

ПРОГНОЗИРОВАНИЕ В ПЕРИОДЫ ЭКОНОМИЧЕСКОЙ НЕСТАБИЛЬНОСТИ: СУЩЕСТВУЮТ ЛИ ПРОСТЫЕ СПОСОБЫ УЛУЧШЕНИЯ КАЧЕСТВА ПРОГНОЗОВ

М.Турунцева, зав. лабораторией, ИЭП им. Е.Т. Гайдара и РАНХиГС
Е.Астафьевая, с.н.с., РАНХиГС

Хорошо известно, что в периоды структурных сдвигов, часто совпадающих с экономическими кризисами, особенно остро встает проблема точности прогнозов экономических показателей. В принципе, на данный момент не существует универсальных методов прогнозирования экономических показателей в периоды структурных изменений. Чаще всего в такие периоды в большей мере приходится опираться на экспертные оценки, чем на прогнозы, построенные на основе каких-либо моделей. Тем не менее, мы попытались понять, существуют ли простые формальные методы, которые позволяют улучшить краткосрочные прогнозы российских макроэкономических показателей в период, когда происходят структурные изменения в экономике. Может ли мы как-то учитывать происходящие в экономике изменения при прогнозировании в режиме «реального времени», используя какие-то достаточно простые математические модели? Может ли помочь учет ошибок прогнозирования в предыдущие моменты времени улучшить прогноз?

В качестве базового массива прогнозов мы использовали прогнозную панель ИЭП им. Е.Т. Гайдара¹ на интервале с августа 2006 г. по декабрь 2013 г. Для возможности сопоставления прогнозов ИЭП с альтернативными прогнозами мы строили альтернативные прогнозы, имитируя режим «реального времени», т.е. так, как мы строили бы их, если бы это происходило одновременно с расчетами прогнозов ИЭП.

Помимо прогнозов ИЭП мы рассматриваем три группы альтернативных прогнозов: простейшие прогнозы, прогнозы с корректировкой на ошибку прогнозирования и прогнозы по моделям, учитывающим структурный сдвиг (или структурные сдвиги).

Простейшие прогнозы. В эту группу включены три вида наивных прогнозов, прогноз, рассчитанный как скользящее среднее, и прогноз по модели детерминированного тренда. Отметим, что прогнозы этой группы строятся с использованием методов, никак не моделирующих наличие структурных сдвигов в данных.

Наивный прогноз (НП) строится по модели случайного блуждания

$$y_t = y_{t-1} + \varepsilon_t,$$

т.е. для любого горизонта прогнозирования прогнозное значение равно значению показателя в последний (в нашем случае – известный) момент времени.

Наивный прогноз с константой (с дрейфом) (НПД) строится по модели случайного блуждания с дрейфом

$$y_t = c + y_{t-1} + \varepsilon_t,$$

т.е. для любого горизонта прогнозирования h прогноз равен значению в последний (известный) момент времени T плюс константа c , умноженная на горизонт прогнозирования h (т.е. прогнозное значение равно $y_T + ch$).

1 См.: http://www.iep.ru/index.php?option=com_biblio&Itemid=124&catid=123&lang=ru&task=showallbib
С декабря 2003 г. по июль 2012 г. – «Бюллетень модельных расчетов краткосрочных прогнозов социально-экономических показателей РФ».

С августа по декабрь 2012 г. – Бюллетень «Модельные расчеты краткосрочных прогнозов социально-экономических показателей РФ».

С января 2013 г. – регулярный раздел «Научного вестника ИЭП им. Гайдара.ру»: <http://www.iep.ru/ru/ob-izdanii.html>

Наивный сезонный прогноз (НСП) строится по модели сезонного случайного блуждания

$$y_t = y_{t-12} + \varepsilon_t,$$

т.е. прогноз равен значению, которое принимает показатель в такой же месяц последнего известного года (в нашем случае – в соответствующий месяц предыдущего года).

Прогноз, рассчитанный как скользящее среднее (СС), строится как среднее значение показателя в последние (известные) двенадцать месяцев. Также как и наивный прогноз, он не зависит от горизонта прогнозирования, а зависит только от момента, в который рассчитывается прогноз.

И, наконец, прогноз, построенный по модели детерминированного линейного тренда (ДТ), получается из оценок модели

$$y_t = \alpha + \beta t + \varepsilon_t.$$

Таким образом, значение прогноза на h шагов вперед, рассчитанное в момент времени T , равно $\alpha + \beta(T + h)$.

Вторая группа прогнозов – *прогнозы, построенные с корректировкой на известные ошибки прогнозирования в предыдущие моменты времени*. Здесь мы корректируем имеющиеся у нас прогнозы ИЭП так, как если бы мы делали это в момент прогнозирования. Для этой цели мы использовали 4 метода корректировки прогнозов, предложенные в работе Турунцевой (2013)¹ и несколько модифицированные для целей настоящего исследования.

Корректировка на величину ошибки одношагового прогноза в последний известный момент времени (К1) предполагает, что все прогнозы в момент времени T (т.е. прогнозы на 1, 2, ..., 6 шагов вперед) корректируются на величину последней известной ошибки прогноза. Такая ошибка рассчитывается по формуле:

$$e_{T-i,1} = y_{T-i+1} - f_{T-i,1},$$

где $i=3$ для большинства рассматриваемых рядов², y_{T-i+1} – истинное значение прогнозируемой переменной в момент $T-i+1$, $f_{T-i,1}$ – прогноз интересующего нас показателя на момент $T-i+1$, сделанный в момент $T-i$.

Тогда скорректированное значение k -шагового прогноза $\tilde{f}_{T,k}$ на момент $T+k$ ($k=1, \dots, 6$), сделанного в момент T , будет равно сумме прогноза $f_{T,k}$ и ошибки прогнозирования на предыдущем шаге $e_{T-i,1}$:

$$\tilde{f}_{T,k} = f_{T,k} + e_{T-i,1}.$$

Корректировка на величину среднего значения последних 6 известных ошибок одношаговых прогнозов (К2) предполагает корректировку прогноза на величину, равную

$$\frac{1}{6} \sum_{j=1}^6 e_{T-j,1},$$

где $e_{T-j,1} = y_{T-j+1} - f_{T-j,1}$, $j = \overline{i, s}$, y_{T-j+1} – истинное значение показателя y в момент времени $T-j+1$, $f_{T-j,1}$ – прогнозное значение показателя y , сделанное в момент времени $T-j$ на один шаг вперед. Таким образом, скорректированный прогноз рассчитывается по формуле:

1 Турунцева М. Нужно ли пытаться исправить потенциальные ошибки прогнозирования? // Научный вестник ИЭП им. Гайдара.ру, № 2 (59), 2013. С. 38–45. Эл. ссылка: http://www.iep.ru/files/Nauchniy_vestnik.ru/2-2013/38-45.pdf
 2 Исключение здесь составляют ряды индексов промышленного производства НИУ ВШЭ, для которых $i=2$, показателей курсов рубля и евро ($i=1$).

$$\tilde{f}_{T,k} = f_{T,k} + \frac{1}{6} \sum_{j=1}^6 e_{T-j,1}.$$

Третий метод – корректировка рассчитываемого k –шагового прогноза на величину последней известной ошибки k –шагового прогноза (К3), т.е. на величину

$$e_{T-i,k} = y_{T-i+k} - f_{T-i,k},$$

где $i=3$ для большинства рассматриваемых рядов, $k=1, \dots, 6$, y_{T-i+k} – истинное значение прогнозируемой переменной в момент прогнозирования $T-i+k$, $f_{T-i,k}$ – прогноз интересующего нас показателя на момент $T-i+k$, сделанный в момент $T-i$. В этом случае скорректированное значение k –шагового прогноза $\tilde{f}_{T,k}$ на момент $T+k$, сделанного в момент T , будет равно сумме прогноза $f_{T,k}$ и ошибки прогнозирования $e_{T-i,k}$, сделанной в момент времени $T-i$, при прогнозировании на k шагов вперед:

$$\tilde{f}_{T,k} = f_{T,k} + e_{T-i,k}.$$

И, наконец, корректировка на величину среднего значения 6 последних известных на момент T ошибок k –шаговых прогнозов (К4). В данном случае прогноз корректируется на величину, равную

$$\frac{1}{6} \sum_{j=1}^6 e_{T-j,k},$$

где $k=1, \dots, 6$, $i=3$ для большинства рассматриваемых рядов (см. выше), $e_{T-j,k} = y_{T-j+k} - f_{T-j,k}$, $j = \overline{i, s}$, y_{T-i+k} – истинное значение показателя y в момент времени $T-j+k$, $f_{T-j,k}$ – прогнозное значение показателя y , сделанное в момент времени $T-j$ на k шагов вперед. В этом случае скорректированный прогноз рассчитывается по формуле:

$$\tilde{f}_{T,k} = f_{T,k} + \frac{1}{6} \sum_{j=1}^s e_{T-j,k}.$$

Третья группа прогнозов – прогнозы, построенные с учетом происходящего структурного сдвига. Главная задача здесь – понять, можем ли мы улучшить прогнозы показателей, зная, что в текущий период времени по каким-то причинам в экономике происходит структурный сдвиг. Прогнозы с учетом структурного сдвига представляют собой комбинации различных способов прогнозирования показателя (1.2–1.7) в окрестностях структурного сдвига и в период после структурного сдвига (1.8–1.15). Кроме того, мы рассматриваем в качестве основной модели (1.1) – модель со структурным сдвигом.

1.1. Модель со структурным сдвигом. Мы вводим в модель структурные сдвиги в константе и тренде. Дата структурного сдвига рассчитывается на всем интервале времени (до декабря 2013 г.) из расчета значимости параметров сдвига. Затем коэффициенты моделей пересчитываются для различных интервалов времени, начиная с момента структурного сдвига каждый раз с добавлением новой точки. Дата сдвига при этом не меняется. Отметим, что исходные даты сдвигов выбирались на основе методики тестирования структурных сдвигов, теоретические аспекты которой рассмотрены в разделе 1. Но, как правило, итоговые даты сдвигов не совпадали (хотя и не сильно отличались от них) с результатами эмпирического приложения этой методики.

Обозначим теперь момент структурного сдвига в динамике показателя как T_0 . Обозначим ${}^0 f_{t,i}$ – прогноз для периода до структурного сдвига; ${}^2 f_{t,i}$ – прогноз для периода после структур-

ногого сдвига по модели ARIMA, включающей структурный сдвиг; ${}^4f_{t,i}$ – прогноз, построенный на базовом интервале $[1, T_0]$ по модели со структурным сдвигом и дамми на три кризисные точки. Тогда мы получаем следующие типы прогнозов. Для модели (1.1) последовательность

прогнозов будет иметь вид: ${}^0f_{[1;T_0-1],i}$; ${}^2f_{[T_0;T],i}$ или ${}^0f_{[1;T_0-1],i}$; ${}^4f_{[T_0;T],i}$ в зависимости от качественных характеристик соответствующих моделей.

Прогнозы в окрестностях структурного сдвига (известные значения ряда $[1, T_0]$, $[1, T_0+1]$, $[1, T_0+2]$) включают в себя:

1.2. Прогнозы по моделям, оцененным по всем известным данным – модель с использованием всех известных данных, без моделирования структурного сдвига:

а) для базового периода $[1, T_0]$ – ${}^1f_{T_0,i}$ – прогноз на момент T_0+i , начиная с момента T_0+1 , построенный на данных до момента структурного сдвига T_0 включительно по модели ARIMA;

б) для базового периода $[1, T_0+1]$ – ${}^1f_{T_0+1,i}$ – прогноз на момент T_0+i , начиная с момента T_0+2 , построенный на данных до момента, следующего за структурным сдвигом T_0+1 включительно по модели ARIMA;

в) для базового периода $[1, T_0+2]$ – ${}^1f_{T_0+2,i}$ – прогноз на момент T_0+i , начиная с момента T_0+3 , построенный на данных до момента, следующего за структурным сдвигом T_0+2 включительно по модели ARIMA;

г) для остальных прогнозируемых моментов времени – прогнозы на моменты T_0+4, T_0+5, \dots , построены по модели ARIMA без структурных сдвигов на данных до T_0+3, T_0+4, \dots , соответственно.

Обозначим это множество значений ${}^1f_{[T_0+3;T],i}$.

Итоговый прогноз выглядит следующим образом (если показатель содержит 1 структурный сдвиг на рассматриваемом нами интервале. Если у показателя 2 структурных сдвига, то в окрестности второго структурного сдвига мы проводили аналогичные расчеты, и прогнозный ряд в этом случае изменится соответствующим образом. Это замечание, примененное соответственно, имеет место для всех моделей ниже): ${}^0f_{[1;T_0-1],i}$; ${}^1f_{T_0,i}$; ${}^1f_{T_0+1,i}$; ${}^1f_{T_0+2,i}$; ${}^1f_{[T_0+3;T],i}$.

1.3. Прогнозы по моделям, оцененным по данным в период $[1, T_0-1]$, и последующей корректировкой на ошибку в момент структурного сдвига – фиксированный прогноз по данным до момента структурного сдвига с корректировкой на ошибку в окрестности структурного сдвига:

а) для базового периода $[1, T_0]$ – ${}^2f_{T_0,i} = f_{T_0-1,i+1} + (y_{T_0} - f_{T_0-1,1})$;

б) для базового периода $[1, T_0+1]$ – ${}^2f_{T_0+1,i} = f_{T_0-2,i+2} + (y_{T_0+1} - f_{T_0-2,2})$;

в) для базового периода $[1, T_0+2]$ – ${}^2f_{T_0+2,i} = f_{T_0-3,i+3} + (y_{T_0+2} - f_{T_0-3,3})$;

г) для остальных прогнозируемых моментов времени – это массив ${}^1f_{[T_0+3;T],i}$ – прогнозы на моменты T_0+4, T_0+5, \dots , построены по модели ARIMA без структурных сдвигов на данных до T_0+3, T_0+4, \dots , соответственно.

Т.е. итоговая последовательность прогнозов в данном случае принимает вид: ${}^0f_{[1;T_0-1],i}$; ${}^2f_{T_0,i}$; ${}^2f_{T_0+1,i}$; ${}^2f_{T_0+2,i}$; ${}^1f_{[T_0+3;T],i}$.

1.4. Прогнозы по моделям, оцененным по всем известным данным и включающим фиктивные переменные для наблюдений после структурного сдвига

а) для базового периода $[1, T_0]$ – ${}^3f_{T_0,i}$ – ARIMA-модель с фиктивной переменной на момент T_0 ;

б) для базового периода $[1, T_0+1] - f_{T_0+1,i}^3$ – ARIMA-модель с фиктивной переменной на моменты T_0 и T_0+1 ;

в) для базового периода $[1, T_0+2] - f_{T_0+2,i}^3$ – ARIMA-модель с фиктивной переменной на момент T_0 , T_0+1 и T_0+2 ;

г) для остальных прогнозируемых моментов времени – это массив ${}^1f_{[T_0+3;T],i}$ – прогнозы на моменты T_0+4 , T_0+5 , ..., построены по модели ARIMA без структурных сдвигов на данных до T_0+3 , T_0+4 , ..., соответственно.

Итоговая последовательность имеет вид: ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^3; f_{T_0+1,i}^3; f_{T_0+2,i}^3; {}^1f_{[T_0+3;T],i}$.

1.5. Для базовых интервалов $[1, T_0]$, $[1, T_0+1]$ и $[1, T_0+2]$ прогнозы оценивались, как в случае (1.2а–в);

г) для базовых интервалов, начиная с $[1, T_0+3]$, как ARIMA-модель с дамми-переменными на три кризисные точки – ${}^3f_{[T_0+3;T],i}$. Тогда последовательность прогнозных значений имеет вид ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^1; f_{T_0+1,i}^1; f_{T_0+2,i}^1; {}^3f_{[T_0+3;T],i}$.

1.6. Модель, являющаяся комбинацией прогнозов (1.2а–в) и (1.5г): ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^2; f_{T_0+1,i}^2; f_{T_0+2,i}^2$ и ${}^3f_{[T_0+3;T],i}$.

1.7. Для всех базовых интервалов прогноз строится по всем имеющимся наблюдениям с включением дамми-переменных на моменты времени T_0 , T_0+1 и T_0+2 . Последовательность прогнозов принимает вид: ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^3; f_{T_0+1,i}^3; f_{T_0+2,i}^3; {}^3f_{[T_0+3;T],i}$.

Прогнозы по моделям, учитывающим структурный(ые) сдвиг(и), составляют следующий массив моделей:

1.8. Для базовых интервалов до $[1, T_0+2]$ включительно прогнозы рассчитываются аналогично (1.2а–в). Далее – по модели со структурным сдвигом ${}^2f_{[T_0+3;T],i}$. В этом случае последовательность прогнозов имеет вид: ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^1; f_{T_0+1,i}^1; f_{T_0+2,i}^1; {}^2f_{[T_0+3;T],i}$.

Пусть $f_{T_0+2,i}^4$ – прогноз, построенный на базовом интервале $[1, T_0+2]$ по модели со структурным сдвигом. Кратко опишем оставшиеся массивы прогнозов:

1.9. Прогнозы рассчитываются как смесь прогнозов из (3.3а–в) и ${}^2f_{[T_0+3;T],i}$, т.е. последовательность прогнозов имеет вид: ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^2; f_{T_0+1,i}^2; f_{T_0+2,i}^2; {}^2f_{[T_0+3;T],i}$.

1.10. Последовательность прогнозов имеет вид: ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^3; f_{T_0+1,i}^3; f_{T_0+2,i}^3; {}^2f_{[T_0+3;T],i}$, т.е. прогнозы рассчитываются с включением дамми на точки околоструктурного сдвига, а после этого – по модели со структурным сдвигом.

1.11. Прогнозы рассчитываются аналогично предыдущему случаю за исключением прогноза на базовом интервале $[1, T_0+2]$, на котором прогноз считается по модели со структурным сдвигом. Последовательность прогнозов в этом случае принимает вид ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^3; f_{T_0+1,i}^3; f_{T_0+2,i}^4; {}^2f_{[T_0+3;T],i}$.

Следующие четыре модели фактически повторяют модели (1.8)–(1.11) с заменой прогноза послеструктурного сдвига на прогноз по модели со структурным сдвигом и дамми на кризис, т.е. на ${}^4f_{[T_0+3;T],i}$. Поэтому последовательности прогнозов принимают вид:

1.12. ${}^0f_{[1;T_0-1],i}; f_{T_0,i}^1; f_{T_0+1,i}^1; f_{T_0+2,i}^1; {}^4f_{[T_0+3;T],i}$.

1.13. ${}^0 f_{[1:T_0-1],i}; f_{T_0,i}^2; f_{T_0+1,i}^2; f_{T_0+2,i}^2; {}^4 f_{[T_0+3;T],i}$.

1.14. ${}^0 f_{[1:T_0-1],i}; f_{T_0,i}^3; f_{T_0+1,i}^3; f_{T_0+2,i}^3; {}^4 f_{[T_0+3;T],i}$.

1.15. ${}^0 f_{[1:T_0-1],i}; f_{T_0,i}^3; f_{T_0+1,i}^3; f_{T_0+2,i}^4; {}^4 f_{[T_0+3;T],i}$.

Таким образом, для каждого показателя было получено по 20 рядов прогнозов, которые сравнивались между собой при помощи простейших характеристик качества.

Для тестирования наличия и датировки структурных сдвигов мы использовали ряд стандартных процедур¹. Понимая то обстоятельство, что непосредственно в периоды структурных изменений мы не можем использовать рассмотренные процедуры, тем не менее, мы уделяем пристальное внимание и этому вопросу, поскольку на исторических данных эти методы могут позволить улучшить качество прогнозов.

В итоге, мы отобрали 30 макроэкономических показателей, у которых были обнаружены структурные сдвиги в интересующий нас период времени. Отметим, что тесты показали наличие структурных сдвигов у большего числа показателей, но для некоторых из них при оценках моделей оказалось, что коэффициенты при параметрах, характеризующих структурные сдвиги, незначимы. Эти показатели были исключены из дальнейшего рассмотрения. Из полученных 30 показателей у 21 показателя был обнаружен один структурный сдвиг на интересующем нас интервале времени, у 9 – по два структурных сдвига. Таким образом, мы получили 39 эпизодов структурных сдвигов.

Анализ влияния структурных сдвигов на качество прогнозов российских макроэкономических показателей и оценку возможности предложенных методов прогнозирования в периоды структурных изменений мы проводили в краткосрочном и долгосрочном аспектах. Исследование в краткосрочном аспекте заключалось в анализе качества всех возможных для данного показателя методов прогнозирования в период времени, близкий к идентифицированному моменту структурного сдвига. Мы рассматривали по два промежутка для каждого структурного сдвига: по 10 и по 17 месяцев. Как показал наш анализ, существует некоторое различие между этими промежутками, но оно не слишком велико. В целом, наибольшую конкурентоспособность на коротких промежутках показали два метода: прогноз по модели линейного тренда и прогноз по модели (1.9), который для момента структурного сдвига и для двух последующих точек строится как на основе данных до момента структурного сдвига, т.е. за период $[1, T_0-1]$, с корректировкой на ошибку в момент структурного сдвига, а начиная с момента T_0+3 – по модели ARMA со структурным сдвигом, определенным выше. На эти два метода приходится около половины всех рассмотренных случаев.

В целом, методы из группы 3 помогают улучшить прогноз в окрестности структурного сдвига для короткого интервала (10 месяцев) в 19 случаях из 39, для длинного интервала (17 месяцев) в 16 случаях. Если рассматривать методы из второй группы (т.е. методы корректировки потенциальной ошибки прогнозирования) как методы корректировки ошибки вследствие наличия структурного сдвига, то число улучшений (в сумме с методами третьей группы) увеличивается до 24 в случае короткого интервала и до 21 для длинного интервала. Таким образом, более чем в половине случаев существует возможность улучшения прогнозов в окрестности структурного сдвига одним из предложенных способов.

1 Мы используем тесты MDF-GLS с одним структурным сдвигом, MDF-GLS с двумя структурными сдвигами, MDF-OLS с одним структурным сдвигом, MDF-OLS с двумя структурными сдвигами, стратегию объединения отвержений, учитывающую только претест на начальное значение и стратегию объединения отвержений, учитывающую оцененное значение числа сдвигов (подробное описание используемых процедур см.: Skrobotov A., 2015, On trend, Breaks and Initial Condition in Unit Root testing, <https://docs.google.com/file/d/0By0reZAB6yFiS2Z0eU9razZsY1U/edit>). Мы предполагали, что в рядах содержится не более двух структурных сдвигов, поскольку тесты на число структурных сдвигов склонны к резкой потере мощности при увеличении числа сдвигов при тестировании, т.е. они все чаще не отвергают нулевую гипотезу о наличии конкретного числа структурных сдвигов.

В долгосрочном периоде – т.е. на всем имеющемся интервале времени – ситуация несколько иная. Можно утверждать, что нет какого-либо метода, который давал бы устойчиво лучшие результаты на всем рассматриваемом периоде времени. Тем не менее, методы третьей группы (с корректировкой на наличие структурных сдвигов) позволяют улучшить прогнозы на всем имеющемся интервале в 17 случаях из 30, методы второй и третьей группы – в 21 случае. Здесь, на наш взгляд, надо отметить, что базовый прогноз (прогноз ИЭП) оказался лучшим на всем интервале для 5 показателей. В отличие от коротких интервалов на длинном интервале прогнозы по модели линейного тренда не позволяют столь часто улучшать прогноз: эти прогнозы оказались лучшими только в двух случаях.

Таким образом, общий вывод заключается в том, что методы учета наличия структурных сдвигов в данных позволяют улучшать прогнозы в режиме «реального времени» более чем в половине случаев как для коротких интервалов (в окрестностях структурного сдвига), так и на длинных интервалах.