

***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
Часть I. Прогнозирование временных рядов традиционными эконометрическими методами и их модификациями	9
Глава 1. Постановка проблемы и инструментарий исследования	9
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
Глава 2. Исследование сравнительного качества прогнозов некоторых макроэкономических рядов РФ, получаемых по фиксированным и рекурсивным TS- и DS-моделям	35
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
--	----

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

Глава 4. Источники ошибок прогнозов и некоторые методы их коррекции	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

Глава 5. Моделирование и прогноз налоговых поступлений в консолидированный и федеральный бюджеты РФ	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
ЧАСТЬ II. Структуры случайных векторов	201
Глава 6. Прогнозирование с использованием структур случайных векторов	201
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
Заключение. Некоторые выводы	
из полученных результатов	286
Приложения	301
Приложение 1. Макросы	301
П1.1. Построение трехмерных информаций	301
П1.2. Построение прогнозов	307
Приложение 2. Исходные данные.	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
Литература	315

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 возростал. 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 ав-топрогноза 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 ав-топрогноза 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 -кратного¹ 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 долгосрочной 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-

¹ 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 одномесячном, шестимесячном и двенадцатимесячном 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 длиной запаздываний, определяемой на основании AIC. Они обнаружили также, что линейные комбинации прогнозов улучшают качество прогнозов несущественно. [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:

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

где y_t – ряд, для которого строится прогноз, h – горизонт прогнозирования, i – индекс модели прогнозирования ($i=1, \dots, 121$), θ_{ih} – вектор неизвестных параметров, u_{it} – ошибка, а Z_t – вектор предикторных переменных. Обычно $Z_t = (y_{t-p}, \dots, y_{t-p}, \Delta y_{t-p}, \dots, \Delta y_{t-p}, 1, t)$ 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 θ_{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-

вышена 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 получают 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^μ для различения моделей с константой, но без тренда, и претестовая статистика DF-GLS^τ – 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-

зрящему 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 η_t , ковариационные матрицы которых различаются для различных моделей коэффициентом пропорциональности λ .

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 λ 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 λ 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 неоднородными.

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 ATVP1, а среди двумерных – VTVP1. При этом ATVP1 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 MSEF (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

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

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 μ 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 предпринимают далее статистическое моделирование, использующее несколько более сложные TS - и DS -модели процесса порождения данных.

В качестве TS DGP используется процесс

$$y_t = \varphi + \gamma + u_t, \text{ где } u_t = \rho u_{t-1} + a_t, a_t \sim NID(0, \sigma_a^2), 0 \leq \rho \leq 1.$$

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

В качестве DS DGP берется процесс

$$y_t = y_{t-1} + \mu + \varepsilon_t, \text{ где } \varepsilon_t = b_t - \psi b_{t-1}, b_t \sim NID(0, \sigma_b), 0 \leq \psi < 1.$$

В этом случае статистическая TS -модель может более успешно аппроксимировать DS DGP при значениях ψ , 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 ρ 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 неперемного 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 неперемного построения прогнозирующих моделей в уровнях.