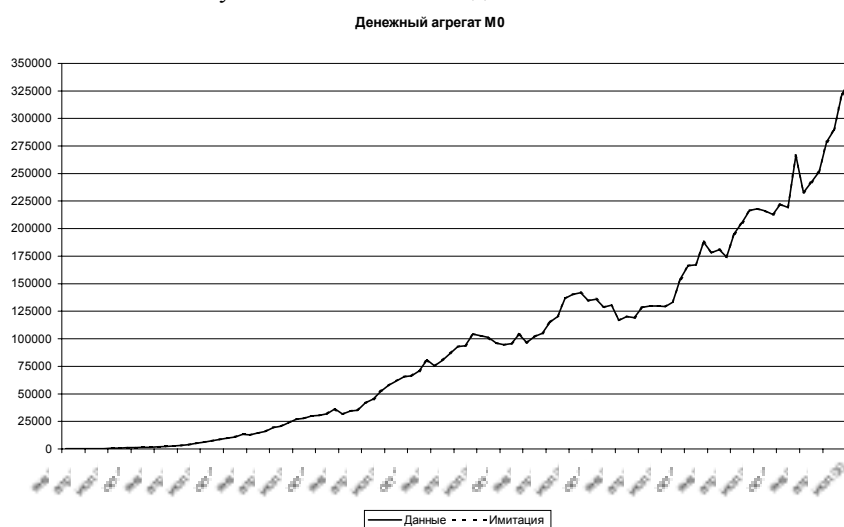
График соответствующего временного ряда также приведен на рис. 6.3.6.

Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель темпа роста  $M0$  и его лаговые значения

до 12 месяцев, т.е. для системы  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя темпа роста  $M0$  от лага в 2 и 12 месяцев с коэффициентом информативности, равным 0,967. Это означает, что значения темпов роста  $M0$  2 и 12 месяцев назад определяют сегодняшний темп роста на 96,7%. Таким образом, модель прогнозирования безработицы в данном случае имеет вид:

$$\tilde{x}_t = x_{t-1} \sum_{y \in Y} yp(y | y_{t-2}, y_{t-12}).$$

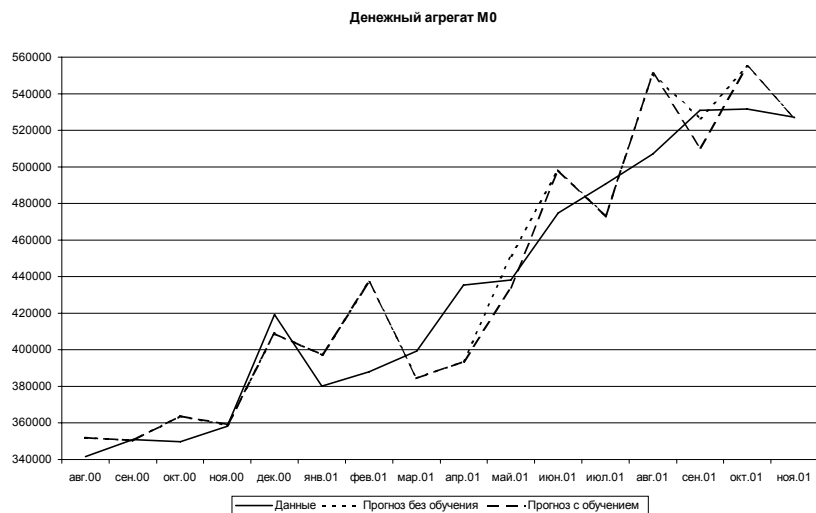
На *рис. 6.3.9.* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации на временном интервале 01.1992 – 07.2000 и в этом случае полностью совпадают.



*Рис. 6.3.9.* Имитация

Для проверки качества прогнозирования использованы данные за период 08.2000 – 11.2001. Построены две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-2}, y_{t-12})$ , построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-2}, y_{t-12})$  уточняется за счет добавления новой точки). В *табл. 6.3.8.* и на *рис. 6.3.10.* приведены результаты од-

ношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (MAPE) без обучения составляет 4,12%, а с обучением 4,19%.



*Рис. 6.3.10. Прогноз*

*Таблица 6.3.8*

**Вероятностный прогноз М0**

Дата	М0	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
авг.00	341627	352103,7	3,07	352103,7	3,07
сен.00	350935	350478,7	0,13	350478,7	0,13
окт.00	349669	363772,3	4,03	363772,3	4,03
ноя.00	358351	358729,1	0,11	359196,4	0,24
дек.00	419262	408772,7	2,50	408772,7	2,50
январь.01	380127	397323,5	4,52	397323,5	4,52
февраль.01	387959	436968,3	12,63	436968,3	12,63
март.01	399395	384384,3	3,76	384384,3	3,76
апрель.01	435350	393586,1	9,59	393586,1	9,59
май.01	438312	451274,3	2,96	433778,2	1,03
июнь.01	474692	498329,0	4,98	498329,0	4,98



Таблица 6.3.8 (продолжение)

Дата	M0	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
июл.01	490598	472936,4	3,60	472936,4	3,60
авг.01	507107	550940,7	8,64	550940,7	8,64
сен.01	530972	526446,3	0,85	510557,3	3,84
окт.01	531481	554666,3	4,36	554666,3	4,36
ноя.01	527287	526652,7	0,12	526652,7	0,12
Средняя ошибка прогноза			4,12		4,19

### 6.3.2.2. Денежный агрегат M1

В качестве исходных данных для построения модели прогнозирования денежного агрегата M1 используются месячные данные за период 6.1995 – 07.2000. Данные за период 08.2000 – 11.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на рис. 6.3.11.

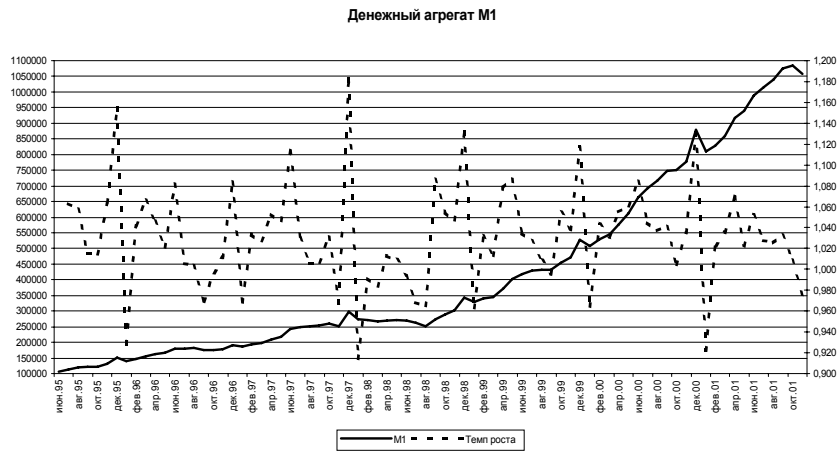


Рис. 6.3.11. Исходные данные

В соответствии с проведенным ранее исследованием ряда M1 на временном интервале 06.1995 – 07.2000 данный ряд, так же как и ряд M0, является DS-рядом. В то же время визуальный анализ графика ряда позволяет

предположить изменение характера его поведения начиная со второй половины 1998 г. Разобьем весь интервал наблюдений на 2 подинтервала: 12.1990 – 08.1998 и 09.1998 – 07.2000. Для каждого из этих интервалов построим детерминированный линейный тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = 104389,501 + 4895,270t, & t = 1, \dots, 38; \\ f_2(t) = -498483,752 + 6869,279t, & t = 39, \dots, 62. \end{cases}$$

Характеристики построенных трендов приведены в табл. 6.3.9.

Таблица 6.3.9

Тренд	R	Нормированный R	F-статистика	t-статистика $s_0$	t-статистика $s_1$
$f_1(t)$	0,974	0,947	659,906	24,486	25,689
$f_2(t)$	0,983	0,964	622,377	-12,093	24,947

Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 6 и 9 месяцев с коэффициентом информативности, равным 0,943. Это означает, что значения  $M1$  6 и 9 месяцев назад определяют сегодняшнее значение  $M1$  на 94,3%. Таким образом, модель прогнозирования инфляции имеет вид

$$\tilde{x}_t = f_2(t) + \sum_{y \in Y} \gamma y(y|y_{t-6}, y_{t-9}).$$

На рис. 6.3.12. приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели на временном интервале 03.1996–07.2000. Как видно из приведенного рисунка, графики исходных данных и имитации полностью совпадают между собой.

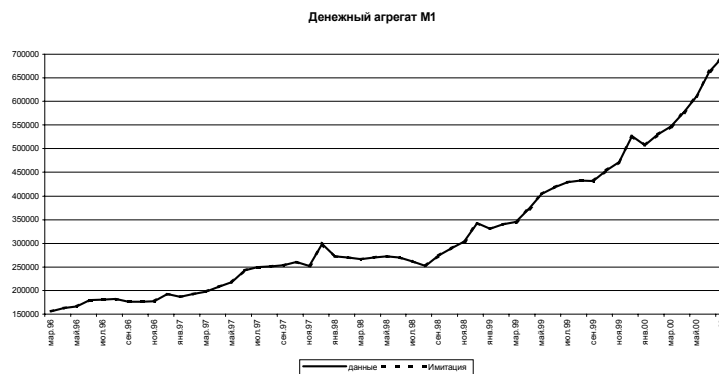


Рис. 6.3.12. Имитация

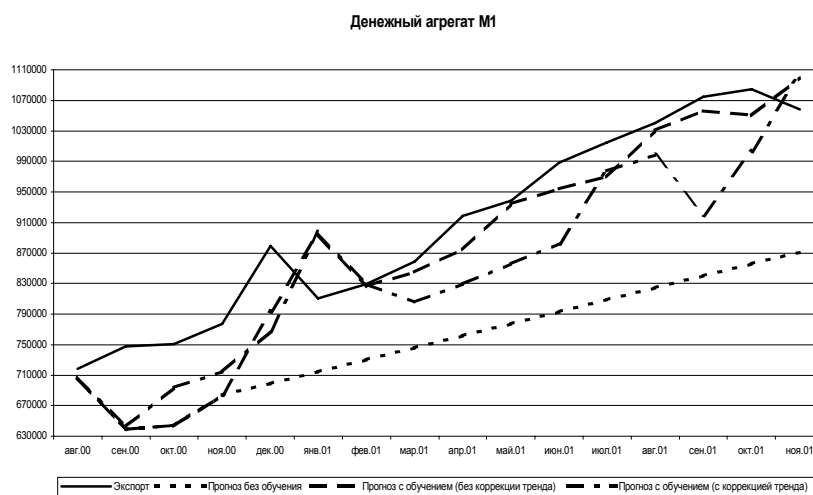


Рис. 6.3.13. Прогноз

Для проверки качества прогнозирования использованы данные за период 09.2000 – 11.2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-6}, y_{t-9})$ , построенное на первоначальных исходных данных) и два с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-6}, y_{t-9})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В табл. 6.3.10 и на рис. 6.3.13 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_2$  меняются от шага к шагу. Соответствующие коэффициенты приведены в табл. 6.3.11. Средняя ошибка прогноза (МАРЕ) без обучения составляет 15,96%, а с обучением 5,46% без коррекции тренда и 7,75% с коррекцией.

Таблица 6.3.10

## Трендовый прогноз М1

Дата	М1	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
авг 00	717950	708005,0	1,39	708005,0	1,39	708005,0	1,39
сен 00	747446	639197,4	14,48	639197,4	14,48	642942,8	13,98
окт 00	750662	643363,1	14,29	643363,1	14,29	693621,3	7,60
ноя 00	777139	683564,4	12,04	683564,4	12,04	714708,3	8,03
дек 00	879310	699181,8	20,49	792756,0	9,84	767490,9	12,72
январь 01	810517	714799,1	11,81	894927,1	10,41	897386,1	10,72
фев 01	829180	730416,4	11,91	826133,8	0,37	828812,7	0,04
мар 01	858381	746033,7	13,09	844797,4	1,58	806149,0	6,08
апр 01	918209	761651,1	17,05	873998,4	4,81	829197,2	9,69
май 01	938533	777268,4	17,18	933826,7	0,50	855851,5	8,81
июнь 01	987901	792885,7	19,74	954150,3	3,42	882195,1	10,70
июль 01	1015090	808503,1	20,35	969767,7	4,46	977309,2	3,72
авг 01	1040765	824120,4	20,82	1030707,6	0,97	998373,8	4,07
сен 01	1074933	839737,7	21,88	1056382,1	1,73	919584,0	14,45
окт 01	1084385	855355,1	21,12	1050370,4	3,14	1003113,6	7,49
ноя 01	1058132	870972,4	17,69	1100002,3	3,96	1105310,5	4,46
Средняя ошибка прогноза			15,96		5,46		7,75

Таблица 6.3.11

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
63	-420876,239	17033,303	67	-508230,235	18889,100
64	-449189,900	17640,024	68	-548606,618	19730,275
65	-476375,476	18218,441	69	-554052,515	19842,949
66	-492468,326	18558,431	70	-558193,841	19928,044

Таблица 6.3.11 (продолжение)

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
71	-564210,774	20050,839	75	-627074,523	21311,866
72	-579725,874	20365,334	76	-642090,608	21608,236
73	-592736,408	20627,291	77	-657534,706	21911,061
74	-610556,706	20983,697	78	-668347,874	22121,707

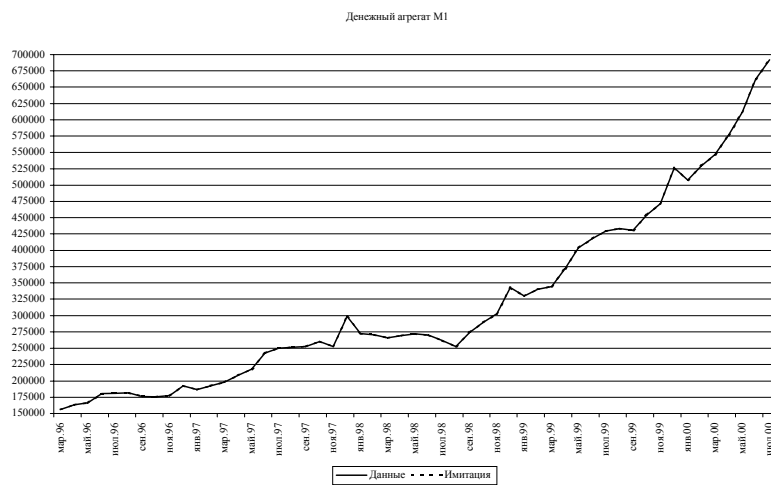
Рассмотрим теперь вероятностную модель прогнозирования. Поскольку в исходных данных явно содержится детерминированный тренд, то преобразуем их в темповые характеристики, т.е. введем новый ряд:

$y_t = \frac{x_t}{x_{t-1}}, t = 1, \dots, T$  График соответствующего временного ряда также приведен на *рис. 6.3.10*.

Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель темпа роста  $M1$  и его лаговые значения до 12 месяцев, т.е. для системы  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя темпа роста  $M1$  от лага в 6 и 7 месяцев с коэффициентом информативности, равным 0,969. Это означает, что значения темпов роста  $M1$  6 и 7 месяцев назад определяют сегодняшний темп роста на 96,9%. Таким образом, модель прогнозирования безработицы в данном случае имеет вид:

$$\tilde{x}_t = x_{t-1} \sum_{y \in Y} yp(y | y_{t-6}, y_{t-7}).$$

На *рис. 6.3.14* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации на временном интервале 03.1996 – 07.2000 и в этом случае полностью совпадают.



*Рис. 6.3.14.* Имитация

Для проверки качества прогнозирования, как и выше, использованы

данные за период 08.2000 – 11.2001. Построены две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-6}, y_{t-7})$ , построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-6}, y_{t-7})$  уточняется за счет добавления новой точки). В *табл. 6.3.12* и на *рис. 6.3.15* приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (MAPE) без обучения составляет 5,22%, а с обучением – 4,75%.

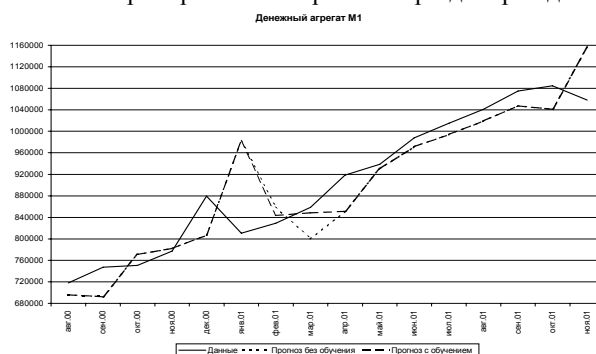
### 6.3.2.3. Денежный агрегат M2

В качестве исходных данных для построения модели прогнозирования денежного агрегата *M2* используются месячные данные за период 12.1990 – 07.2000. Данные за период 08.2000 – 11.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на *рис. 6.3.16*.

В соответствии с проведенным ранее исследованием ряда *M2* на временном интервале 06.1995 – 07.2000 данный ряд, также как и два предыдущих, является *DS*-рядом. В то же время визуальный анализ графика ряда позволяет предположить изменение характера его поведения начиная со второй половины 1998 г. Разобьем весь интервал наблюдений на 2 подинтервала: 12.1990 – 08.1998 и 09.1998 – 07.2000. Для каждого из этих интервалов построим детерминированный линейный тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = -94369,716 + 4970,452t, & t = 1, \dots, 92; \\ f_2(t) = -1888494,862 + 23763,636t, & t = 93, \dots, 116. \end{cases}$$

Статистические характеристики построенных трендов приведены в *табл. 6.3.13*.



*Рис. 6.3.15. Прогноз*

Таблица 6.3.12

## Вероятностный прогноз М1

Дата	М1	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
авг.00	717950	696241,9	3,02	696241,9	3,02
сен.00	747446	692445,2	7,36	692445,2	7,36
окт.00	750662	770796,8	2,68	770796,8	2,68
ноя.00	777139	781978,5	0,62	781978,5	0,62
дек.00	879310	806620,7	8,27	806620,7	8,27
январь.01	810517	982236,3	21,19	982236,3	21,19
фев.01	829180	859480,3	3,65	843815,9	1,77
мар.01	858381	799723,9	6,83	848273,5	1,18
апр.01	918209	851240,0	7,29	851240,0	7,29
май.01	938533	930178,3	0,89	930178,3	0,89
июнь.01	987901	971262,1	1,68	971262,1	1,68
июль.01	1015090	993563,9	2,12	993563,9	2,12
авг.01	1040765	1018845,0	2,11	1018845,0	2,11
сентябрь.01	1074933	1047381,9	2,56	1047381,9	2,56
октябрь.01	1084385	1041058,3	4,00	1041058,3	4,00
ноябрь.01	1058132	1156347,1	9,28	1156347,1	9,28
Средняя ошибка прогноза			5,22		4,75

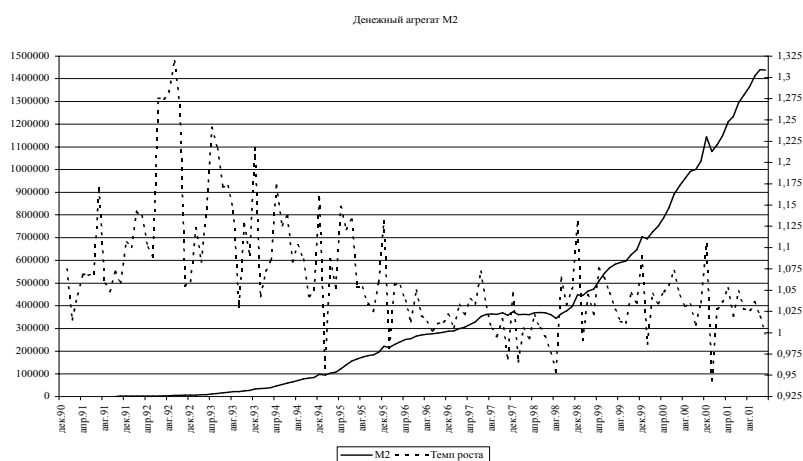


Рис. 6.3.16. Исходные данные

Таблица 6.3.13

Тренд	R	Нормированный R <sup>2</sup>	F-статистика	t-статистика c <sub>0</sub>	t-статистика c <sub>1</sub>
$f_1(t)$	0,953	0,907	898,541	-10,514	29,976
$f_2(t)$	0,988	0,976	893,163	-22,574	29,886

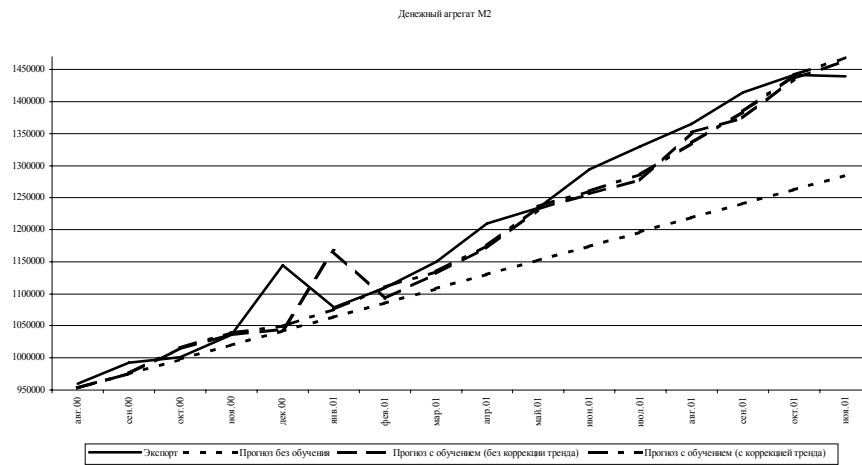
Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 1 и 8 месяцев с коэффициентом информативности, равным 0,924. Это означает, что значения M2 1 и 8 месяцев назад определяют сегодняшнее значение M2 на 92,4%. Таким образом, модель прогнозирования инфляции имеет вид :

$$\tilde{x}_t = f_2(t) + \sum_{y \in Y} y p(y | y_{t-1}, y_{t-8}).$$

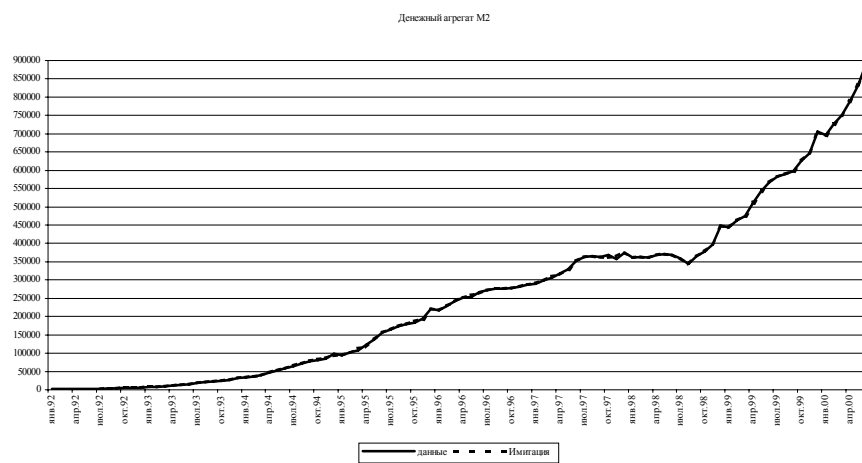
На *рис. 6.3.17* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации почти полностью совпадают между собой. Среднее абсолютное отклонение (MAPE) имитационных значений от истинных составляет 0,51%.

Для проверки качества прогнозирования использованы данные за период 08.2000 – 11.2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-8})$ , построенное на первоначальных исходных данных) и два с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-8})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В *табл. 6.3.14* и на *рис. 6.3.18* приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_2$  меняются от шага к шагу. Соответствующие коэффициенты приведены в *табл. 6.3.15*. Средняя ошибка прогноза (MAPE) без обучения составляет 6,20%, а с обучением – 2,47% без коррекции тренда и 1,84% – с коррекцией.





*Рис. 6.3.17. Имитация*



*Рис. 6.3.18. Прогноз*

Таблица 6.3.14

## Трендовый прогноз М2

Дата	М2	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
авг 00	960100	953310,0	0,71	953310,0	0,71	953310,0	0,71
сен 00	992400	975420,1	1,71	975420,1	1,71	976656,0	1,59
окт 00	1001200	997530,1	0,37	1014510,0	1,33	1015716,3	1,45
ноя 00	1036400	1019640,1	1,62	1036620,1	0,02	1039880,1	0,34
дек 00	1144300	1041750,2	8,96	1045420,1	8,64	1049511,2	8,28
январь 01	1079300	1063860,2	1,43	1166410,0	8,07	1075650,3	0,34
фев 01	1109700	1085970,3	2,14	1092760,2	1,53	1110230,2	0,05
мар 01	1149500	1108080,3	3,60	1131810,0	1,54	1134921,7	1,27
апр 01	1210000	1130190,3	6,60	1171610,0	3,17	1174965,8	2,90
май 01	1233700	1152300,4	6,60	1232110,0	0,13	1235826,1	0,17
июн 01	1294300	1174410,4	9,26	1255810,0	2,97	1259826,9	2,66
июль 01	1330200	1196520,4	10,05	1277920,1	3,93	1286713,9	3,27
авг 01	1365500	1218630,5	10,76	1352310,0	0,97	1335690,2	2,18
сен 01	1414400	1240740,5	12,28	1374420,1	2,83	1384629,6	2,10
окт 01	1441200	1262850,6	12,38	1436510,0	0,33	1441979,4	0,05
ноя 01	1439100	1284960,6	10,71	1463310,0	1,68	1469101,0	2,08
Средняя ошибка прогноза			6,20		2,47		1,84

Таблица 6.3.15

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
117	-498483,752	6869,279	121	-589828,211	7773,412
118	-532312,207	7205,323	122	-638019,335	8245,874
119	-560816,089	7487,540	123	-646040,710	8324,259
120	-576472,511	7642,044	124	-652322,303	8385,443

Таблица 6.3.15 (продолжение)

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
125	-659299,727	8453,185	129	-745936,517	9287,212
126	-680298,565	8656,399	130	-770499,627	9521,892
127	-694883,545	8797,091	131	-797546,854	9779,485
128	-720806,658	9046,351	132	-816958,807	9963,775

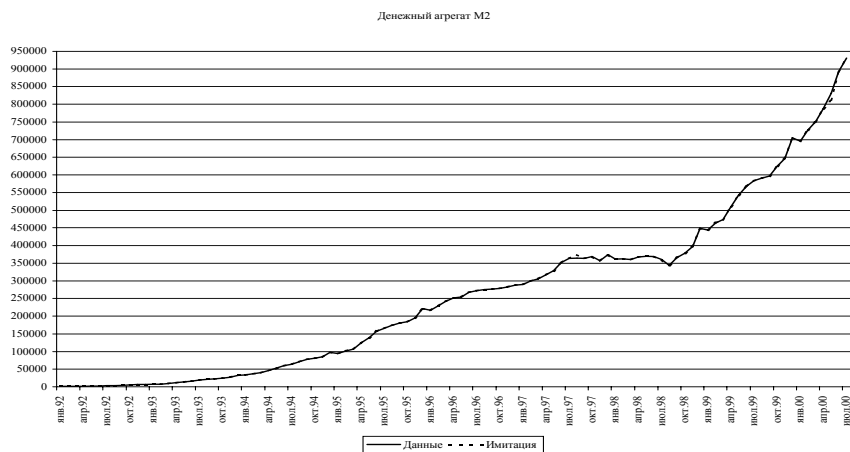
Рассмотрим теперь вероятностную модель прогнозирования. Поскольку в исходных данных явно содержится детерминированный тренд, то преобразуем их в темповые характеристики, т.е. введем новый ряд:

$y_t = \frac{x_t}{x_{t-1}}, t = 1, \dots, T$ . График соответствующего временного ряда также приведен на *рис. 6.3.16*.

Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель темпа роста  $M2$  и его лаговые значения до 12 месяцев, т.е. для системы  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя темпа роста  $M2$  от лага в 6 и 9 месяцев с коэффициентом информативности, равным 0,940. Это означает, что значения темпов роста  $M2$  6 и 9 месяцев назад определяют сегодняшний темп роста на 94,0%. Таким образом, модель прогнозирования безработицы в данном случае имеет вид:

$$\tilde{x}_t = x_{t-1} \sum_{y \in Y} yp(y | y_{t-6}, y_{t-9}).$$

На *рис. 6.3.19* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации на временном интервале 01.1992 – 07.2000 в этом случае практически полностью совпадают – среднее расхождение (MAPE) между исходными данными и их имитацией составляет 0,11%.



*Рис. 6.3.19.* Имитация

Для проверки качества прогнозирования, как и выше, использованы данные за период 08.2000 – 11.2001. Построено две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-6}, y_{t-9})$ ), построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-6}, y_{t-9})$  уточняется за счет добавления новой точки). В табл. 6.3.16 и на рис. 6.3.20 приведены результаты одношагового прогнозирования по приведенной модели. В этом случае результаты прогнозирования с обучением и без совпадают между собой. Средняя ошибка прогноза (MAPE) составляет 4,96%.

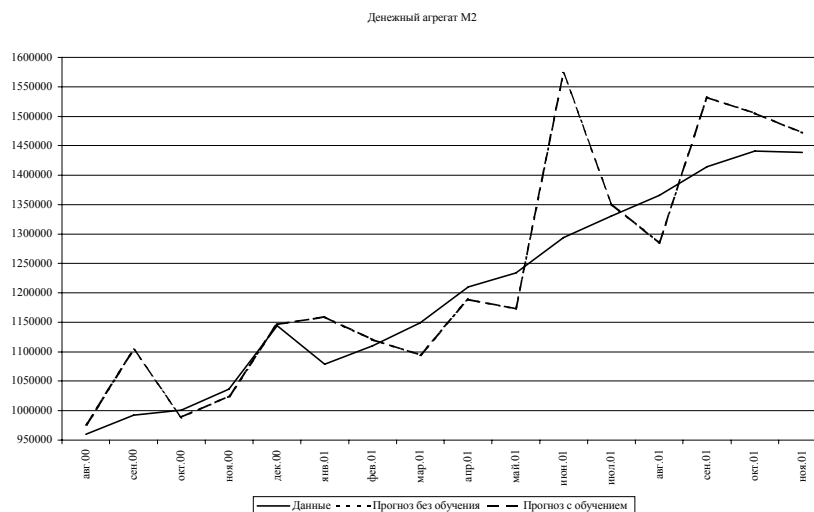


Рис. 6.3.20. Прогноз

Таблица 6.3.16

### Вероятностный прогноз М2

Дата	М2	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	Ошибка
авг.00	960100	976434,0	1,70	976434,0	1,70
сен.00	992400	1102370,2	11,08	1102370,2	11,08
окт.00	1001200	988645,0	1,25	988645,0	1,25
ноя.00	1036400	1024720,0	1,13	1024720,0	1,13
дек.00	1144300	1146105,5	0,16	1146105,5	0,16

Таблица 6.3.16 (продолжение)

Дата	M2	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	Ошибка
январь.01	1079300	1159212,0	7,40	1159212,0	7,40
февраль.01	1109700	1120365,2	0,96	1120365,2	0,96
март.01	1149500	1094425,3	4,79	1094425,3	4,79
апрель.01	1210000	1188734,2	1,76	1188734,2	1,76
май.01	1233700	1172597,6	4,95	1172597,6	4,95
июнь.01	1294300	1573512,6	21,57	1573512,6	21,57
июль.01	1330200	1351701,4	1,62	1351701,4	1,62
август.01	1365500	1284331,0	5,94	1284331,0	5,94
сентябрь.01	1414400	1531754,1	8,30	1531754,1	8,30
октябрь.01	1441200	1505436,7	4,46	1505436,7	4,46
ноябрь.01	1439100	1471591,6	2,26	1471591,6	2,26
Средняя ошибка прогноза			4,96		4,96

### 6.3.3. Динамика внешнеторговых характеристик

#### 6.3.3.1. Экспорт

В качестве исходных данных для построения модели прогнозирования экспорта используются месячные данные за период 01.1994 – 04.2000. Данные за период 05.2000 – 12.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на рис. 6.3.21.

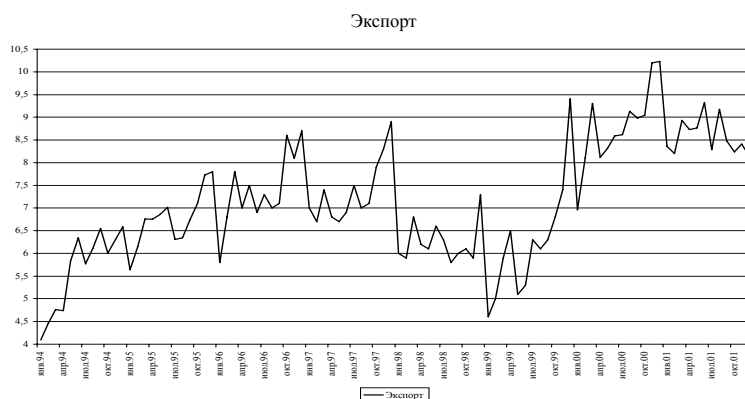


Рис. 6.3.21. Исходные данные

В соответствии с проведенным ранее исследованием ряд экспорта представляет собой стационарный относительно тренда ряд. Построим детерминированный линейный тренд ряда на временном промежутке 01.1994 – 04.2000. Получаем

$$f(t) = 6,194 + 0,013t.$$

Статистические характеристики построенного тренда приведены в табл. 6.3.17.

Таблица 6.3.17

Тренд	R	Нормированный R <sup>2</sup>	F-статистика	Значимость F	t-статистика c <sub>0</sub>	t-статистика c <sub>1</sub>
$f(t)$	0,260	0,055	5,379	0,023	25,666	2,319

Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 1 и 12 месяцев с коэффициентом информативности, равным 0,874. Это означает, что значения экспорта 1 и 12 месяцев назад определяют сегодняшний экспорт на 87,4%. Таким образом, модель прогнозирования инфляции имеет вид :

$$\tilde{x}_t = f(t) + \sum_{y \in Y} yp(y|y_{t-1}, y_{t-12}).$$

На рис. 6.3.19 приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации за период 01.1995 – 04.2000 полностью совпадают между собой.

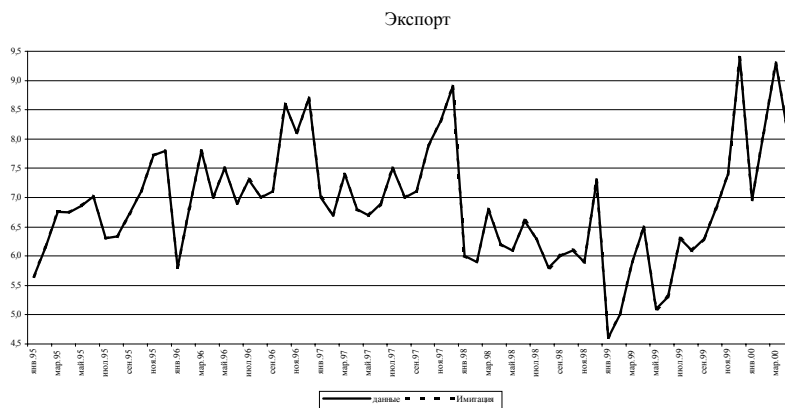


Рис. 6.3.22. Имитация

Для проверки качества прогнозирования использованы данные за период 05.2000 – 12.2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-12})$ , построенное на первоначальных исходных данных) и два с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-12})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В *табл. 6.3.18* и на *рис. 6.3.23* приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f$  меняются от шага к шагу. Соответствующие коэффициенты приведены в *табл. 6.3.19*. Средняя ошибка прогноза (МАРЕ) без обучения составляет 11,24%, а с обучением – 9,06% без коррекции тренда и 7,06% – с коррекцией.

Рассмотрим теперь вероятностную модель прогнозирования. Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель инфляции и лаговые значения до 12 месяцев. Получаем зависимость показателя экспорта от лага в 1 и 8 месяцев с коэффициентом информативности, равным 0,836. Это означает, что значения экспорта 1 и 8 месяцев назад определяют сегодняшнюю инфляцию на 83,6%. Таким образом, модель прогнозирования инфляции имеет вид:

$$\tilde{x}_t = \sum_{x \in X} xp(x | x_{t-1}, x_{t-8}).$$

На *рис. 6.3.24* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации полностью совпадают между собой на временном интервале 01.1995 – 04.2000 (так же, как и в случае имитации с трендом).

Экспорт



Рис. 6.3.23. Прогноз

Таблица 6.3.18

**Трендовый прогноз экспорта**

Дата	Экс-порт	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
май.00	8,32	7,52	9,55	7,52	9,55	7,52	9,55
июн.00	8,58	9,34	8,78	8,33	2,99	8,33	2,97
июл.00	8,61	6,48	24,73	6,48	24,73	6,61	23,19
авг.00	9,13	6,49	28,89	8,62	5,57	8,63	5,53
сен.00	8,98	7,56	15,89	7,56	15,89	8,20	8,72
окт.00	9,04	8,19	9,47	8,19	9,47	6,90	23,72
ноя.00	10,20	6,43	36,94	6,43	36,94	9,01	11,62
дек.00	10,23	8,72	14,74	8,72	14,74	9,21	9,92
январь.01	8,35	8,22	1,56	8,22	1,56	8,19	2,02
фев.01	8,20	7,61	7,22	7,61	7,22	8,43	2,77
мар.01	8,93	9,39	5,22	9,39	5,22	8,93	0,06
апр.01	8,73	9,36	7,16	9,36	7,16	9,41	7,78
май.01	8,77	9,37	6,88	9,37	6,88	8,09	7,70
июн.01	9,33	9,38	0,61	9,38	0,61	9,13	2,04
июль.01	8,29	8,81	6,28	8,81	6,28	8,58	3,59
авг.01	9,18	8,87	3,35	8,99	2,03	8,86	3,41
сен.01	8,48	9,47	11,70	8,31	1,94	8,68	2,36
окт.01	8,23	8,89	8,03	9,20	11,76	8,20	0,42
ноя.01	8,41	8,52	1,30	9,03	7,35	9,28	10,34
дек.01	8,16	9,51	16,56	8,42	3,27	8,44	3,46
Средняя ошибка прогноза			11,24		9,06		7,06



Таблица 6.3.19

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
77	6,194	0,013	82	6,018	0,019	87	5,851	0,025	92	5,778	0,028
78	6,164	0,014	83	5,982	0,021	88	5,831	0,026	93	5,760	0,028
79	6,130	0,015	84	5,922	0,023	89	5,816	0,026	94	5,758	0,028
80	6,097	0,016	85	5,865	0,025	90	5,803	0,028	95	5,762	0,028
81	6,054	0,018	86	5,855	0,025	91	5,778	0,028	96	5,762	0,028

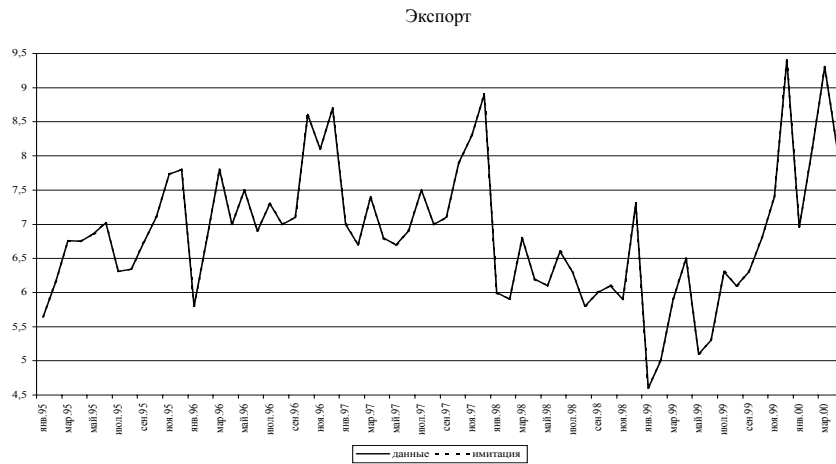


Рис. 6.3.24. Имитация  
Экспорт

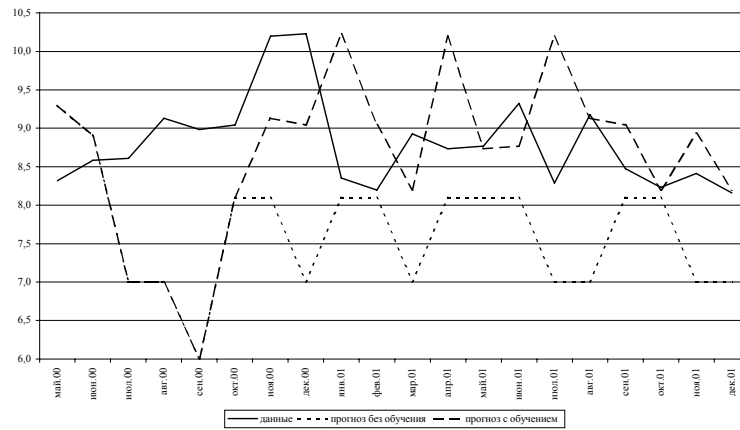


Рис. 6.3.25. Прогноз  
Экспорт

Для проверки качества прогнозирования по модели без тренда использованы данные за период 05.2000 – 12.2001. Построены две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(x_t | x_{t-1}, x_{t-8})$ , построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(x_t | x_{t-1}, x_{t-8})$  уточняется за счет добавления новой точки). В табл. 6.3.20 и на рис. 6.3.25 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (МАРЕ) без обучения составляет 14,17%, а с обучением 11,22%.

Таблица 6.3.20

**Вероятностный прогноз экспорта**

Дата	Экспорт	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
май.00	8,3	9,30	11,85	9,30	11,85
июн.00	8,6	8,90	3,68	8,90	3,68
июл.00	8,6	7,00	18,72	7,00	18,72
авг.00	9,1	7,00	23,35	7,00	23,35
сен.00	9,0	6,00	33,21	6,00	33,21
окт.00	9,0	8,10	10,42	8,10	10,42
ноя.00	10,2	8,10	20,56	9,13	10,43
дек.00	10,2	7,00	31,55	9,04	11,58
январь.01	8,4	8,10	3,04	10,23	22,41
фев.01	8,2	8,10	1,21	9,04	10,28
мар.01	8,9	7,00	21,59	8,20	8,16
апр.01	8,7	8,10	7,24	10,20	16,77
май.01	8,8	8,10	7,60	8,73	0,39
июн.01	9,3	8,10	13,14	8,77	5,99
июл.01	8,3	7,00	15,53	10,20	23,04
авг.01	9,2	7,00	23,71	9,13	0,47
сен.01	8,5	8,10	4,45	9,04	6,67
окт.01	8,2	8,10	1,62	8,20	0,41
ноя.01	8,4	7,00	16,77	8,93	6,15
дек.01	8,2	7,00	14,17	8,20	0,53
Средняя ошибка прогноза			14,17		11,22

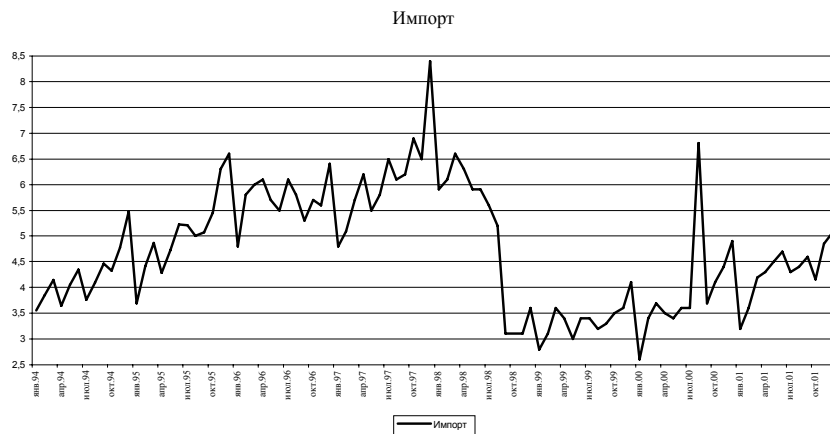
### 6.3.3.2. Импорт

В качестве исходных данных для построения модели прогнозирования экспорта используются месячные данные за период 01.1994 – 04.2000. Данные за период 05.2000 – 12.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на *рис. 6.3.26*.

В соответствии с проведенным ранее исследованием ряда импорта он стационарен относительно линейного тренда на временном интервале 01.1994 – 04.1998. Разобьем весь интервал наблюдений на 2 подинтервала: 01.1994 – 08.1998; 09.1998 – 04.2000. Для каждого из этих интервалов построим детерминированный линейный тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = 4,078 + 0,046t, & t = 1, \dots, 56; \\ f_2(t) = 1,950 + 0,021t, & t = 57, \dots, 76. \end{cases}$$

Статистические характеристики построенных трендов приведены в *табл. 6.3.21*.



*Рис. 6.3.26. Исходные данные*

*Таблица 6.3.21*

Тренд	R	Нормированный R	F-статистика	Значимость F	t-статистика $c_0$	t-статистика $c_1$
$f_1(t)$	0,769	0,583	78,133	0,000	24,074	8,839
$f_2(t)$	0,361	0,082	2,703	0,118	2,323	1,644

Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 6 и 12 месяцев с коэффициентом информативности, равным 0,803. Это означает, что значения инфляции 6 и 12 месяцев назад определяют сегодняшнюю инфляцию на 80,3%. Таким образом, модель прогнозирования инфляции имеет вид:

$$\tilde{x}_t = f_2(t) + \sum_{y \in Y} y p(y | y_{t-1}, y_{t-5}).$$

На *рис. 6.3.27* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации полностью совпадают между собой (так же как и при имитации экспорта).

Для проверки качества прогнозирования использованы данные за период 05.2000 – 12.2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-5})$ , построенное на первоначальных исходных данных) и два с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-5})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В *табл. 6.3.21* и на *рис. 6.3.28* приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_2$  меняются от шага к шагу. Соответствующие коэффициенты приведены в *табл. 6.3.22*. Средняя ошибка прогноза (МАРЕ) без обучения составляет 11,28%, а с обучением – 9,79% без коррекции тренда и 13,32% – с коррекцией. В данном случае (в отличие от предыдущих) прогноз с обучением и коррекцией тренда является худшим, а не лучшим. Это обусловлено всплеском объема импорта в августе 2000 г.

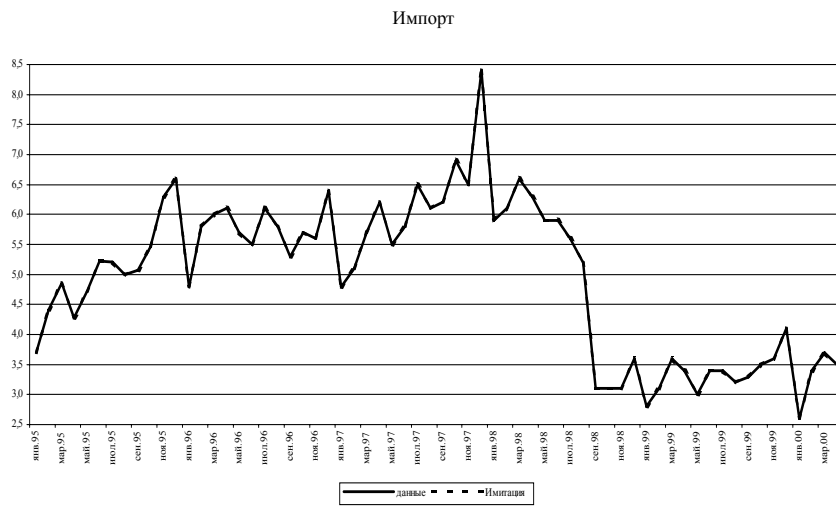


Рис. 6.3.27. Имитация

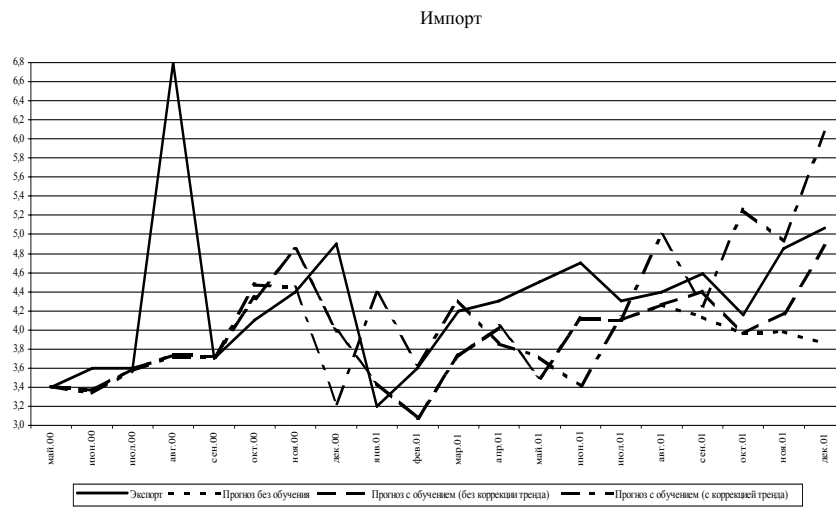


Рис. 6.3.28. Прогноз

Таблица 6.3.22

## Трендовый прогноз импорта

Дата	Импорт	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
Май.00	3,40	3,41	0,21	3,41	0,21	3,41	0,21
Июн.00	3,60	3,37	6,52	3,37	6,52	3,34	7,27
Июл.00	3,60	3,59	0,15	3,59	0,15	3,58	0,63
Авг.00	6,80	3,73	45,13	3,73	45,13	3,72	45,30
Сен.00	3,70	3,73	0,74	3,73	0,74	3,72	0,53
Окт.00	4,10	4,33	5,70	4,33	5,70	4,47	9,08
Ноя.00	4,40	4,84	9,93	4,84	9,93	4,45	1,11
Дек.00	4,90	3,98	18,76	3,98	18,76	3,23	34,15
Янв.01	3,20	3,44	7,64	3,44	7,64	4,38	36,79
Фев.01	3,60	3,08	14,38	3,08	14,38	3,64	1,06
Мар.01	4,20	3,73	11,25	3,73	11,25	4,31	2,58
Апр.01	4,30	4,03	6,30	4,03	6,30	3,86	10,34
Май.01	4,50	3,50	22,13	3,50	22,13	3,72	17,34
Июн.01	4,70	4,11	12,61	4,11	12,61	3,41	27,35
Июл.01	4,30	4,11	4,52	4,11	4,52	4,12	4,11
Авг.01	4,40	4,26	3,14	4,26	3,14	4,99	13,33
Сен.01	4,59	4,13	10,01	4,40	4,11	4,26	7,29
Окт.01	4,16	3,96	4,65	3,96	4,65	5,26	26,61
Ноя.01	4,85	3,98	17,91	4,18	13,92	4,93	1,62
Дек.01	5,07	3,86	23,92	4,87	3,91	6,07	19,67
Средняя ошибка прогноза			11,28		9,79		13,32

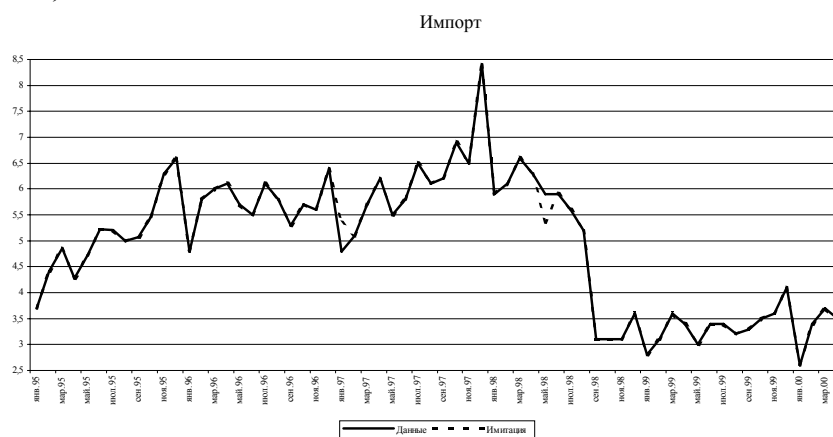
Таблица 6.3.23

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
77	1,950	0,021	82	0,197	0,048	87	0,616	0,042	92	0,502	0,043
78	2,067	0,019	83	0,212	0,048	88	0,631	0,041	93	0,529	0,043
79	2,018	0,020	84	0,093	0,050	89	0,622	0,042	94	0,509	0,043
80	1,995	0,020	85	-0,221	0,054	90	0,557	0,043	95	0,625	0,042
81	-0,069	0,052	86	0,323	0,046	91	0,449	0,044	96	0,551	0,043

Рассмотрим теперь вероятностную модель прогнозирования. Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель объема импорта и лаговые значения этого показателя до 12 месяцев, т.е. для системы  $\{x_t, x_{t-1}, \dots, x_{t-12}\}$ . Получаем зависимость показателя инфляции от лага в 9 и 11 месяцев с коэффициентом информативности, равным 0,867. Это означает, что значения объема импорта 9 и 11 месяцев назад определяют сегодняшний объем импорта на 86,7%. Таким образом, модель без тренда для прогнозирования импорта в данном случае имеет вид:

$$\tilde{x}_t = \sum_{x \in X} xp(x|x_{t-9}, x_{t-11}).$$

На *рис. 6.3.29* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации близки между собой. Средняя относительная ошибка (MAPE) имитации за период 01.1995 – 04.2000 составляет 0,54%.



*Рис. 6.3.29.* Имитация



Рис. 6.3.30. Прогноз

Для проверки качества прогнозирования, как и выше, использованы данные за период 05.2000 – 12.2001. Построено две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(x_t | x_{t-1}, x_{t-12})$ , построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(x_t | x_{t-1}, x_{t-12})$  уточняется за счет добавления новой точки). В табл. 6.3.24 и на рис. 6.3.30 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (MAPE) без обучения в этом случае составляет 20,60%, а с обучением 16,09%. Таким образом, в данном случае прогнозные свойства модели оказались крайне плохими.

Таблица 6.3.24

**Вероятностный прогноз импорта**

Дата	Импорт	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
май.00	3,4	3,70	8,82	3,70	8,82
июн.00	3,6	3,70	2,78	3,55	1,39
июл.00	3,6	2,95	18,06	2,95	18,06
авг.00	6,8	3,30	51,47	3,60	47,06
сен.00	3,7	5,48	48,11	5,48	48,11



Таблица 6.3.24 (продолжение)

Дата	Импорт	Прогноз без обучения	Прогноз с обучением	3,40	17,07
		результат	ошибка	результат	ошибка
дек.00	4,9	4,10	16,33	4,10	16,33
янв.01	3,2	3,70	15,63	3,70	15,63
фев.01	3,6	3,70	2,78	3,70	2,78
мар.01	4,2	3,70	11,90	3,20	23,81
апр.01	4,3	3,70	13,95	4,73	10,08
май.01	4,5	6,40	42,22	6,40	42,22
июн.01	4,7	3,70	21,28	4,20	10,64
июл.01	4,3	5,70	32,56	5,70	32,56
авг.01	4,4	4,86	10,45	4,86	10,45
сен.01	4,6	6,00	30,66	6,00	30,66
окт.01	4,2	5,46	31,41	4,40	5,90
ноя.01	4,9	5,46	12,55	5,46	12,55
дек.01	5,1	5,48	8,09	3,70	27,02
Средняя ошибка прогноза			20,60		16,09

#### 6.3.4. Безработица

В качестве исходных данных для построения модели прогнозирования безработицы используются месячные данные за период 01.1994 – 08.2000. Данные за период 09.2000 – 12.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на рис. 6.3.31.

В соответствии с проведенным ранее исследованием ряда безработицы на временном интервале 01.1994 – 04.1998 данный ряд является стационарным относительно линейного тренда. Разобьем весь интервал наблюдений на 2 подинтервала: 01.1994 – 08.1998 и 09.1998 – 08.2000. Для каждого из этих интервалов построим детерминированный линейный тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = 5,051 + 0,064t, & t = 1, \dots, 56; \\ f_2(t) = 15,749 - 0,102t, & t = 57, \dots, 80. \end{cases}$$

Статистические характеристики построенных трендов приведены в табл. 6.3.25.

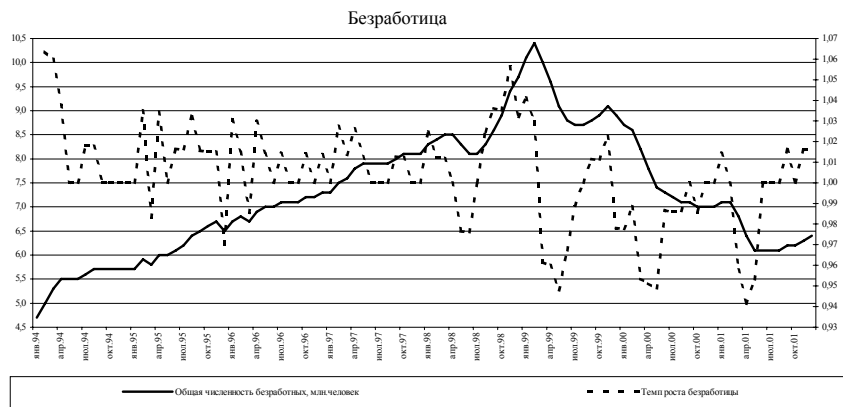


Рис. 6.3.31. Исходные данные

Таблица 6.3.25

Тренд	R	Нормированный R <sup>2</sup>	F-статистика	Значимость F	t-статистика c <sub>0</sub>	t-статистика c <sub>1</sub>
$f_1(t)$	0,989	0,978	2435,141	0,000	119,237	49,347
$f_2(t)$	0,800	0,623	38,987	0,000	13,979	-6,244

Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 1 и 12 месяцев с коэффициентом информативности, равным 0,876. Это означает, что значения инфляции 1 и 12 месяцев назад определяют сегодняшнюю инфляцию на 87,6%. Таким образом, модель прогнозирования инфляции имеет вид:

$$\tilde{x}_t = f_2(t) + \sum_{y \in Y} yp(y|y_{t-1}, y_{t-12}).$$

На рис. 6.3.32 приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из рисунка, графики исходных данных и имитации практически совпадают между собой. Средняя относительная ошибка (MAPE) имитации за период 01.1995–08.2000 составляет 0,02%.

Для проверки качества прогнозирования использованы данные за период 09.2000 – 12.2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятнос-

тей  $p(y_t | y_{t-1}, y_{t-12})$ , построенное на первоначальных исходных данных) и два с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-12})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В табл. 6.3.26 и на рис. 6.3.33 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_2$  меняются от шага к шагу. Соответствующие коэффициенты приведены в табл. 6.3.27. Средняя ошибка прогноза (МАРЕ) без обучения составляет 3,26%, а с обучением 3,67% без коррекции тренда и 4,01% – с коррекцией.

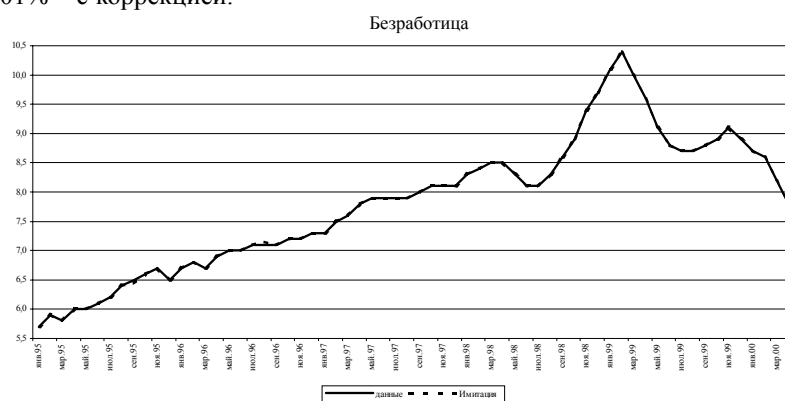


Рис. 6.3.32. Имитация

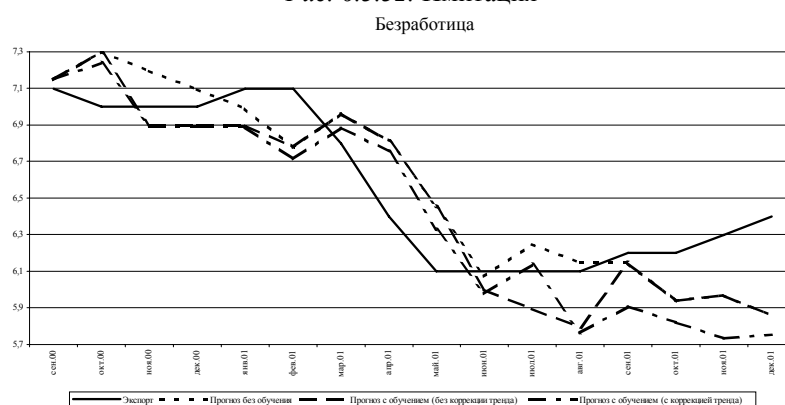


Рис. 6.3.33. Прогноз

Рассмотрим теперь построение вероятностной модели прогнозирования. В этом случае исходные данные считаются реализациями некоторой случайной величины. Поскольку в исходных данных явно содержится линейный тренд, то преобразуем их в темповые характеристики, т.е. введем новый ряд:  $y_t = \frac{x_t}{x_{t-1}}, t = 1, \dots, T$ . График соответствующего временного ряда также приведен на *рис.6.3.31*.

Таблица 6.3.26

**Трендовый прогноз безработицы**

Дата	Безработица	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
сен.00	7,10	7,15	0,68	7,15	0,68	7,15	0,68
окт.00	7,00	7,30	4,27	7,30	4,27	7,24	3,41
ноя.00	7,00	7,20	2,81	6,90	1,46	6,89	1,55
дек.00	7,00	7,09	1,35	6,90	1,46	6,89	1,56
январь.01	7,10	6,99	1,52	6,90	2,85	6,89	2,95
фев.01	7,10	6,78	4,53	6,78	4,53	6,72	5,41
мар.01	6,80	6,96	2,43	6,96	2,43	6,89	1,29
апр.01	6,40	6,81	6,43	6,81	6,43	6,75	5,53
май.01	6,10	6,45	5,77	6,45	5,77	6,33	3,72
июнь.01	6,10	6,07	0,43	6,00	1,67	5,98	1,99
июль.01	6,10	6,25	2,42	5,90	3,35	6,14	0,62
авг.01	6,10	6,15	0,75	5,79	5,02	5,76	5,51
сен.01	6,20	6,15	0,88	6,15	0,88	5,91	4,71
окт.01	6,20	5,94	4,21	5,94	4,21	5,82	6,13
ноя.01	6,30	5,97	5,26	5,97	5,26	5,73	9,01
дек.01	6,40	5,86	8,47	5,86	8,47	5,76	10,07
Средняя ошибка прогноза			3,26		3,67		4,01

Таблица 6.3.27

t	c <sub>0</sub>	c <sub>1</sub>	t	c <sub>0</sub>	c <sub>1</sub>	t	c <sub>0</sub>	c <sub>1</sub>	t	c <sub>0</sub>	c <sub>1</sub>
81	15,749	-0,102	85	16,239	-0,110	89	16,117	-0,108	93	16,392	-0,112
82	15,971	-0,106	86	16,153	-0,108	90	16,267	-0,110	94	16,328	-0,111
83	16,144	-0,108	87	16,037	-0,107	91	16,355	-0,111	95	16,242	-0,110
84	16,226	-0,110	88	16,020	-0,106	92	16,394	-0,112	96	16,113	-0,108

Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель темпа роста безработицы и его лаговые значения до 12 месяцев, т.е. для системы  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя темпа роста безработицы от лага в 1 и 10 месяцев с коэффициентом информативности, равным 0,906. Это означает, что значения темпов роста безработицы 1 и 10 месяцев назад определяют сегодняшний темп роста на 90,6%. Таким образом, модель прогнозирования безработицы в данном случае имеет вид:

$$\tilde{x}_t = x_{t-1} \sum_{y \in Y} yp(y | y_{t-1}, y_{t-10}).$$

На рис. 6.3.34 приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации и в этом случае почти совпадают. Средняя относительная ошибка (MAPE) имитации за период 01.1995 – 08.2000 составляет 0,06%.

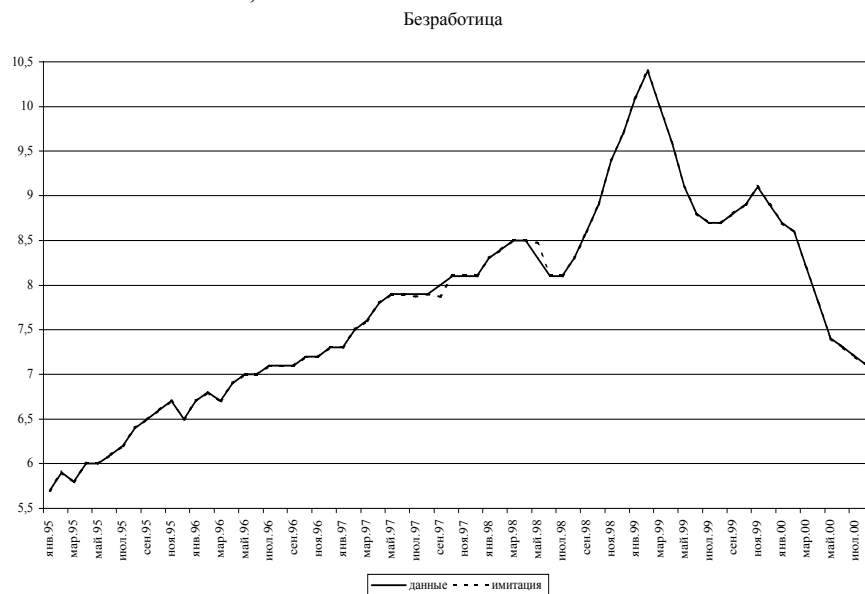


Рис. 6.3.34. Имитация

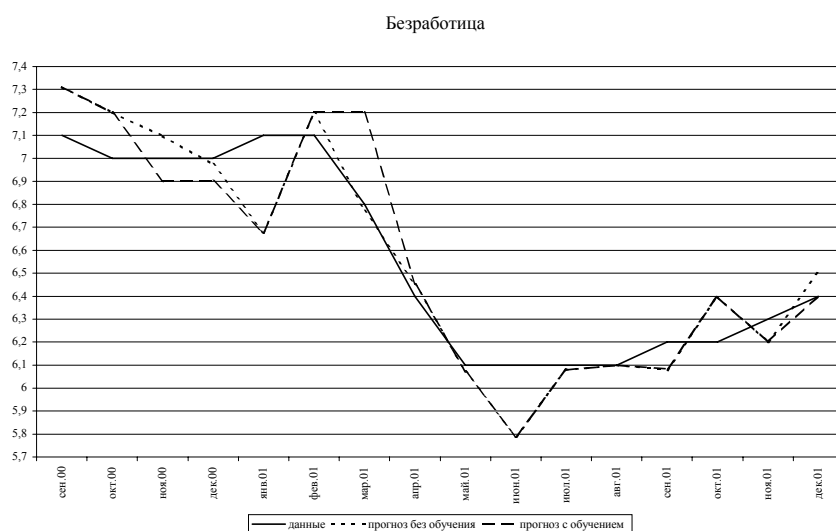


Рис. 6.3.35. Прогноз

Таблица 6.3.28

**Вероятностный прогноз безработицы**

Дата	Безработица	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
сен.00	7,1	7,31	2,99	7,31	2,99
окт.00	7,0	7,20	2,86	7,20	2,86
ноя.00	7,0	7,10	1,41	6,90	1,41
дек.00	7,0	6,97	0,36	6,90	1,41
январь.01	7,1	6,67	5,99	6,67	5,99
февраль.01	7,1	7,20	1,41	7,20	1,43
март.01	6,8	6,77	0,44	7,20	5,90
апрель.01	6,4	6,45	0,80	6,45	0,80
май.01	6,1	6,07	0,46	6,07	0,46
июнь.01	6,1	5,79	5,13	5,79	5,13
июль.01	6,1	6,08	0,36	6,08	0,36
август.01	6,1	6,10	0,00	6,10	0,00
сентябрь.01	6,2	6,08	1,97	6,08	1,88
октябрь.01	6,2	6,40	3,23	6,40	3,23
ноябрь.01	6,3	6,20	1,59	6,20	1,59
декабрь.01	6,4	6,50	1,61	6,40	0,03
Средняя ошибка прогноза			1,91		2,22

Для проверки качества прогнозирования, как и выше, использованы данные за период 09.2000 – 12.2001. Построено две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-10})$ ), построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-10})$  уточняется за счет добавления новой точки). В *табл. 6.3.28* и на *рис. 6.3.35* приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (MAPE) без обучения составляет 1,91%, а с обучением – 2,22%.

### 6.3.5. Доходы федерального бюджета

#### 6.3.5.1. Налоговые доходы федерального бюджета

В качестве исходных данных для построения модели прогнозирования налоговых доходов федерального бюджета используются месячные данные за период 01.1992 – 05.2000. Данные за период 06.2000 – 12.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на *рис. 6.3.36*.

В соответствии с проведенным ранее исследованием ряда налоговых доходов федерального бюджета он является стационарным в разностях на подпериоде 01.1996 – 05.2000. В то же время визуальный анализ графика ряда позволяет предположить наличие излома тренда во второй половине 1998 г. Разобьем весь интервал наблюдений на 2 подинтервала: 01.1992 – 08.1998 и 09.1998 – 05.2000. Для каждого из этих интервалов построим детерминированный линейный тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = -2797,443 + 335,083t, & t = 1, \dots, 80; \\ f_2(t) = -246389,160 + 3215,930t, & t = 81, \dots, 101. \end{cases}$$

Статистические характеристики построенных трендов приведены в *табл. 6.3.29*.

Таблица 6.3.29

Тренд	R	Нормированный R <sup>2</sup>	F-статистика	Значимость F	t-статистика c <sub>0</sub>	t-статистика c <sub>1</sub>
$f_1(t)$	0,836	0,695	181,141	0,000	-2,410	13,459
$f_2(t)$	0,944	0,885	155,253	0,000	-10,467	12,460

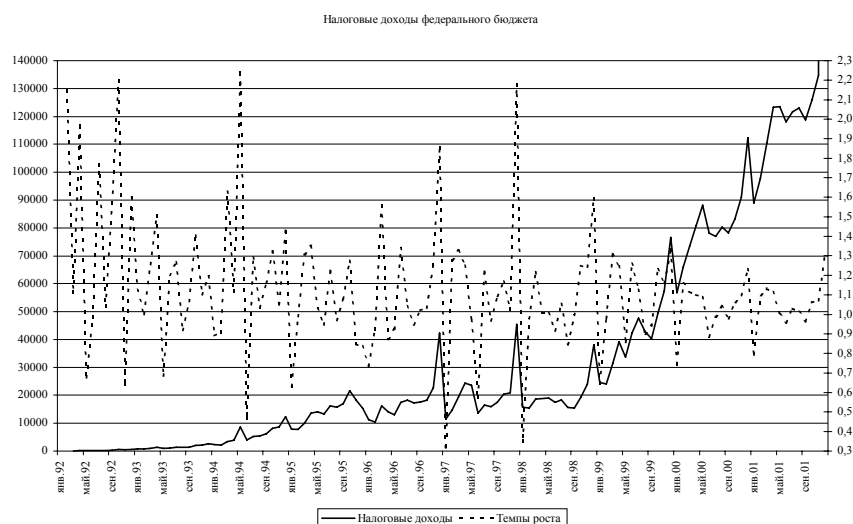


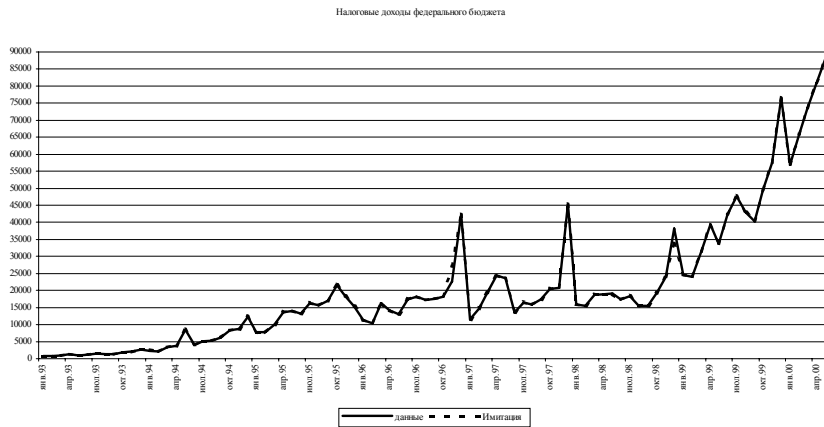
Рис. 6.3.36. Исходные данные

Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 3 и 9 месяцев с коэффициентом информативности, равным 0,902. Это означает, что значения налоговых доходов 3 и 9 месяцев назад определяют сегодняшний объем налоговых доходов федерального бюджета на 90,2%. Таким образом, модель прогнозирования налоговых доходов федерального бюджета имеет вид:

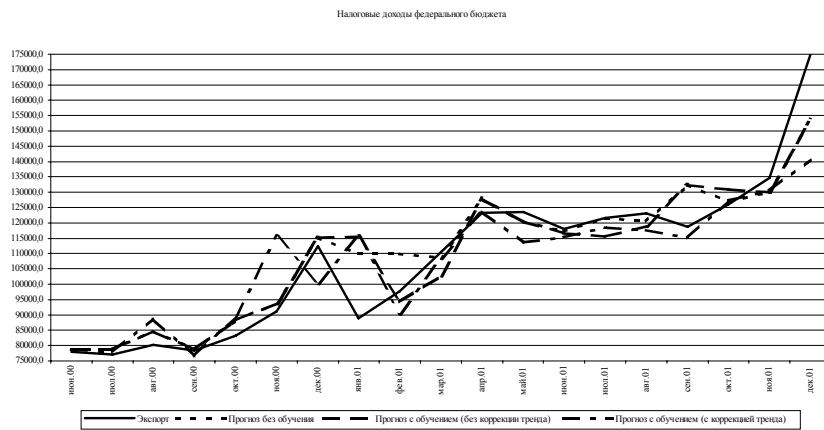
$$\tilde{x}_t = f_2(t) + \sum_{y \in Y} yp(y | y_{t-3}, y_{t-9})$$

На рис. 6.3.37 приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации близки между собой. Средняя относительная ошибка (MAPE) имитации за период 01.1993–05.2000 составляет 0,60%.





*Рис. 6.3.37. Имитация*



*Рис. 6.3.38. Прогноз*

Для проверки качества прогнозирования использованы данные за период 06.2000 – 12.2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-3}, y_{t-9})$ , построенное на первоначальных исходных данных) и два с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-3}, y_{t-9})$  уточняется за счет добавления новой точки, но в

одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В табл. 6.3.30 и на рис. 6.3.38 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_2$  меняются от шага к шагу. Соответствующие коэффициенты приведены в табл. 6.3.31. Средняя ошибка прогноза (МАРЕ) без обучения составляет 5,03%, а с обучением – 5,70% без коррекции тренда и 7,59% – с коррекцией.

Таблица 6.3.30

**Трендовый прогноз налоговых доходов федерального бюджета**

Дата	Налоговые доходы федерального бюджета	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
Июн.00	78032,05	78715,21	0,88	78715,21	0,88	78715,21	0,88
Июл.00	77035,19	78730,34	2,20	78730,34	2,20	78003,91	1,26
Авг.00	80207,21	84756,66	5,67	84756,66	5,67	88946,79	10,90
Сен.00	78308,11	78830,00	0,67	78830,00	0,67	76673,11	2,09
Окт.00	83175,24	88255,55	6,11	88255,55	6,11	89374,90	7,45
Ноя.00	91118,25	93760,65	2,90	93760,65	2,90	115637,57	26,91
Дек.00	112423,64	115075,10	2,36	115075,10	2,36	100166,54	10,90
Янв.01	88911,28	109833,24	23,53	115639,57	30,06	115441,47	29,84
Фев.01	97636,17	109944,84	12,61	94387,76	3,33	90254,42	7,56
Мар.01	110422,92	108626,58	1,63	102718,72	6,98	108484,01	1,76
Апр.01	123290,04	127938,82	3,77	127938,82	3,77	123806,63	0,42
Май.01	123489,51	120405,86	2,50	120405,86	2,50	113735,39	7,90
Июн.01	118092,78	117306,37	0,67	116623,21	1,24	115297,26	2,37
Июл.01	121577,55	121490,30	0,07	115626,35	4,89	118564,95	2,48
Авг.01	123081,87	120537,43	2,07	118842,28	3,44	117572,77	4,48
Сен.01	118676,28	132456,35	11,61	132456,35	11,61	115040,88	3,06
Окт.01	125962,34	126846,71	0,70	130956,50	3,96	127365,83	1,11
Ноя.01	134743,06	130062,64	3,47	130062,64	3,47	130363,23	3,25
Дек.01	175067,74	153666,26	12,22	153666,26	12,22	140685,36	19,64
Средняя ошибка прогноза			5,03		5,70		7,59

Таблица 6.3.31

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
102	-246389,160	3215,930	107	-217493,796	2888,962
103	-242643,055	3173,199	108	-217134,418	2884,939
104	-235793,167	3095,359	109	-229050,348	3017,831
105	-230566,490	3036,189	110	-222233,993	2942,093
106	-222443,028	2944,571	111	-220042,628	2917,835

Таблица 6.3.31 (продолжение)

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
112	-223654,253	2957,669	117	-238159,189	3116,752
113	-231764,318	3046,790	188	-235064,999	3083,362
114	-237120,517	3105,435	119	-234007,240	3071,988
115	-237672,240	3111,454	120	-235150,126	3084,233
116	-238299,867	3118,276			

Рассмотрим теперь вероятностную модель прогнозирования. Поскольку в исходных данных явно содержится тренд, то преобразуем их в темповые характеристики, т.е. введем новый ряд:  $y_t = \frac{x_t}{x_{t-1}}$ ,  $t = 1, \dots, T$ . График соответствующего временного ряда также приведен на *рис. 6.3.36*.

Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель темпа роста налоговых доходов федерального бюджета и его лаговые значения до 12 месяцев, т.е. для системы  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя темпа роста налоговых доходов бюджета от лага в 4 и 11 месяцев с коэффициентом информативности, равным 0,914. Это означает, что значения темпов роста налоговых доходов 4 и 11 месяцев назад определяют сегодняшний темп роста на 91,4%. Таким образом, модель прогнозирования безработицы в данном случае имеет вид:

$$\tilde{x}_t = x_{t-1} \sum_{y \in Y} yp(y | y_{t-4}, y_{t-11}).$$

На *рис. 6.3.39* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели за период 01.1993 – 05.2000. Как видно из приведенного рисунка, графики исходных данных и имитации в этом случае совпадают.

Налоговые доходы федерального бюджета

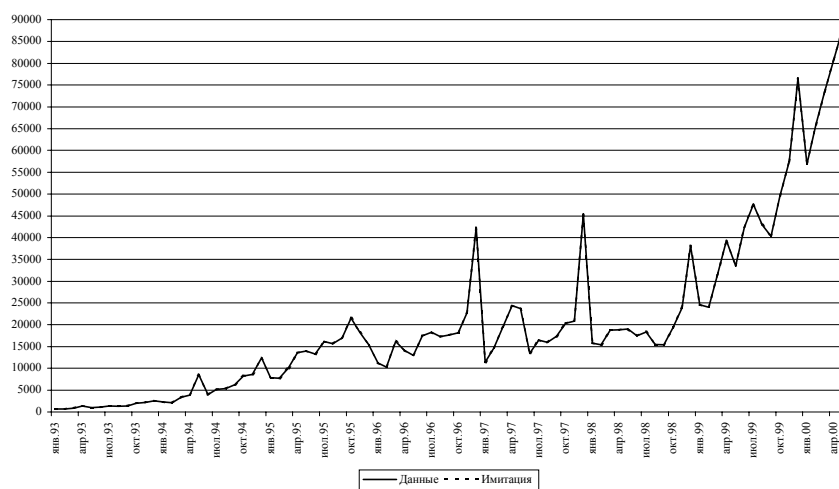


Рис. 6.3.39. Имитация

Налоговые доходы федерального бюджета

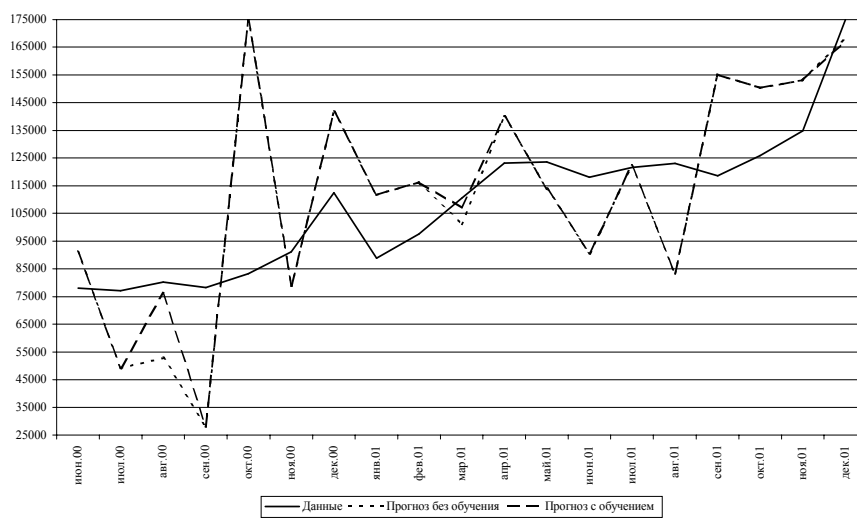


Рис. 6.3.40. Прогноз

Для проверки качества прогнозирования, как и выше, использованы данные за период 06.2000 – 12.2001. Построено две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-4}, y_{t-11})$ ), построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-4}, y_{t-11})$  уточняется за счет добавления новой точки). В табл. 6.3.32 и на рис. 6.3.40 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (MAPE) без обучения составляет 26,35%, а с обучением – 24,56%.

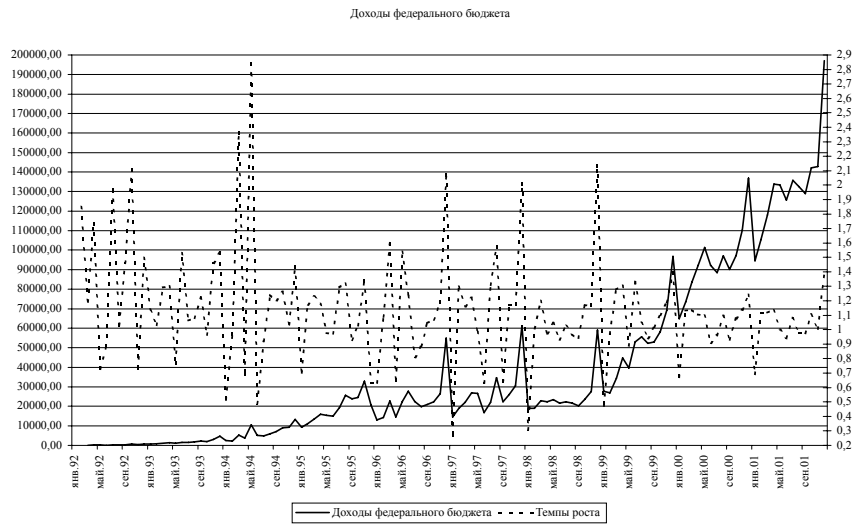
Таблица 6.3.32

**Вероятностный прогноз налоговых доходов федерального бюджета**

Дата	Налоговые доходы федерального бюджета	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
июн.00	78032,0	91210,0	16,89	91210,0	16,89
июл.00	77035,2	49169,8	36,17	49169,8	36,17
авг.00	80207,2	52731,8	34,26	76051,1	5,18
сен.00	78308,1	27936,7	64,32	27936,7	64,32
окт.00	83175,2	175701,8	111,24	175701,8	111,24
ноя.00	91118,3	78696,6	13,63	78696,6	13,63
дек.00	112423,6	141675,5	26,02	141675,5	26,02
янв.01	88911,3	111704,7	25,64	111704,7	25,64
фев.01	97636,2	116215,3	19,03	116215,3	19,03
мар.01	110422,9	101048,7	8,49	107217,2	2,90
апр.01	123290,0	139901,5	13,47	139901,5	13,47
май.01	123489,5	113849,5	7,81	113849,5	7,81
июн.01	118092,8	90623,8	23,26	90623,8	23,26
июл.01	121577,6	122220,3	0,53	122220,3	0,53
авг.01	123081,9	83221,7	32,39	83221,7	32,39
сен.01	118676,3	155141,6	30,73	155141,6	30,73
окт.01	125962,3	150399,6	19,40	150399,6	19,40
ноя.01	134743,1	153207,6	13,70	153207,6	13,70
дек.01	175067,7	168575,2	3,71	167412,1	4,37
Средняя ошибка прогноза			26,35		24,56

### 6.3.5.2. Доходы федерального бюджета

В качестве исходных данных для построения модели прогнозирования общих доходов федерального бюджета используются месячные данные за период 01.1992 – 05.2000. Данные за период 06.2000 – 12.2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на *рис. 6.3.41*.



*Рис. 6.3.41.* Исходные данные

В соответствии с проведенным ранее исследованием ряда доходов федерального бюджета на подпериоде 01.1996 – 05.2000 он является  $DS$ -рядом. В то же время визуальный анализ графика ряда позволяет предположить наличие излома тренда во второй половине 1998 г. Разобьем весь интервал наблюдений на 2 подинтервала: 01.1992 – 08.1998 и 09.1998 – 05.2000. Для каждого из этих интервалов построим детерминированный линейный тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = -3854,891 + 436,823t, & t = 1, \dots, 80; \\ f_2(t) = -246389,160 + 3215,930t, & t = 81, \dots, 101. \end{cases}$$

Статистические характеристики построенных трендов приведены в *табл. 6.3.33*.

Таблица 6.3.33

Тренд	R	Нормированный R <sup>2</sup>	F-статистика	Значимость F	t-статистика c <sub>0</sub>	t-статистика c <sub>1</sub>
f <sub>1</sub> (t)	0,818	0,665	157,917	0,000	-2,379	12,567
f <sub>2</sub> (t)	0,906	0,810	86,526	0,000	-7,710	9,302

Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя  $y_t$  от лага в 1 и 7 месяцев с коэффициентом информативности, равным 0,711. Это означает, что значения доходов федерального бюджета 1 и 7 месяцев назад определяют сегодняшний объем налоговых доходов федерального бюджета на 71,1%. Таким образом, модель прогнозирования инфляции имеет вид:

$$\tilde{x}_t = f_2(t) + \sum_{y \in Y} \gamma P(y | y_{t-1}, y_{t-7}).$$

На рис. 6.3.42 приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из приведенного рисунка, графики исходных данных и имитации не так близки между собой, как во всех предыдущих случаях. Средняя относительная ошибка (MAPE) имитации за период 01.1993–05.2000 составляет 6,23%.

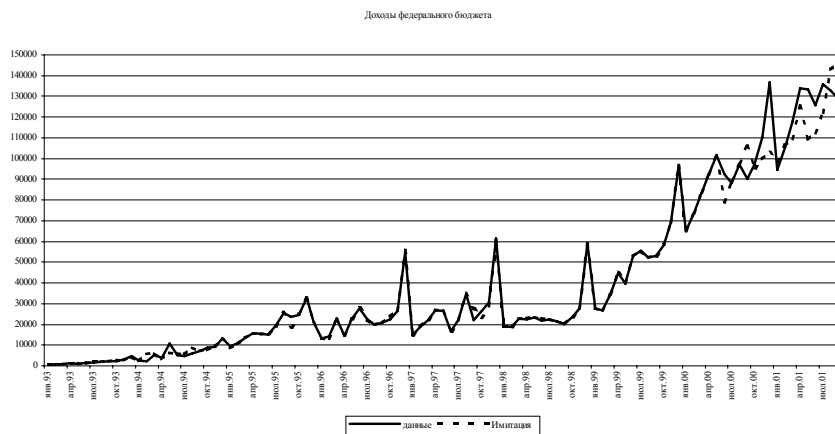


Рис. 6.3.42. Имитация

Для проверки качества прогнозирования использованы данные за период 06.2000 – 12.2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-7})$ , построенное на первоначальных исходных данных) и два с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-1}, y_{t-7})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В табл. 6.3.34 и на рис. 6.3.43 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_2$  меняются от шага к шагу. Соответствующие коэффициенты приведены в табл. 6.3.35. Средняя ошибка прогноза (МАРЕ) без обучения составляет 10,27%, а с обучением – 9,16% без коррекции тренда и 9,90% – с коррекцией.

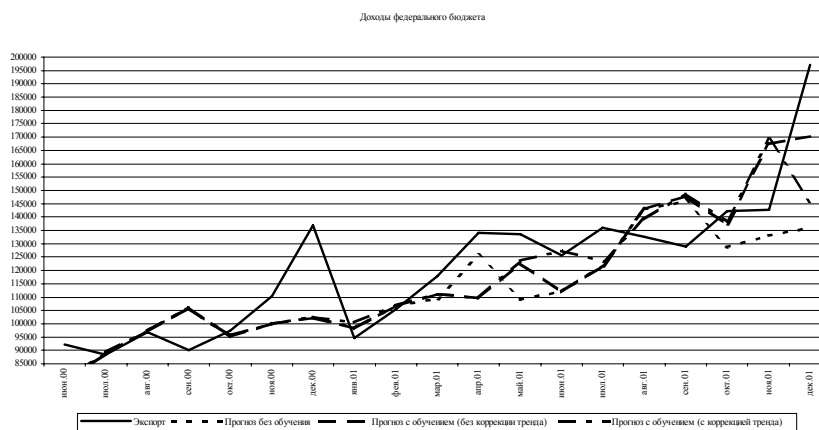


Рис. 6.3.43. Прогноз

Таблица 6.3.34

#### Трендовый прогноз доходов федерального бюджета

Дата	Доходы федерального бюджета	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
Июн.00	92225	78775,1	14,58	78775,1	14,58	78775,1	14,58
Июл.00	88387	88813,3	0,48	88813,3	0,48	89392,9	1,14



Таблица 6.3.34 (продолжение)

Дата	Доходы федерального бюджета	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
Авг.00	96925	97386,9	0,48	97386,9	0,48	97398,4	0,49
Сен.00	90147	106084,9	17,68	106084,9	17,68	106249,5	17,86
Окт.00	97210	95566,6	1,69	95566,6	1,69	95187,1	2,08
Ноя.00	110286	99903,3	9,41	99982,3	9,34	99918,3	9,40
Дек.00	136894	102675,7	25,00	102249,8	25,31	102287,5	25,28
Янв.01	94602	98182,4	3,78	98182,4	3,78	100417,3	6,15
Фев.01	105459	106656,6	1,14	106656,6	1,14	107051,1	1,51
Мар.01	117906	109429,0	7,19	111072,2	5,80	110978,2	5,88
Апр.01	133929	125492,2	6,30	109554,1	18,20	109589,9	18,17
Май.01	133450	109272,3	18,12	122722,0	8,04	123747,8	7,27
Июн.01	125606	112044,8	10,80	125494,5	0,09	127325,4	1,37
Июл.01	135925	122083,1	10,18	121657,1	10,50	123434,2	9,19
Авг.01	132578	143037,2	7,89	143037,2	7,89	139471,9	5,20
Сен.01	128974	145809,7	13,05	147791,0	14,59	148685,3	15,28
Окт.01	142335	128836,4	9,48	137313,5	3,53	138332,0	2,81
Ноя.01	142814	133173,0	6,75	167391,0	17,21	169191,3	18,47
Дек.01	197117	135945,4	31,03	170163,5	13,67	145755,5	26,06
Средняя ошибка прогноза			10,27		9,16		9,90

Таблица 6.3.35

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
102	-270413,670	3577,432	107	-241473,723	3250,102
103	-268064,691	3550,638	108	-244309,058	3281,840
104	-259203,816	3449,946	109	-262041,907	3479,605
105	-256849,456	3423,293	110	-247993,809	3323,515
106	-246659,941	3308,374	111	-240922,558	3245,235

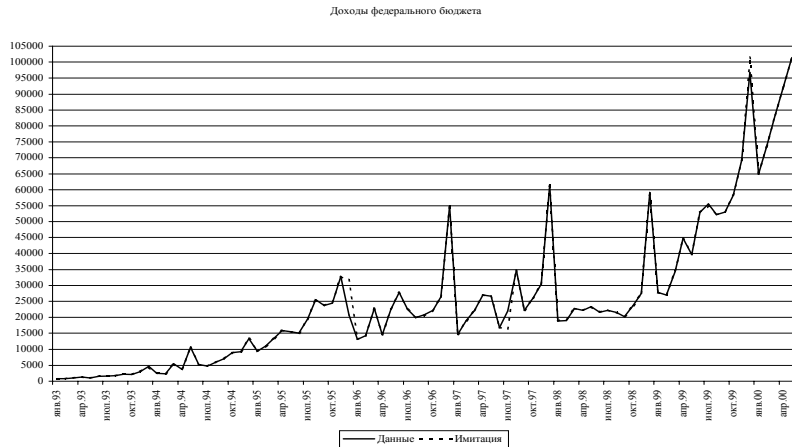
$t$	$c_0$	$c_1$	$t$	$C_0$	$c_1$
112	-240159,029	3236,814	117	-245650,493	3297,443
113	-246138,409	3302,521	118	-241230,811	3249,749
114	-249266,028	3336,765	119	-241266,754	3250,135
115	-246715,101	3308,937	120	-240302,609	3239,805
116	-247640,641	3318,997			

Рассмотрим теперь вероятностную модель прогнозирования. Поскольку в исходных данных явно содержится тренд, то преобразуем их в темповые характеристики, т.е. введем новый ряд:  $y_t = \frac{x_t}{x_{t-1}}, t = 1, \dots, T$ .

График соответствующего временного ряда также приведен на *рис. 6.3.41*. Построим информативную структуру 3-го порядка для системы, содержащей 13 переменных: показатель темпа роста доходов федерального бюджета и его лаговые значения до 12 месяцев, т.е. для системы  $\{y_t, y_{t-1}, \dots, y_{t-12}\}$ . Получаем зависимость показателя темпа роста доходов бюджета от лага в 5 и 8 месяцев с коэффициентом информативности, равным 0,873. Это означает, что значения темпов роста доходов бюджета 5 и 8 месяцев назад определяют сегодняшний темп роста на 87,3%. Таким образом, модель прогнозирования безработицы в данном случае имеет вид:

$$\tilde{x}_t = x_{t-1} \sum_{y \in Y} yp(y | y_{t-5}, y_{t-8}).$$

На *рис. 6.3.44* приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели за период 01.1993 – 05.2000. Как видно из приведенного рисунка, графики исходных данных и имитации в этом случае достаточно близки. Средняя относительная ошибка (MAPE) имитации за анализируемый период составляет 1,15%.



*Рис. 6.3.44.* Имитация

Для проверки качества прогнозирования, как и выше, использованы данные за период 06.2000 – 12.2001. Построено два одношаговых прогноза: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-5}, y_{t-8})$ , построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-5}, y_{t-8})$  уточняется за счет добавления новой точки). В табл. 6.3.36 и на рис. 6.3.45 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (MAPE) без обучения составляет 27,97%, а с обучением 27,32%.

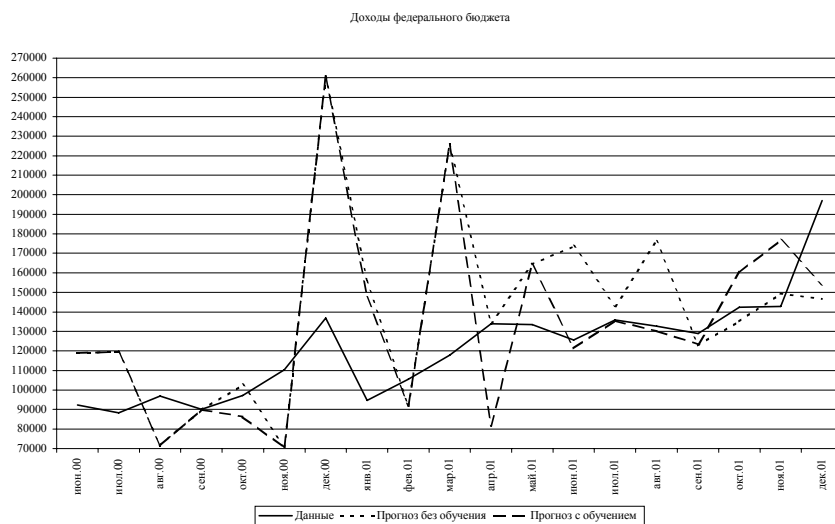


Рис. 6.3.45. Прогноз

Таблица 6.3.36

**Вероятностный прогноз налоговых доходов федерального бюджета**

Дата	Доходы федерального бюджета	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
июн.00	92224,8	118875,39	28,90	118875,39	28,90
июл.00	88387,4	119659,80	35,38	119659,80	35,38
авг.00	96925,3	71616,34	26,11	71616,34	26,11
сен.00	90146,7	89927,48	0,24	89927,48	0,24
окт.00	97209,8	102599,57	5,54	86395,79	11,12

Таблица 6.3.36 (продолжение)

Дата	Доходы федерального бюджета	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	Ошибка
ноя.00	110286,2	70618,59	35,97	70618,59	35,97
дек.00	136893,8	260729,22	90,46	260729,22	90,46
январь.01	94601,9	155804,27	64,69	147619,60	56,04
февр.01	105459,5	91764,47	12,99	91764,47	12,99
мар.01	117906,2	225713,88	91,44	225713,88	91,44
апр.01	133928,6	134193,75	0,20	81480,37	39,16
май.01	133450,1	164458,54	23,24	164458,54	23,24
июн.01	125605,9	173707,92	38,30	121314,88	3,42
июль.01	135925,3	142957,11	5,17	135447,32	0,35
авг.01	132578,1	176360,36	33,02	130269,56	1,74
сентябрь.01	128974,0	123006,21	4,63	123306,09	4,39
октябрь.01	142335,0	135366,58	4,90	160090,10	12,47
ноябрь.01	142813,5	149389,85	4,60	176674,60	23,71
декабрь.01	197117,4	146438,65	25,71	154003,14	21,87
Средняя ошибка прогноза			27,97		27,32

### 6.3.6. Валовой внутренний продукт

В качестве исходных данных для построения модели прогнозирования валового внутреннего продукта используются квартальные данные за период I/1994 – II/2000. Данные за период III/2000 – IV/2001 используются для проверки построенных в разделе моделей прогнозирования. График этих данных приведен на *рис. 6.3.46*.

В соответствии с проведенным ранее исследованием ряда ВВП на временном интервале I/1994 – II/2000 нельзя сделать вывод о его принадлежности к *DS*- или *TS*-рядам. В то же время визуальный анализ графика ряда позволяет выдвинуть гипотезу о наличии излома тренда начиная с IV квартала 1998 г. Разобьем весь интервал наблюдений на 2 подинтервала: I/1994 – III/1998 и IV/1998 – II/2000. Для каждого из этих интервалов построим детерминированный линейный тренд. Получаем

$$f(t) = \begin{cases} f_1(t) = 128,737 + 32,978t, & t = 1, \dots, 19; \\ f_2(t) = -1953,286 + 139,000t, & t = 20, \dots, 25. \end{cases}$$

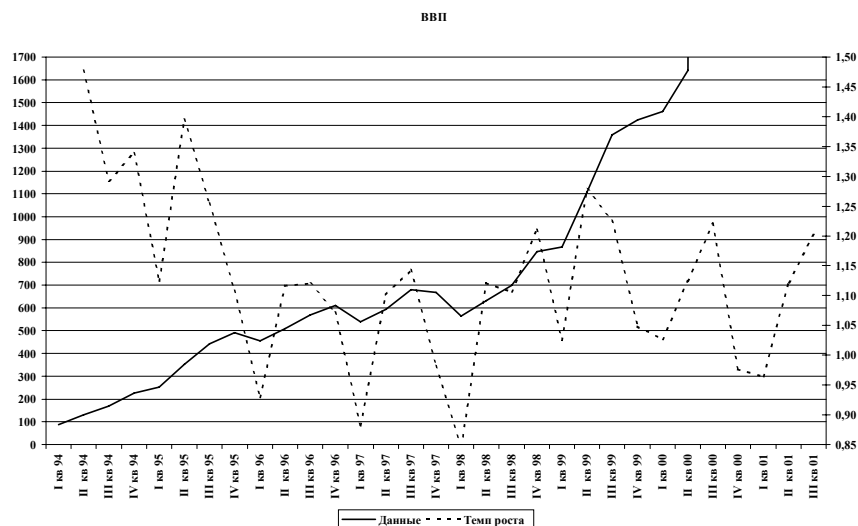


Рис. 6.3.46. Исходные данные

Статистические характеристики построенных трендов приведены в табл. 6.3.37.

Таблица 6.3.37

Тренд	R	Нормированный R <sup>2</sup>	F-статистика	Значимость F	t-статистика c <sub>0</sub>	t-статистика c <sub>1</sub>
f <sub>1</sub> (t)	0,940	0,876	128,311	0,0000	3,904	11,327
f <sub>2</sub> (t)	0,974	0,939	93,501	0,0002	-5,886	9,670

Рассмотрим временной ряд  $\{y_t\} = \{x_t - f(t)\}$ . Построим информативную структуру 3-го порядка для системы, содержащей 5 переменных:  $\{y_t, y_{t-1}, \dots, y_{t-4}\}$ . Получаем зависимость показателя  $y_t$  от лага в 2 и 4 квартала с коэффициентом информативности, равным 0,900. Это означает, что значения ВВП 2 и 4 квартала назад определяют сегодняшний ВВП на 90,0%. Таким образом, модель прогнозирования ВВП имеет вид:

$$\tilde{x}_t = f_2(t) + \sum_{y \in Y} y p(y | y_{t-2}, y_{t-4}).$$

На рис. 6.3.47 приведены результаты имитации (ретропрогноз) исходных данных по приведенной модели. Как видно из рисунка, графики исход-

ных данных и имитации за период I/1995 – II/2000 полностью совпадают между собой.

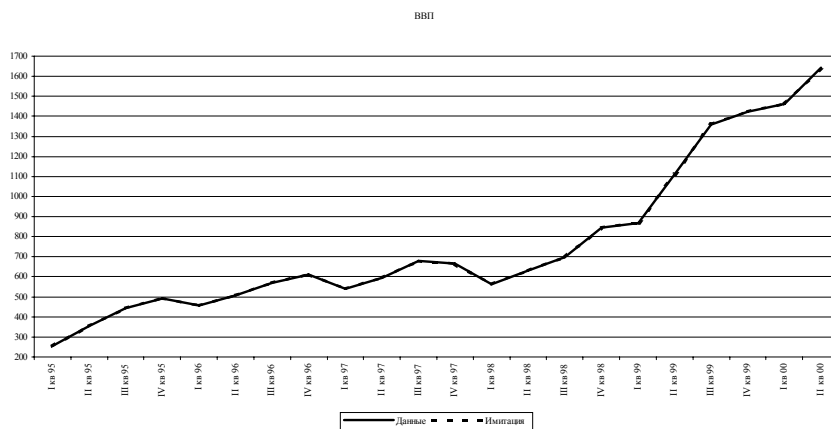


Рис. 6.3.47. Имитация

Для проверки качества прогнозирования использованы данные за период III/2000 – IV/2001. Построены три последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(y_t | y_{t-2}, y_{t-4})$ , построенное на первоначальных исходных данных) и два с обучением (в обоих случаях на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-2}, y_{t-4})$  уточняется за счет добавления новой точки, но в одном используется один и тот же тренд, а во втором тренд корректируется на каждом шаге). В табл. 6.3.38 и на рис. 6.3.48 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. При этом для случая прогноза с обучением и коррекцией тренда коэффициенты функции  $f_2$  меняются от шага к шагу. Соответствующие коэффициенты приведены в табл. 6.3.39. Средняя ошибка прогноза (МАРЕ) без обучения составляет 5,93%, а с обучением – 4,47% без коррекции тренда и 4,63% – с коррекцией.

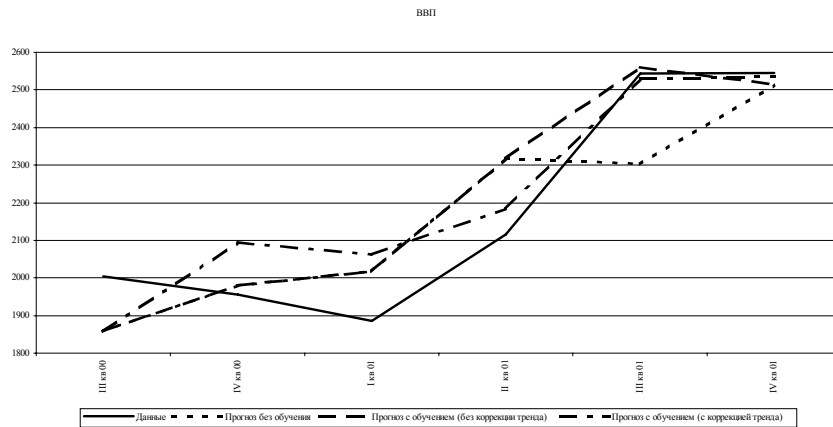


Рис. 6.3.48. Прогноз

Таблица 6.3.38

**Трендовый прогноз ВВП**

Дата	ВВП	Прогноз без обучения		Прогноз с обучением (без коррекции тренда)		Прогноз с обучением (с коррекцией тренда)	
		результат	ошибка	результат	ошибка	результат	ошибка
III кв 00	2004	1858,5	7,26	1858,5	7,26	1858,5	7,26
IV кв 00	1956	1980,0	1,23	1980,0	1,23	2094,2	7,06
I кв 01	1886	2017,0	6,95	2017,0	6,95	2062,5	9,36
II кв 01	2116	2316,9	9,49	2316,9	9,49	2183,9	3,21
III кв 01	2543	2303,5	9,42	2560,0	0,67	2530,5	0,49
IV кв 01	2545	2514,0	1,22	2514,0	1,22	2535,4	0,38
Средняя ошибка прогноза			5,93		4,47		4,63

Таблица 6.3.39

$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$	$t$	$c_0$	$c_1$
27	-1953,286	139,000	29	-2201,467	150,367	31	-1775,636	131,636
28	-2327,810	156,024	30	-1863,733	135,467	32	-1989,145	140,787

Рассмотрим теперь построение вероятностной модели прогнозирования. В этом случае исходные данные считаются реализациями некоторой случайной величины. Поскольку в исходных данных явно содержится тренд, то преобразуем их в темповые характеристики, т.е. введем новый

$$\text{ряд: } y_t = \frac{x_t}{x_{t-1}}, t = 1, \dots, T.$$

График соответствующего временного ряда также приведен на рис. 6.3.46. Построим информативную структуру 3-го порядка для системы, содержащей 5 переменных: показатель темпа роста ВВП и его лаговые значения до 4 кварталов, т.е. для системы  $\{y_t, y_{t-1}, \dots, y_{t-4}\}$ . Получаем зависимость показателя темпа роста доходов бюджета от лага в 3 и 4 квартала с коэффициентом информативности, равным 0,850. Это означает, что значения темпов роста ВВП 3 и 4 квартала назад определяют сегодняшний темп роста на 85,0%. Таким образом, модель прогнозирования безработицы в данном случае имеет вид:

$$\tilde{x}_t = x_{t-1} \sum_{y \in Y} yp(y | y_{t-3}, y_{t-4}).$$

На рис. 6.3.49 представлены результаты имитации (ретропрогноз) исходных данных по приведенной модели за период II/1995 – II/2000. Как видно из рисунка, графики исходных данных и имитации в этом случае полностью совпадают.

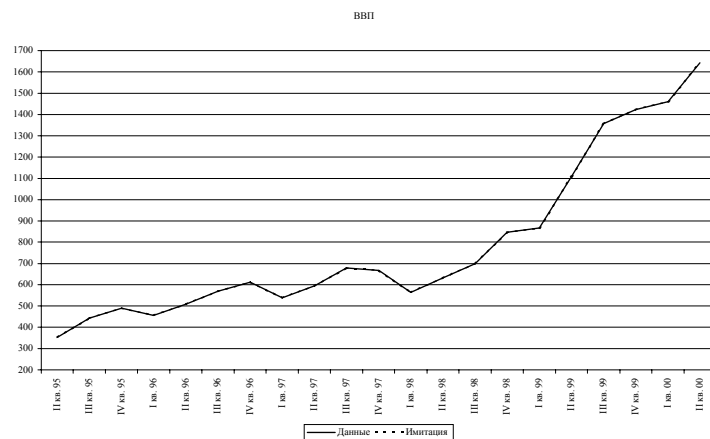


Рис. 6.3.49. Имитация



Для проверки качества прогнозирования, как и выше, использованы данные за период III/2000 – IV/2001. Построено две последовательности одношаговых прогнозов: без обучения (используется условное распределение вероятностей  $p(z_t | z_{t-3}, z_{t-4})$ ), построенное на первоначальных исходных данных) и с обучением (на каждом шаге условное распределение вероятностей  $p(y_t | y_{t-5}, y_{t-8})$  уточняется за счет добавления новой точки). В табл. 6.3.40 и на рис. 6.3.50 приведены результаты одношагового прогнозирования (без обучения и с обучением) по приведенной модели. Средняя ошибка прогноза (MAPE) без обучения составляет 15,06%, а с обучением – 11,98%.

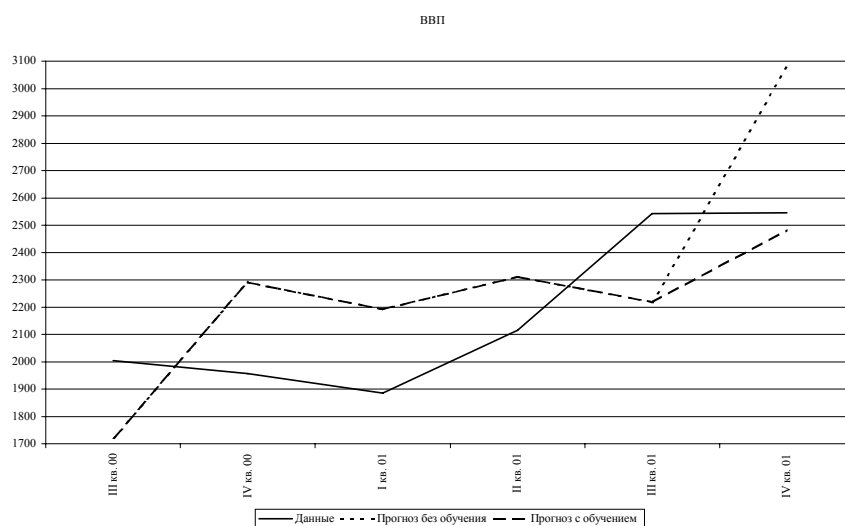


Рис. 6.3.50. Прогноз

Таблица 6.3.40

### Вероятностный прогноз ВВП

Дата	ВВП	Прогноз без обучения		Прогноз с обучением	
		результат	ошибка	результат	ошибка
III кв. 00	2004	1721,8	14,08	1721,8	14,08
IV кв. 00	1956	2290,8	17,11	2290,8	17,11
I кв. 01	1886	2191,8	16,22	2191,8	16,22
II кв. 01	2116	2311,5	9,24	2311,5	9,24
III кв. 01	2543	2218,8	12,75	2218,8	12,75
IV кв. 01	2545	3077,8	20,93	2482,1	2,47
Средняя ошибка прогноза			15,06		11,98

Проведем краткий анализ полученных результатов. Практически все построенные модели хорошо описывают динамику изменения исходных данных (см. табл. 6.3.41). Исключение составляет лишь трендовая модель доходов федерального бюджета, для которой средняя ошибка (МАРЕ) имитации составляет 6,23%. В то же время для денежных агрегатов *M0* и *M1*, динамики экспорта и ВВП обе модели дали полное совпадение имитации с исходными данными; для динамики импорта полностью совпала с исходными данными имитация по трендовой модели, а для динамики ВВП – по вероятностной модели.

Таблица 6.3.41

**Средние ошибки имитации (в %)**

Показатель	Трендовая модель	Вероятностная модель
Инфляция	0,40	0,94
M0	0,00	0,00
M1	0,00	0,00
M2	0,51	0,11
Экспорт	0,00	0,00
Импорт	0,00	0,54
Безработица	0,02	0,06
Налоговые доходы федерального бюджета	0,60	0,00
Доходы федерального бюджета	6,23	1,15
ВВП	0,00	0,00

Рассмотрим так называемую «наивную модель», отвечающую случайному блужданию, т.е. ситуации, когда будущее значение временного ряда является значением некоторой случайной величины, симметрично распределенной вокруг текущего значения. В этом случае наилучшим прогнозом будущего значения является текущее значение.

В табл. 6.3.42 приведены средняя ошибка (МАРЕ) имитации по такой «наивной модели» и величина относительной дисперсии по каждому из рассматриваемых рядов. Как правило, ошибка имитации по «наивной модели» тем лучше, чем меньше разброс исходных данных. Из общего правила резко выпадают результаты, относящиеся к рядам «Инфляция» и «ВВП». Что касается ряда «ВВП», то это скорее всего результат малой длины ряда: для имитации используется всего 22 точки. Для ряда «Инфляция» два значения в самом начале имитируемых данных (январь и февраль 1992 г.) резко выпадают из общей динамики. Если убрать эти две точки, то величина относительной дисперсии ряда разностей составляет 0,27%.

Таблица 6.3.42

**«Наивная модель». Средние ошибки имитации  
и дисперсии ряда разностей (в %)**

Показатель	Период имитации	Ошибки	Дисперсии
Инфляция	01.1992 $\pm$ 08.2000	4,28	7,92
M0	01.1992 $\pm$ 07.2000	9,61	0,88
M1	03.1996 $\pm$ 07.2000	4,45	0,27
M2	01.1992 $\pm$ 07.2000	7,63	0,28
Экспорт	01.1995 $\pm$ 04.2000	9,95	1,90
Импорт	01.1995 $\pm$ 04.2000	10,75	2,34
Безработица	01.1995 $\pm$ 08.2000	1,90	0,07
Налоговые доходы федерального бюджета	01.1993 $\pm$ 05.2000	22,95	13,37
Доходы федерального бюджета	01.1993 $\pm$ 05.2000	28,29	19,20
ВВП	I.1995 $\pm$ II.2000	13,46	1,22

Как видно из сравнения *табл. 6.3.41* и *6.3.42*, во всех случаях имитация рассматриваемых процессов с помощью трендовых и вероятностных моделей, использующих лаговую структуру, дает существенное улучшение точности получаемого результата. В том числе и для самого плохого случая: трендовая модель доходов федерального бюджета имитирует реальный процесс в 4,5 раза точнее, чем «наивная модель».

Все это свидетельствует о том, что построенные модели достаточно хорошо отражают внутренние закономерности изменения исследуемых рядов на анализируемых периодах времени (периодах имитации). Однако прогностические способности построенных моделей не слишком удовлетворительны. В *табл. 6.3.43* приведены абсолютные процентные ошибки (максимальные и средние) прогнозов по трендовым (без обучения, с обучением без коррекции тренда и с коррекцией тренда) и вероятностным (без обучения и с обучением) моделям по всем анализируемым временным рядам.

Результаты, приведенные в *табл. 6.3.43*, свидетельствуют о том, что не имеется надежных априорных оснований для выбора модели прогнозирования. Для ряда налоговых доходов федерального бюджета лучший результат (средняя ошибка прогноза 5,0%) получен по трендовой модели без обучения. Для рядов импорта, доходов федерального бюджета и ВВП лучший результат (средняя ошибка прогноза 9,8%, 9,2%, 4,5%, соответственно) получен по трендовой модели с обучением без коррекции тренда. Для рядов M2 и экспорта лучший прогноз (средняя ошибка 1,8% и 7,1%) получен по трен-

довой модели с обучением и коррекцией тренда. Для рядов *М0* и безработицы лучший результат (средняя ошибка прогноза 4,1%, 1,9%) получен по вероятностной модели без обучения. Наконец, для рядов инфляции и *М1* лучший прогноз (средняя ошибка 0,4%, 4,8%) получен по вероятностной модели с обучением. В табл. 6.3.44 приведены сравнительные данные о средней ошибке прогноза (МАРЕ) по лучшей из пяти рассмотренных моделей и «наивного прогноза».

Таблица 6.3.43

**Максимальные и средние ошибки различных методов прогнозирования**

Показатель	Трендовый прогноз						Вероятностный прогноз			
	без обучения		с обучением				без обучения		с обучением	
			без коррекции тренда		с коррекцией тренда					
	макс.	сред.	макс.	сред.	макс.	сред.	макс.	сред.	макс.	сред.
Инфляция	18,39	6,21	2,71	1,00	1,95	0,76	1,49	0,45	<b>1,41</b>	<b>0,41</b>
<i>М0</i>	33,45	23,99	18,86	4,18	25,82	5,14	<b>12,63</b>	<b>4,12</b>	<b>12,63</b>	4,19
<i>М1</i>	21,88	15,96	14,48	5,46	<b>14,45</b>	7,75	21,19	5,22	21,19	<b>4,75</b>
<i>М2</i>	12,38	6,20	8,64	2,47	<b>8,28</b>	<b>1,84</b>	21,57	4,96	21,57	4,96
Экспорт	36,94	11,24	36,94	9,06	<b>23,72</b>	<b>7,06</b>	33,21	14,17	33,21	11,22
Импорт	<b>45,13</b>	11,28	<b>45,13</b>	<b>9,79</b>	45,30	13,32	51,47	20,60	48,11	16,09
Безработица	8,47	3,26	8,47	3,67	10,07	4,01	<b>5,99</b>	<b>1,91</b>	<b>5,99</b>	2,22
Налоговые доходы федерального бюджета	<b>23,53</b>	<b>5,03</b>	30,06	5,70	29,84	7,59	111,24	26,35	111,24	24,56
Доходы федерального бюджета	31,03	10,27	<b>25,31</b>	<b>9,16</b>	26,06	9,90	91,44	27,97	91,44	27,32
ВВП	9,49	5,93	9,49	<b>4,47</b>	<b>9,36</b>	4,63	20,93	15,06	17,11	11,98

Таблица 6.3.44

**Средние ошибки прогноза**

Показатель	«Наивный прогноз»	Прогноз по лучшей
Инфляция	0,46	0,41
<i>М0</i>	4,34	4,12
<i>М1</i>	4,05	4,75
<i>М2</i>	3,53	1,84

Таблица 6.3.44 (продолжение)

Показатель	«Наивный прогноз»	Прогноз по лучшей
Экспорт	5,15	7,06
Импорт	14,70	9,79
Безработица	1,40	1,91
Налоговые доходы федерального бюджета	8,90	5,03
Доходы федерального бюджета	11,13	9,16
ВВП	10,08	4,47

Данные табл. 6.3.44 свидетельствуют о невысоком качестве прогнозирования по построенным моделям. Даже при наиболее качественном прогнозе по ряду инфляции качество прогноза по вероятностной модели с обучением (средняя ошибка 0,41%) незначительно превосходит качество «наивного прогноза» (средняя ошибка 0,46%). Лишь в четырех случаях (для рядов *M2*, импорта, налоговых доходов федерального бюджета и ВВП) качество прогноза по лучшей модели заметно превосходит качество «наивного прогноза». При этом в трех случаях (для рядов *M1*, экспорта и безработицы) «наивный прогноз» дает лучший результат чем лучшая из построенных моделей.

Представляются целесообразными два пути совершенствования построенных в данной работе моделей.

Первый путь наиболее простой. По мере накопления информации можно перейти к построению информативных структур более высоких порядков. В этом случае можно получить зависимость исследуемого показателя не от двух, а от трех (для информативных структур четвертого порядка), четырех (для информативных структур пятого порядка) и более лагов. При этом, естественно, точность модели возрастает, но снижается ее надежность<sup>13</sup>.

Второй путь заключается в переходе к рассмотрению взаимосвязанных временных рядов и их моделированию векторными случайными процессами. В этом случае получается зависимость исследуемого показателя не только от его лаговых значений, но и от связанных с ним показателей и их запаздывающих значений.

<sup>13</sup> См.: Юдин А.Д. Сложность статистических систем // Доклады АН СССР, 1982. Т. 266. № 5.

## Заключение. Некоторые выводы из полученных результатов

Для сравнения результатов прогнозирования при помощи эконометрических моделей сравним полученные результаты. Приведенная ниже таблица показывает ранжирование различных моделей прогнозирования, использовавшихся в главах 2 – 4. В каждой строке ранг 1 соответствует модели с наименьшим значением MAPE (вычисленным, соответственно, по 7, 8 или 16 точкам), а ранг 8 – модели с наибольшим значением MAPE (MAPE – средняя абсолютная процентная ошибка последовательности одношаговых прогнозов). Таким образом, прогнозы по модели, имеющей ранг 1, наилучшие (с точки зрения MAPE), а прогнозы по модели, имеющей ранг 8, наихудшие в соответствующей строке.

Для краткости в таблице обозначено:

- Fix – прогнозирование по фиксированной модели;
- Rec – прогнозирование по рекурсивной модели;
- Track – коррекция прогнозов методом «back-on-track»;
- Ave – коррекция прогнозов методом «back-on-ave».

Ряд/Тип	Кол-во точек	DS-модели				TS-модели			
		Fix	Track	Ave	Rec	Fix	Track	Ave	Rec
M1/DS	7	2	8	3	4	6	7	1	5
Аддит.	16	2	4	1	3	8	5	7	6
M1/DS	7	1			3	2			4
Инов.	16	1			2	4			3
M0/DS	7	4			1	3			2
Аддит.	16	5	6	4	1	8	3	7	2
M0/DS	7	1-2			3	4			1-2
Инов.	16	3			1	4			2
M2/DS	7	2			3	1			4
Аддит.	16	2	8	6	1	4	7	5	3
M2/DS	7	1			3	4			2
Инов.	16	2			1	3			4
Эксп/TS	8	2			1	4			3
	16	2	8	5	1	6	7	4	3
Наддох/DS	7	3			1	4			2
Аддит.	16	7	4	2	5	8	3	1	6
Наддох/DS	7	1			2	4			3
Инов.	16	3			2	4			1
Безраб/TS	7	3	5-6	7-8	4	1	5-6	7-8	2
Пром/DS	7	5	8	6-7	3-4	2	6-7	1	3-4
Цвет/DS (ск)	7	2	1	4	6	7	3	5	8
Цвет/TS (н/с)	7	2	8	7	3	4	6	1	5

В следующей таблице указано поведение ошибок последовательных прогнозов по оцененным фиксированным моделям (наличие или отсутствие выраженного систематического смещения прогнозов):

Ряд/Тип	Кол-во точек	Наличие смещения	
		DS-модели	TS-модели
<b>M1/DS</b>	7	–	+
<b>Аддит.</b>	16	–	+
<b>M1/DS</b>	7	–	+
<b>Иннов.</b>	16	–	+
<b>M0/DS</b>	7	–	+
<b>Аддит.</b>	16	+	+
<b>M0/DS</b>	7	–	–
<b>Иннов.</b>	16	+	+
<b>M2/DS</b>	7	–	–
<b>Аддит.</b>	16	–	–
<b>M2/DS</b>	7	–	–
<b>Иннов.</b>	16	–	–
<b>Эксп/TS</b>	8	–	+
	16	–	+
<b>Налдох/DS</b>	7	+	+
<b>Аддит.</b>	16	+	+
<b>Налдох/DS</b>	7	–	+
<b>Иннов.</b>	16	–	+
<b>Безраб/TS</b>	7	–	–
<b>Пром/DS</b>	7	–	+
<b>Цвет/DS (ск)</b>	7	–	+
<b>Цвет/TS (н/с)</b>	7	?	?

По данным двух последних таблиц можно составить таблицы предпочтений в парах использованных прогностических моделей при наличии и при отсутствии (выраженного) систематического смещения прогнозов.

Если выраженное систематическое смещение прогнозов есть, то имеем следующую картину предпочтений:

	<b>Фиксир.</b>	<b>Рекурс.</b>	<b>Back-on-track</b>	<b>Back-on-ave</b>
<b>Фиксир.</b>		Рекурс.	?	Back-on-ave
<b>Рекурс.</b>	Рекурс.		?	?
<b>Back-on-track</b>	?	?		?
<b>Back-on-ave</b>	Back-on-ave	?	?	

Если выраженного систематического смещения прогнозов нет, то картина предпочтений такова:

	Фиксир.	Рекурс.	Back-on-track	Back-on-ave
Фиксир.		?	Фиксир.	Фиксир.
Рекурс.	?		Рекурс.	?
Back-on-track	Фиксир.	Рекурс.		?
Back-on-ave	Фиксир.	?	?	

В обеих таблицах знаки вопроса (?) означают отсутствие более или менее определенных указаний на предпочтительность выбора одной из двух соответствующих моделей.

На основании первой из приведенных четырех таблиц можно также сделать заключение о том, что с точки зрения MAPE при выборе (для целей прогнозирования) между фиксированными *TS*- и *DS*-моделями последние (*DS*-модели) выглядят предпочтительнее как в тех случаях, когда прогнозируемый ряд классифицирован как *DS*-ряд (15 из 18 рассмотренных случаев), так и в тех случаях, когда прогнозируемый ряд классифицирован как *TS* ряд (3 из 4 случаев). Это подтверждает уже приводившееся в главе 1 замечание о том, что *DS*-процесс способен быстрее адаптироваться к изменениям структурного параметра, по крайней мере на одношаговых прогнозах.

На основании проведенного анализа макроэкономических рядов РФ можно сделать следующие выводы.

- Результаты прогнозирования с использованием одномерных эконометрических моделей временных рядов существенным образом зависят как от выбора модели, по которой производится прогнозирование (фиксированная или рекурсивная модель, модель в уровнях – *TS*-модель или модель в разностях – *DS*-модель), так и от поведения самого ряда за пределами интервала, на котором эта модель оценивалась. При этом качество прогнозов изменяется в весьма широких пределах.

Если точность последовательности одношаговых прогнозов определять величиной средней абсолютной процентной ошибки (MAPE), то по этой характеристике лучшие результаты достигаются при прогнозировании рядов, представляющих денежный агрегат *M1* и нескорректированный на сезонность индекс интенсивности производства цветных металлов (MAPE порядка 2% – 3%), а также сезонно скорректированные ряды, представляющие индекс производства цветных металлов (MAPE порядка 0,06% – 0,1%) и индекс интенсивности промышленного производства в целом (MAPE порядка 0,1%). (Заметим, однако, что результаты, полученные для двух пос-



ледних рядов, не являются удивительными: при построении сезонно скорректированного варианта ряда учитываются его действительные «будущие» значения, так что имитация прогнозирования в текущем времени для таких реализаций не имеет особого смысла.)

В то же время при построении 16 последовательных одношаговых прогнозов для ряда, представляющего объемы экспорта, даже в наилучшем варианте значение MAPE оказывается равным 7%.

Более того, в последнем случае наилучшие результаты дает использование «наивных» прогнозов, соответствующих использованию модели простого случайного блуждания (без сноса), т.е. прогнозирующих последующее значение ряда как равное текущему значению ряда. При использовании таких прогнозов значение MAPE удастся снизить до 5%. Впрочем, это единственный ряд, для которого подобные прогнозы оказались наилучшими. С другой стороны, использование другого типа «наивных» прогнозов, ориентирующихся на оцененный линейный тренд (см. разд. 2.8), дает несколько лучший результат (MAPE = 0,704%) по сравнению с наилучшим результатом (MAPE = 0,724%), полученным по эконометрическим моделям, для ряда, представляющего данные об общей численности безработных (и только для этого ряда).

- Полная модель не обязательно дает лучшее качество одношаговых прогнозов при выходе за пределы интервала, на котором модель оценивалась, по сравнению с редуцированной моделью.

Так, при анализе одношаговых прогнозов для ряда значений денежного агрегата M1 по *DS*-модели с аддитивным выбросом, датированным 02.1999, мы получили в разд. 2.1.1 для использованной там фиксированной *DS*-модели следующие показатели качества одношаговых прогнозов, соответствующих периоду с 08.2000 по 02.2001:

**RMSE=30788.31, MAE=25004.23, MAPE=3.119278.**

В то же время, если исключить из этой модели составляющие AR(2), ..., AR(7), то показатели качества получаемой при этом редуцированной модели равны, соответственно,

**RMSE=30337.52, MAE=24768.65, MAPE=3.090592**

и оказываются лучшими по сравнению с аналогичными показателями для полной модели.

- Редуцированная модель, оказывающаяся предпочтительнее полной по критериям  $R^2_{adj}$  (скорректированный коэффициент детерминации), AIC (информационный критерий Акаике), SIC (информационный критерий Шварца), не обязательно дает лучшее качество одношаговых прогнозов при выходе за пределы интервала, на котором модель оценивалась, по сравнению с полной моделью.

Если из той же полной модели (что и в предыдущем пункте) исключить только составляющую AR(7), то получим следующие показатели качества прогнозов на один шаг вперед по таким образом редуцированной модели:

**RMSE=30897.48, MAE=25365.34, MAPE=3.168368.**

Все три показателя у редуцированной модели хуже, чем у полной, хотя для редуцированной модели значение SIC=22.277 меньше, чем для полной модели (22.354).

- Рекурсивная модель (коэффициенты которой переоцениваются при поступлении каждого нового наблюдения) не обязательно дает лучшее качество одношаговых прогнозов по сравнению с фиксированной моделью (коэффициенты которой, оцененные на базовом интервале, не переоцениваются при поступлении новых наблюдений). Однако при увеличении количества последовательных одношаговых прогнозов предпочтение смещается в сторону рекурсивной модели, на что указывают следующие результаты.

При расчете характеристик качества прогнозов по 7 (или 8) последовательным прогнозам рекурсивные модели оказались лучшими только в 41,7% случаев, а при расчете характеристик качества прогнозов по 16 последовательным прогнозам рекурсивные модели оказались лучшими в 90,9% случаев. Так, для ряда, представляющего значения денежного агрегата M0, значения средней абсолютной процентной ошибки (MAPE), вычисленной по 16 последовательным прогнозам (модель с аддитивным выбросом), равны 5,13% при использовании фиксированной DS-модели и 3,31% – при использовании рекурсивной DS-модели.

- Для рядов, классифицированных при предварительном тести-

ровании как *DS*-ряды, лучшее качество одношаговых прогнозов среди фиксированных моделей могут давать как модели в разностях (в 15 из 18 рассмотренных случаев), так и модели в уровнях (3 случая из 18). Однако при сравнении качества по 16 последовательным прогнозам, для всех рассмотренных *DS*-рядов модели в разностях оказались предпочтительнее.

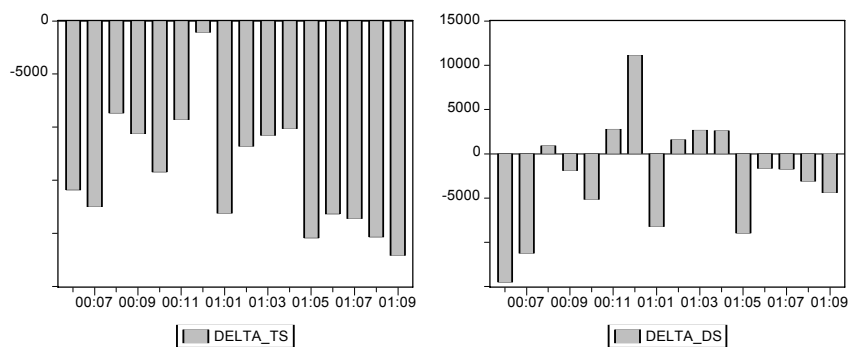
В качестве примера здесь можно привести ряд, представляющий значения денежного агрегата *M1*, классифицированный как *DS*-ряд. Значения средней абсолютной процентной ошибки (*MAPE*), вычисленной в разд. 3.1.1 по 16 последовательным прогнозам (модель с инновационным выбросом), равны 9,02% при использовании модели в уровнях и 3,05% – при использовании модели в разностях.

- Для рядов, классифицированных при предварительном тестировании как *TS*-ряды, лучшее качество одношаговых прогнозов среди фиксированных моделей могут давать модели в разностях («передифференцированные» ряды) – 3 случая из 4.

Примером здесь служит ряд, представляющий объемы экспорта (раздел 3.2) и классифицированный как *TS*-ряд. Значения средней абсолютной процентной ошибки (*MAPE*), вычисленной по 16 последовательным прогнозам, равны 8,39% при использовании модели в уровнях и 6,99% – при использовании модели в разностях.

- Наличие систематического смещения последовательных одношаговых прогнозов в основном связано со сдвигами детерминированных (нестохастических) составляющих ряда. При этом среди прогнозов по фиксированным моделям наибольшему смещению подвержены прогнозы по моделям в уровнях.

Так, при прогнозировании ряда налоговых доходов федерального бюджета в разд. 3.3 прогнозы по фиксированной модели в разностях (инновационный выброс) не обнаруживали систематического смещения, тогда как прогнозы по фиксированной модели в уровнях имели весьма сильное смещение:



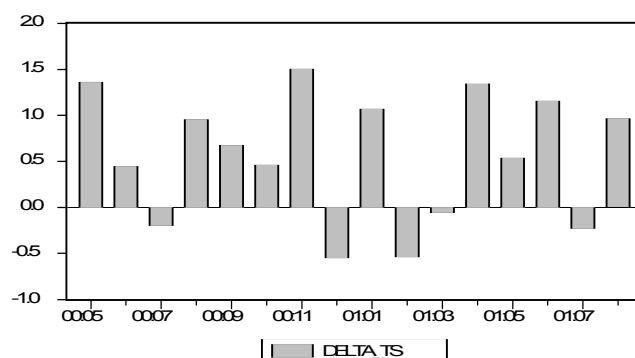
(здесь DELTA\_TS – ошибки прогнозов по модели в уровнях, DELTA\_DS – ошибки прогнозов по модели в разностях).

- В ряде случаев улучшение одношаговых прогнозов по фиксированной модели может быть получено коррекцией систематического смещения методом «back-on-track», согласно которому очередной одношаговый прогноз, сделанный на основании оцененной модели, корректируется с учетом ошибки предыдущего одношагового прогноза (разд. 4.3).

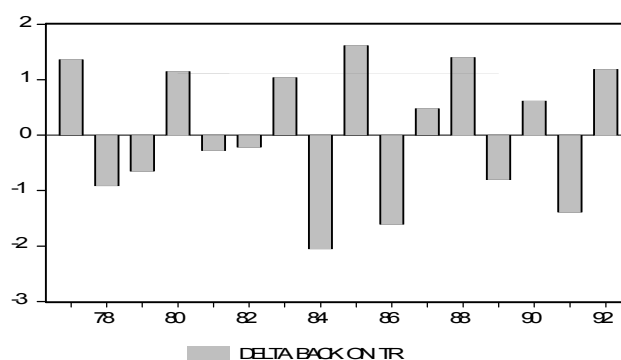
Примерами могут служить последовательности 16 одношаговых прогнозов для рядов M1, M0 и ряда налоговых доходов федерального бюджета по моделям в уровнях с аддитивным выбросом. Так, для ряда M1 значения средней абсолютной процентной ошибки (MAPE), вычисленной по 16 последовательным прогнозам (разд. 4.3.1), равны 9,37% при использовании фиксированной TS-модели и 3,63% – при использовании прогнозов, скорректированных методом «back-on-track».

- При использовании метода «back-on-track» предотвращение систематического смещения прогнозов достигается за счет увеличения варибельности прогнозов, что для некоторых рядов приводит к значительному возрастанию средней процентной ошибки прогноза.

Примером здесь служит ряд, представляющий объемы экспорта (раздел 4.3.5). Последовательность ошибок прогнозов по фиксированной TS-модели для этого ряда имеет некоторое смещение.



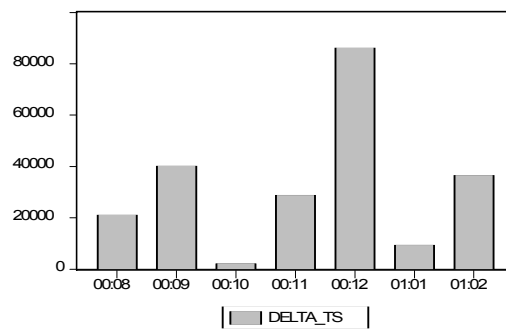
Коррекция прогнозов методом «back-on-track» полностью устраняет это смещение:



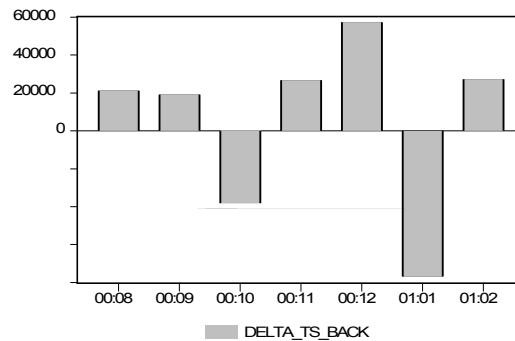
Однако значения средней абсолютной процентной ошибки (MAPE), вычисленной по 16 последовательным прогнозам, равное 6,99% при использовании фиксированной *TS*-модели, возрастает до 12,17% в результате коррекции прогнозов методом «back-on-track».

- В реальных экономическо-правовых условиях предстоящий сдвиг детерминированной составляющей ряда в некоторых случаях может быть предсказан заранее. Модификация модели, учитывающая наличие такого сдвига, может приводить к значительному улучшению качества прогнозов – устранению систематического смещения прогнозов без существенного возрастания их вариабельности.

Пример, в котором подобная модификация оказывается успешной, в то время как процедура «back-on-track», устраняя систематическое смещение, приводит к ухудшению характеристик качества последовательности прогнозов, приведен в разд. 4.3.1. Этот пример связан с прогнозированием ряда *M1* по *TS*-модели с аддитивным выбросом. Прогнозы имеют систематическое смещение; ошибки прогнозов имеют вид:

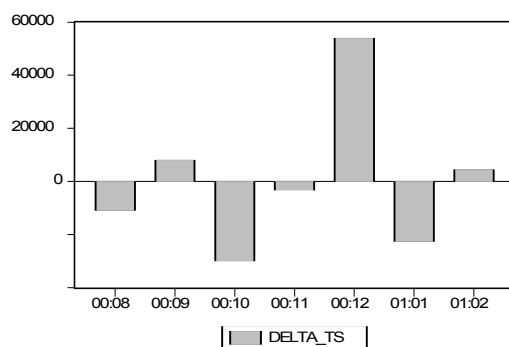


Процедура «back-on-track» устраняет систематическое смещение: ошибки скорректированных прогнозов имеют вид:



Однако в результате коррекции значение *MAPE* возрастает с 3,96% до 4,75%.

При прогнозе по модифицированной модели, учитывающей ожидаемый сдвиг уровня ряда *M1*, смещение прогнозов также устраняется: ошибки прогнозов имеют вид:



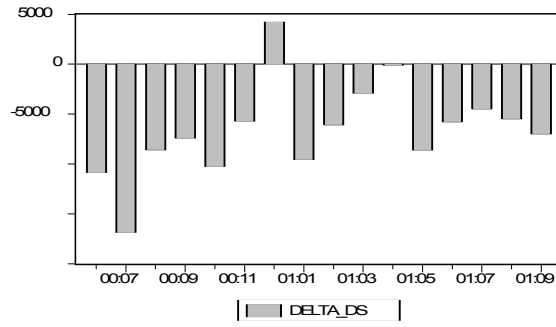
Однако при этом MAPE равна всего лишь 2,36%.

- Модификация процедуры «back-on-track» («back-on-average»), при которой каждый очередной одношаговый прогноз, сделанный на основании оцененной фиксированной модели, корректируется с учетом ошибок предыдущих одношаговых прогнозов (разд 4.3.), также может приводить к значительному улучшению точностных характеристик прогнозов.

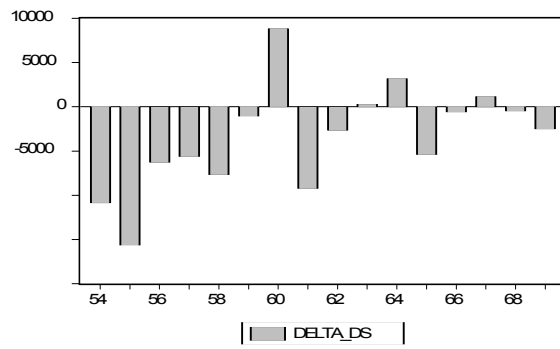
Например, для ряда налоговых доходов федерального бюджета (разд. 4.3.4) значения средней абсолютной процентной ошибки (MAPE), вычисленной по 16 последовательным прогнозам, равны 7,76% при использовании фиксированной *DS*-модели (с аддитивным выбросом), 5,48% – при коррекции прогнозов методом «back-on-track» и 4,59% – при коррекции прогнозов методом «back-on-ave».

- При наличии выраженного систематического смещения прогнозов по фиксированным моделям более приемлемыми являются прогнозы по рекурсивным моделям или прогнозы, скорректированные методом «back-on-ave».

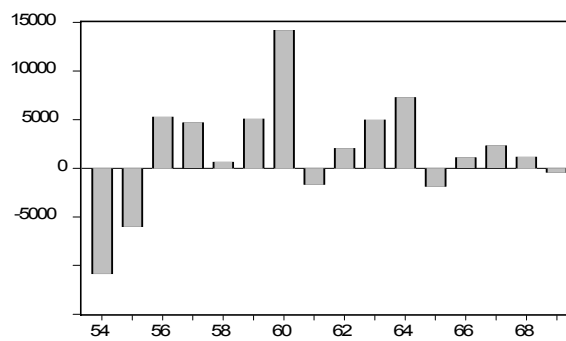
Например, для ряда налоговых доходов федерального бюджета прогнозы по фиксированной *DS*-модели с аддитивным выбросом (разд. 3.3) имеют значительное смещение:



В то же время при прогнозировании по рекурсивной *DS*-модели систематическое смещение прогнозов в значительной мере устраняется:



Использование же коррекции прогнозов методом «back-on-ave» приводит к следующей последовательности ошибок прогнозов:

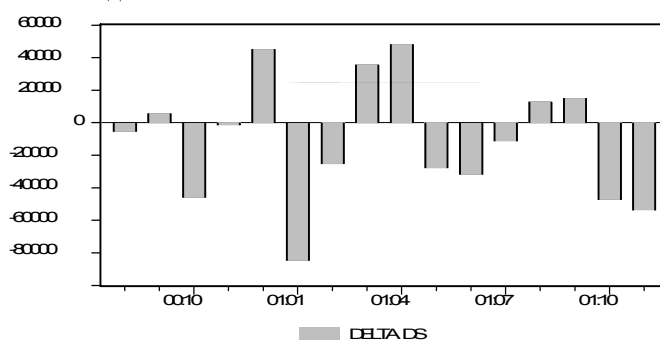




В итоге значение  $MAPE = 7,76\%$  для фиксированной  $DS$ -модели уменьшается до значения  $5,74\%$  при использовании рекурсивной модели и до значения  $4,59$  – при использовании коррекции прогнозов методом «back-on-ave».

- При отсутствии выраженного систематического смещения прогнозов по фиксированным моделям коррекция методами «back-on-track» и «back-on-average» не приводит к улучшению прогнозов, а прогнозы по рекурсивным моделям предпочтительнее прогнозов, скорректированных по методу «back-on-track».

Например, для ряда М2 при прогнозировании по фиксированной  $DS$ -модели с аддитивным выбросом систематическое смещение прогнозов отсутствует: ошибки прогнозов имеют вид:



При прогнозировании по  $DS$ -моделям с аддитивным выбросом получают следующие результаты (разд. 4.3.3). Значения средней абсолютной процентной ошибки ( $MAPE$ ), вычисленной по 16 последовательным прогнозам, равны  $2,60\%$  при использовании фиксированной модели,  $3,86\%$  – при коррекции прогнозов методом «back-on-track»,  $3,26\%$  – при коррекции прогнозов методом «back-on-ave» и  $2,46\%$  – при использовании рекурсивной модели.

- Прогнозирование с использованием методов, предложенных во второй части работы, дало следующие результаты. Для ряда налоговых доходов федерального бюджета лучший результат (средняя ошибка прогноза  $5,0\%$ ) получен по трендовой модели без обучения. Для рядов импорта, доходов федерального

бюджета и ВВП лучший результат (средняя ошибка прогноза 9,8%, 9,2%, 4,5%, соответственно) получен по трендовой модели с обучением без коррекции тренда. Для рядов М2 и экспорта лучший прогноз (средняя ошибка 1,8% и 7,1%) получен по трендовой модели с обучением и коррекцией тренда. Для рядов М0 и безработицы лучший результат (средняя ошибка прогноза 4,1%, 1,9%) получен по вероятностной модели без обучения. Наконец, для рядов инфляции и М1 лучший прогноз (средняя ошибка 0,4%, 4,8%) получен по вероятностной модели с обучением.

- При сравнении по средней абсолютной процентной ошибке (МАРЕ) наилучших последовательностей одношаговых прогнозов, построенных в двух частях настоящего отчета для одних и тех же рядов, получаются следующие результаты. Для рядов М0, М1, Экспорт, Налоговые доходы федерального бюджета, Безработица лучшие результаты дают эконометрические модели, представленные в Части 1 (3% против 4,75% – для ряда М1; 3,31% против 4,12% – для ряда М0, 6,69% – для ряда Экспорт; 4,50% против 5,03% – для ряда Налоговые доходы федерального бюджета; 0,72% против 1,91% – для ряда Безработица). Для ряда М2 лучшие результаты дает модель, представленная в Части 2 (1,84% против 2,46%).

Проведенное исследование помимо анализа и сравнения различных методов прогнозирования предоставляет базовую информацию, на основе которой следует строить методику прогнозирования макроэкономических показателей. Исследование показало, что для рассматриваемых макроэкономических показателей не удастся выработать единую универсальную методику прогнозирования. Исключение составляет только общая рекомендация, что выбор между *DS*- и *TS*-моделью, осуществляемый на основе теста на единичный корень, включает в себя лучшую или близкую к лучшей модель. Для улучшения оценок с использованием новых данных имеет смысл использовать рекурсивные модели, особенно, если имеют место структурные сдвиги в динамике показателей (или эти структурные сдвиги могут быть в некоторой степени предсказаны на основе содержательных соображений или дополнительной информации).

Приведенные выше результаты и соображения означают, по-видимому, что прогнозы следует делать на основе нескольких моделей с последующим анализом отклонений прогнозов от фактических значений с тем, чтобы в перс-

пективе выбрать модель, которая в большинстве случаев дает удовлетворительный прогноз на несколько периодов вперед. Для того чтобы это реализовать, необходимо на регулярной основе прогнозировать некоторый набор основных макроэкономических показателей (промышленное производство, инфляцию, денежные агрегаты, налоговые поступления и др.), публикуя и открывая для обсуждения получаемые результаты одновременно с анализом качества предыдущих прогнозов. При этом можно предложить следующую последовательность действий:

- Начинать следует с подбора моделей, учитывающих поведение ряда на достаточно протяженном интервале вплоть до момента построения прогноза и, в то же время, не углубляясь слишком далеко в прошлое, поскольку структура рассмотренных рядов весьма сильно различается на различных временных промежутках.
- Если в распоряжении нет достаточно протяженного периода, на котором структура ряда не претерпевает серьезных изменений, можно построить модель для более протяженного периода, учитывающую структурные изменения в процессе порождения данных, в форме аддитивного или инновационного выброса.
- Определенную пользу можно получить от предварительного тестирования и отнесения ряда к классу  $TS$ - или  $DS$ -рядов. Для рядов, классифицированных как  $DS$ -ряды, прогнозирование следует производить скорее по модели в разностях. Для рядов, классифицированных как  $TS$ -ряды, в некоторых случаях предпочтительнее оказывается модель в уровнях. При использовании моделей в уровнях для рядов, классифицированных как  $DS$ -ряды, как правило, возникает систематическое смещение получаемых прогнозов.
- В процессе регулярного прогнозирования с использованием новых данных имеет смысл использовать рекурсивные модели, особенно если имеют место структурные сдвиги в динамике показателей.
- При отсутствии систематического смещения регулярных прогнозов можно в течение некоторого времени использовать одну и ту же модель для процесса порождения данных. При выявлении смещения регулярных прогнозов следует либо переходить к прогнозированию по рекурсивной модели или по уточнен-

ной модели, учитывающей изменение структуры процесса порождения данных, либо, продолжая использовать исходную модель, производить коррекцию прогнозов методами, описанными в главе 4. В то же время применять такие методы при отсутствии заметного систематического смещения прогнозов не следует – это может привести к уменьшению точности прогнозов за счет возрастания их вариабельности.

- В реальных экономическо-правовых условиях предстоящий сдвиг детерминированной составляющей ряда в некоторых случаях может быть предсказан заранее. Модификация модели, учитывающая наличие такого сдвига, может приводить к значительному улучшению качества прогнозов – устранению систематического смещения прогнозов без существенного возрастания их вариабельности.
- Наряду с построением прогнозов по эконометрическим моделям, интересно проводить параллельное построение прогнозов с использованием нетрадиционных методов, описанных во второй части работы, и сравнивать получаемые результаты.

## ПРИЛОЖЕНИЯ

### Приложение 1 . Макросы

#### *П1.1. Построение трехмерных информаций*

```
Sub Информации3()  
,  
  
Dim n As Integer, nn As Integer, tau As Integer  
Dim id(150, 450), idn(150, 450), ro(450, 450), ind(450), nnn(450), im(50)  
Dim h1(150), h2(150, 150), h3(150, 150, 150)  
Dim i2(5, 150), ii2(5, 150), i3(5, 150, 150), ii3(5, 150, 150)  
Sheets("Данные").Select  
tau = 12  
nn = 101  
eps = 0.025  
n1 = n * (tau + 1)  
'Подготовка данных  
For i = 1 To n  
    im(i) = Cells(1, 1 + i).Value  
    For j = 1 To nn  
        id(i, j) = Cells(1 + j, 1 + i).Value  
    Next  
Next  
For i = 1 To n  
    idmax = id(i, 1)  
    idmin = id(i, 1)  
    For j = 2 To nn  
        If id(i, j) > idmax Then  
            idmax = id(i, j)  
        End If  
        If id(i, j) < idmin Then  
            idmin = id(i, j)  
        End If  
    Next  
    For j = 1 To nn  
        idn(i, j) = (id(i, j) - idmin) / (idmax - idmin)  
    Next  
Next
```

```

If tau <> 0 Then
  For tau1 = 1 To tau
    For i = 1 To n
      im(n * tau1 + i) = im(i) + Str(tau1)
      For j = tau1 To nn
        idn(n * tau1 + i, j) = idn(n * (tau1 - 1) + i, j - 1)
      Next
    Next
  Next
End If
'Одномерные энтропии
For i = 1 To n1
  For j1 = 1 To nn - 1
    For j2 = j1 + 1 To nn
      ro(j1, j2) = Abs(idn(i, j1) - idn(i, j2))
    Next
  Next
  h1(i) = Log(nn)
  For t = 1 To nn
    ind(t) = 1
    nnn(t) = 1
  Next
  For t = 1 To nn - 1
    If ind(t) <> 0 Then
      ind(t) = 0
      For t1 = t + 1 To nn
        If ind(t1) <> 0 Then
          If ro(t, t1) <= eps Then
            ind(t1) = 0
            nnn(t) = nnn(t) + 1
            nnn(t1) = 0
          End If
        End If
      Next
    End If
  Next
End If
If nnn(t) <> 0 Then
  h1(i) = h1(i) - nnn(t) * Log(nnn(t)) / nn
End If

```

```

Next
h1(i) = h1(i) / Log(2)
For tau1 = 1 To tau
  h1(i + tau1) = h1(i)
Next
Next
'Двухмерные энтропии
For i = 1 To n1 - 1
  For tau1 = 0 To tau
    For j = 1 To n
      i1 = n * tau1 + j
      For j1 = tau1 + 1 To nn - 1
        For j2 = j1 + 1 To nn
          ro(j1, j2) = ((idn(i, j1) - idn(i, j2)) ^ 2 + (idn(i1, j1) - idn(i1, j2)) ^ 2)
            ^ 0.5
        Next
      Next
      h2(i, i1) = Log(nn)
      For t = tau1 + 1 To nn
        ind(t) = 1
        nnn(t) = 1
      Next
      For t = tau1 + 1 To nn - 1
        If ind(t) <> 0 Then
          ind(t) = 0
          For t1 = t + 1 To nn
            If ind(t1) <> 0 Then
              If ro(t, t1) <= eps Then
                ind(t1) = 0
                nnn(t) = nnn(t) + 1
                nnn(t1) = 0
              End If
            End If
          Next
        End If
      Next
      End If
      If nnn(t) <> 0 Then
        h2(i, i1) = h2(i, i1) - nnn(t) * Log(nnn(t)) / (nn - tau1)
      End If
    End If
  End If
End If

```

```

Next
h2(i, i1) = h2(i, i1) / Log(2)
h2(i1, i) = h2(i, i1)
Next
Next
Next
"Трёхмерные энтропии
For i = 1 To n1 - 2
For i1 = i + 1 To n1 - 1
For tau1 = 0 To tau
For j = 1 To n
ii = n * tau1 + j
For j1 = tau1 + 1 To nn - 1
For j2 = j1 + 1 To nn
ro(j1, j2) = ((idn(i, j1) - idn(i, j2)) ^ 2 + (idn(i1, j1) - idn(i1, j2)) ^
2 + (idn(ii, j1) - idn(ii, j2)) ^ 2) ^ 0.5
Next
Next
Next
h3(i, i1, ii) = Log(nn)
For t = tau1 + 1 To nn
ind(t) = 1
nnn(t) = 1
Next
For t = tau1 + 1 To nn - 1
If ind(t) <> 0 Then
ind(t) = 0
For t1 = t + 1 To nn
If ind(t1) <> 0 Then
If ro(t, t1) <= eps Then
ind(t1) = 0
nnn(t) = nnn(t) + 1
nnn(t1) = 0
End If
End If
Next
End If
If nnn(t) <> 0 Then
h3(i, i1, ii) = h3(i, i1, ii) - nnn(t) * Log(nnn(t)) / (nn - tau1)

```



```

        End If
    Next
    h3(i, i1, ii) = h3(i, i1, ii) / Log(2)
    h3(i, ii, i1) = h3(i, i1, ii)
    h3(i1, i, ii) = h3(i, i1, ii)
    h3(i1, ii, i) = h3(i, i1, ii)
    h3(ii, i, i1) = h3(i, i1, ii)
    h3(ii, i1, i) = h3(i, i1, ii)
Next
Next
Next
'Dвухмерные информации и коэффициенты информативности
For i = 1 To n
    For j = 1 To n1
        i2(i, j) = h1(i) + h1(j) - h2(i, j)
        ii2(i, j) = i2(i, j) / h1(i)
    Next
Next
'Трехмерные информации и коэффициенты информативности
For i = 1 To n
    For j = 1 To n1 - 1
        For j1 = j + 1 To n1
            i3(i, j, j1) = h1(i) + h2(j, j1) - h3(i, j, j1)
            ii3(i, j, j1) = i3(i, j, j1) / h1(i)
        Next
    Next
Next
Next
Sheets("3").Select
For i = 1 To n
    Cells(1, 1 + i).Value = im(i)
    Cells(1, 2 + n + i).Value = im(i)
    Cells(1, 3 + 2 * n + i).Value = im(i)
    Cells(2, 1 + i).Value = h1(i)
    For j = 1 To n1
        If i = 1 Then
            Cells(2 + j, 1).Value = im(j)
        End If
    Next
Next

```

```

If j = i Then
    Cells(2 + j, 1 + i).Value = "*"
    Cells(2 + j, 2 + n + i).Value = "*"
    Cells(2 + j, 3 + 2 * n + i).Value = "*"
Else
    Cells(2 + j, 1 + i).Value = h2(i, j)
    Cells(2 + j, 2 + n + i).Value = i2(i, j)
    Cells(2 + j, 3 + 2 * n + i).Value = ii2(i, j)
End If
Next
k = 0
For j = 1 To n1 - 1
    For j1 = j + 1 To n1
        k = k + 1
        If i = 1 Then
            Cells(2 + n + k, 1).Value = im(j) + "," + im(j1)
        End If
        If j = i Then
            Cells(2 + n + k, 1 + i).Value = "*"
            Cells(2 + n + k, 2 + n + i).Value = "*"
            Cells(2 + n + k, 3 + 2 * n + i).Value = "*"
        Else
            If j1 = i Then
                Cells(2 + n + k, 1 + i).Value = "*"
                Cells(2 + n + k, 2 + n + i).Value = "*"
                Cells(2 + n + k, 3 + 2 * n + i).Value = "*"
            Else
                Cells(2 + n + k, 1 + i).Value = h3(i, j, j1)
                Cells(2 + n + k, 2 + n + i).Value = i3(i, j, j1)
                Cells(2 + n + k, 3 + 2 * n + i).Value = ii3(i, j, j1)
            End If
        End If
    Next
Next
Next
End Sub

```

### *П1.2. Построение прогнозов*

```
Sub Прогноз3()  
,  
Dim id(1000), idn(1000), pr(1000)  
Dim dat As Date  
Sheets("Данные").Select  
tau1 = 1  
tau2 = 7  
t = 101  
tt = 13  
ttt = 120  
eps = 0.005  
For t1 = 1 To ttt  
    id(t1) = Cells(1 + t1, 2).Value  
Next  
idmax = id(tau2)  
idmin = id(tau2)  
For t1 = tau2 + 1 To t  
    If id(t1) > idmax Then  
        idmax = id(t1)  
    End If  
    If id(t1) < idmin Then  
        idmin = id(t1)  
    End If  
Next  
For t1 = 1 To ttt  
    idn(t1) = (id(t1) - idmin) / (idmax - idmin)  
Next  
For t2 = tt To ttt  
    pr(t2 + 1 - tt) = 0
```

```

pr1 = 0
eps1 = eps
While pr1 = 0
  For t1 = tau2 + 1 To t
    ro = ((idn(t1 - tau1) - idn(t2 - tau1)) ^ 2 + (idn(t1 - tau2) - idn(t2 -
tau2)) ^ 2) ^ 0.5
    If ro <= eps1 Then
      pr(t2 + 1 - tt) = pr(t2 + 1 - tt) + id(t1)
      pr1 = pr1 + 1
    End If
  Next
  eps1 = eps1 + 0.1 * eps
Wend
pr(t2 + 1 - tt) = pr(t2 + 1 - tt) / pr1
Next
'Печать результата
Sheets("Прогноз3").Select
For t2 = tt To ttt
  Cells(t2 + 2 - tt, 2).Value = id(t2)
  Cells(t2 + 2 - tt, 3).Value = pr(t2 - tt + 1)
  Cells(t2 + 2 - tt, 4).Value = pr(t2 - tt + 1) - id(t2)
Next
End Sub

```

## Приложение 2. Исходные данные

### П2.1. Инфляция

В качестве исходных данных для инфляции используется *темпы прироста индекса потребительских цен, %* – месячные данные с 01.1991 по 08.2000 (для построения модели) и с 09.2000 по 12.2001 (для построения прогноза); источник – Госкомстат РФ.

**Индекс потребительских цен (ИПЦ)** измеряет отношение стоимости фактического фиксированного набора товаров и услуг в текущем периоде к его стоимости в базисном периоде. ИПЦ является важнейшим показателем, характеризующим уровень инфляции, и используется для целей государственной политики, анализа и прогноза ценовых процессов в экономике, пересмотра минимальных социальных гарантий, решения правовых споров.

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Январь	6,2	245,0	25,8	17,9	17,8	4,1	2,3	1,4	8,4	2,3	2,8
Февраль	4,8	38,3	24,7	10,8	11,0	2,8	1,5	0,9	4,1	1,0	2,3
Март	6,3	29,9	20,1	7,4	8,9	2,8	1,4	0,6	2,8	0,6	1,9
Апрель	63,5	21,7	18,8	8,5	8,5	2,2	1,0	0,4	3,0	0,9	1,8
Май	3,0	12,0	18,1	6,9	7,9	1,6	0,9	0,5	2,2	1,8	1,8
Июнь	1,2	18,6	19,9	6,0	6,7	1,2	1,1	0,1	1,9	2,6	1,6
Июль	0,6	11,0	22,4	5,3	5,4	0,7	0,9	0,2	2,8	1,8	0,5
Август	0,5	8,6	25,8	4,6	4,6	-0,2	-0,1	3,7	1,2	1,0	0,0
Сентябрь	1,1	11,5	23,1	7,7	4,5	0,3	-0,3	38,4	1,5	1,3	0,6
Октябрь	3,5	22,9	19,5	15,0	4,7	1,2	0,2	4,5	1,4	2,1	1,1
Ноябрь	8,9	26,1	16,4	15,0	4,5	1,9	0,6	5,7	1,2	1,5	1,4
Декабрь	12,1	25,4	12,5	16,4	3,2	1,4	1,0	11,6	1,3	1,6	1,6

## П2.2. Денежные агрегаты (M0, M1, M2)

**Денежный агрегат M0** – наличные деньги в обращении.

В качестве исходной информации используются данные: денежный агрегат M0, млрд руб. (с 1998 г. млн руб.) – месячные данные с 12.1990 по 07.2000 (для построения модели) и с 08.2000 по 11.2001 (для построения прогноза); источник – Центральный банк РФ.

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Январь		73	199	1903	12967	31802	75371	96372	116672	178014	232852	380127
Февраль		76	216	2278	14573	34381	80371	102079	120255	180781	242046	387959
Март		82	255	2559	15939	35240	86735	105213	119147	174132	251531	399395
Апрель		85	321	3309	19411	41639	93112	115227	128606	195246	279064	435350
Май		89	369	4020	20669	45459	93652	120369	129856	205285	289272	438312
Июнь		100	458	5111	23811	52520	104368	136851	129806	216388	321766	474692
Июль		112	645	6259	27048	57970	102850	140397	129326	218163	334037	490598
Август		123	830	7305	27918	61658	101117	141621	133377	216181	341627	507107
Сентябрь		133	998	8408	30016	65495	96218	134873	154212	212804	350935	530972
Октябрь		142	1196	9826	30522	66285	94398	135795	166451	221959	349669	531481
Ноябрь		153	1449	10952	31982	70700	95801	128817	167269	219325	358351	527287
Декабрь	79	191	1716	13278	36482	80815	103824	130540	187843	266544	419262	

**Денежный агрегат M1** – сумма денег вне банков и депозитов до востребования в банковской системе (без депозитов органов государственного управления), т.е. представляет собой все денежные средства в экономике страны, которые могут быть использованы как средство платежа.

В качестве исходной информации используются данные: денежный агрегат M1, млрд руб. (с 1998 г. млн руб.) – месячные данные с 06.1995 по 07.2000 (для построения модели) и с 08.2000 по 11.2001 (для построения прогноза); источник – Центральный банк РФ.

	1995	1996	1997	1998	1999	2000	2001
Январь		140347	186311	272669	329986	508059	810517
Февраль		146202	192515	270401	340333	529876	829180
Март		155905	197754	266022	344782	546439	858381
Апрель		163174	208172	269489	371892	576378	918209
Май		166538	217770	271839	403982	611197	938533
Июнь	106311	180106	242496	270256	418070	662680	987901
Июль	112990	181138	249777	261569	429382	692388	1015090
Август	119494	181808	251167	252356	432913	717950	1040765
Сентябрь	121301	176079	252764	274112	430987	747446	1074933
Октябрь	123054	175546	260660	289197	454337	750662	1084385
Ноябрь	130920	177482	252213	302828	471573	777139	1058132
Декабрь	151267	192402	298289	342817	526772	879310	

**Денежный агрегат M2** – объем наличных денег в обращении (вне банков) и остатков средств в национальной валюте на расчетных, текущих счетах и депозитах нефинансовых предприятий, организаций и физических лиц, являющихся резидентами Российской Федерации. В этот агрегат не включаются депозиты в иностранной валюте. Начиная с 1 января 1998 г. в состав денежной массы не включаются данные по кредитным организациям с отозванной лицензией.

В качестве исходной информации используются данные: денежный агрегат M2, млрд руб. (с 1998 г. млн руб.) – месячные данные с 12.1990 по 07.2000 (для построения модели) и с 08.2000 по 11.2001 (для построения прогноза); источник – Центральный банк РФ.

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Январь		455,5	1054	7187	33980	93800	216700	289900	361200	444200	695000	1079300
Февраль		462,4	1204	7782	36439	101900	229200	299500	362900	463900	726600	1109700

Март		482,9	1369	8913	39550	107300	241800	305800	360400	473800	751400	1149500
Апрель		516,1	1506	11063	46401	123200	251000	317800	368000	509600	787900	1210000
Май		550,9	1641	13460	52253	138200	254200	328400	370000	542400	831600	1233700
Июнь		589,0	2093	15765	59414	156600	266900	352000	368600	567700	892200	1294300
Июль		689,0	2668	18482	64363	165000	271900	363000	360000	583200	931200	1330200
Август		728,4	3422	21121	70970	173800	275300	364600	343600	590800	960100	1365500
Сентябрь		763,6	4515	21771	77063	179700	276000	363000	365800	597400	992400	1414400
Октябрь		818,6	5722	24554	80359	184200	278800	368800	377600	625100	1001200	1441200
Ноябрь		866,3	6038	26788	84348	195200	282300	357400	396900	646500	1036400	1439100
Декабрь	424	958,0	6400	32601	97800	220800	288300	374100	448300	704700	1144300	

### П2.3. Экспорт и импорт

**Экспорт** – вывоз из страны товаров отечественного производства, а также реэкспорт товаров. К товарам отечественного производства относятся также товары иностранного происхождения, ввезенные в страну и подвергшиеся существенной переработке, изменяющей основные качественные или технические характеристики товаров. К реэкспортным товарам относятся товары, ввезенные в страну, а затем вывезенные за границу без переработки.

В качестве исходной информации используются данные: объем экспорта (во все страны), *млрд долл.* – месячные данные с 01.1994 по 04.2000 (для построения модели) и с 05.2000 по 12.2001 (для построения прогноза); источник – Госкомстат РФ.

	1994	1995	1996	1997	1998	1999	2000	2001
Январь	4,09	5,64	5,80	7,00	6,00	4,60	7,0	8,4
Февраль	4,46	6,15	6,80	6,70	5,90	5,00	8,1	8,2
Март	4,76	6,76	7,80	7,40	6,80	5,90	9,3	8,9
Апрель	4,74	6,75	7,00	6,80	6,20	6,50	8,1	8,7
Май	5,83	6,86	7,50	6,70	6,10	5,10	8,3	8,8
Июнь	6,34	7,02	6,90	6,90	6,60	5,30	8,6	9,3
Июль	5,77	6,31	7,30	7,50	6,30	6,30	8,6	8,3
Август	6,11	6,34	7,00	7,00	5,80	6,10	9,1	9,2
Сентябрь	6,55	6,73	7,10	7,10	6,00	6,30	9,0	8,5
Октябрь	6,01	7,10	8,60	7,90	6,10	6,80	9,0	8,2
Ноябрь	6,31	7,73	8,10	8,30	5,90	7,40	10,2	8,4
Декабрь	6,59	7,80	8,70	8,90	7,30	9,40	10,2	8,2

**Импорт** – ввоз товаров в страну. В импорт включаются ввезенные товары, предназначенные для потребления в экономике страны, реэкспорта, и товары, закупаемые для отечественных организаций за границей, для потребления на месте.

В качестве исходной информации используются данные: объем импорта (во все страны), *млрд долл.* – месячные данные с 01.1994 по 04.2000 (для построения модели) и с 05.2000 по 12.2001 (для построения прогноза); источник – Госкомстат РФ.

	1994	1995	1996	1997	1998	1999	2000	2001
Январь	3,55	3,70	4,80	4,80	5,90	2,80	2,60	3,20
Февраль	3,85	4,41	5,80	5,10	6,10	3,10	3,40	3,60
Март	4,14	4,86	6,00	5,70	6,60	3,60	3,70	4,20
Апрель	3,64	4,28	6,10	6,20	6,30	3,40	3,50	4,30
Май	4,06	4,72	5,70	5,50	5,90	3,00	3,40	4,50
Июнь	4,35	5,22	5,50	5,80	5,90	3,40	3,60	4,70
Июль	3,76	5,21	6,10	6,50	5,60	3,40	3,60	4,30
Август	4,09	5,00	5,80	6,10	5,20	3,20	6,80	4,40
Сентябрь	4,46	5,07	5,30	6,20	3,10	3,30	3,70	4,60
Октябрь	4,33	5,46	5,70	6,90	3,10	3,50	4,10	4,20
Ноябрь	4,77	6,30	5,60	6,50	3,10	3,60	4,40	4,90
Декабрь	5,48	6,60	6,40	8,40	3,60	4,10	4,90	5,10

## П2.4. Безработица

В качестве исходной информации используются данные: общая численность безработных (на конец года), млн человек – месячные данные с 01.1994 по 08.2000 (для построения модели) и с 09.2000 по 12.2001 (для построения прогноза); источник – Госкомстат РФ.

	1994	1995	1996	1997	1998	1999	2000	2001
Январь	4,7	5,7	6,7	7,3	8,3	10,1	8,7	7,1
Февраль	5,0	5,9	6,8	7,5	8,4	10,4	8,6	7,1
Март	5,3	5,8	6,7	7,6	8,5	10,0	8,2	6,8
Апрель	5,5	6,0	6,9	7,8	8,5	9,6	7,8	6,4
Май	5,5	6,0	7,0	7,9	8,3	9,1	7,4	6,1
Июнь	5,5	6,1	7,0	7,9	8,1	8,8	7,3	6,1
Июль	5,6	6,2	7,1	7,9	8,1	8,7	7,2	6,1
Август	5,7	6,4	7,1	7,9	8,3	8,7	7,1	6,1
Сентябрь	5,7	6,5	7,1	8,0	8,6	8,8	7,1	6,2
Октябрь	5,7	6,6	7,2	8,1	8,9	8,9	7,0	6,2
Ноябрь	5,7	6,7	7,2	8,1	9,4	9,1	7,0	6,3
Декабрь	5,7	6,5	7,3	8,1	9,7	8,9	7,0	6,4

## П2.5. Доходы федерального бюджета

**Доходы бюджета** – денежные средства, поступающие в безвозмездном и безвозвратном порядке в соответствии с бюджетным и налоговым законодательством Российской Федерации в распоряжение органов государственной власти Российской Федерации. В доходах бюджетов могут быть частично централизованы доходы, зачисляемые в бюджеты других уровней бюджетной системы Российской Федерации для целевого финансирования централизованных мероприятий, а также безвозмездные перечисления. В составе доходов бюджетов обособленно учитываются доходы целевых бюджетных фондов.

**К налоговым доходам федерального бюджета** относятся предусмотренные налоговым законодательством Российской Федерации федеральные налоги и сборы, а также пени и штрафы.

В качестве исходной информации налоговых доходов используются следующие данные: все налоговые доходы федерального бюджета, млрд руб. (с 1998 г. млн руб.) – месячные данные с 01.1992 по 05.2000 (для построения модели) и с 06.2000 по 12.2001 (для построения прогноза); источник – Министерство финансов РФ.

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Январь	32,0	729,0	2316,0	7798,4	11252,4	11460,1	15792,6	24579,3	56841,4	88911,3
Февраль	68,8	728,6	2099,0	7698,0	10359,6	14666,8	15411,7	24059,5	65851,5	97636,2
Март	76,3	914,4	3421,0	10062,0	16172,0	19486,8	18745,2	31444,3	73470,0	110422,9
Апрель	150,2	1382,0	3824,0	13580,5	14117,8	24375,4	18882,1	39246,0	80889,9	123290,0
Май	101,4	946,0	8580,0	14008,5	13036,7	23683,3	19049,2	33584,7	88129,7	123489,5
Июнь	99,6	1135,0	4017,0	13254,2	17473,3	13522,3	17423,2	42332,7	78032,0	118092,8
Июль	176,2	1438,0	5192,0	16215,7	18213,8	16504,8	18379,2	47658,2	77035,2	121577,6
Август	183,0	1320,0	5373,5	15709,6	17241,0	15935,0	15520,2	42869,2	80207,2	123081,9
Сентябрь	286,9	1407,0	6232,3	16935,8	17616,6	17409,4	15420,9	40238,8	78308,1	118676,3
Октябрь	630,6	1977,0	8244,7	21578,7	18188,7	20351,7	19284,9	49601,9	83175,2	125962,3
Ноябрь	404,9	2174,0	8654,5	18288,6	22755,6	20811,2	23928,0	57408,7	91118,3	134743,1
Декабрь	647,5	2599,0	12376,0	15333,2	42294,0	45341,0	38146,7	76483,9	112423,6	175067,7

В качестве исходной информации доходов бюджета используются следующие данные: совокупные (налоговые и неналоговые доходы) федерального бюджета, млрд руб. (с 1998 г. млн руб.) – месячные данные с 01.1992 по 05.2000 (для построения модели)



и с 06.2000 по 12.2001 (для построения прогноза); источник – Министерство финансов РФ.

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Январь	40,00	756,00	2377,00	9326,90	13091,31	14660,34	18902,14	27757,83	64912,72	94601,94
Февраль	73,95	778,60	2285,00	10953,40	14244,49	19021,50	19033,98	26925,27	73386,83	105459,46
Март	88,20	1006,40	5402,00	13450,30	22777,80	22111,65	22832,02	34403,32	83524,48	117906,21
Апрель	153,48	1310,00	3735,00	15855,40	14644,05	26926,85	22236,57	44830,05	92222,53	133928,63
Май	111,61	1001,00	10604,00	15490,60	22532,89	26583,40	23330,67	39746,22	101450,10	133450,12
Июнь	97,96	1531,00	5174,00	15037,40	27780,87	16821,07	21697,54	52920,62	92224,77	125605,93
Июль	193,23	1626,00	4758,00	19421,88	22509,60	21979,16	22248,31	55543,63	88387,38	135925,29
Август	198,87	1753,00	5903,00	25615,93	19928,20	34725,33	21525,06	52214,41	96925,27	132578,05
Сентябрь	290,10	2145,00	7107,69	23766,52	20857,61	22175,35	20248,14	53010,89	90146,72	128973,97
Октябрь	611,60	2066,00	8944,39	24424,18	22264,55	26005,26	23689,62	58279,83	97209,84	142335,01
Ноябрь	444,30	3020,00	9252,80	32809,71	26409,18	30471,98	27590,58	69257,43	110286,24	142813,51
Декабрь	665,73	4635,00	13308,20	20724,71	54909,13	61352,32	59051,86	96819,86	136893,79	197117,39

### П2.6. Валовой внутренний продукт

**Валовой внутренний продукт (ВВП)** – представляет собой на стадии производства сумму добавленных стоимостей отраслей экономики, а на стадии использования – стоимость товаров и услуг, предназначенных для конечного потребления, накопления и экспорта. ВВП рассчитывается в текущих основных и рыночных ценах (номинальный ВВП). Для изучения динамики ВВП применяются постоянные цены. Реальный ВВП рассчитывается методом дефлятирования.

В качестве исходной информации используются данные: номинальный объем валового внутреннего продукта, *трлн руб. (с 1998 г млрд руб.)* – квартальные данные с 01.1994 по 02.2000 (для построения модели) и с 03.2000 по 04.2001 (для построения прогноза); источник – Госкомстат РФ.

	1994	1995	1996	1997	1998	1999	2000	2001
I квартал	88	253	456	539	564	867	1461	1886
II квартал	130	353	509	594	632	1108	1642	2116
III квартал	168	443	570	679	699	1358	2004	2543
IV квартал	225	491	611	667	846	1424	1956	2545

### П2.7. Индекс интенсивности промышленного производства

В качестве исходной информации используются сезонно скорректированные месячные данные по индексу интенсивности промышленного производства с 12.1990 по 07.2000, рассчитываемые ЦЭК при Правительстве РФ.

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Январь	94,42	80,53	69,57	50,42	47,26	42,48	40,34	41,90	40,96	46,26
Февраль	93,03	79,78	69,78	48,95	46,61	42,27	40,55	41,49	41,41	46,48
Март	91,35	79,10	69,42	48,06	45,89	42,08	40,77	41,17	41,96	46,48
Апрель	89,58	78,02	68,64	47,32	45,46	41,85	40,97	40,80	42,71	46,49
Май	88,24	76,27	67,50	46,60	45,48	41,61	41,24	40,18	43,59	46,66
Июнь	87,59	74,03	65,92	46,13	45,79	41,42	41,67	39,32	44,34	47,06
Июль	87,50	71,92	64,01	45,97	45,97	41,29	42,19	38,48	44,81	47,68
Август	87,40	70,34	62,06	46,05	45,73	41,21	42,62	38,00	44,94	
Сентябрь	86,74	69,31	60,20	46,31	45,05	41,08	42,82	38,10	44,87	
Октябрь	85,36	68,71	58,12	46,75	44,16	40,83	42,82	38,73	44,85	
Ноябрь	83,52	68,61	55,54	47,23	43,35	40,51	42,67	39,60	45,16	
Декабрь	81,77	69,00	52,75	47,48	42,80	40,30	42,35	40,38	45,73	

### ***П2.8. Индекс интенсивности производства цветных металлов (SA)***

В качестве исходной информации используются сезонно скорректированные месячные данные по индексу интенсивности промышленного производства с 12.1990 по 07.2000, рассчитываемые ЦЭК при Правительстве РФ.

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Январь	100.00	93.14	83.62	72.17	72.04	75.02	79.88	81.26	85.71	87.78	95.94	100.03
Февраль	100.46	92.17	82.42	70.53	72.80	76.07	80.69	81.57	84.82	88.53	96.24	100.74
Март	100.39	91.02	81.59	70.52	72.72	76.51	81.49	81.82	84.17	89.67	96.11	
Апрель	99.86	89.82	81.03	71.94	71.91	76.67	81.71	82.10	84.13	90.91	95.94	
Май	98.83	88.81	80.58	73.71	70.83	77.14	81.32	82.36	84.61	91.99	95.97	
Июнь	97.47	88.03	80.17	74.56	70.13	78.18	80.53	82.67	85.08	93.03	96.31	
Июль	96.27	87.54	80.02	73.95	70.10	79.39	79.48	83.23	85.23	94.02	97.01	
Август	95.71	87.37	80.18	72.37	70.45	80.27	78.53	84.20	85.27	94.52	97.96	
Сентябрь	95.71	87.33	80.29	70.70	70.83	80.61	78.08	85.39	85.57	94.35	98.95	
Октябрь	95.61	87.09	79.62	69.67	71.37	80.45	78.39	86.37	86.19	94.06	99.65	
Ноябрь	94.96	86.32	77.77	69.77	72.27	79.96	79.37	86.77	86.85	94.31	99.92	
Декабрь	94.03	85.06	74.95	70.79	73.60	79.62	80.51	86.47	87.34	95.14	99.91	

### ***П2.9. Индекс интенсивности производства цветных металлов (NSA)***

В качестве исходной информации используются месячные данные по индексу интенсивности промышленного производства с 12.1990 по 07.2000, рассчитываемые ЦЭК при Правительстве РФ.

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Январь	100.63	93.73	81.41	70.04	72.05	73.00	78.39	79.74	85.55	88.16	97.07	100.96
Февраль	91.60	86.43	79.54	64.59	67.26	71.36	77.05	74.80	78.55	79.81	91.03	91.90
Март	102.91	90.72	82.50	72.88	73.22	78.08	84.64	83.05	85.58	90.87	97.19	
Апрель	99.47	90.50	80.44	67.72	73.24	76.88	80.85	82.36	81.92	92.40	95.84	
Май	103.03	91.33	84.52	80.35	73.57	77.59	83.87	85.29	88.28	94.33	98.36	
Июнь	97.11	89.55	81.41	76.87	69.22	79.891	81.27	82.75	86.49	91.24	95.91	
Июль	97.62	87.36	78.64	73.02	71.07	81.25	82.31	84.14	86.01	96.42	99.04	
Август	93.93	87.23	82.308	73.67	73.45	81.27	76.86	85.84	86.39	97.79	99.41	
Сентябрь	94.88	87.52	78.15	70.10	67.90	78.28	78.10	82.73	83.51	92.18	96.91	
Октябрь	99.37	87.67	81.89	69.36	72.62	82.21	78.79	87.82	87.21	94.27	101.53	
Ноябрь	94.04	86.46	78.25	68.44	70.95	79.64	75.79	86.84	85.90	92.20	99.47	
Декабрь	92.41	86.05	75.14	71.22	74.93	79.46	85.46	86.94	89.34	97.56	100.11	

## Литература

1. *Айвазян С.А., Мхитарян В.С.* (1998): Прикладная статистика и основы эконометрики. М.: ЮНИТИ, 1998.
2. *Бокс Дж., Дженкинс Г.* (1974): Анализ временных рядов. Прогноз и управление. М.: Мир, 1974. Вып. 1, 2.
3. *Гаврилец Ю.Н.* (1974): Социально-экономическое планирование (системы и модели). М.: Экономика, 1974.
4. *Заруцкий В.И.* (1980): О выделении некоторых графов связей для нормальных векторов большой размерности / Алгоритмическое и программное обеспечение прикладного статистического анализа. М.: Наука, 1980.
5. Эконометрический анализ динамических рядов основных макроэкономических показателей (2001): Институт экономики переходного периода, Научные труды № 34Р. М.: ИЭПП, 2001.
6. *Юдин А.Д.* Об информативных структурах многомерных случайных величин // Известия АН СССР, 1977. №6.
7. *Юдин А.Д.* (1976): Структуры наборов псевдонезависимых случайных величин / Модели и методы исследования социально-экономических процессов. М.: ЦЭМИ АН СССР, 1976.
8. *Юдин А.Д.* (1979): О выделении существенных связей в многомерной случайной величине / Модели социально-экономических процессов и социальное планирование. М.: Наука, 1979.
9. *Юдин А.Д.* (1982): Сложность статистических систем. Доклады АН СССР, 1982. Т. 266. № 5.
10. *Юдин Д.Б., Юдин А.Д.* (1985): Математики измеряют сложность. М.: Знание, 1985 (сер. «Число и мысль». Вып. 8).
11. *Яглом А.М., Яглом И.М.* (1973): Вероятность и информация. М.: Наука, 1973.
12. *Batchelor Roy, P. Dua* (1990): Forecaster Ideology, Forecasting Technique, and the Accuracy of Economic Forecasts. International Journal of Forecasting, 6, № 1, 3–10.
13. *Bates J.M., C.W.J. Granger* (1969): The Combination of Forecasts, Operation Research Quarterly, 20, 451–468.
14. *Box J.E.P., G.M. Jenkins* (1976): Time Series Analysis, Forecasting and Control, Revised edition, San Francisco, Holden Day.
15. *Brown R.G.* (1963): Smoothing, Forecasting and Prediction. Prentice-Hall, Englewood Cliffs, N.Y.

16. *Campbell John J., Pierre Perron* (1991): Pitfalls and Opportunities: What Macroeconomists should know about Unit Roots, *Macroeconomics Annual*, 1991, NBER, 141–201.
17. *Christiano L.J., M. Eichenbaum* (1990): Unit Roots in Real GDP: Do We Know, and Do We Care? *Carnegie-Rochester Conference Series on Public Policy*, 32, 7–62.
18. *Christoffersen Peter F., Francis X. Diebold* (1998): Cointegration and Long-Horizon Forecasting, *Journal of Business and Economic Statistics*, 16, № 4, 450–458.
19. *Clements Michael P., D.F. Hendry* (1996): Multi-step estimation for forecasting, *Oxford Bulletin of Economics and Statistics*, 58, № 4, 657–684.
20. *Clements Michael P., D.F. Hendry* (1998a): Forecasting Economic Processes, *International Journal of Forecasting*, 14, № 1, 111–131.
21. *Clements Michael P., D.F. Hendry* (1998b): *Forecasting Economic Time Series*. Cambridge: Cambridge University Press. (The Marshall Lectures on Economic Forecasting.)
22. *Clements Michael P., D.F. Hendry* (2000): Forecasting with difference-stationary and trend-stationary models. Discussion Paper Series, № 5. Department of Economics, University of Oxford.
23. *Clements Michael P., D.F. Hendry* (2001): *Forecasting Non-stationary Economic Time Series*. Cambridge, Massachusetts: The MIT Press.
24. *Dickey D.A., W.R. Bell, R.B. Miller* (1986): Unit Roots in Time Series Models: Tests and Implications, *American Statistica*, 40, 12–26.
25. *Diebold F.X., Senhadji A.* (1996): The Uncertain Unit Root in U.S. GNP: Comment, *American Economic Review*, 86, 1291–1298.
26. *Diebold Fransis X.* (1998): The Past, Present and Future of Macroeconomic Forecasting, *Journal of Economic Perspectives*, 12, № 2, 175–192.
27. *Diebold Fransis X., Lutz Kilian* (2000): Unit-Root Tests Are Useful for Selecting Forecasting Models, *Journal of Business and Economic Statistics*, 18, № 3, 265-273.
28. *Elliott G., T.J. Rothenberg, J.H. Stock* (1996): Efficient Tests for an Autoregressive Unit Root, *Econometrica*, 64, 813–836.
29. *Franses P.H., F. Kleibergen* (1996): Unit Roots in the Nelson-Plosser Data: Do They Matter for Forecasting? *International Journal of Forecasting*, 12, № 2, 283–288.
30. *Granger C.W.J., Newbold P.* (1986): *Forecasting Economic Time Series*, Second Edition. New York: Academic Press.

31. *Granger C.W.J., R. Ramanathan* (1984): Improved Methods of Combining Forecasting, *Journal of Forecasting*, 3, 197–204.
32. *Gregory Allan W., Gregor W. Smith, James Yetman* (2001): Testing for Forecast Consensus, *Journal of Business and Economic Statistics*, 19, № 1, 34–43.
33. *Harvey David I., Stephen J. Leybourne, Paul Newbold* (1998): Tests for Forecast Encompassing, *Journal of Business and Economic Statistics*, 16, № 2, 254.
34. *Klein L.R.* (1984): The Importance of the Forecast, *Journal of Forecasting*, 3, № 1, 1–9.
35. *Klein L.R.* (1991): *Comparative Performance of US Econometric Models*. Oxford: Oxford University Press.
36. *Klein L.R., E. Burmeister* [eds] (1976): *Econometric Modeling Performance*. Philadelphia: University of Pennsylvania Press.
37. *Kolmogoroff A.* (1939): Sur L'interpolation et L'extrapolation des Suites Stationnaires, *Compt. Rend.*, 208, 2043.
38. *Kumar K., D.S. Gill* (2000): On Forecasting Economic Time Series Data: A Comparative Study, *Indian Journal of Economics*, 81, 265–273.
39. *Markidakis S.A., A. Anderson, R. Carbonne, R. Flides, M. Hibon, R. Lewandowski, J. Newton, E. Parsen, R. Winkler* (1982): The Accuracy of Extrapolation (Time-Series) Methods: Results of a Forecasting Competition, *Journal of Forecasting*, 1, 111–153.
40. *McNees S.K.* (1979): The Forecasting Record of the 1970's, *New England Economic Review*, Sep./Oct., 15–21.
41. *McNees S.K., J. Ries* (1983): The Track Record of Macroeconomic Forecasters, *New England Economic Review*, Nov./Dec., 5–18.
42. *Meese R., J. Geweke* (1984): A Comparison of Autoregressive Univariate Forecasting Procedures for Macroeconomic Time Series, *Journal of Business and Economic Statistics*, 2, 191–200.
43. *Nelson C.R., C.I. Plosser* (1982): Trends and Random Walks in Macroeconomic Time Series, *Journal of Monetary Economics*, 10, 139–162.
44. *Nyblom J.* (1989): Testing for Constancy of Parameters over Time, *Journal of the American Statistical Association*, 84, 223–230.
45. *Quandt R.E.* (1960): Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes, *Journal of the American Statistical Association*, 55, 324–330.
46. *Rudebusch G.D.* (1993): The Uncertain Root in Real GNP, *American Economic Review*, 83, 264–272.

47. *Snyder Ralph D., J. Keith Ord, Anne B. Koehler* (2001): Prediction Intervals for ARIMA Models, *Journal of Business and Economic Statistics*, 19, № 2, 217–225.
48. *Staiger D., Stock J.H., M.W. Watson* (2001): Prices, Wages and the U.S. NAIRU in the 1990s, NBER Working Paper No. 8320, NBER, Jun. 2001.
49. *Stock J.H.* (1996): VAR, Error Correction and Pretest Forecasts at Long Horizons, *Oxford Bulletin of Economics and Statistics*, 58, № 4, 685–701.
50. *Stock J. H., M.W. Watson* (1994): Evidence on Structural Instability in Macroeconomic Time Series Relations, NBER Technical Working Paper No.164, NBER, Sept. 1994
51. *Stock J.H., Watson M.W.* (1996): A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series, *Journal of Business and Economic Statistics*, 14, № 1, 11–30.
52. *Stock J. H., M.W. Watson* (1998a): Business Cycle Fluctuations in U.S. Macroeconomic Time Series, NBER Working Paper No. 6528, NBER, Apr. 1998.
53. *Stock J.H., M.W. Watson* (1998b): Diffusion Indexes, NBER Working Paper No. 6702, NBER, Aug. 1998.
54. *Stock J.H., Watson M.W.* (1999): Forecasting Inflation, NBER Working Paper No.7023, NBER, March 1999.
55. *Stock J.H., M.W. Watson* (2001): Forecasting Output and Inflation: the Role of Asset Prices, NBER Working Paper No. 8180, NBER, Mar. 2001.
56. *Swanson N.R., H. White* (1995): A Model Selection Approach to Assessing the Information in Term Structure Using Linear Models and Artificial Neural Networks, *Journal of Business and Economic Statistics*, 13, 265–275.
57. *Swanson N.R., H. White* (1997): Forecasting Economic Time Series Using Flexible versus Fixed Specification and Linear versus Nonlinear Econometric Models, *International Journal of Forecasting*, 13, № 4, 439–461.
58. *Swanson N.R., H. White* (1997): A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks, *Review of Economics and Statistics*, 79, 540–550.
59. *Wallis K.F.* (1989): Macroeconomic forecasting: A survey, *Economic Journal*, 99, 28–61.
60. *Wallis K.F., M.J. Andrews, D.N.F. Bell, P.G. Fisher, J.D. Whitley* (1984): *Models of the UK Economy, A Review by the ESRC Macroeconomic Modeling Bureau.* Oxford: Oxford University Press.
61. *Wallis K.F., M.J. Andrews, P.G. Fisher, J. Longbottom, J.D. Whitley* (1986): *Models of the UK Economy: A Third Review by the ESRC Macroeconomic Modeling Bureau.* Oxford: Oxford University Press.

62. *Wallis K.F., J.D. Whitley* (1991): Sources of Error in Forecasts and Expectations: UK economic models 1984-8, *Journal of Forecasting*, 10, 231–253.
63. *Weigand A.S., Gershenfeld* (1994): *Time Series prediction: Forecasting the Future and Understanding the Past*. Addison-Wesley for Santa Fe Institute: Reading.
64. *West Kenneth D.* (2001): Tests for Forecast Encompassing When Forecasts Depend on Estimated Regression, *Journal of Business and Economic Statistics*, 19, № 1, 29–33.
65. *Wiener N.* (1949): *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. John Wiley, New York.
66. *Zarnowitz Victor* (1978): On the Accuracy and Properties of Recent Macroeconomic Forecasts, NBER Working Paper No 0229.
67. *Zarnowitz Victor* (1991): Has Macro-Forecasting failed? NBER Working Paper No 3867.
68. *Zarnowitz Victor, Phillip Braun* (1992): Twenty-two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Combinations of Forecasting Performance, NBER Working Paper No 3965.