Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»

Е. Астафьева
М.Турунцева

«Предсказуемость предварительных оценок макроэкономических показателей: обзор подходов»

(препринт)

Москва 2020

Работа сделана в рамках научно-исследовательской работы, выполненной в соответствии с Государственным заданием РАНХиГС при Президенте Российской Федерации на 2020 год

Аннотация

В данной работе приводится обзор методов тестирования предсказуемости макроэкономических показателей на основе пересмотров данных в реальном времени. Рассматриваются различные подходы к их прогнозированию и обсуждаются их свойства.

Ключевые слова: пересмотр данных, временные ряды, прогнозирование, качество прогнозов.

JEL: C12, C22.

Принятие своевременных экономических и политических решений, планирование бюджета и бизнес-планирование всегда основывалось на анализе текущей ситуации и прогнозировании путей развития экономической активности и политической обстановки, базу которых составляют официальные оценки значений различных показателей. Поэтому пересмотры данных и прогнозы, основанные на более ранних/поздних оценках, могут повлечь за собой различные экономико-политические решения.

Тестирование предсказуемости может проводиться различными способами. Один из способов был описан в работе (Berger, A. N., & Krane, S. D., 1985). Предполагается, что исследователь может предсказать значения будущих пересмотров на основе уже произошедших, то есть предполагается автокоррелированность пересмотров. Рассматривается модель, которая записывается с помощью формулы (1):

$$\text{пересмотр}_t = \alpha + \beta * \text{пересмотр}_{t-1} + e_t, \quad (1)$$

где пересмотр_t – разница между изначальными данными и данными, полученными в результате последнего пересмотра.

Если коэффициент β статистически значим, то пересмотр считается не эффективным, так как его можно предсказать на основе прошлых значений. Включение константы α позволяет учесть смещенность изначальных данных, так как если α статистически значима, то среднее значение пересмотра отличается от 0.

Второй способ тестирования предсказуемости основан на использовании изначальных данных в качестве потенциальной объясняющей переменной для пересмотра. Рассматривается модель, которая записывается формулой (2):

$$\text{пересмотр}_t = \alpha + \beta * \text{изначальные данные}_t + e_t. \quad (2)$$

Согласно третьему подходу к учету предсказуемости, на будущие значения пересмотров может влиять состояние экономики. Этот способ позволяет проверить наличие различий в пересмотре данных в зависимости от того, в фазе рецессии или подъема находится экономика. Модель записывается с помощью формулы (3):

$$\text{пересмотр}_t = \alpha + \beta * \text{дамми на спад}_t + e_t, \quad (3)$$

где дамми на спад_t – дамми переменная, принимающая значения 1, если темп экономического роста снижается, и 0 – если повышается.

Авторы статьи (Strohsal, T., Wolf, E, 2019) ссылаются на подход к определению эффективности пересмотров, предложенный в работе (Aruoba, 2008). Согласно этому подходу, пересмотр считается эффективным, если он соответствует трем критериям:

– Несмещенность, $E(r_t) = 0$;

– Дисперсии, $V(r_t)$, должны быть маленькими;

– Непредсказуемость, $E(r_t|I_{initial}) = 0$, $I_{initial}$ – первоначальные значения показателя.

Для тестирования гипотезы непредсказуемости авторы используют ряд моделей. В основе многих из них лежит гипотеза о том, что ряд данных описывается AR(1) процессом с помощью формулы (4):

$$x_t = \beta x_{t-1} + \upsilon_t.$$

(4)

Статистические агентства наблюдают истинное значение только с ошибкой измерения η_t (формула (5)):

$$w_t = x_t + \eta_t.$$

(5)

Затем статистические агентства оценивают w_t и на основе полученной оценки публикуют значения показателей y_t.

В качестве наивного прогноза финального значения $\widehat{x_t}$ можно проигнорировать возможные пересмотры и предположить $\widehat{x_t} = y_t$.

Второй вариант основывается на фильтре Калмана. Согласно этому способу, значительное отличие y_t от предыдущих значений x_{t-1} может возникать либо в результате существенного шока υ_t в формуле (6), либо в результате значительной ошибки измерения η_t. Для учета этого исследователь обновляет свои ожидания $\widehat{x_T} = \hat{\beta} x_{T-1}$ до того момента, как появятся статистические данные. Тогда предлагается оценивать $\widehat{x_t}$ по формуле (6):

$$\widehat{x_T} = \hat{\beta} x_{T-1} + \hat{\gamma}_K (y_T - \hat{\beta} x_{T-1}), \tag{6}$$

где $\hat{\gamma}_K = \widehat{\sigma_\upsilon^2} / (\widehat{\sigma_\upsilon^2} + \widehat{\sigma_\eta^2})$.

Идея этого подхода заключается в том, что исследователи включают в прогноз как ожидания, сделанные на основе предыдущей информации, так и статистические данные, несвязанные с этими ожиданиями.

В работе (Howrey, 1978) предлагается моделировать разницу между реальными данными и тем, что публикуют статистические агентства (формула (7)):

$$y_t - x_t = h(y_{t-1} - x_{t-1}) + \varphi_t.$$

(7)

Тогда прогноз аналогичен прогнозу по модели Калмана (8):

$$\widehat{x_T} = \hat{\beta}x_{T-1} + \hat{\gamma}_H\left(y_T - \hat{\beta}x_{T-1} - h(y_{T-1} - x_{T-1})\right),$$

(8)

где $\hat{\gamma}_H = \widehat{\sigma_\upsilon^2}/(\widehat{\sigma_\upsilon^2} + \widehat{\sigma_\varphi^2})$.

В модели Сарджента (Sargent, 1987) предполагается, что статистические агентства сами фильтруют переменную шума w_t. Однако после фильтра все равно может возникнуть белый шум ζ_t, связанный с техническими ошибками или опечатками. В модели считается, что опубликованные данные моделируются с помощью формулы (9):

$$y_t = \hat{\beta}x_{t-1} + g\left(w_t - \hat{\beta}x_{t-1}\right) + \zeta_t,$$

(9)

А прогноз записывается с помощью формулы (10):

$$\widehat{x_T} = \hat{\beta}x_{T-1} + \hat{\gamma}_S(y_T - \hat{\beta}x_{T-1}),$$

(10)

где $\hat{\gamma}_S = (\widehat{\sigma_\upsilon^2} + \widehat{Cov}(\upsilon, y - x))/(\widehat{\sigma_\upsilon^2} + 2\widehat{Cov}(\upsilon, y - x) + \widehat{\sigma_{y-x}^2})$.

Разница между моделями (Howrey, 1978) и (Sargent, 1987) заключается в том, что в модели (Howrey, 1978) не предполагается корреляция между шоками υ_t и φ_t (формулы (6) и (9) соответственно), но предполагается автокорреляция в пересмотрах. В модели (Howrey, 1978), наоборот, пересмотры не автокоррелированы, но есть корреляция между пересмотрами и шоком φ_t, который влияет на истинное значение x_t.

Подход, описанный в работе (Kishor, Koenig, 2012) включает в себя модели (Howrey, 1978) и (Sargent, 1987). Процесс пересмотра описывается формулой (11):

$$y_t - x_t = h(y_{t-1} - x_{t-1}) + \varepsilon_t - (1 - g)\upsilon_t,$$

(11)

где $\varepsilon_t = \zeta_t + g\eta_t$.

Уравнение для прогноза описывается формулой (12):

$$\widehat{x_T} = \hat{\beta}x_{T-1} + \hat{\gamma}_K[y_T - \hat{k}(y_{T-1} - x_{T-1}) - \hat{\beta}x_{T-1}],$$

(12)

где $\hat{\gamma}_K = g\widehat{\sigma_\upsilon^2}/(g^2\widehat{\sigma_\upsilon^2} + \widehat{\sigma_\varepsilon^2})$.

При $g = 1$ модель превращается в модель (Howrey, 1978), а при $\hat{k} = 0$ – в модель (Sargent, 1987).

Наиболее полная эмпирическая проверка предсказуемости пересмотров представлена в работе (Clements, Galvão, 2018) на примере публикуемых оценок темпов роста ВВП США. Их анализ учитывает систему пересмотров значений показателя и включает рассмотрение предварительной оценки $y_t^{t+1/3}$ (публикуемой в среднем через 30 дней после окончания квартала), второй оценки $y_t^{t+2/3}$ (публикуемой в среднем через 60 дней после окончания наблюдательного квартала, обозначенная) и третьей оценки y_t^{t+1} (публикуемой в среднем на 30 дней позже второй оценки).

В работе изучается относительное качество прогнозов второй и третьей оценок ВВП (для $v = \frac{2}{3}$ и $v = 1$) на основании различных моделей в сравнении с наивным прогнозом (предполагающего, что новая оценка равна предыдущей). Простейшие модели оценивают пересмотры в зависимости от уже опубликованных оценок показателя и включают следующие спецификации, представленные в таблице 1.

Полученные по данным в период с 02.1966 г. по 01.2014 г. относительные RMSFE (root mean squared forecast error) показателей в ретроспективном периоде с 02.2001 г. по 12.2013 г. свидетельствуют о несущественности преимуществ простейших моделей перед наивным прогнозом, подтверждая выводы, полученные в более ранних работах (например, (Mankiw, Shapiro, 1986), (Faust, Rogers, Wright, 2005)) об ограниченной предсказуемости первых оценок темпов роста ВВП США.

Таблица 1 – Модели пересмотров

Описание модели	Формула
1) Модель среднего пересмотра строится в предположении о некоррелированности пересмотров, но допускает возможность их смещения.	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \beta_0 + \varepsilon_t$
2) Авторегрессионная модель допускает корреляцию пересмотров между оценками одинаковых уровней и наличие у них ненулевого смещения.	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \beta_0 + \beta_1\left(y_{t-1}^{t+v-1} - y_{t-1}^{t+v-\frac{4}{3}}\right) + \varepsilon_t$
3) модель предполагает наличие зависимости пересмотров от предшествующей опубликованной оценки показателя. По построению эти три модели совпадают с моделями, используемыми при проверке гипотез по классификации природы пересмотров (на новости и шум).	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \beta_0 + \beta_1 y_t^{t+v-\frac{1}{3}} + \varepsilon_t$
4) модель рассматривает зависимость пересмотров от предшествующей опубликованной оценки показателя для q кварталов. Данная спецификация является упрощением «многовинтажной» модели, рассматриваемой в работах [23], [24]	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \beta_0 + \sum_{i=0}^{q} \beta_{i+1} y_{t-i}^{t+v-\frac{1}{3}} + \varepsilon_t$
5) Пороговая модель является расширением третьей модели, допускающим различные реакции пересмотров в зависимости от значения и знака предшествующей опубликованной оценки показателя. Пороговое значение оценивается совместно с коэффициентами регрессии методом	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \left[\beta_0 + \beta_1 y_t^{t+v-\frac{1}{3}}\right] \mathrm{I}\left(y_t^{t+v-\frac{1}{3}} \le c\right) + \left[\beta_2 + \beta_3 y_t^{t+v-\frac{1}{3}}\right] \mathrm{I}\left(y_t^{t+v-\frac{1}{3}} > c\right) + \varepsilon_t$

сеточного поиска при условии, что выше и ниже порога лежит не менее 15% наблюдений.	

Примечание – Источник: (Clements, Galvão, 2018)

Спецификация следующей группы моделей учитывает то факт, что в период между различными оценками темпа роста ВВП становятся доступной информация о значениях (окончательных или предварительных) других показателей $X_{i,t}^{t+v}$, характеризующих экономическую ситуацию и коррелирующих с ВВП или его компонентами. В своей работе Клементс и Гальвао рассматривают следующие показатели: ИПП, численность занятых, розничный товарооборот, жилое строительство, заказы на товары длительного пользования, торговый баланс, ИПЦ, индекс производства ISM, потребительские настроения. В этих моделях в качестве объясняющих переменных используются пересмотры различных вариантов квартальных различий ежемесячных значений показателей $x_t^{t+v} = X_t^{t+v} - X_{t-1}^{t+v}$, приведенные в таблице 2.

Таблица 2 – Описание моделей

Описание модели	Формула
6) Модель (именуемая авторами «новый пересмотр») рассматривает последние опубликованные пересмотры регрессоров.	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \delta_0 + \sum_{i=1}^{n} \delta_{\mathrm{i}} \left(x_{i,t}^{t+v} - x_{i,t}^{t+v-\frac{1}{3}} \right) + \varepsilon_t$
7) Модель (именуемая авторами «новая оценка») рассматривает последние опубликованные значения регрессоров.	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \delta_0 + \sum_{i=1}^{n} \delta_{\mathrm{i}} x_{i,t+v-1/3}^{t+v} + \varepsilon_t$
8) Модель (именуемая авторами «обновленная оценка») рассматривает последние опубликованные значения регрессоров для прогнозируемого периода.	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \delta_0 + \sum_{i=1}^{n} \delta_{\mathrm{i}} x_{i,t}^{t+v} + \varepsilon_t$
9) Модель (именуемая авторами «предыдущая публикация») рассматривает значения регрессоров, опубликованные месяцем ранее	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \delta_0 + \sum_{i=1}^{n} \delta_{\mathrm{i}} x_{i,t}^{t+v-1/3} + \varepsilon_t$
10) В качестве регрессоров последней в этой группе модели выступают данные обследований.	$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \delta_0 + \sum_{i=1}^{n} \delta_{\mathrm{i}} x_{i,t+v-2/3} + \varepsilon_t$

Примечание – Источник: (Clements, Galvão, 2018)

Последняя модель отличается от предыдущих. Во-первых, она использует данные, которые не подлежат пересмотрам. Во-вторых, она имеет непосредственное отношение к эффективности прогнозов y_t^{t+v}. Если оказывается, что эта модель превосходит по качеству наивные прогнозы, то первые оценки показателя не являются эффективными, поскольку используют не всю доступную информацию.

По результатам эмпирических оценок в работе (Clements, Galvão, 2018) вторая оценка ВВП предсказуема для модели (8 – «обновленная оценка») и для модели (9 –

«предыдущая публикация»). Кроме этого, обе оценки (и вторая, и третья) предсказуемы на основе ежемесячных показателей в периоды экономических спадов. Но, если для второй оценки точность прогнозов моделей 6-10, оцененная по RMSFE, может почти вдвое превышать точность прогнозов наивной модели, то для третьей оценки выигрыш в точности существенно меньше (и, как правило, реализуется в условиях рецессии).

При рассмотрении третьей группы моделей авторы (Clements, Galvão, 2018) дополнили учитываемый при прогнозировании информационный набор ежедневными финансовыми переменными. Для обеспечения временных рядов достаточной длины они используют только пять показателей $x_{i,t+v-l-j/m}$: фондовый индекс SP500, индекс DJIA (Dow Jones industrial average) вместе с краткосрочной ставкой (трехмесячных казначейских векселей) и двумя процентными спрэдами.

(11) Первой изучается обычная MIDAS – модель, которая оценивает квартальные пересмотры на основе ежедневных данных формула (135)):

$$y_t^{t+v} - y_t^{t+v-\frac{1}{3}} = \alpha_0 + \alpha_1 \sum_{i=1}^{n} \sum_{j=0}^{K-1} w_{\mathrm{i,j}}\,(\theta_i\,, K) x_{i,t+v-l-j/m} + \varepsilon_t,$$

(13)

где m = 60 – число наблюдений за квартал,

K – число включаемых в модель ежедневных лагов,

$w_{\mathrm{i,j}}(\theta_i, K)$ – весовые функции, причем агрегированные веса w_j (θ,m) в сумме дают 1, что обеспечивает возможность оценки параметра α_1 . Простейший вид весов состоит в «плоском агрегировании», когда $w_{\mathrm{i,j}}(\theta_i, K) = \frac{1}{K}\ \ \forall j, i.$

Параметры весовой функции оцениваются одновременно с коэффициентами модели α_0 и α_1 с помощью нелинейного метода наименьших квадратов.

(12) Другой вариант модели – STMIDAS (Smooth Transition MIDAS) строится в предположении, что коэффициенты модели могут меняться в зависимости от рыночной ситуации и денежно-кредитной политики (формула (14)):

$$\begin{aligned} y_t^{t+v} - y_t^{t+v-\frac{1}{3}} &= \alpha_0 + \alpha_1 x_{t+v}(\theta, K, l, m)\left[1 - G\left(\gamma, c, x_{t+v}(\lambda, K, l, m)\right)\right] \\ &+ \alpha_2 x_{t+v}(\theta, K, l, m)\left[G\left(\gamma, c, x_{t+v}(\lambda, K, l, m)\right)\right] + \varepsilon_t \end{aligned} \quad ,$$

(14)

где

$x_{t+v}(\lambda, K, l, m) = \sum_{j=0}^{K-1} w_{\mathrm{j}}(\lambda, K) x_{i,t+v-l-j/m}$, $x_{t+v}(\theta, K, l, m) = \sum_{j=0}^{K-1} w_{\mathrm{j}}(\theta_i, K) x_{i,t+v-l-j/m} + \varepsilon_t$, коэффициенты регрессии теперь зависят от разности между агрегированными высокочастотными данными и порогом C и определяются путем взвешивания на основе логистической функции $G\big(\gamma, c, x_{t+v}(\lambda, K, l, m)\big) = \big[1 + exp\big(-\gamma(x_{t+v}(\lambda, K, l, m) - C)\big)\big]^{-1}$ (γ задает гладкость функции).

Результаты сравнения точности прогнозов второй и третьей оценок ВВП на основе ежедневных финансовых данных с наивным прогнозом свидетельствуют, что показатели доходности акций позволяют получить значимые улучшения качества прогнозов. Эмпирические выводы, полученные в работе (Clements, Galvão, 2018), совпадают с результатами многих других работ. Например, авторы (Andreou, Ghysels, Kourtellos, 2011) показали, что лучшие качественные характеристики прогнозов роста выпуска на один квартал вперед достигаются при использовании моделей показателя на основе процентных ставок, спредов по облигациям и доходности акций. Так что ежедневные финансовые переменные несут в себе новую информацию о пересмотрах данных о выпуске. Еще раньше Гилберт (Gilbert, 2011) утверждал, что показатели доходности акций могут улучшить прогнозы второй и третьей оценок ВВП вследствие их отклика на изменения данных об экономической активности, реагирующих на информацию об ожидаемых объемах выпуска.

С учетом отмеченной в последней работе информационной взаимосвязью между ежедневными финансовыми показателями и ежемесячными экономическими показателями Клементс и Гальвао (Clements, Galvão, 2018) дополнили свой анализ рассмотрением комбинированных моделей второй и третьей групп. Для этого авторы включают показатели доходности акций в модели (8) и (9) и оценивают MIDAS – модели с использованием бета-функций весов, оцениваемых совместно с параметрами модели с помощью нелинейного метода наименьших квадратов. Полученные результаты показывают, что комбинирование моделей приводит к значимому сокращению ошибки прогнозирования третьей оценки роста ВВП независимо от фазы экономического цикла, и сокращению ошибки прогнозирования второй оценки в период рецессии. Так что финансовые показатели несут информацию, дополняющую ту, что содержится в экономических переменных.

Прогнозирование с использованием данных, подверженных пересмотрам

Задача, которая стоит перед специалистами, занимающимися прогнозированием, состоит в том, чтобы в момент времени T построить прогноз истинного значения некоторой макроэкономической переменной $\tilde{y}_t$ на h периодов вперед $\{\hat{y}_T, \hat{y}_{T+1}, \hat{y}_{T+2}, \dots, \hat{y}_{T+h}\}$. Здесь перед исследователем встает два вопроса: какую спецификацию использовать для моделирования переменных с пересмотрами и какие данные следует рассматривать при оценке модели.

В работе (Croushore, 2011) автор отмечает, что пересмотры данных могут оказывать значительное влияние на прогнозы, так как модели, оцененные на одной и той же выборке, но в разные моменты времени (например, модели, построенные в 1 кв. 2014 и 2 кв.2015 на данных за 2000-2013 гг.), могут отличаться из-за пересмотра данных.

(Croushore, 2011) рассматривает стандартную модель пересмотра данных, которая описывается уравнением (15):

$$y_t^v = y_t^{v-1} + r_t^{v,v-1}, \tag{15}$$

где $r_t^{v,v-1}$ – разница между данными из винтажа данных (выпуска статистического отчета) v и данными из винтажа $v-1$.

(Croushore, 2011) предполагает, что предсказать пересмотр данных очень сложно, и общие свойства пересмотров могут проявиться только через десятилетия. Однако он также ссылается на работу (Faust, Rogers, Wright, 2005), авторы которой пришли к выводу, что пересмотры данных можно предсказать для всех стран G-7 за исключением США.

Предположения относительно ошибки, включенной в модель, зависят от стратегии пересмотра данных статистическими агентствами. Если агентство просто предоставляет выборочную информацию, то ошибка представляет собой ошибку измерения, некоррелированную с истинным значением показателя, и, следовательно, при дальнейших пересмотрах значение шума в оценке будет уменьшаться. Если же агентство комбинирует выборочную информацию с другой полезной информацией, чтобы получить наиболее близкую оценку к реальному значению, то пересмотр является оптимальным прогнозом будущих пересмотров, то есть при каждом пересмотре к предыдущей оценке добавляются новости.

В статье (Croushore, 2011) два этих случая моделируются уравнениями (16) и (17):

$y_t^* = y_t^v + e_t^v, \quad y_t^v \perp e_t^v$ – пересмотры, связанные с новостями, (16)

$y_t^* = y_t^v - u_t^v, \quad y_t^v \perp u_t^v$ – пересмотры, связанные с шумом. (17)

В уравнении (16) пересмотры непредсказуемы, так как они не коррелируют с ошибкой. В модели (17) пересмотры коррелируют с предыдущими данными, потому что ошибки винтажей данных ортогональны только истинному значению переменных, а не друг другу. Таким образом, в случае (17) пересмотры предсказуемы, и каждый предыдущий пересмотр не является оптимальным прогнозом для последующих.

(Croushore, 2011) отмечает, что для того, чтобы смоделировать пересмотры данных, необходимо понимать, добавляет ли статистическое агентство новости в данные или использует только некий шум. Однако вместо того, чтобы использовать предположения о стратегии агентства, можно оценивать факторную модель. Предполагается, что пересмотры данных по большому количеству переменных нивелируют друг друга, и факторная модель дает достоверные прогнозы. В таком случае одна или несколько ненаблюдаемых переменных (факторов) вызывает изменения в различных показателях. Данные модели оцениваются методом главных компонент, и, если пересмотры различных данных некоррелированы между собой, то такой подход позволит полностью убрать влияние пересмотров на прогнозы.

Однако Фауст и Райт (Faust, Wright, 2009) показали, что использование нескольких двумерных моделей превосходит по качеству факторные модели. При этом предполагается, что шум, возникающий при использовании большого количества факторов, ухудшает прогнозы сильнее, чем избавление от ошибок, появляющихся в результате пересмотров.

Еще один способ избежать влияния пересмотров основан на идее оценивать модель только на пересмотренных данных, не учитывая недавно опубликованные данные, которые еще не были пересмотрены. Однако в работе (Howrey, 1978) был получен вывод, что данные, скорректированные с помощью фильтра Калмана, дают более точные прогнозы, чем полное игнорирование недавних данных. В работе (Harvey, McKenzie, Blake, Desai, 1983) отмечается, что модели пространства состояний также дают достаточно точные прогнозы. Также в качестве способов учета пересмотров данных рассматриваются следующие подходы:

– учет ожиданий (см, например, (Lee, Kevin, Nilss Olekalns, and Kalvinder Shields, 2008). Помимо моделирования процесса порождения данных и процесса измерения данных, исследователь также моделирует процесс ожиданий;

– коинтеграция и общие тренды. (Garratt, A., Lee, K., Mise, E., and Shields, K, 2008) используют коинтегрированные VAR для прогнозирования тренда выпуска, чтобы оценить разрыв выпуска;

– комбинация прогнозов. (Altavilla, Carlo, and Matteo Ciccarelli, 2014) рекомендуют использовать комбинации прогнозов и для моделей, и для наборов информации;

– использование одного уравнения. В работе (Koenig, E. F., Dolmas, S., and Piger, J., 2003) предлагается использовать первоначальные данные, так как, согласно выводам из работы, первоначальные данные дают более точные прогнозы, чем пересмотренные или спрогнозированные данные.

Наиболее часто используемая структура, рассматриваемая для моделирования пересмотра данных, предложена в работе (Jacobs, van Norden, 2011). Она основана на непосредственном различении новостей и шума, так что публикуемое в момент $(t + s)$ значение показателя для периода t y_t^{t+s}, определяется, как сумма истинного значения и ошибок, определяемых новостными и шумовыми изменениями (по формуле (18)):

$$y_t^{t+s} = \tilde{y}_t + v_t^{t+s} + \varepsilon_t^{t+s},$$

18)

где $\tilde{y}_t$ – истинное значение рассматриваемой переменной;

v_t^{t+s} и ε_t^{t+s} – компоненты новостей и шума.

Если данные подвергаются более, чем одному пересмотру, то уравнение (18) переписывается в матричном виде (19):

$$y_t = 1 \cdot \tilde{y}_t + v_t + \varepsilon_t,$$

(19)

где все составляющие представляют собой k – мерные вектора (k – число пересмотров): $y_t = \left(y_t^{t+1}, y_t^{t+2}, \ldots, y_t^{t+k}\right)'$, $\varepsilon_t = \left(\varepsilon_t^{t+1}, \varepsilon_t^{t+2}, \ldots, \varepsilon_t^{t+k}\right)'$, $v_t = \left(v_t^{t+1}, v_t^{t+2}, \ldots, v_t^{t+k}\right)'$, $1 - k$-мерный вектор единиц.

Относительно истинного значения рассматриваемой переменной предполагается, что она представляет собой стационарный процесс и подчиняется авторегрессионной модели порядка p с k новостными компонентами, так что уравнение записывается формулой (20):

$$\tilde{y}_t = \rho_0 + \sum_{i=1}^{p} \rho_i \tilde{y}_{t-i} + R_1 \eta_{1t} + \sum_{i=1}^{k} v_{i,t},$$

(20)

где $v_{i,t} = \mu_{v_i} + \sigma_{v_i} \eta_{2t,i}$,

η_{1t} и $\eta_{2t,i}$ являются независимыми, одинаково распределенными случайными величинами с нулевым математическим ожиданием и единичной дисперсией.

В этом случае компоненты векторного уравнения, характеризующие новостную и шумовую составляющие публикуемого для момента t значения переменной, записываются с помощью формулы (21):

$$v_t = \begin{bmatrix} v_t^{t+1} \\ v_t^{t+2} \\ \vdots \\ v_t^{t+k} \end{bmatrix} = -\begin{bmatrix} \sum_{i=1}^{k} \mu_{v_i} \\ \sum_{i=2}^{k} \mu_{v_i} \\ \vdots \\ \mu_{v_k} \end{bmatrix} - \begin{bmatrix} \sum_{i=1}^{k} \sigma_{v_i} \eta_{2t,i} \\ \sum_{i=2}^{k} \sigma_{v_i} \eta_{2t,i} \\ \vdots \\ \sigma_{v_k} \eta_{2t,k} \end{bmatrix}; \quad \boldsymbol{\varepsilon_t} = \begin{bmatrix} \varepsilon_t^{t+1} \\ \varepsilon_t^{t+2} \\ \vdots \\ \varepsilon_t^{t+k} \end{bmatrix} = -\begin{bmatrix} \mu_{\varepsilon_1} \\ \mu_{\varepsilon_2} \\ \vdots \\ \mu_{\varepsilon_k} \end{bmatrix} + \begin{bmatrix} \sigma_{\varepsilon_1} \eta_{3t,1} \\ \sigma_{\varepsilon_2} \eta_{3t,2} \\ \vdots \\ \sigma_{\varepsilon_k} \eta_{3t,k} \end{bmatrix},$$

(21)

где $\eta_{3t,i}$ также являются независимыми, одинаково распределенными случайными величинами с нулевым математическим ожиданием и единичной дисперсией. И кроме этого, для $\eta_t = [\eta_{1t}, \eta_{2t}^T, \eta_{3t}^T]$ выполняется $\begin{cases} M\eta_t = 0 \\ M(\eta_t \eta_t^{/}) = I \end{cases}$.

Статистическая модель пересмотров описывается уравнением состояния (21) и уравнением перехода, задаваемым (20) и (21), и может быть оценена методом максимального правдоподобия с помощью фильтра Калмана (см. (Jacobs, van Norden, 2011)).

На основе анализа пересмотров реального ВВП США авторы показывают, что их формулировка модели пространства состояний, обладающая достаточной гибкостью, позволяет оценивать сложную динамику показателя даже при использовании небольшого количества наблюдений и винтажей. Еще одно преимущество модели состоит в относительной легкости дифференциации новостной и шумовой составляющих в ошибках измерений.

Простой пример приводится в (Clements, Galvão, 2018). Пусть первое публикуемое значение переменной не включает новостную составляющую: $y_t^{t+1} = \tilde{y}_t + v_t^{t+1} + \varepsilon_t^{t+1} = \rho_0 + \sum_{i=1}^{p} \rho_i \tilde{y}_{t-i} + R_1 \eta_{1t} + \sigma_{\varepsilon_1} \eta_{3t,1}$. Вторая публикация равна $y_t^{t+2} = \tilde{y}_t + v_t^{t+2} + \varepsilon_t^{t+2} = \rho_0 + \sum_{i=1}^{p} \rho_i \tilde{y}_{t-i} + R_1 \eta_{1t} + \sigma_{v_i} \eta_{2t,1} + \sigma_{\varepsilon_2} \eta_{3t,2}$. И если шумовая составляющая второй оценки отсутствует ($\sigma_{\varepsilon_2} \eta_{3t,2} = 0$), то эта вторая оценка более точно, в сравнении с первой оценкой, отражает истинное значение $\tilde{y}_t$. Кроме этого, т.к. $Cov(y_t^{t+2} - \tilde{y}_t, y_t^{t+2}) = Cov\left(\sum_{i=1}^{k} v_{i,t}, v_{1,t}\right) = 0$, т.е. пересмотр данных не коррелирует с y_t^{t+2}, то вторая публикация является эффективной оценкой $\tilde{y}_t$. Если у второй оценки отсутствует новостная составляющая ($R_1 \eta_{1t} + \sigma_{v_i} \eta_{2t,1} = 0$), то $Cov(y_t^{t+2} - \tilde{y}_t, y_t^{t+2}) = \sigma_{\varepsilon_2}^2$ и пересмотры являются предсказуемыми на основе второй оценки.

При ответе на второй вопрос у исследователя существует две альтернативы. Как правило, для прогнозирования показателя оценивают модель по данным, доступным на момент прогнозирования, т.е. последним доступным данным: $\{y_p^T, \ldots; y_{T-k+1}^T, \ldots, y_{T-1}^T\}$. Такой подход получил название EOS – End Of Sample – метод (см, Clements, Galvão, 2013) [23]). Основной аргумент в пользу его использования состоит в том, что более поздние оценки данных (построенные в условиях наличия большего объема исходной информации) являются более точными, в сравнении с ранними оценками. Так что EOS – стратегия опирается на "лучшие" (из доступных в момент составления прогноза) оценки данных.

Другой вариант состоит в использовании данных реального винтажа (получившего название, RTV – Real Time Vintage - метод), т.е. наблюдений показателя одного винтажа. Обоснование этого подхода состоит в том, что при прогнозировании значений будущих периодов они будут обусловлены известными на момент составления прогноза винтажами объясняющих переменных. Как будет показано ниже, свойства оценок параметров моделей, полученных этими методами, могут существенно отличаться.

A. Прогнозирование на основе авторегрессионных моделей

Если для моделирования истинного значения переменной, описываемого уравнением (22) в соответствии с моделью Джейкобса и Ван Нордена:

$$\tilde{y}_t = \rho_0 + \sum_{i=1}^{p} \rho_i \tilde{y}_{t-i} + R_1 \eta_{1t} + \sum_{i=1}^{k} v_{i,t}$$

(22)

используется авторегрессия порядка p, то оценки задаются, как сумма $\phi_0 + \phi^{/} y^T$, где $y^T = \left(y_p^T, \dots, \dots, y_{T-2}^T, y_{T-1}^T\right)^{/}$, ϕ_0 – константа и $\phi^{/} = \left(\phi_0, \dots, \phi_p\right)$ – вектор параметров. Тогда оптимальные значения параметров (ϕ_0^*, ϕ^*), минимизирующие квадрат ошибки являются решением: $(\phi_0^*, \phi^*) = \arg\min M\left\{\left(y_T^{T+f} - \phi_0 - \phi^{/} y^{T+1}\right)^2\right\}$.

Клементс и Гальвао (Clements, Galvão, 2013) показали, что при $f = 1$, эти оптимальные значения задаются с помощью формулы (23):

$$\begin{gathered}\phi^* = \left(\Sigma_{\tilde{y}} + \Sigma_\nu + \Sigma_{\tilde{y}\nu} + \Sigma^{/}{}_{\tilde{y}\nu} + \Sigma_\varepsilon\right)^{-1}\left(\Sigma_{\tilde{y}} + \Sigma^{/}{}_{\tilde{y}\nu}\right)\rho \\ \phi_0^* = (1 - \phi^* \mathbf{1})\mu_{\tilde{y}} - \sum_{j=1}^{k} \mu_{\nu_j} - \mu_{\varepsilon_1} - \phi^* \mu_\varepsilon - \phi^* \mu_\nu\end{gathered},$$

(23)

где Σ_ν и Σ_ε - матрицы вторых моментов новостной и случайной компонент;

$\Sigma_{\tilde{y}\nu}$ - матрица вторых моментов новостной компоненты и реального значения;

$\mu_{\tilde{y}} = \left(1 - \rho(1)\right)^{-1}\left[\rho_0 + \sum_{j=1}^{k} \mu_{\nu_j}\right]$;

$E(\varepsilon_t^{t+1}) = -\mu_{\varepsilon_1}$, $E(\nu_t^{t+1}) = -\sum_{j=1}^{k} \mu_{\nu_j}$.

Если пересмотры характеризуются нулевым математическим ожиданием, т.е. $\forall j: \mu_{\nu_j} = 0$, $\forall k: \mu_{\varepsilon_k} = 0$, полученные выражения принимают более простой вид:

для пересмотров, являющихся новостями $\phi^* = \left(\Sigma_{\tilde{y}} + \Sigma_\nu + \Sigma_{\tilde{y}\nu} + \Sigma^{/}{}_{\tilde{y}\nu}\right)^{-1}\left(\Sigma_{\tilde{y}} + \Sigma^{/}{}_{\tilde{y}\nu}\right)\rho$

для пересмотров, являющихся шумом $\phi^* = \left(\Sigma_{\tilde{y}} + \Sigma_\varepsilon\right)^{-1}\Sigma_{\tilde{y}}\rho$

И, кроме этого, определенные таким образом параметры являются оптимальными для любых значений $f \geq 1$.

В случае EOS–метода AR-модель оценивается в виде формулы (24):

$$y_t^T = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i}^T + e_{t,EOS}\,, \quad t = p, \dots, T-2, T-1,$$

(24)

которая в матричном виде переписывается в виде (25):

$$y^T = 1 \cdot \beta_0 + y_{-1}^T \beta + \varepsilon_{EOS},$$

(25)

где $y^T = \left(y_p^T, \dots, y_{T-2}^T, y_{T-1}^T\right)^{/}$;

$y_{-1}^T = \left[y_{T-1}^T, y_{T-2}^T, \dots, y_{T-p}^T\right]$, а $y_{T-i}^T = \left(y_{p-i}^T, \dots, y_{T-i-2}^T, y_{T-i-1}^T\right)^{/}$;

1 – (T–p)-мерный вектор единиц;

$\varepsilon_{EOS} = \left(\varepsilon_p^T, \dots, \varepsilon_{T-2}^T, \varepsilon_{T-1}^T\right)^{/}$.

Для случая, когда процесс порождения данных описывается моделью Джейкобса и Ван Нордена, Клементс и Галвао (Clements, Galvão, 2013) показали, что выборочные МНК – оценки параметров уравнения определяются по формуле (26):

$$\begin{aligned} \beta^* &= \left(\Sigma_{\tilde{y}} + \Sigma_\nu + \Sigma_{\tilde{y}\nu} + \Sigma^{/}{}_{\tilde{y}\nu} + \Sigma_\varepsilon\right)^{-1}\left(\Sigma_{\tilde{y}} + \Sigma^{/}{}_{\tilde{y}\nu}\right)\rho \\ \beta_0^* &= (1 - \beta^* \mathbf{1})\left(\mu_{\tilde{y}} - \mu_{\nu_k} - \mu_{\varepsilon_k}\right) \end{aligned}, \qquad (26)$$

где $\mu_{\tilde{y}} = \left(1 - \rho(1)\right)^{-1}\left[\rho_0 + \sum_{j=1}^{l} \mu_{\nu_j}\right]$;

$E\left(\varepsilon_t^{t+k}\right) = -\mu_{\varepsilon_k}$, $E\left(\nu_t^{t+k}\right) = -\mu_{\nu_k}$.

Полученные результаты свидетельствуют о том, что EOS–метод не позволяет получить оптимальных оценок параметров модели ($\beta^* \neq \phi^*$ и $\beta_0^* \neq \phi_0^*$) и, как следствие, оптимальных (в смысле среднего квадратического отклонения) прогнозов истинного значения показателя при наличии пересмотров данных. При этом отсутствие оптимальности не зависит от природы пересмотров, и выполняется как в случае новостных изменений, так и в случае шума.

Природа пересмотров оказывает влияние на масштабы отклонения EOS–оценок параметров от оптимальных. Например, если пересмотры являются шумом, то расхождения возрастают с увеличением значения ρ и различий дисперсий пересмотров различных публикаций $\sigma_{\varepsilon_i}^2$.

Кроме этого, прогнозы, полученные EOS–методом, нельзя считать непредвзятыми, т.к. они являются смещенными (например, для прогнозов первого публикуемого значения в соответствии с оценками параметров (β_0^*, β^*) смещение составит $\left(\mu_{\nu_l} - \sum_{j=1}^{k} \mu_{\nu_j}\right) + \left(\mu_{\varepsilon_k} - \mu_{\varepsilon_1}\right) - \beta^*\left[\left(\mu_\nu + 1\mu_{\nu_k}\right) + \left(\mu_\varepsilon + 1\mu_{\varepsilon_k}\right)\right]$).

Использование данных RTV–метода может преодолеть недостатки использования EOS – данных. Точность прогноза повышается за счет реорганизации данных реального времени, используемых при оценке модели прогнозирования, таким образом, чтобы данные, используемые при оценке, были аналогичны данным, от которых зависит прогноз. В рамках RTV AR-модель оценивается в виде модели (27):

$$y_{t-1}^{t} = \gamma_0 + \sum_{i=1}^{p} \gamma_i y_{t-i-1}^{t-1} + e_{t-1,RTV}\,, \quad t = p, \dots, T-2, T-1, \tag{27}$$

которое в матричном виде переписывается в виде (28):

$$y^{t} = 1 \cdot \gamma_0 + \gamma_{-1}^{t}\gamma + \varepsilon_{RTV}, \tag{28}$$

где $y^{t} = \left(y_{p}^{p+1}, \dots, y_{T-2}^{T-1}, y_{T-1}^{T}\right)^{/}$;

$y_{-1}^{t} = \left[y_{t-1}^{t}, \dots, y_{t-2}^{t}, y_{t-p}^{t}\right]$;

$y_{t-i}^{t} = \left(y_{p-i}^{p}, \dots, y_{T-i-2}^{T-2}, y_{T-1-i}^{T-1}\right)^{/}$;

1 – (T–p)-мерный вектор единиц;

$\varepsilon_{RTV} = \left(\varepsilon_{p}^{P+1}, \dots, \varepsilon_{T-2}^{T-1}, \varepsilon_{T-1}^{T}\right)^{/}$.

Оценивание уравнения методом наименьших квадратов предполагает, что оценки параметров находятся как решение задачи минимизации: $(\gamma_0^*, \gamma^*) = \arg\min M\{(y_{t-1}^{t} -$ $\gamma 0 - \gamma y t - 2 t - 1 2$, формулировка которой совпадает с задачей минимизации суммы квадратов отклонений прогнозов в реальном времени для f = 1. Так что оценки параметров, полученные RTV–методом, не отличаются от оптимальных ($\gamma^* = \phi^*$ и $\gamma_0^* = \phi_0^*$), и, для f = 1, они являются несмещенными независимо от природы пересмотров и значений их математических ожиданий.

Эмпирическая проверка, представленная в работе (Clements, Galvão, 2013) подтверждает, что для авторегрессионных моделей использование RTV-метода повышает точность прогнозирования показателей в сравнении с EOS-подходом, даже если среднее значение пересмотра данных не равно нулю. Различия качественных характеристик RTV и EOS прогнозов увеличиваются для переменных, пересмотры которых являются преимущественно новостями, и для очень устойчивых авторегрессионных процессов.

Б. Интервальные оценки прогнозов

Выбор используемых при моделировании данных находит отражение и в свойствах интервальных оценок прогнозов. В рамках EOS – метода *AR*-модель включает *p* авторегрессионных членов и константу и оценивается в виде (29):

$$y_t^T = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i}^T + e_{t,EOS}\,, \quad t = p, \dots, T-2, T-1.$$

(29)

В случае, когда для оценивания используется достаточно длинный временной ряд (так что $T \gg k$) EOS – оценки можно считать приблизительно равными оценкам, полученным по полностью пересмотренным данным ($\beta_i = \rho_i, \;\; i = 0, \dots, p$). Тогда выборочная оценка среднего квадратического отклонения рассчитывается по формуле (30):

$$\hat{\sigma}_{T-1,EOS} = \sqrt{M\left[\left(R_1\eta_{1t} + \sum_{i=1}^{l} \upsilon_{i,t}\right)^2\right]} = \sqrt{R_1^2 + \sum_{i=1}^{l} \sigma_{\upsilon_i}^2} \qquad .$$

(30)

Вместе с тем, в предположении, что $k \geq p$, ожидаемая квадратическая ошибка одношагового прогноза будет равна (31):

$$M\left(y_T^{T+1} - \hat{y}_{T,EOS}\right)^2 = M\left(\sum_{i=1}^{p} \alpha_i (y_{T-i} - y_{T-i}^T) + R_1\eta_{1T} + \varepsilon_T^{T+1}\right)^2 = \sum_{i=1}^{p} \alpha_i^2 \left(\sigma_{\varepsilon_i}^2 + \right.$$

$+j=ik\sigma vj2+R12+\sigma\varepsilon 12.$

(31)

Сравнительные результаты внутривыборочной дисперсии и квадрата отклонения прогнозируемого значения показателя от публикуемого зависят от природы проводимых пересмотров. Если пересмотры включают в себя лишь одну составляющую (являются либо чисто новостями, либо чисто шумом), то выражение для ошибки прогноза перепишется в сокращенном виде, как:

$$M\left(y_T^{T+1} - \hat{y}_{T,EOS}\right)^2 = \sum_{i=1}^{p} \rho_i^2 \sum_{j=i}^{k} \sigma_{\nu_j}^2 + R_1^2 \qquad \text{– для новостей}$$

$$M\left(y_T^{T+1} - \hat{y}_{T,EOS}\right)^2 = \sum_{i=1}^{p} \rho_i^2 \sigma_{\varepsilon_i}^2 + R_1^2 + \sigma_{\varepsilon_1}^2 \qquad \text{– для шума}$$

В результате, если пересмотры являются шумом, то $M\left(y_T^{T+1} - \hat{y}_{T,EOS}\right)^2 > \sigma_{T-1,EOS}^2 = R_1^2$, и внутривыборочная дисперсия ошибки оказывается меньше квадрата отклонения для прогнозируемых значений. В присутствии новостных пересмотров наблюдается обратная ситуация: если $\sum_{i=1}^{p} \rho_i^2 \sum_{j=i}^{k} \sigma_{\nu_j}^2 < \sum_{j=1}^{k} \sigma_{\nu_j}^2$, то внутривыборочная дисперсия ошибки превышает соответствующее значение для прогнозов.

В работе Клементса (Clements, 2017) подробно рассмотрены интервальные оценки, как для частных случаев спецификации авторегрессионной модели, так и для модели в общем виде. Полученные автором результаты свидетельствуют, что при $p > 1$ ожидаемая квадратическая ошибка прогноза может превышать внутривыборочную дисперсию, когда $|\rho_i|$ велики, но в целом можно ожидать, что интервалы прогнозирования, основанные на

EOS-методе, будут иметь тенденцию к завышению неопределенности, связанной с будущими наблюдениями, когда пересмотры являются новостями. Если пересмотры являются шумом, то доверительные интервалы прогноза по EOS-данным, напротив, будут слишком узки, и занижать неопределенность вне выборки.

В соответствии с RTV – методом оценки параметров модели и прогнозы рассчитываются на основе данных последней доступной на момент прогнозирования публикации. Преимущество такого подхода состоит в том, что исходная выборка является однородной с точки зрения обработки информации: не происходит смешивания полностью пересмотренных и малопересмотренных данных.

В этом случае AR-модель оценивается по формуле (32):

$$y_{t-1}^{t} = \gamma_0 + \sum_{i=1}^{p} \gamma_i y_{t-i-1}^{t-1} + e_{t-1,RTV}\,, \quad t = p, \dots, T-2, T-1 \tag{32}$$

и одношаговый прогноз y_T определяется, как $\hat{y}_{T,RTV} = \beta_0 + \beta_1 y_{T-1}^{T} + \dots + \beta_p y_{T-p}^{T}$.

На примере модели AR(1): $y_{t-1}^{t} = \gamma y_{t-2}^{t-1} + e_t$, где $\gamma = cov\,(y_{t-1}^{t}, y_{t-2}^{t-1})/var\,(y_{t-2}^{t-1}) = \rho$, $cov\,(y_{t-1}^{t}, y_{t-2}^{t-1}) = M[(\rho y_{t-2} - \eta_{t-1}), (y_{t-2} - \nu_{t-2})] = \rho\sigma_y^2 - \rho\sigma_\nu^2$, $var\,(y_{t-2}^{t-1}) = var\,(y_{t-2} - \upsilon_{t-2}) = \sigma_y^2 + \sigma_\nu^2 - 2cov\,(y_{t-2} - \nu_{t-2}) = \sigma_y^2 - \sigma_\nu^2$ легко показать, что в присутствии новостных пересмотров оценка внутривыборочной дисперсии ошибки составляет: $\sigma_{RTV}^2 = var\,(y_{t-1}^{t} - \rho y_{t-2}^{t-1}) = \sigma_\eta^2 + \alpha\sigma_\nu^2$, что совпадает с ошибкой прогноза: $M\left[\left(y_T^{T+1} - \hat{y}_{T,RTV}\right)^2\right] = M[y_T - \nu_T - \alpha(y_{T-1} - \nu_{T-1})^2] = \sigma_\eta^2 + \alpha\sigma_\nu^2$.

Если пересмотры данных являются шумом, то $\beta = \phi^* = \rho\sigma_y^2/\left(\sigma_y^2 + \sigma_\varepsilon^2\right)$. Дисперсия ошибки в выборке равна $var\,(y_{t-1}^{t} - \gamma y_{t-2}^{t-1}) = var[y_{t-1} + \varepsilon_{t-1} - \gamma(y_{t-2} + \varepsilon_{t-2})]$, а квадратическая ошибка прогноза, оценивающая вневыборочную неопределенность равна $var\left(y_T^{T+1} - y_{T,RTV}\right)^2 = var[y_T + \varepsilon_T - \gamma(y_{T-1} + \varepsilon_{T-1})]$. Так что и для изменений, представляющих собой шум, они равны.

Таким образом, применение RTV-метода, в отличие от использования EOS, позволяет обеспечить правильные интервальные оценки прогнозов независимо от природы пересмотров. Впервые этот факт был отмечен в работе (Koenig, Dolmas, Piger, 2003), а позже обстоятельно описан в работе (Clements, 2017).

(Clements, 2017) также осуществил эмпирическую проверку представленных результатов. На примере 25 макропеременных переменных США, публикации которых предполагают наличие пересмотров, в период со II квартала 1996 г. по I квартал 2011 г. он осуществил всесторонний анализ временных рядов. Клементс провел тестирование свойств пересмотров данных и классифицировал переменные на показатели, подверженные новостным изменениям, изменениям, содержащим шум, и показатели, относительно которых невозможно однозначно отнести к первым двум категориям. Дальнейший анализ показал, что гипотеза о правильном безусловном покрытии интервальных оценок, полученных EOS-методом, отвергается для большего числа переменных, чем в случае доверительных интервалов, полученных RTV-методом, и отклонения в основном связаны со смещением. Кроме этого, более высокая точность определения интервалов для RTV-данных определяется более точными оценками дисперсии прогноза, а не точностью точечных оценок.

B. Линейная регрессионная модель

Выбор данных для оценки прогнозирующей модели показателя отражается на ее прогностических способностях не только в случае ARIMA – спецификации. В работе (Koenig, Dolmas, Piger, 2003) рассмотрены результаты оценки параметров линейной модели взаимосвязи исследуемого показателя $\tilde{y}_t$ и вектора объясняющих переменных $\tilde{x}_t$, которые могут включать, в том числе, и лаги y (формула (33)):

$$\tilde{y}_t = \alpha\tilde{x}_t + \varepsilon_t. \tag{33}$$

В соответствии с EOS – подходом при оценивании параметров модели $\hat{\alpha}_1$ и для моделируемого показателя и для объясняющих переменных используются последние опубликованные на момент построения оценок данные (формула (34)):

$$y_t^T = \hat{\alpha}_1 x_t^T + \varepsilon_{t,1}. \tag{34}$$

Такая спецификация приводит к несостоятельным МНК-оценкам параметра α. Смещение оценки зависит от природы пересмотров используемых переменных.

Если $\xi_t^T = x_t^T - x_t^{t+1}$ является случайной составляющей, некоррелированной с x_t^T и y_t^T, то вероятностный предел параметра составит $plim\hat{\alpha}_1 = \alpha + \Sigma_{X_T X_T}^{-1}\Sigma_{\xi\xi}\alpha$, где $\Sigma_{X_T X_T} = plim\frac{X_{T-1}^{/}X_{T-1}}{T-1}$; $\Sigma_{\xi\xi} = plim\frac{\Xi/\Xi}{T-1}$ (X_T и Ξ – матрицы размером (T – 1) x k, в которых t-строки равны x_t^T и ξ_t^T соответственно).

Если ξ_t^T характеризуют новую информацию, некоррелированную со всеми переменными, то $plim\hat{\alpha}_1 = \alpha + \left(\Sigma_{XX} + \Sigma_{\xi\xi}\right)^{-1}\left[\Sigma_{\xi\nu} - \Sigma_{\xi\xi}\alpha\right]$, где $\Sigma_{XX} = plim\frac{X/X}{T}$; $\Sigma_{\xi\nu} = plim\frac{\Xi/\mathrm{N}}{T}$, ($X$ матрица размером (T – 1) x k, в которой t-строка равна $\tilde{x}_t$, а N – вектор размера T x 1, состоящий из $\nu_t^T = y_t^T - y_t^{t+1}$.

Проблема использования последних опубликованных оценок показателей состоит в том, что получаемые временные ряды являются неоднородными с точки зрения пересмотра данных: в начале периода оценки показателей являются многократно или окончательно пересмотренными, а в конце периода – предварительными или частично пересмотренными с учетом доступной на последний момент времени информации. В результате, связи между переменными левой и правой частей уравнения в начале выборочного периода могут сильно отличаться от связей в конце выборочного периода. Но именно последние связи актуальны для построения точного текущего прогноза.

RTV – подход может быть реализован двумя способами. Во-первых, прогнозирующее уравнение (35) может быть оценено по данным реального винтажа (значение в каждой точке в пределах выборки является последним доступным на тот момент времени) в обеих частях уравнения:

$$y_t^{t+1} = \hat{\alpha}_2 x_t^{t+1} + \varepsilon_{t,2}. \quad (35)$$

Во-вторых, в левой части уравнения могут быть использованы последние опубликованные на данный момент данные, а в правой части – данные реального винтажа (уравнение (36)):

$$y_t^T = \hat{\alpha}_3 x_t^{t+1} + \varepsilon_{t,3}.$$

(36)

Для первого RTV – подхода ошибка уравнения $\varepsilon_{t,2} = \varepsilon_t - \nu_t$, где $\nu_t = \tilde{y}_t - y_t^{t+1}$. Некоррелированость x_t^{t+1} и ν_t позволяет использовать для оценивания метод наименьших квадратов, который приводит к несмещенным оценкам параметра α. Для второго RTV – подхода ошибка равна $\varepsilon_{t,3} = \varepsilon_t - (\tilde{y}_t - y_t^T)$. И при условии того, что x_t^{t+1} и $(\tilde{y}_t - y_t^T)$ некоррелированы, МНК-оценки параметра α также являются несмещенными.

Ошибки двух RTV – подходов связаны соотношением (37):

$$\varepsilon_{t,3} = \varepsilon_{t,2} + \nu_t^T = \varepsilon_{t,2} + (y_t^T - y_t^{t+1}),$$

(37)

где $\nu_t^T = y_t^T - y_t^{t+1}$ – разница между последней и начальной оценками $\tilde{y}_t$.

Так как $\varepsilon_{t,2} = y_t^{t+1} - \hat{\alpha}_2 x_t^{t+1}$, то $\varepsilon_{t,2}$ некоррелирован с ν_t^T при условии, что y_t^{t+1} является эффективной оценкой y_t^T. Следовательно, ошибка второго способа реализации RTV – подхода будет иметь большую дисперсию, чем ошибка первого способа: $D(\varepsilon_{t,3}) = D(\varepsilon_{t,2} + \nu_t^T) \geq D(\varepsilon_{t,2})$. При этом, дисперсия ошибки второго способа будет уменьшаться по t, но всегда оставаться не меньше дисперсии ошибки первого способа. В итоге оценивание по модели $y_t^{t+1} = \hat{\alpha}_2 x_t^{t+1} + \varepsilon_{t,2}$ дает более точные оценки параметров по конечной выборке, обеспечивая большую эффективность прогнозирования.

Преимущество первого RTV – подхода просто объяснить, не только с точки зрения статистики, но и чисто логически. Ошибка этой модели представляет собой разность между оценкой показателя, предоставляемой статистическим агентством, и оценкой, полученной исследователем по имеющимся у него данным. Так что дисперсия этой ошибки определяется различиями их информационных наборов. Ошибка модели второго RTV – подхода формируется, кроме этого, еще и разностью между оценками агентства разных винтажей, характеризующей расширения информационного набора агентства с течением времени. Если официальные прогнозы эффективны, то эти две составляющие не коррелированы, так что дисперсия ошибки первого RTV – подхода не превосходит ошибки второго RTV – подхода.

В качестве эмпирического доказательства различий эффективности прогнозов в зависимости от включаемых в модель данных авторы (Koenig, Dolmas, Piger, 2003) приводят результаты рекурсивных прогнозов темпов роста ВВП в период с I квартала 1990 г. по IV квартал 1997 г. и их сравнение с фактическими значениями. Расчеты авторов подтверждают, что при построении прогнозов показателей нельзя не учитывать факт пересмотра данных. Качественные характеристики прогнозов, полученных в рамках EOS – подхода существенно уступают RTV – оценкам. Так что при построении прогнозирующей модели в правую часть уравнения необходимо включать данные реального винтажа, т.е. данные, которые были опубликованы в тот момент времени. В большинстве случаев использование RTV – данных для оценки параметров прогнозирующей модели оказываются более эффективным даже в случае, когда задача состоит в прогнозировании пересмотренных данных.

Кроме этого, если оценки, публикуемые статистическими агентствами, полностью используют имеющуюся на момент построения оценки информацию, то более высокой эффективности прогнозирования можно достигнуть, если использовать данные реального винтажа и в правой части уравнения, т.к. их использование в качестве зависимой переменной устраняет непредсказуемый шум из уравнения и дает более точные оценки коэффициентов.

Г. VAR - модели

Кишор и Кениг (Kishor, Koenig, 2012) отмечают, что пересмотры данных представляют существенную проблему и в VAR – прогнозировании. Как правило, показатели, включаемые в VAR – модель для формирования прогноза, представляют собой либо первые оценки, либо незначительно пересмотренные данные, а выборка, используемая для оценки VAR, состоит в основном из пересмотренных данных.

Для решения этой проблемы они предлагают легко реализуемую методологию оценки и прогнозирования VAR – модели, основанную на нескольких предположениях относительно характера пересмотров данных. Авторы допускают, что публикуемые оценки являются эффективными после определенного числа пересмотров (эти эффективные прогнозы рассматриваются в дальнейшем, как истинные значения показателя). Для меньшего числа пересмотров ошибки могут представлять собой сумму истинного значения и опечаток или сумму истинного значения и погрешностей измерения, причем для погрешностей допускается возможность коррелированности.

Процедура оценивания Кишора и Кенига состоит из трех этапов. Первый предполагает построение OLS оценок параметров VAR – модели по полностью пересмотренным данным. Продолжая свою аналогию с яблоками (полностью пересмотренными данными) и апельсинами (данными, которые еще подвержены переоценке), авторы связывают яблоки с яблоками на основе уравнения состояния VAR – модели. На втором этапе с помощью OLS оцениваются коэффициенты уравнений для данных более поздних периодов. Так получают уравнения наблюдений, определяющие зависимость апельсинов от яблок. Третий этап состоит в получении оценок истинных значений показателей для более поздних периодов на основе применения фильтра Калмана к уравнениям состояния и наблюдений и построении прогнозов на будущие периоды. Во фруктовой терминологии осуществляется переход от текущих апельсинов к яблокам.

Эмпирические примеры, приведенные Кишором и Кенигом, свидетельствует, что их процедура позволяет получить эффективные оценки в случае регулярных пересмотров и незначительных методологических изменений. Крупные изменения в методологии получения публикуемых статистическим агентством оценок, плохо отражаются моделью и должны рассматриваться как структурные разрывы, обрабатываемые другими средствами (например, фиктивными переменными).

Вместе с тем, расчеты, представленные в работе (Clements, Galvão, 2013) показывают, что модель Кишора – Кенига уступает по прогностическим способностям ARIMA – моделям при краткосрочном прогнозировании, но превосходит одномерные модели – в долгосрочной перспективе.

Проблем, возникающих в результате пересчета данных, можно избежать, если рассматривать VAR – модель для показателей одного и того же винтажа (см. (Clements, Galvao, 2012), называемой V-VAR (формула (38)):

$$y^{t+1} = c + \sum_{i=1}^{p} \Gamma_i y^{t+1-i} + \varepsilon^{t+1} \quad , \quad (38)$$

где $y^{t+1} = \left(y_t^{t+1}, y_{t-1}^{t+1}, \dots, y_{t-q+1}^{t+1}\right)^{/}$, $y^{t+1-i} = \left(y_{t-i}^{t+1-i}, y_{t-1-i}^{t+1-i}, \dots, y_{t-q+1-i}^{t+1-i}\right)^{/}$;

переменные c и ε^{t+1} имеют размерность $q \times 1$;

матрица Γ_i имеет размерность $q \times q$;

V-VAR – модель оценивает динамику последовательных наборов публикуемых данных, которые включают как первые опубликованные оценки y_t^{t+1}, так и пересмотренные оценки "прошлых" наблюдений $y_{t-1}^{t+1}, \dots, y_{t-q+1}^{t+1}$.

В большинстве случаев в присутствии пересмотров VAR – модель первого порядка (p = 1) достаточна для характеристики динамики y^{t+1}. Причем при p = 1 она тесно связана с одномерными моделями, в частности ее первое уравнение совпадает с ARIMA – моделью, рассмотренной ранее. И на него распространяются все выводы относительно прогностических свойств оценок, полученных для разных типов данных. Однако, VAR – модель позволяет получить более полную картину прогнозов в сравнении с одномерной моделью за счет прогнозирования значений показателя различных винтажей.

Помимо самой V-VAR – модели Клементс и Гальвао (Clements, Galvao, 2012) в своей работе рассматривают две ее модификации. Первая – SV-VAR (seasonal vintage-based VAR) – позволяет учесть периодичность публикаций пересмотренных данных. С этой целью в модель включаются дамми – переменные и модель преобразуется к виду (39) (например, при p = 1):

$$y^{t+1} = [\tilde{c} + \tilde{\Gamma}_1 y^t](1 - D_s^{t+1}) + [c + \Gamma_1 y^t] D_s^{t+1} + \varepsilon^{t+1},$$

(39)

где $\tilde{\Gamma}_1 = \begin{bmatrix} & \gamma_{2\times q} & \\ 0_{(q-2)\times 1} & I_{(q-2)\times(q-2)} & 0_{(q-2)\times 1} \end{bmatrix}$;

$\tilde{c} = (c_1, c_2, 0, \dots, 0)^{/}$.

Вторая модификация RV-VAR (news-restricted vintage-based VAR) предполагает, что после небольшого числа пересмотров дальнейшие переоценки становятся непредсказуемыми. Т.е. после проведения (n – 1) пересмотра данных следующая публикуемая оценка показателя y_t^{t+n+1} является эффективным прогнозом в том смысле, что пересмотр от y_t^{t+n} до y_t^{t+n+1} непредсказуем, т.е. $M[y_t^{t+n+1} - y_t^{t+n} | \boldsymbol{y}^{t+n}] = 0$, тогда как $\left[y_t^{t+i+1} - y_t^{t+i} \middle| \boldsymbol{y}^{t+i}\right] \neq 0$ для $i < n$. В рамках этого предположения $M[y_{t-n}^{t+1} | y_{t-n}^{t}] = y_{t-n}^{t}$, что позволяет уменьшить число оцениваемых коэффициентов. Ограничения, накладываемые на модель $y^{t+1} = c + \sum_{i=1}^{p} \Gamma_i y^{t+1-i} + \varepsilon^{t+1}$ реализуются в матрицах Γ_i вида (40):

$$\Gamma_1 = \begin{bmatrix} & \gamma_{n\times q} & \\ 0_{(q-n)\times(n-1)} & I_{(q-n)\times(q-n)} & 0_{(q-n)\times 1} \end{bmatrix}, \ \Gamma_i = \begin{bmatrix} \gamma_{i,n\times q} \\ 0 \end{bmatrix} \ (i = 2, \dots, p).$$

(40)

Полученные по данным в период с VI квартала 1965 г. по III квартал 2006 г результаты (Clements, Galvao, 2012) подтверждают, что путем использования многовинтажных моделей качество прогнозных оценок показателей могут быть улучшены для прогнозирования как пересмотров данных, так и пересмотренных значений будущих наблюдений. Прогностические свойства различных спецификаций VAR – моделей свидетельствуют о предсказуемости пересмотров прошлых данных (модели VAR в большинстве случаев более эффективны в сравнении с базовыми одномерными моделями). При сравнении между VAR – моделями в большинстве случаев SV-VAR уступает V-VAR и RV-VAR, а RV-VAR уступает V-VAR.

Еще один положительный аспект прогнозирования по VAR – моделям отмечается в другой работе тех же авторов (Clements, Galvao, 2013), в которой Клементс и Гальвао рассматривают VAR – прогнозы на h шагов вперед. В момент времени (T + 1) информационная база, используемая при оценивании, включает данные всех винтажей до (T + 1) включительно, т. е. наборы y^{t+1} для t = 1, 2, …T. Прогнозы на h шагов вперед вектора y^{T+1+h} определяются как условное ожидание, для заданной модели и располагаемой информации (формула (41)):

$$y^{T+1+h|T+1} = M(y^{T+1+h}|y^{T+1}, y^{T}, \dots).$$

(41)

Элементы вектора $y^{T+1+h|T+1} = \left(y_{T+h}^{T+1+h|T+1}, \dots, y_{T+h-q+1}^{T+1+h|T+1}\right)$, представляют собой прогнозы первой оценки показателей, наблюдаемых в момент $(T+h)$, прогнозы второй оценки показателей, наблюдаемых в момент $(T+h-1)$, и т.д. до q - й оценки показателей, наблюдаемых в момент $(T+h-q+1)$.

Если данные подвергаются q пересмотрам (причем, чтобы учесть все регулярные изменения оценок q может быть достаточно большим), то $\overline{y}_t = y_t^{t+q}$. Поэтому если исследователю требуется получить прогнозы "истинных" (полностью пересмотренных) значений показателя, т.е. $\overline{y}_{T-1}, \overline{y}_{T}, \overline{y}_{T+1}, \overline{y}_{T+2}$, то он должен рассмотреть прогнозы подмножества элементов $y^{T+1+h|T+1}$, а именно последних элементов каждого вектора для h = 1; 2; 3; … ; h . Для винтажа $(T+1)$ это приводит к множеству прогнозов: $y_{T+2-q}^{T+2|T+1}, y_{T+3-q}^{T+3|T+1}, \dots, y_{T+h-q+1}^{T+1+h|T+1}$, которые являются оценками q-го уровня. Некоторые из них будут относиться к периодам, предшествующим винтажному $(T+1)$, а другие – к самому винтажному периоду. Когда $h-q+1 \leq 0$, исследователь получает прогнозы прошлых наблюдений (или "обратные трансляции"), $\overline{y}_{T}, \overline{y}_{T-1}, \overline{y}_{T-2}$, ..., а для $h-q+1 > 0$ – прогнозов "полностью пересмотренных" будущих наблюдений: $\overline{y}_{T+1}, \overline{y}_{T+2}$, ...

Д. Пересмотры и прогнозирование на основе DSGE-моделей

В настоящее время для прогнозирования последствий систематических политических и экономических изменений часто применяют динамические стохастические модели общего равновесия (DSGE), при оценке параметров которых, как правило, используются временные ряды показателей, доступные на момент оценивания (EOS – подход). Но эти данные, как и в случае более простых моделей, подвержены пересмотрам, что ведет к ухудшению свойств оценок.

Пространство состояний модели DSGE описывается уравнением состояния (42):

$$y_t = F(\theta)y_{t-1} + G(\theta)\nu_t,$$

(42)

где y_t – вектор эндогенных переменных модели DSGE размерности $n \times 1$ (y_t представляют собой детрендированные значения показателей в отклонениях от значений стационарного состояния и могут включать в себя запаздывающие переменные);

ν_t – вектор структурных шоков размерности $r \times 1$, таких что $\nu_t \sim iidN(0, Q)$, а Q является диагональной матрицей;

θ – вектор структурных параметров;

$G(\theta)$ – матрица размерности $n \times r$.

Уравнение измерений задается в виде (43):

$$\tilde{Y}_t = d(\theta) + H(\theta)y_t,$$

(43)

где $\tilde{Y}_t$ – вектор истинных значений эндогенных переменных размерности $m \times 1$ (как правило, $m < n$ и $m \leq r$), т. е. набор наблюдаемых переменных измеряется без ошибок, хотя Галвао (Galvão, 2017) показывает, что этот подход будет работать, и при наличии ошибок измерения.

Если динамические стохастические модели общего равновесия оцениваются на основе показателей, публикуемые оценки которых предполагают проведение q пересмотров, то есть, когда $\tilde{Y}_t = Y_t^{t+q}$, то временные ряды состоят из полностью пересмотренных наблюдений и наблюдений, которые только опубликованы или пересмотрены частично. Чтобы избежать проблем, возникающих при использовании неоднородных (с точки зрения пересмотров) данных, можно сократить длину временных рядов и рассматривать только полностью пересмотренные наблюдения. Такой подход применяется Крушором и Силом в работе (Croushore, Sill, 2014), которые оценивают уравнение измерений в виде (44):

$$Y_{t-q}^T = d(\theta) + H(\theta)y_{t-q},\ t = 1, \dots, T-1. \quad (44)$$

Но в этом случае отсутствие информации о последних значениях показателей приводит к существенному снижению точности прогнозирования будущих значений.

Второй подход состоит в оценивании модели реальных винтажей в виде (45):

$$Y_{t-q}^{t} = d(\theta) + H(\theta) y_{t-q}, \; t = 1, \dots, T-1.$$

(46)

Но и в этом случае q наблюдений должны быть исключены из выборки.

Альтернативой является метод, предложенный Гальвао (Galvão, 2017) для совместной оценки параметров модели DSGE (с пересмотренными данными) и параметров, описывающих сам процесс пересмотра данных. На основе этой модели можно вычислить обратные трансляции и прогнозы для данных, подлежащих пересмотру, включая их базовую прогностическую плотность.

Автор определяет стандартизированные наблюдаемые изменения между первыми опубликованными оценками Y_t^{t+1} и истинными значениями Y_t^{t+q} как (46):

$$rev_t^{t+q,1} = \left(Y_t^{t+1} - Y_t^{t+q}\right) - M_1, \; t = 1, \dots, T-q,$$

(46)

а стандартизированные наблюдаемые изменения между оценками, опубликованными после k пересмотров Y_t^{t+k} и истинными значениями Y_t^{t+q} как (47):

$$rev_t^{t+q+1-k,k} = \left(Y_t^{t+k} - Y_t^{t+q}\right) - M_k, \; t = 1, \dots, T-q+\nu, k = 1, \dots, q-1.$$

(47)

где M_1 и M_k - векторы средних пересмотров размерности $m \times 1$.

Для сохранения длины временного ряда (включения всех имеющихся на момент оценивания точек данных) Гальвао переписывает уравнение измерений в виде (48):

$$\begin{bmatrix} Y_t^{t+1} \\ Y_{t-1}^{t+1} \\ \vdots \\ Y_{t-q+1}^{t+1} \end{bmatrix} = \begin{bmatrix} d(\theta) + M_1 \\ d(\theta) + M_2 \\ \vdots \\ d(\theta) \end{bmatrix} + \begin{bmatrix} H(\theta) & 0_m & \dots & 0_m & I_m & 0_m & \dots & 0_m \\ 0_m & H(\theta) & \dots & 0_m & 0_m & I_m & \dots & 0_m \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_m & 0_m & \dots & H(\theta) & 0_m & 0_m & \dots & 0_m \end{bmatrix} \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-q+1} \\ rev_t^1 \\ rev_{t-1}^2 \\ \vdots \\ rev_{t-q+2}^{q-1} \end{bmatrix},$$

(48)

где $rev_t^k = \left(X_t^{t+k} - X_t\right) - M_k, \; k = 1, \dots, q-1$ определяют пересмотры данных.

Уравнения состояния дополняются уравнениями процессов пересмотра данных следующим образом (формула (49)):

$$rev_t^k = K_k rev_{t-1}^k + \xi_t^k + A_k \nu_t, \; \xi_t^k \sim N(0, R_k).$$

(49)

Такая спецификация пересмотров соответствует структуре, предложенной Джейкобсом и Ван Норденом (Jacobs, van Norden, 2011), и учитывает как новостные, так и шумовые составляющие пересмотров. Слагаемое ξ_t^k (own innovation term) допускает наличие изменений оценок, связанных со снижением погрешностей измерений, относительно которых полагается, что они некоррелированы с переменными, так что матрица R_k является диагональной.

Последнее слагаемое $A_k \nu_t$ позволяет учесть пересмотры данных, вызванные наличием новой информации, не доступной на момент текущей публикации оценки показателя, но включенной в пересмотренные данные, используемых для оценки величины структурных шоков. Для новостных пересмотров предполагается их коррелированность с истинными значениями показателей и допускается зависимость от методов обработки информации статистическим агентством, приводящих к коррелированности обновлений оценок и структурных шоков.

Кроме этого, спецификация модели допускает наличие автокорреляции пересмотров, отмеченной Кишором и Кенигом (Kishor, N. K. and Koenig, 2012), если матрица K_k (размерности $m \times m$) ненулевая.

Предложенный Гальвао метод оценки DSGE – модели основан на подходе Metropolis-in-Gibbs, и предполагает совместную оценку DSGE параметров $\{\theta, Q\}$ и параметров процесса пересмотра: $\{M_1, \dots, M_v; K_1, \dots, K_k; A_1, \dots, A_k; R_1, \dots, R_k\}$. Метод реализуется в четыре этапа. На первом шаге вычисляются апостериорных распределений для θ, с применением RWMH – алгоритма (Random-Walk Metropolis-Hasting) для пространства состояний, определяемого уравнением состояния и уравнениями измерений в виде $Y_{t-q}^t = d(\theta) + H(\theta) y_{t-q}$ (т.е. для оценки ковариационной матрицы $\hat{\Sigma}_\theta$ используется не все наблюдения временных рядов, а только те, которые полностью пересмотрены). Второй шаг состоит в построении оценок временного ряда структурных шоков ν_t на основе значений показателей и полученных ранее оценок θ с помощью рекурсивного алгоритма сглаживания для модели пространства состояний. На третьем этапе оцениваются условные (при заданных значениях переменных и структурных шоков) распределения параметров, определяющих пересмотры, в предположении нормального и обратного гамма законов распределения. На четвертом шаге, применяя алгоритм сглаживания к пространству состояний, строится полный (для всего рассматриваемого периода времени) вектор показателей.

Результаты эмпирической проверки, представленные Гальвао (Galvão, 2017), полученные по данным в период с I квартала 1984 г. по VI квартал 2008 г, показывают, что краткосрочные прогнозы показателей, полученные в соответствии с DSGE – моделью для последних опубликованных данных (EOS – подход) уступают по качеству прогнозам по RTV – данным. Прогностические свойства DSGE – моделей могут быть улучшены (особенно в отношении интервальных оценок) если при оценивании по RTV – данным учитывается структура пересмотров.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Altavilla, Carlo, and Matteo Ciccarelli. Monetary Policy Analysis in Real-Time. Vintage Combination from a Real-Time Dataset//CESifo Working Paper Series 3372, CESifo Group Munich - 2014.Andreou, E., Ghysels, E., and Kourtellos, A. Forecasting With Mixed-Frequency Data//The Oxford Handbook of Economic Forecasting, eds. M. P. Clements and D. F. Hendry, New York: Oxford University Press - 2011 - P.225–246.
2. Aruoba, S. B. Data Revisions are Not Well-Behaved//Journal of Money, Credit and Banking - 2008 - Vol. 40. P. 319–340.
3. Berger, A. N., & Krane, S. D. The informational efficiency of econometric model forecasts//The Review of Economics and Statistics - 1985 - Vol. 6. P. 128–134.
4. Clements, M.P. Assessing Macro Uncertainty in Real-Time When Data Are Subject to Revision//Journal of Business & Economic Statistics - 2017 - P. 420-433
5. Clements, M. P., and Galvao, A. B. Improving Real-Time Estimates of Output Gaps and Inflation Trends With Multiple-Vintage VAR Models//Journal of Business and & Economic Statistics - 2012 - Vol. 30. P. 554–562.
6. Clements, M. P., and Galvao, A. B. Forecasting With Vector Autoregressive Models of Data Vintages: US Output Growth and Inflation//International Journal of Forecasting – 2013 - Vol. 29. P. 698–714.
7. Clements, M.P., Galvão, A.B. Data Revisions and Real-Time Forecasting, Centre - Discussion Paper// ICMA, Henley Business School, Reading - 2018.
8. Croushore, D. Forecasting with real-time data vintages//chapter 9. In Clements, M. P., and Hendry, D. F. (eds.), The Oxford Handbook of Economic Forecasting - 2011 - P. 247.267.
9. Croushore, D. and Sill, K. Analysing data revisions with a dynamic stochastic general equilibrium model//Federal Reserve Bank of Philadelphia Working Paper - 2014.
10. Faust, J., Rogers, J. H., and Wright, J. H. News and noise in G-7 GDP announcements//Journal of Money, Credit and Banking - 2005 - Vol. 37 (3). P. 403.417.

11. Faust, Jon, and Jonathan H. Wright. Comparing Greenbook and Reduced Form Forecasts using a Large Realtime Dataset//Journal of Business and Economic Statistics - 2009 - Vol. 27. P. 468–479.

12. Galvão, A. B. Data revisions and DSGE models//Journal of Econometrics - 2017 - Vol. 196 (1). P. 215.232.

13. Garratt, A., Lee, K., Mise, E., and Shields, K. Real Time Representations of the Output Gap//Review of Economics and Statistics - 2008 - Vol. 90. P. 792–804.Gilbert, T. Information Aggregation Around Macroeconomic Announcements: Revisions Matter//Journal of Financial Economics - 2011 - Vol. 101. P. 114–131.

14. Harvey, A.C., C.R. McKenzie, D.P.C. Blake, and M.J. Desai. Irregular Data Revisions//in Arnold Zellner, ed., Applied Time Series Analysis of Economic Data. Washington, D.C.: U.S. Department of Commerce, Economic Research Report ER-5 - 1983 - P. 329–347.

15. Howrey, E. P. The use of preliminary data in econometric forecasting//The Review of Economics and Statistics - 1978 - Vol. 60(2). P. 193-200.

16. Jacobs, J. P. A. M., and van Norden, S. Modeling Data Revisions: Measurement Error and Dynamics of ‘True’ Values//Journal of Econometrics - 2011 - Vol. 161. P. 101-109.

17. Kishor, N. K. and Koenig, E. F. VAR estimation and forecasting when data are subject to revision//Journal of Business & Economic Statistics - 2012 - Vol. 30(2). P. 181–190.

18. Koenig, E. F., Dolmas, S., and Piger, J. The Use and Abuse of Real-Time Data in Economic Forecasting//The Review of Economics and Statistics - 2003 - Vol. 8. P. 618–628.

19. Lee, Kevin, Nilss Olekalns, and Kalvinder Shields. Nowcasting, Business Cycle Dating and the Interpretation of New Information when Real Time Data are Available//Working paper, University of Melbourne - 2008.

20. Mankiw, N. G., and Shapiro, M. D. News or Noise: An Analysis of GNP Revisions//Survey of Current Business US Department of Commerce, Bureau of Economic Analysis - 1986. P. 20–25.

21. Sargent, Thomas J. Macroeconomic Theory, 2nd ed. Boston: Academic Press - 1987.

22. Strohsal, T., Wolf, E. Data revisions to German national accounts: Are initial releases good nowcasts//Diskussionsbeiträge, Freie Universität Berlin, Fachbereich Wirtschaftswissenschaft, Berlin - 2019.