

Различные индексы прогнозирования экономической активности в России*

Олег Демидов[†]

Droege & Comp., Москва, Россия

Рассматриваются различные способы вычисления индексов прогнозирования экономической активности в России. Первый способ – это методика, используемая российским Центром Развития и основанная на концепции «циклов роста». Второй подход объединяет в себе динамический метод главных компонент и динамический факторный анализ. Третий подход – это методика американского Национального Бюро Экономических Исследований, состоящая в построении диффузионных индексов с использованием динамической факторной модели. Данная работа является попыткой раскрыть преимущества и недостатки этих трех методов в применении к российским данным и предложить наилучшую методику для прогнозирования экономической активности в России.

Ключевые слова: индекс экономической активности, опережающие и текущие индикаторы, метод динамических главных компонент, факторная модель, Россия

Классификация JEL: C32, E37, O11

1 Введение

Система текущих и опережающих индикаторов – один из самых широко используемых методов прогнозирования будущей экономической активности. Этот метод впервые был применен в Соединенных Штатах в 1930-х годах. Регулярные публикации по используемым индикаторам были начаты в конце 1960-х и продолжаются по сей день (см. Stock & Watson, 1989, 2002). В 1980-х в OECD Statistics Directorate начали строить опережающие индикаторы для стран Евросоюза (см. Forni, Hallin, Lippi & Reichlin, 2001). Что касается России, то такие индикаторы строятся в Центре Развития с использованием своей собственной методологии (см. Smirnov, 2000, 2006). Однако методология Центра Развития имеет несколько недостатков. Данная статья является попыткой приложить европейскую процедуру и процедуру Стока и Уотсона к российским данным, а также улучшить некоторые аспекты данных методологий.

Обычно текущие и опережающие индикаторы экономической активности – это взвешенные средние определенных временных рядов, и необходим критерий того, как взвешивать различные компоненты индексов, чтобы учесть временные сдвиги между этими компонентами. Это порождает методологические проблемы: какой статистический критерий надо выбрать для построения этих средних, и какой метод необходим, чтобы реализовать оценивание для больших панелей временных рядов?

Метод, предложенный Центром Развития, состоит из следующих шагов. Во-первых, проводится датирование поворотных точек бизнес-циклов. Во-вторых, выбираются начальные опережающие индикаторы согласно тому критерию, что их поворотные точки предшествуют поворотным точкам бизнес-циклов. В-третьих, единый опережающий индикатор вычисляется как взвешенное среднее всех выбранных опережающих индикаторов с весами, обратно пропорциональными волатильностям этих индикаторов.

*Цитировать как: Демидов, Олег (2008). «Различные индексы прогнозирования экономической активности в России», Квантиль, №5, стр. 83–102. Citation: Demidov, Oleg (2008). “Different indexes for forecasting economic activity in Russia,” *Quantile*, No.5, pp. 83–102.

[†]Адрес: 115054, Россия, г. Москва, ул. Бахрушина, 32/1. Электронная почта: demidovs@inbox.ru

У этой процедуры есть несколько очевидных недостатков. Во-первых, требуется датирование поворотных точек всех переменных, что порой вызывает достаточно большие трудности из-за высокой волатильности макроэкономических переменных. Во-вторых, процедура отбора имеет неопределенность, поскольку некоторые переменные могут быть как опережающими, так и запаздывающими в различных точках. Наконец, простое усреднение по опережающим индикаторам не принимает во внимание временные сдвиги и взаимную корреляцию между переменными.

В то же время метод, предложенный в Forni, Hallin, Lippi & Reichlin (2001), позволяет избежать некоторые из этих недостатков. Европроцедура использует основную идею динамической факторной модели: благодаря сильным совместным движениям макроэкономических временных рядов динамика каждой переменной может быть представлена в виде суммы малоразмерной компоненты, общей для всех переменных, и ортогонального идиосинкретического остатка (см. Sargent & Sims, 1977). Факторные модели дают статистическую основу для построения индексов из большого количества временных рядов. Модель, использованная в работе Forni, Hallin, Lippi & Reichlin (2001), носит название обобщенной динамической факторной модели. Она объединяет в себе динамический факторный анализ и метод динамических главных компонент. Так что оценщик строится таким образом, чтобы принять во внимание временные сдвиги между переменными, правильно взвешивая опережающие и запаздывающие переменные. Это обстоятельство позволяет оценивать модель, используя все переменные – опережающие, текущие и запаздывающие – без необходимости классифицировать их априори.

Однако метод, предложенный в Forni, Hallin, Lippi & Reichlin (2001), имеет ряд недостатков, которые будут обсуждаться позже, и будут предложены возможные усовершенствования этого метода. Также необходимо принять во внимание то, что построение индексов для России имеет ряд особенностей по сравнению с построением индексов для Еврозоны, таких как отсутствие достаточного объема качественных данных и специфика российских бизнес-циклов.

Более того, вероятно, европроцедура не является оптимальной с точки зрения прогнозирования, так как требует использования приближенной динамической факторной модели. Поэтому также необходимо изучить предсказательные свойства методологии Stock & Watson (2002), которая использует точную статическую модель для прогнозирования. Все методы прогнозирования сравниваются с помощью среднеквадратичной ошибки предсказания (СКОП) вневыборочного прогноза для нахождения наилучшего способа прогнозирования будущего ВВП.

Работа организована следующим образом. Раздел 2 обсуждает теоретические основы циклических индикаторов Центра Развития, опережающих индикаторов европроцедуры и диффузионных индексов Стока и Уотсона. Процедура оценивания и результаты для России представлены в разделе 3. Завершает работу Заключение.

2 Теоретическая часть

2.1 Система циклических индикаторов Центра Развития

Как уже было сказано, методология Центра Развития включает следующие шаги. Сначала поворотные точки бизнес-циклов датируются по поворотным точкам индекса промышленного производства с удаленной сезонностью. Затем рассматриваются поворотные точки 40 индикаторов, чтобы решить, являются ли эти индикаторы опережающими. После тестирования этих переменных 7 из них включены в систему опережающих индикаторов для России (детали см. в Smirnov, 2006):

- Средняя месячная цена нефти марки Urals;

- Доля предприятий с увеличившимся или неизменным спросом;
- Доля предприятий, не имеющих избыточных запасов конечной продукции;
- Денежный агрегат М2, включая рубли и иностранную валюту;
- Обратный индекс реального эффективного обменного курса рубля, вычисленный ЦБ РФ, против корзины валют российских торговых партнеров;
- Индекс РТС;
- Межбанковская процентная ставка.

Наконец, единый опережающий индекс вычислен по следующей процедуре:

1. Для всех подобранных опережающих индикаторов X^i (i – номер ряда) вычисляются симметричные ежемесячные приросты (t – текущий момент времени):

$$x_t^i = 200 \frac{X_t^i - X_{t-1}^i}{X_t^i + X_{t-1}^i}. \quad (1)$$

2. Оцениваются средние значения x_{av}^i и стандартные отклонения s^i полученных приростных рядов:

$$x_{av}^i = \frac{\sum_{t=1}^T x_t^i}{T}, \quad s^i = \sqrt{\frac{\sum_{t=1}^T (x_t^i - x_{av}^i)^2}{T-1}}, \quad (2)$$

где T – число периодов наблюдений.

3. Для каждого t рассчитываются усредненный прирост g_t , а также среднее g_{av} и стандартное отклонение s^g ряда g_t :

$$g_t = \frac{\sum_{i=1}^n (x_t^i / s^i)}{n}, \quad g_{av} = \frac{\sum_{t=1}^T g_t}{T}, \quad s^g = \sqrt{\frac{\sum_{t=1}^T (g_t - g_{av})^2}{T-1}}, \quad (3)$$

где n – число исходных индикаторов.

4. Шаги 1–2 выполняются для индексов промышленного производства y . Результат – среднее y_{av} и стандартное отклонение s^y приростного ряда.
5. Показатель g_t корректируется так, чтобы его волатильность была равна волатильности приростов индекса промышленного производства:

$$G_t = g_t \frac{s^y}{s^g}. \quad (4)$$

6. По рекурсивной формуле рассчитываются значения сводного опережающего индекса Z_t путем возврата от приростов к агрегату:

$$Z_1 = \frac{200 + G_1}{200 - G_1}, \quad Z_t = Z_{t-1} \frac{200 + G_t}{200 - G_t}. \quad (5)$$

7. Полученный индекс Z приводится к той же базе, которую имеет индекс промышленного производства (1996 = 100). Для этого все значения Z_t делятся на среднемесячный уровень 1996 г. и умножаются на 100. В итоге получается, что сводный опережающий индекс имеет ту же базу, что и индекс промышленного производства, а симметричные приросты двух этих показателей имеют одинаковую волатильность.

2.2 Построение опережающего и текущего индикаторов по методике Forni, Hallin, Lippi & Reichlin (2001)

В работе Forni, Hallin, Lippi & Reichlin (2001) было предложено оценивать обобщенную динамическую факторную модель по панели, содержащей макроэкономические переменные всех стран Еврозоны, и извлекать ненаблюдаемую общую компоненту из каждого ряда панели. Для этого требуется «очищать» переменные как от ошибок измерения, так и от локальных, то есть присущих только этой переменной, компонент. Тогда текущий индикатор для Еврозоны определяется как взвешенное среднее общих («очищенных») компонент ВВП всех стран. Поэтому так построенный индекс со стохастической размерностью возможно больше единицы учитывает взаимную корреляцию внутри и между странами и суммирует ту часть динамики ВВП, которая наиболее коррелирована с остальной экономикой и поэтому наиболее полезна для оценивания бизнес цикла Еврозоны.

Информацию о ведуще-запаздывающих связях между переменными в панели и текущим индикатором можно затем восстановить по оценкам ковариаций между общими компонентами переменных и текущим индикатором. Эта информация позволяет определить опережающие переменные во всех экономиках Еврозоны, которые затем могут быть агрегированы в единый опережающий индикатор. Наконец, опережающий индикатор можно использовать для прогнозирования текущего индикатора.

Полностью теоретическую основу обобщенной динамической факторной модели можно найти в оригинальной статье Forni, Hallin, Lippi & Reichlin (2001). В данном разделе будут представлены только общее описание метода, его критика и возможные пути улучшения.

«Очищение» переменных с помощью метода динамических главных компонент

Главная идея здесь заключается в том, чтобы разделить каждую переменную на две ортогональные компоненты, первая из которых представляет часть индивидуальной динамики, сильно коррелированную с остальной панелью, а вторая – не интересующую нас идиосинкратическую часть.

Предположим, что все макроэкономические ряды, преобразованные должным образом, представляют собой реализацию стационарного в широком смысле n -мерного векторного процесса $x_t = (x_{1t}, \dots, x_{nt})'$ с нулевым средним. Необходимо выделить с помощью малого числа q «агрегированных индексов» часть динамики, общую для всех процессов x_{it} для $i = 1, \dots, n$:

$$z_{ht} = p_h(L)x_t, \quad h = 1, \dots, q,$$

где q процессов z_{ht} , $h = 1, \dots, q$, являются линейными комбинациями различных лагов переменных из x_t , L – оператор лагирования и $p_h(L) - 1 \times n$ вектор двустороннего линейного фильтра.

Наконец, рассмотрим декомпозицию

$$x_t = \gamma_t^q + \zeta_t^q = C^q(L)z_t^q + \zeta_t^q = K^q(L)x_t + \zeta_t^q, \quad (6)$$

где $\gamma_t^q = (\gamma_{1t}^q, \dots, \gamma_{nt}^q)$ – проекция x_t на настоящее, прошлое и будущее $z_t^q = (z_{1t} \dots z_{qt})'$, и ζ_t^q – остаточный вектор. Потребуем, чтобы фильтр $p_h(L)$ и соответствующие процессы z_{ht} , $h = 1, \dots, q$, удовлетворяли условию максимизации объясненной дисперсии для данного q

$$\sum_{j=1}^n \mathbb{V}[\gamma_{jt}^q]. \quad (7)$$

Процессы z_{1t}, \dots, z_{nt} называются «главными компонентами временного ряда» или «динамическими главными компонентами» процесса x_t . Эти компоненты так же связаны с собственными числами и векторами матрицы спектральной плотности процесса x_t , как статистические главные компоненты связаны с собственными числами и векторами ковариационной матрицы.

Точнее, пусть $P_x(k) = \mathbb{E}[x_t x'_{t-k}]$ – матрица корреляционной функции стационарного процесса x_t , и $\Sigma(\theta)$, $-\pi < \theta < \pi$, – матрица спектральной плотности x_t . Матрицы корреляционной функции и спектральной плотности связаны преобразованием Фурье, то есть

$$P_x(k) = \int_{-\infty}^{\infty} e^{i\theta k} \Sigma(\theta) d\theta, \quad \Sigma(\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\theta k} P_x(k) dk.$$

Пусть вектор $p_h(e^{-i\theta})$ будет собственным вектором, соответствующим h -му собственному числу матрицы $\Sigma(\theta)$, у которой собственные числа упорядочены по убыванию. Кроме того, обозначим это собственное число как $\lambda_h(\theta)$ и положим $\lambda_h = \frac{1}{2\pi} \int_{-\pi}^{\pi} \lambda_h(\theta) d\theta$. Тогда максимальная объясненная дисперсия (7) равна $\lambda_1 + \dots + \lambda_q$, и процент объясненной дисперсии дается отношением

$$\frac{\lambda_1 + \dots + \lambda_q}{\lambda_1 + \dots + \lambda_n}. \quad (8)$$

Далее, есть возможность выписать явные выражения для фильтров $C^q(L)$ и $K^q(L)$, заданных в (6). Эти выражения выглядят так:

$$\begin{aligned} C^q(L) &= (p_1(F)' \cdots p_q(F)'), \\ K^q(L) &= C^q(L) C^q(F)' = p_1(F)' p_1(L) + \cdots + p_q(F)' p_q(L), \end{aligned}$$

в тесной аналогии со статическим случаем. Поэтому фильтр $K^q(L)$ можно найти следующим образом. Во-первых,

$$K^q(e^{-i\theta}) = p_1(e^{-i\theta})' p_1(e^{-i\theta}) + \cdots + p_q(e^{-i\theta})' p_q(e^{-i\theta}).$$

Во-вторых, используя обратное преобразование Фурье, можно вычислить

$$K^q(k) = \int_{-\infty}^{\infty} e^{i\theta k} K^q(e^{-i\theta}) d\theta.$$

Наконец, получаем выражение для $K^q(L)$:

$$K^q(L) = \int_{-\infty}^{\infty} K^q(k) L^k dk.$$

Также необходимо найти матрицу спектральной плотности главных компонент $\Sigma^\gamma(\theta)$, которая задается выражением

$$\Sigma^\gamma(\theta) = K^q(e^{-i\theta}) \Sigma(\theta) K^q(e^{i\theta})'.$$

Следует отметить, что все фильтры и матрица спектральной плотности должны оцениваться из конечной выборки процесса x_t длины T . Более того, все непрерывные модели должны быть заменены на дискретные аналоги. Оценитель, используемый здесь (который обозначен как χ_{nt}^T в силу причин, которые будут ясны ниже), детально описан в Приложении.

Главные компоненты и обобщенная динамическая факторная модель

Процедура «очистения» основана на выборе малого числа q и поэтому может содержать существенную неопределенность. Однако, если предположить, что x_t генерируется факторной моделью, то можно построить достаточно четкий критерий выбора q . В динамическом факторном подходе переменные представляются в виде суммы двух ненаблюдаемых компонент: общей компоненты, формируемой малым числом факторов, присущих всем переменным системы, и идиосинкратической компоненты, которая некоррелирована с общими компонентами и присуща отдельной переменной. Если придерживаться этой модели, то отделение идиосинкратической части от общих компонент представляется достаточно естественной процедурой «очистения». Чтобы лучше понять данную факторную модель, будет удобно полагать, что вектор x_t формируется первыми n элементами бесконечной последовательности $x_{jt}, j = 1, \dots, \infty$. Чтобы подчеркнуть зависимость от n , будем писать x_{nt} вместо x_t . В данной модели

$$x_{nt} = \chi_{nt} + \xi_{nt} = B_n(L)u_t + \xi_{nt}, \quad (9)$$

где χ_{nt} – общая компонента, $u_t = (u_{1t}, \dots, u_{qt})'$ – вектор общих шоков, ковариационно стационарный q -векторный процесс с невырожденной матрицей спектральной плотности, $B_n(L)$ – матрица двустороннего квадратично суммируемого фильтра, и идиосинкратические компоненты ξ_{nt} ортогональны к u_{t-k} для каждого k .

Используя некоторые предположения, Forni, Hallin, Lippi & Reichlin (2001) показывают глубокую связь между представлениями (6) и (9). Основная интуиция здесь в следующем. При вычислении главных компонент мы берем среднее по переменным, входящим в x_t . Когда n большое, мы получаем некий результат типа закона больших чисел. Идиосинкратические компоненты исчезают, а остаются только линейные комбинации различных лагов главных компонент. Такие линейные комбинации охватывают почти то же динамическое пространство, что и общие факторы.

Данный результат дает простой критерий для выбора числа используемых главных компонент. Если модель (9) справедлива, то собственные числа $\lambda_{hn} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \lambda_{hn}(\theta) d\theta$ ограничены для $h > q$ и расходятся для $h \leq q$ при $n \rightarrow \infty$. Следовательно, при больших n можно ожидать большой «скачок» при переходе от λ_{qn} к $\lambda_{q+1,n}$. Поэтому главные компоненты стоит добавлять в процедуру оценивания до тех пор, пока увеличение в объясненной вариации не станет меньше некоторого выбранного значения. Точнее, если обозначить через λ_{hn}^T , где T – число наблюдений, оценку λ_{hn} , и выбрать $\alpha \in (0, 1)$, то критерий выбора $q = q^*$ будет следующим:

$$\frac{\lambda_{q^*n}^T}{\sum_{h=1}^n \lambda_{hn}^T} > \alpha, \quad \frac{\lambda_{q^*+1,n}^T}{\sum_{h=1}^n \lambda_{hn}^T} < \alpha. \quad (10)$$

Процедура оценивания

Теперь, после описания теоретической основы метода, можно представить полную процедуру для построения текущего и опережающего индикаторов и идентификации поворотных точек бизнес-цикла. В данном разделе обсуждается подход Forni, Hallin, Lippi & Reichlin (2001) и предлагаются возможные усовершенствования этого подхода.

Процедура оценивания состоит из следующих шагов.

Шаг 1: выбор переменных для включения в панель На первом шаге процедуры необходимо решить, какие переменные стоит включать в массив данных, по которому будут вычисляться общие компоненты. Как было показано, с теоретической точки зрения все

доступные переменные должны быть включены в процедуру оценивания. На практике, однако, не рекомендуется использовать переменные с большой идиосинкратической компонентой, поскольку она может быть неправильно интерпретирована как дополнительная общая компонента. Для выбора массива данных Forni, Hallin, Lippi & Reichlin (2001) предлагают следующую последовательность действий. Они выделяют 2 набора переменных: «ядро», содержащее переменные, которые точно надо использовать, и набор «кандидатов» на включение. Затем они преобразовывают переменные так, чтобы получить стационарность, и нормализуют их, вычитая среднее и деля на стандартное отклонение. Стационарность достигается взятием первой разности (или разности в логарифмах). Очевидно, однако, что в данных, по крайней мере российских, присутствуют большие сезонные компоненты, которые также должны быть исключены. В данной статье сезонность удаляется с помощью регрессии на сезонные фиктивные переменные, хотя есть и другие, возможно более лучшие, способы.

Используя ядро, Forni, Hallin, Lippi & Reichlin (2001) фиксируют q^* согласно критерию (10) и вычисляют соответствующую степень общности как процент объясненной дисперсии (8). Как только эта степень μ зафиксирована, они оценивают по очереди каждого кандидата. Для оценивания формируется расширенная система, состоящая из ядра и кандидата, для нее вычисляется значение (8) с $q = q^*$, и кандидат проходит экзамен, если значение (8) больше, чем μ . Успешные кандидаты не добавляются к ядру до тех пор, пока не будут оценены все кандидаты. Итоговый массив данных включает в себя ядро и всех успешных кандидатов.

На этом этапе также оцениваются вектор общих компонент χ_{nt} и матрица его спектральной плотности $\Sigma^{\chi}(\theta)$ с $q = q^*$, как это описано в Приложении.

Следует отметить, что процедура отбора переменных, предложенная в Forni, Hallin, Lippi & Reichlin (2001), имеет ряд недостатков. Во-первых, имеется неопределенность в выборе ядра и кандидатов. Одна из проблем состоит в том, что данные для некоторых переменных из ядра могут быть недостаточно хороши, что может привести к получению малого μ и включению переменных, слабо связанных с ВВП. Другая проблема возникает, если, наоборот, выбрано слишком хорошее ядро с высоким μ , тогда можно потерять переменные с немного меньшей связью с ВВП, которые имеют опережающую динамику и могут быть эффективно использованы при построении опережающего индекса. Во-вторых, предложенный критерий включения переменных в итоговое оценивание не принимает во внимание временные сдвиги и корреляции переменных относительно ВВП. Чтобы по возможности исправить эти недостатки, была разработана альтернативная процедура отбора, которая будет представлена в последнем параграфе этого раздела.

Шаг 2: текущий индикатор Текущий индикатор ΔC_t равен общей компоненте ВВП, умноженной на вычисленное стандартное отклонение ВВП. Пусть ВВП соответствует первой компоненте x_{nt} . Тогда текущий индикатор есть

$$\Delta C_t = \chi_{1nt} \sigma_1.$$

Агрегированный индикатор в уровнях вычисляется путем добавления дрейфа δ к текущему индикатору, как кумулятивная сумма $\Delta C_t + \delta$:

$$C_t = C_0 + \delta t + \sum_{\tau=1}^t \Delta C_{\tau}.$$

Шаг 3: процикличные и противоцикличные переменные Теперь можно классифицировать общие компоненты χ_{jt} как находящиеся в фазе или противофазе по отношению к текущему индикатору. Для этого с использованием оценки $\Sigma^{\chi}(\theta)$ вычисляется спектральная плотность каждой общей компоненты по отношению к ΔC_t . Затем вычисляется аргумент каждой из этих плотностей, который является фазовым сдвигом переменной относительно

российского ВВП. Обозначим фазовый сдвиг общей компоненты χ_{jt} как $\phi_j(\theta)$, $-\pi < \theta \leq \pi$. На нулевой частоте фазовый сдвиг может быть либо 0, либо π в зависимости от того, положительна или отрицательна долгосрочная корреляция. Пусть долгосрочная корреляция j -й компоненты с текущим индикатором равна ρ_j , тогда

$$\rho_j = \Sigma_{1j}^X(0).$$

Если она положительна, то переменная проциклична, иначе – противоциклична.

Долгосрочную корреляцию также можно использовать в процедуре отбора переменных как индикатор связи переменной с ВВП.

Шаг 4: опережающие, текущие и запаздывающие переменные Для такой классификации переменных необходимо вычислить фазовый сдвиг переменной χ_{jt} , $j = 1, \dots, n$, относительно ΔC_t на типичной частоте бизнес цикла $\theta^* > 0$. Обозначим этот фазовый сдвиг как $\psi_j(\theta^*)$. Тогда классифицируем χ_{jt} как текущую, если $|\psi_j(\theta^*)|$ меньше, чем заданное значение τ , опережающую, если $\psi_j(\theta^*) < -\tau$, и запаздывающую, если $\psi_j(\theta^*) > \tau$. Фазовый сдвиг $\psi_j(\theta^*)$ равен

$$\psi_j(\theta^*) = \text{Arg}(\Sigma_{1j}^X(\theta^*)),$$

где Arg означает аргумент комплексного числа. Он также может быть использован в процедуре отбора переменных. Например, опережающие переменные иногда можно включить в финальное оценивание, даже если они не очень сильно коррелированы с ВВП.

Шаг 5: опережающий и запаздывающий индикаторы Построение опережающего индикатора как обычного среднего ведущих переменных, предложенное Forni, Hallin, Lippi & Reichlin (2001), не учитывает фазовые сдвиги и корреляцию переменных относительно ВВП. Похоже, что корреляции переменных с ВВП могут быть частично учтены усреднением с весами, пропорциональными долгосрочным корреляциям ρ_j :

$$\Delta L_t = \frac{\sum_{j \in \ell} \rho_j \chi_{jt}}{\sum_{j \in \ell} |\rho_j|}.$$

Здесь ℓ обозначает набор j таких, что χ_{jt} являются опережающими.

Аналогичное выражение имеет место для запаздывающего индикатора за исключением того, что вместо опережающих переменных берутся запаздывающие.

Шаг 6: Добавление значений в крайних точках Следует отметить, что общие компоненты χ_{jnt}^T вычисляются с помощью двустороннего фильтра $K_n^T(L)$ длины $2M + 1$ (см. Приложение). Поэтому в начале и в конце периода наблюдений, то есть при $t = 1, \dots, M$ и $t = T - M + 1, \dots, T$, невозможно построить оценки общих компонент стандартным образом. По этой причине Forni, Hallin, Lippi & Reichlin (2001) заполняют оценки в крайних точках проекциями χ_{jnt}^T на текущие и опережающие либо запаздывающие переменные, соответственно.

Такие проекции χ_{jnt}^T не используют все возможности, которые дает процедура оценивания. Здесь можно использовать весь двусторонний фильтр и все переменные для заполнения значений в крайних точках.

Если быть более точным, то в начале интервала наблюдений при $t = 1, \dots, M$ необходимо сделать следующие шаги. Сначала находим x_0 как проекцию χ на момент 0, используя проекцию фильтра $K_n^T(L)$

$$x_0 = \chi_0^p = \sum_{k=-M}^{-1} K_k x_{0-k}$$

(здесь отброшены индексы T и n , индекс p обозначает проекцию). Затем вычисляем χ_M , используя полный фильтр $K_n^T(L)$:

$$\chi_M = \sum_{k=-M}^M K_k x_{M-k}.$$

Далее аналогичная процедура повторяется для $t = M - 1, \dots, 1$.

В конце периода наблюдений при $t = T - M + 1, \dots, T$ используется такой же метод добавления значений общих компонент, как и в начале периода.

Шаг 7: поворотные точки Поворотные точки наилучшим образом идентифицируются по текущему индикатору в уровнях без дрейфа. Поворотные точки определяются как даты t^* , в которые текущий индикатор C_t достигает локальных минимумов или максимумов.

Шаг 8: прогноз Проекция ΔC_{t+1} на опережающие и текущие переменные является прогнозом текущего индикатора на один шаг вперед. В работе Forni, Hallin, Lippi & Reichlin (2001) он вычисляется с использованием той же модели, которая применяется для добавления значений в крайних точках.

Критика здесь та же, что и для шагов 5 и 6. Поэтому для прогноза текущего индикатора на период вперед можно использовать метод из шага 6 с добавлением одной дополнительной точки в конце интервала наблюдений. Эту точку также можно рассматривать в качестве опережающего индикатора.

Альтернативная процедура отбора переменных

Во-первых, все правильно преобразованные переменные включаются в оценивание, и находятся соответствующие q^* , μ , ρ_j , $\psi_j(\theta^*)$.

Во-вторых, выбираются некие μ^* и ρ^* , и находятся переменные с $\rho_j < \rho^*$. Тогда эти переменные исключаются по очереди, и вычисляются новые μ_j . Каждый раз исключаем только одну переменную из начальной панели. После этого находим наименьший μ_j , и если $\mu - \mu_j > \mu^*$, то, возможно, есть необходимость исключить эту переменную. Единственное, что может спасти переменную от исключения в данной ситуации – это ее опережающая динамика относительно ВВП. Здесь следует принимать во внимание временной сдвиг переменной и экономическую интуицию относительно связи этой переменной с ВВП.

Наконец, после решения об исключении переменной берется новая панель и повторяется заново та же последовательность действий. Процедура заканчивается, когда больше не остается кандидатов на исключение.

2.3 Диффузионные индексы

Здесь приводится общее описание методологии Stock & Watson (2002).

Пусть y_t – скалярный временной ряд, а X_t – N -мерный векторный временной ряд. Переменная y_t является переменной для прогнозирования, в то время как X_t содержит полезную информацию для прогнозирования y_{t+1} . Предполагается, что X_t может быть представлен факторной структурой

$$X_t = \Lambda_t F_t + e_t, \tag{11}$$

где $F_t - r \times 1$ вектор общих факторов, $\Lambda_t - N \times r$ матрица коэффициентов при общих факторах и $e_t - N \times 1$ вектор идиосинкратических шумов. В общем случае идиосинкратические шумы коррелированы по времени и по рядам.

Главная цель – это оценивание $\mathbb{E}[y_{t+1}|X_t]$. Здесь y_{t+1} моделируется как

$$y_{t+1} = \beta_t' F_t + \epsilon_{t+1}, \quad (12)$$

где $\mathbb{E}[\epsilon_{t+1}|X_t, y_t, \beta_t, X_{t-1}, y_{t-1}, \beta_{t-1}, \dots] = 0$. Это выражение включает в себя три предположения: что $\mathbb{E}[y_{t+1}|X_t, y_t, \beta_t, X_{t-1}, y_{t-1}, \beta_{t-1}, \dots]$ зависит от X_t только через F_t и никак иначе, что лаги F_t не входят в (12), и что лаги y_t не входят в (12).

Матрица коэффициентов Λ_t ($N \times r$) и коэффициенты β_t ($r \times 1$) меняются со временем согласно AR(1)-процессам

$$\begin{aligned} \Lambda_t &= \Lambda_{t-1} + h\zeta_t, \\ \beta_t &= \beta_{t-1} + \eta_t, \end{aligned} \quad (13)$$

где h – диагональная $N \times N$ матрица, а η_t и ζ_t – соответственно $r \times 1$ и $N \times r$ стохастические шумы.

Следуя Стоку и Уотсону, динамическая факторная модель оценивается в ее статической форме. Здесь используется подход квазиправдоподобия, в том смысле, что оценивание мотивировано при сильных параметрических предположениях, но состоятельность оцениваемых факторов показана при более слабых непараметрических предположениях (детали см. в Stock & Watson, 2002). Для определения стратегии оценивания предположим, что $h = 0$, то есть $\Lambda_t = \Lambda_0 = \Lambda$, и e_{it} – IID $N(0, \sigma_e^2)$ и независимы между рядами. Также отклонимся от рассмотрения F_t в рамках динамической факторной модели, в которой F_t подчиняется стохастическому процессу. Вместо этого будет рассматривать F_t как оцениваемую $r \times T$ матрицу неизвестных неслучайных параметров. В таких предположениях метод максимального правдоподобия находит (F, Λ) как решение задачи нелинейных наименьших квадратов с целевой функцией

$$V_{NT}(F, \Lambda) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (X_{it} - \Lambda_i' F_t)^2. \quad (14)$$

Пусть (F^*, Λ^*) минимизирует значение $V_{NT}(F, \Lambda)$. Тогда удовлетворяются условия первого порядка

$$\Lambda_i^* = \left(\sum_{t=1}^T F_t^* F_t^{*'} \right)^{-1} \sum_{t=1}^T F_t^* X_{it}, \quad (15)$$

$$F_t^* = \left(\sum_{i=1}^N \Lambda_i^* \Lambda_i^{*'} \right)^{-1} \sum_{i=1}^N \Lambda_i^* X_{it}. \quad (16)$$

Неизвестные параметры можно оценить с помощью решения двух задач на собственные числа. Первая задача ставится заменой (15) в (14) и решается выбором в качестве F собственных векторов, соответствующих k наибольшим собственным числам $T \times T$ матрицы $\frac{1}{N} \sum_{i=1}^N X_i X_i'$, где $X_i = (X_{i1}, X_{i2}, \dots, X_{iT})'$.

Вторая задача на собственные числа ставится заменой (16) в (14) и решается выбором в качестве Λ собственных векторов, соответствующих k наибольшим собственным числам $N \times N$ матрицы $\frac{1}{T} \sum_{t=1}^T X_t X_t'$. Эти собственные вектора являются первыми k главными компонентами X_t .

Однако, есть неопределенность в выборе числа оцениваемых факторов k . В идеале это число должно быть равно реальному числу факторов r . Следуя Стоку и Уотсону, будем рассматривать задачу оценивания числа факторов, входящих в уравнение для прогноза, используя информационный критерий в форме

$$BIC_k = \ln \frac{SSR(k)}{T} + \frac{\ln T}{T} k. \quad (17)$$

Здесь $SSR(k)$ – сумма квадратов остатков оценивания (12) с помощью МНК, используя k оцениваемых факторов. Информационный критерий оценивает r , решая задачу $\min_k BIC_k$.

Помимо байесовского информационного критерия, который может давать смещенную оценку r , возможно также использование критерия из Bai & Ng (2002).

3 Процедуры оценивания и результаты

Данные

Изначальная панель данных, использованная в данной работе, состояла из 24 временных рядов для России. Среди них были переменные из реального сектора, рынка труда и рынка капитала, ценовые и финансовые индексы, денежные агрегаты, обменные курсы и процентные ставки. Использованы два основных источника данных: база Международного Валютного Фонда (МВФ) и база данных Сергея Смирнова из Центра Развития. Выбран временной интервал квартальных данных базы МВФ, начиная с 1-го квартала 1995 г. и заканчивая 1-м кварталом 2007 г. (в общей сложности 49 точек). Месячные данные из базы Сергея Смирнова были трансформированы в квартальные данные на том же интервале времени. Более подробное описание переменных, источников и трансформации данных приведено в таблице 1.

Таблица 1: Описание переменных, источников и трансформации данных

Переменная	Пояснение	Источник	Трансформация
GDP	ВВП	IMF	Разность в логарифмах
Inv	Инвестиции	IMF	Разность в логарифмах
GovCons	Госрасходы	IMF	Разность в логарифмах
PrivCons	Частное потребление	IMF	Разность в логарифмах
IndEmploy	Занятость	IMF	Разность в логарифмах
Wages	Изменения зарплаты (проценты)	IMF	Нет
CPI	Потребительские цены (2000=100)	IMF	Разность в логарифмах
PPI	Изменение в PPI (проценты)	IMF	Нет
M+QM	Деньги + квази-деньги	IMF	Разность в логарифмах
Madj	Деньги без сезонности	IMF	Разность в логарифмах
CapAccounts	Баланс капитала	IMF	Разность в логарифмах
BankReserves	Банковские резервы	IMF	Разность в логарифмах
REER	Реальный эффективный обменный курс	IMF	Разность в логарифмах
NEER	Номинальный эффективный обменный курс	IMF	Разность в логарифмах
MMrate	Межбанковская процентная ставка	IMF	Разность
UralsPrice	Цена на нефть Urals	Smirnov	Разность в логарифмах
ShareNoExc	Доля предприятий без избыточных запасов готовой продукции	Smirnov	Разность в логарифмах
RTS	Индекс RTS	Smirnov	Разность в логарифмах
Agriculture	Сельское хозяйство Y-o-Y	Smirnov	Логарифм
Construction	Строительство Y-o-Y	Smirnov	Логарифм
TransTurnover	Объем грузоперевозок	Smirnov	Разность в логарифмах
Retail	Торговля Y-o-Y	Smirnov	Логарифм
Services	Услуги Y-o-Y	Smirnov	Логарифм
ShareIorS	Доля предприятий с увеличившимся или неизменным спросом	Smirnov	Разность в логарифмах

Замечания: Y-o-Y означает относительное изменение величины по сравнению с прошлым годом.

Циклические индикаторы для России

Циклический индикатор строится на основе метода Центра Развития как взвешенное среднее семи временных рядов с опережающей динамикой (см. подраздел 2.1). Рис. 1 показывает в сравнении этот индикатор и российский ВВП в логарифмах. Можно увидеть, что циклический индикатор не является хорошим предсказателем будущего ВВП, однако в целом тренды достаточно схожи.

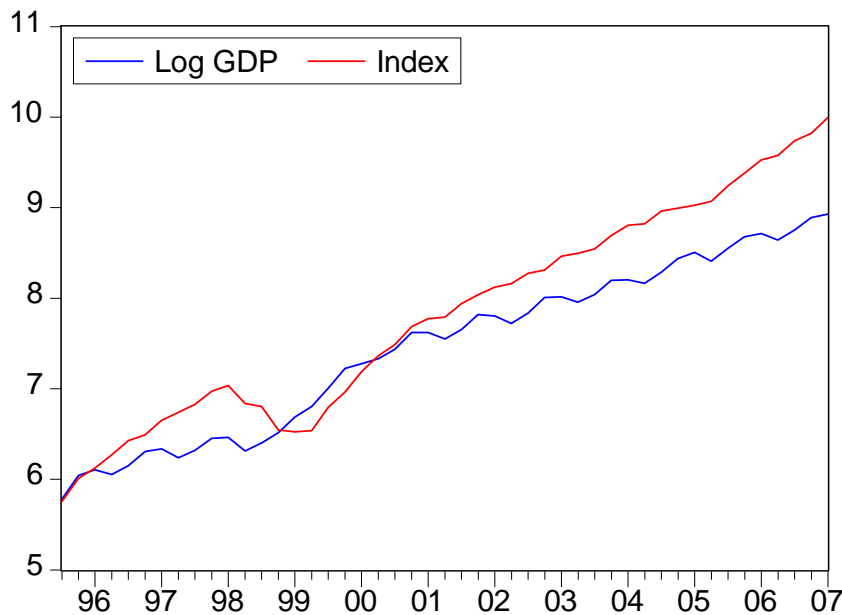


Рис. 1: ВВП в логарифмах и циклический индикатор Центра Развития

Среднеквадратичное отклонение прогноза (СКОП) процентного изменения ВВП для циклического индикатора составляет 0,00888, что соответствует стандартному отклонению прогноза около 0,0942. Такое стандартное отклонение выглядит достаточно большим, поскольку оно предполагает среднюю ошибку прогноза более 9 процентов квартального изменения ВВП. Поэтому необходимо искать другие пути прогнозирования будущей экономической активности для России.

Авторегрессия

Простейший путь прогнозирования любого временного ряда – это использование одноперенного авторегрессионного прогноза, основанного на модели

$$\Delta \ln y_{t+1} = \mu + \sum_{j=0}^p \gamma_j \Delta \ln y_{t-j} + \epsilon_t.$$

Здесь $\Delta \ln y_t$ – разность в логарифмах или процентное изменение ВВП в момент t , μ и γ_j – коэффициенты авторегрессии, ϵ_t – инновация, а p выбирается, исходя из минимизации СКОП.

Вычисления проводились для двух типов преобразования данных – данных с удаленной сезонностью и данных, в которых удалены сезонность и тренд. Сравнение этих двух типов мотивировано наличием структурного сдвига в российских данных в 3-м квартале 1998 г. и, следовательно, двух различных трендов в логарифме ВВП. Поэтому может оказаться неправильным вычитание одинакового тренда из обеих частей наблюдений.

Также вычисления проводились для двух различных интервалов времени – полного интервала (1995Q1:2007Q1 – 49 наблюдений) и укороченного (1998Q3:2007Q1 – 35 наблюдений). В первом интервале больше наблюдений, однако внутри него есть структурный сдвиг, второй – короче, но, возможно, имеет более лучшие структурные характеристики на всем периоде.

Для всех случаев был вычислен СКОП, и лучшим по этому показателю стал прогноз по данным с удаленной сезонностью на полном интервале времени с 7-ю лагами. С точки зрения СКОП, этот прогноз превзошел почти в 10 раз прогноз, основанный на методологии Центра Развития.

Опережающий и текущий индикаторы

Теперь оценим текущий, опережающий и запаздывающий индикаторы для России, следуя процедуре, описанной в подразделе 2.2. Сначала используем полный интервал наблюдений, начиная с 1-го квартала 1995 г. и заканчивая 1-м кварталом 2007 г. (всего 49 точек).

Используем альтернативную процедуру отбора переменных из раздела 2.2 для получения переменных, которые войдут в финальное оценивание. Выбираем $\alpha = 0,2$ и находим $q = 2$, т.е. используем два общих фактора. Для всех 24 переменных степень общности равна 0,71. Выбираем параметры альтернативной процедуры следующим образом: $\rho^* = 0,05$, $\mu^* = 0,01$. Четыре переменные были кандидатами на исключение из-за малой корреляции с общей компонентой ВВП:

- Реальный эффективный обменный курс,
- Межбанковская процентная ставка,
- Доля предприятий с увеличившимся или неизменным спросом,
- Доля предприятий, не имеющих избыточных запасов конечной продукции.

Среди этих переменных только реальный эффективный обменный курс не был исключен, во-первых, потому что степень общности увеличивалась менее, чем на μ^* , при исключении REER, и во-вторых, потому что REER имеет опережающую динамику по отношению к ВВП. Все три исключенные переменные входят в метод построения циклического индикатора Центра Развития; возможно, поэтому такой индикатор не дает удовлетворительного прогноза ВВП.

В итоге, в окончательный массив данных вошла 21 переменная, степень общности составила 0,78.

После построения массива данных необходимо найти параметры процедуры оценивания, оптимальные для прогнозирования ВВП. Делается следующая последовательность шагов. Строится вневыборочный прогноз на один период вперед, основанный на двустороннем фильтре, и находится его СКОП. Сравниваются СКОП прогнозов с различными значениями параметров M и q (здесь ширина фильтра M меняется от 1 до 10, а число общих факторов q – от 1 до 7). Наконец, выбирается прогноз с наименьшим СКОП. Эта процедура проводится для тех же четырех случаев трансформации данных, что рассматривались в пункте с авторегрессионным прогнозом.

Лучшим прогнозом с точки зрения СКОП является прогноз для данных с удаленными сезонностью и трендом на полном интервале времени с одним общим фактором и шириной фильтра, равной восьми. Это говорит о том, что один общий фактор дает наилучшую аппроксимацию компоненты ВВП, и в данных присутствует сильная зависимость от прошлого. СКОП у этого прогноза немного меньше, чем у авторегрессионного прогноза, однако использование данного метода не дает существенных улучшений по сравнению с обычной авторегрессией.

Также следует отметить, что европроцедура работает гораздо лучше для данных с удаленным трендом, чем без удаления тренда. Возможно, это объясняется тем, что корреляционная теория очень строго требует стационарности используемых временных рядов.

Таблица 2: Результаты оценивания по методике Forni, Hallin, Lippi & Reichlin (2001)

Переменная	Отношение дисперсий	Корреляция с ВВП	Временной сдвиг в месяцах
GDP	0,979	1,000	0,00
Inv	0,751	0,157	2,23
GovCons	0,733	0,143	2,34
PrivCons	0,959	0,312	-1,50
IndEmploy	0,295	-0,246	-1,60
Wages	0,937	0,413	2,07
CPI	1,010	0,437	-2,35
PPI	0,978	0,396	0,73
M+QM	0,845	0,220	0,26
Madj	0,799	0,230	1,78
CapAccounts	0,606	0,138	9,52
BankReserves	0,677	0,144	1,67
REER	0,732	-0,034	17,4
NEER	0,862	-0,187	7,27
UralsPrice	0,540	-0,059	18,6
RTS	0,560	-0,092	16,5
Agriculture	0,689	0,422	2,82
Construction	0,652	0,605	-0,98
TransTurnover	0,726	0,139	-6,51
Retail	0,906	0,640	0,46
Services	0,722	0,563	-1,10

Замечания: Объяснения переменных см. в таблице 1.

Таблица 2 представляет результаты процедуры оценивания с оптимальными параметрами для финального массива данных. Она показывает отношение между дисперсией общей компоненты и дисперсией самой переменной, корреляцию переменной с ВВП и временной сдвиг каждой переменной по отношению к ВВП на частоте $\theta^* = \pi/16$, соответствующей бизнес-циклу длиной 8 лет. Опережающие переменные определяются как те, у которых временное опережение больше месяца (0,33 квартала), что соответствует фазовому опережению $\tau = \pi/48$; запаздывающие переменные определяются как отстающие более чем на один месяц; остальные переменные являются текущими. Сдвиги представлены в виде временных сдвигов, числа даны в месяцах, положительные числа означают, что переменная опережает ВВП.

Эта таблица показывает, что переменные рынка труда, цены, денежные агрегаты и номинальный обменный курс имеет высокую степень общности с реальным сектором экономики России. В то же время реальный обменный курс, цена на нефть Urals, индекс RTS и процентные ставки имеют более слабую корреляцию с реальным сектором вопреки обычным представлениям.

Все переменные, кроме занятости, обменных курсов, цены на нефть и индекса RTS, цикличны. Имеется 11 опережающих переменных, среди которых самые очевидные – это инвестиции, госрасходы, обменные курсы, цена на нефть и индекс RTS. Также присутствуют 5 запаздывающих переменных, среди которых частное потребление, занятость и объем грузоперевозок. В целом, большая часть результатов интуитивно достаточно понятна.

Рис. 2 показывает результирующие текущий, опережающий и запаздывающий индикаторы для России по сравнению с логарифмом ВВП. С одной стороны, можно увидеть, что

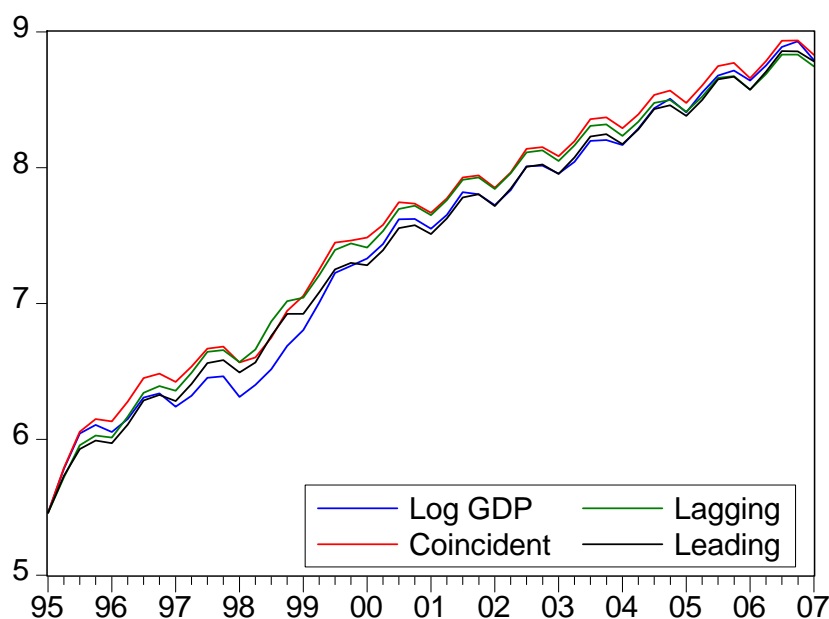


Рис. 2: ВВП в логарифмах и текущий, опережающий и запаздывающий индикаторы европроцедуры.

все индикаторы гораздо лучше коррелируют с ВВП, чем циклический индикатор Центра Развития. Поэтому прогнозные характеристики европроцедуры гораздо лучше. Однако, с другой стороны, из Рис. 2 видно, что текущий индикатор далеко не полностью повторяет логарифм ВВП. Когда мы делаем прогноз с помощью двустороннего фильтра, то прогнозируется текущий индикатор, а не сам логарифм ВВП. Поэтому точность прогнозирования данным методом не такая высокая, как если бы мы строили прогноз самого ВВП.

Таким образом, методология Forni, Hallin, Lippi & Reichlin (2001) дает очень полезную информацию о связи различных переменных с ВВП. Она также дает неплохой метод прогнозирования ВВП, который, однако, имеет определенный недостаток, описанный выше. Достаточно естественная идея, которая здесь возникает, – это соединить два метода: европроцедуру и регрессионный анализ. Это означает, что регрессия ВВП на общие факторы и дальнейшее прогнозирование на основе этой регрессии могут дать существенные улучшения точностных характеристик прогноза ВВП. Данный подход используется Стоком и Уотсоном в построении их диффузионных индексов.

Методология Стока и Уотсона

Будем рассматривать прогноз диффузионного индекса в виде

$$\Delta \ln y_{t+1} = \mu + \sum_{i=1}^q \beta_i F_{i,t} + \sum_{j=0}^p \gamma_j \Delta \ln y_{t-j} + \epsilon_t,$$

где $F_{i,t}$ – оцениваемые факторы в момент времени t , q – число факторов, p – число лагов авторегрессионной компоненты, и q и p выбираются из условия минимизации СКОП.

При поиске наилучшего прогноза q пробегает значения от 1 до 7, p – от 0 до 8. Здесь также производится оценивание для всех четырех типов трансформации данных. Наилучший прогноз получается с одним общим фактором и семью лагами авторегрессионной компоненты для данных с удаленными сезонностью и трендом на полном временном интервале. Этот прогноз ощутимо превосходит прогноз, основанный на европроцедуре, поскольку в данном случае мы прогнозируем сам логарифм ВВП, а не его общую компоненту.

Объединенная методология

Вполне возможно, что лучшая точность предсказания процедуры Стока и Уотсона может быть достигнута, если учитывать фазовый сдвиг переменных относительно ВВП. Действительно, запаздывающие переменные могут приносить с собой запаздывающую динамику в общие факторы, которые используются как регрессоры для будущего ВВП. Однако очевидно, что качество регрессоров тем лучше, чем больше в них опережающей динамики. Поэтому возникает идея, что исключение запаздывающих переменных из массива данных может привести к более качественному прогнозу.

Здесь рассмотрены два случая. В первом случае в процедуру оценивания включены только опережающие переменные, во втором – к опережающим переменным добавлены еще и текущие. В обоих случаях были рассмотрены все типы трансформации данных. Типы переменных для полного интервала времени определялись из таблицы 2, для укороченного интервала – из европроцедуры с оптимальными параметрами на укороченном интервале. Стоит отметить, что почти все переменные определялись одинаково на обоих интервалах времени.

Использование текущих переменных мотивировано тем, что они также могут нести часть ведущей динамики. Однако результаты оценивания показали, что лучше использовать только опережающие переменные. Наилучший прогноз получается для данных с удаленной сезонностью на полном интервале с тремя общими факторами и шестью лагами в авторегрессионной части. СКОП этого прогноза почти в 20 раз лучше СКОП, полученного по методологии Центра Развития.

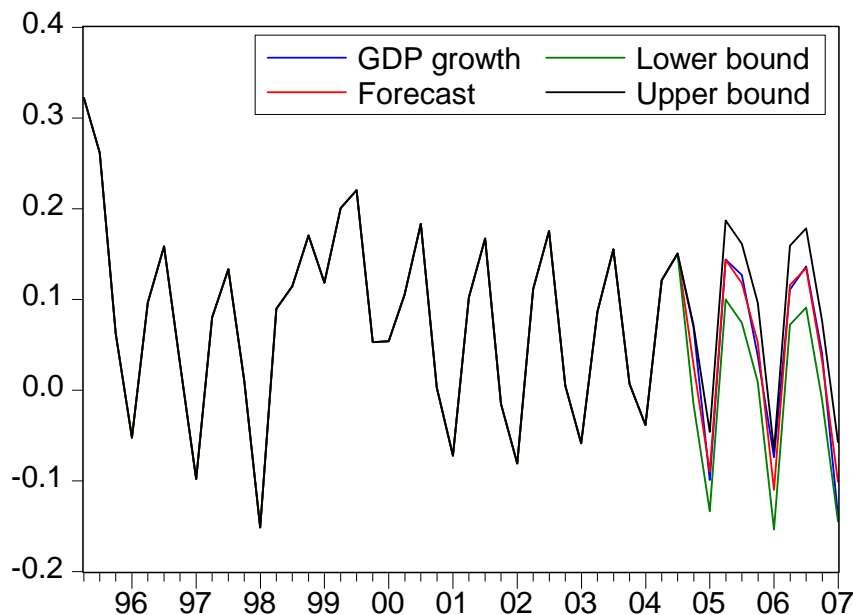


Рис. 3: Разность в логарифмах ВВП, прогноз, нижняя и верхняя границы 95%-ного доверительного интервала объединенной процедуры.

Рис. 3 показывает в сравнении наилучший прогноз и процентное изменение ВВП. Также показаны верхняя и нижняя границы 95%-ного доверительного интервала, и можно видеть, что логарифм ВВП остается внутри этого интервала в течение всего времени.

Результаты

Результаты для всех процедур прогнозирования представлены в таблице 3. В ней показаны отношения СКОП каждого прогноза к СКОП прогноза по методу Центра Развития. Это

означает, что относительное СКОП прогноза Смирнова равно 1, и если СКОП прогноза меньше 1, то это прогноз превосходит прогноз по методу Центра Развития.

Таблица 3: Результаты расчетов по прогнозированию

СКОП	SA DM	SA DR	SA DM S	SA DR S
AR	0,1665 (10)	0,1047 (7)	0,2170 (4)	0,2101 (6)
Euro	0,1012 (1,8)	0,7068 (1,8)	0,1560 (1,8)	0,7867 (1,4)
Stock & Watson Full	0,0823 (1,7)	0,0834 (2,7)	0,1490 (2,5)	0,1479 (1,4)
Stock & Watson Leading	0,0578 (3,6)	0,0557 (3,6)	0,0894 (5,5)	0,0906 (5,5)
Stock & Watson Lead&Coin	0,0701 (2,7)	0,0704 (4,6)	0,0927 (6,5)	0,0900 (6,5)

Замечания: Результаты представлены для четырех типов трансформации данных: с удаленными сезонностью и трендом для полного периода наблюдений (1995Q1:2007Q1), с удаленной сезонностью для полного периода (1995Q1:2007Q1), с удаленными сезонностью и трендом для укороченного периода (1998Q3:2007Q1), с удаленной сезонностью для укороченного периода (1998Q3:2007Q1). Числа без скобок показывают отношение СКОП каждого прогноза к СКОП прогноза по методу Центра Развития. Для авторегрессионного прогноза (AR) число в скобках означает используемое число лагов ВВП. Для европроцедуры (Euro) числа в скобках соответствуют числу общих факторов и ширине двустороннего фильтра. Для прогноза по Стоку и Уотсону числа в скобках показывают выбранные числа общих факторов и лагов ВВП. Full обозначает то, что процедура сделана по всем переменным, Leading – только по опережающим, Lead&Coin – по опережающим и текущим.

Можно видеть, что для прогнозирования лучше использовать полный интервал наблюдений, чем укороченный. Видимо, это связано с тем, что проблема короткой выборки более серьезная, чем проблема, связанная со структурным сдвигом в 1998Q3.

Первый основной результат состоит в том, что установлен порядок среди существующих статистических методов с точки зрения точности прогноза: самый лучший – это Stock & Watson (2002), затем европроцедура и авторегрессия, имеющие почти одинаковые свойства, и, наконец, метод Центра Развития, обладающий наименьшей точностью.

Второй результат состоит в том, что использование метода, который объединяет методологию Стока и Уотсона с европроцедурой, может дать значительное улучшение точности прогноза. Этот объединенный метод дает прогноз со стандартным отклонением около 0,0222, что соответствует средней ошибке прогноза процентного изменения ВВП на следующий квартал порядка двух процентов.

4 Заключение

В данной статье рассмотрены различные типы индексов для прогнозирования экономической активности в России. В частности, обсуждены три методики: циклический индикатор Центра Развития, методология, предложенная в Forni, Hallin, Lippi & Reichlin (2001), и диффузионные индексы Стока и Уотсона (Stock & Watson, 2002).

Результаты, основанные на дальнейшем развитии процедуры Forni, Hallin, Lippi & Reichlin (2001), показали, что текущий индикатор российского бизнес-цикла сильно коррелирован с некоторыми переменными рынка труда, ценами, денежными агрегатами и номинальным обменным курсом, в то время как он слабее коррелирован с реальным обменным курсом, ценой на нефть Urals, индексом RTS и процентными ставками. Текущий индикатор достаточно хорошо прогнозируется взвешенным средним опережающих переменных, таких как инвестиции, госрасходы, обменные курсы, цена на нефть и индекс РТС.

Все методы прогнозирования будущего ВВП были сравнены, используя среднее квадратичное отклонение прогноза. Среди существующих методов наиболее точный прогноз дает метод Стока и Уотсона. Главное достижение данной работы – это построение методологии, которая

объединила в себе диффузионные индексы Стока и Уотсона с процедурой Forni, Hallin, Lippi & Reichlin (2001). Она дала наилучший прогноз с точки зрения СКОП среди всех опробованных методов.

Благодарности

Автор благодарит Олег Замулина, Кирилла Сосунова, Константина Стырина и Станислава Анатольева за полезные обсуждения и комментарии. Автор также благодарит Сергея Смирнова, Кирилла Сосунова и Олега Замулина за предоставленные данные.

Список литературы

- Bai, J. & S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.
- Brillinger, D.R. (1981). *Time Series Data Analysis and Theory*. New York: Holt, Rinehart and Winston.
- Forni, M., M. Hallin, M. Lippi & L. Reichlin (2001). Coincident and leading indicators for the Euro area. *Economic Journal* 111, C62–85.
- Forni, M., M. Hallin, M. Lippi & L. Reichlin (2000). The generalized factor model: Identification and estimation. *Review of Economic and Statistics* 82, 540–554.
- Forni, M. & M. Lippi (2001). The generalized dynamic factor model: Representation theory. *Econometric Theory* 17, 1113–1141.
- Geweke, J. (1977). The dynamic factor analysis of economic time series. Глава 19 в *Latent Variables in Socio-Economic Models* под редакцией D.J. Aigner & A.S. Golberger. Amsterdam: North-Holland.
- Sargent, T.J. & C.A. Sims (1977). Business cycle modeling without pretending to have too much a priori economic theory. Глава 19 в *New Methods in Business Research* под редакцией C.A. Sims. Minneapolis: Federal Reserve Bank of Minneapolis.
- Smirnov, S. (2000). A system of leading indicators for Russia. Working paper, Центр Развития (Development Center, <http://dcenter.ru>).
- Smirnov S. (2006). A new system of cyclical indicators for Russia. 28th CIRET Conference, Rome.
- Stock, J.H. & M.H. Watson (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomic Annual* 1989, 351–394.
- Stock, J.H. & M.H. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics* 20, 147–162.

Приложение

Данное приложение содержит некоторые исправления и дополнения к процедуре оценивания Forni, Hallin, Lippi & Reichlin (2001). Здесь показывается, как оцениваются общие компоненты χ_{nt} и матрица их спектральной плотности $\Sigma_n^{\chi^T}(\theta)$. Используется выражение (9), то есть опускается индекс q и принимается явная зависимость от n .

Процедура оценивания состоит из четырех шагов. Во-первых, вычисляется матрица спектральной плотности $\Sigma(\theta)$ вектора x_{nt} на нескольких частотах, используя окно Бартлетта размера $M = M(T)$. Точнее, вычисляется выборочная ковариационная матрица Γ_k^T векторов x_{nt} и $x_{n,t-k}$ для $k = 0, \dots, M$:

$$\Gamma_k^T = \frac{1}{T} \sum_{t=k+1}^T x_t x_{t-k}.$$

Затем вычисляются $2M + 1$ точек дискретного преобразования Фурье усеченной двусторонней последовательности $\Gamma_{-M}^T, \dots, \Gamma_M^T$, где $\Gamma_{-k} = \Gamma_k'$, то есть вычисляется

$$\Sigma_n^T(\theta_s) = \frac{1}{2\pi} \sum_{k=-M}^M \Gamma_k^T \omega_k e^{-ik\theta_s},$$

где

$$\theta_s = \pi s/S, \quad s = -S + 1, \dots, S, \quad \theta_s \in (-\pi, \pi],$$

и $\omega_k = 1 - |k|/(M + 1)$ – веса, соответствующие окну Бартлета размера M . Стоит отметить, что $\Sigma_n^T(\theta_s)$ – это комплексная эрмитова матрица по построению. Состоятельное оценивание $\Sigma(\theta_s)$ обеспечивается, когда $M(T) \rightarrow \infty$ и $M(T)/T \rightarrow 0$ при $T \rightarrow \infty$. Правило $M = \lfloor \sqrt{T}/4 \rfloor$ действует достаточно хорошо для большинства MA и AR моделей малого порядка в симуляциях (см. Forni, Hallin, Lippi & Reichlin, 2000).

Во-вторых, вычисляются собственные числа $\lambda_h(\theta_s)$ матрицы $\Sigma_n^T(\theta_s)$ в убывающем порядке для всех s , и формируются

$$\lambda_h = \frac{1}{2S} \sum_{s=-S+1}^S \lambda_h(\theta_s).$$

Затем используется критерий (10) с некоторым α для определения q , и фиксируется соответствующая степень общности (8). Используя эту степень, решается вопрос о том, какие переменные включать в панель для оценивания.

В-третьих, вычисляются первые q собственных векторов $p_{hn}^T(\theta_s)$, $h = 1, \dots, q$, матрицы $\Sigma_n^T(\theta_s)$ для $s = -S + 1, \dots, S$. В случае, когда $M = 0$, $p_{hn}^T(\theta_s)$ является просто h -м собственным вектором ковариационной матрицы x_{nt} : динамические главные компоненты редуцируются к статическим главным компонентам. С помощью собственных векторов вычисляются

$$\Phi_n^T(\theta_s) = \sum_{h=1}^q p_{hn}^T(\theta_s)' p_{hn}^T(\theta_s),$$

где $'$ обозначает сопряжение и транспонирование.

Наконец, оценивается $n \times n$ матрица фильтра $K_n(L)$ как обратное дискретное преобразование Фурье от

$$(\Phi_n^T(\theta_{-S+1}), \dots, \Phi_n^T(\theta_S)).$$

Точнее, вычисляются

$$K_{kn}^T = \frac{1}{2S} \sum_{s=-S+1}^S \Phi_n^T(\theta_s) e^{ik\theta_s}$$

для $k = -M, \dots, M$. Оценка фильтра задается выражением

$$K_n^T(L) = \sum_{k=-M}^M K_{kn}^T L^k.$$

Оценка общих компонент

$$\chi_{nt}^T = K_n^T(L) x_t = \sum_{k=-M}^M K_{kn}^T x_{t-k}.$$

Матрица спектральной плотности общих компонент может быть оценена как

$$\Sigma_n^{\chi^T}(\theta_s) = \Phi_n^T(\theta_s) \Sigma_n^T(\theta_s) \Phi_n^T(\theta_s).$$

Different indexes for forecasting economic activity in Russia

Oleg Demidov

Droege & Comp., Moscow, Russia

This paper considers different ways of computing indexes for forecasting economic activity in Russia. The first is the methodology used by the Russian Development Centre based on the concept of “growth cycles”. The second combines the dynamic principal components and dynamic factor analyses. The third approach is the NBER methodology based of diffusion indexes constructed using a dynamic factor model. This paper is an attempt to reveal strengths and weaknesses of the three methods in application to Russian data and to develop a better methodology for forecasting economic activity in Russia.

Keywords: index of economic activity, leading and coincident indicators, dynamic principal components, factor model, Russia

JEL Classification: C32, E37, O11