

УДК 338.27

ОЦЕНКА ЗОЛОВОАЛЮТНЫХ РЕЗЕРВОВ В РАМКАХ ДИНАМИЧЕСКОЙ МОДЕЛИ ДЛЯ ПАНЕЛЬНЫХ ДАННЫХ

Бабешко Л.О., Орлова И.В., Бронников Е.В.

ФГОБУ ВО «Финансовый университет при Правительстве РФ», Москва, e-mail: LBabeshko@fa.ru

Целью данного исследования является анализ и моделирование золотовалютных резервов восьми стран: Германии, Франции, США, Китая, России, Нидерландов, Швейцарии и Японии. В качестве эконометрического инструментария выбран аппарат моделей для панельных данных, позволяющий учитывать индивидуальную динамику объектов и повышающий точность оценок. Апробация моделей выполнена по панельным данным запасов золота в восьми странах за семнадцать лет с 2000 по 2016 г., в программной среде R: замена пропущенных наблюдений – с помощью пакета Multivariate Imputation by Chained Equations, оценка параметров моделей и тестирование характера эффектов с учетом иерархической структуры эконометрических моделей для панельных данных – при помощи пакета plm. В качестве регрессоров модели использованы лаговые значения ВВП каждой страны и запас золота. В динамических моделях для панельных данных проблема эндогенности регрессоров присутствует независимо от природы индивидуального эффекта. Поэтому для решения этой проблемы в работе применен метод инструментальных переменных, не требующий внешних инструментов (подход Хаусмана – Тейлора). В результате эмпирического тестирования подтверждена гипотеза положительной связи запасов золота от выбранных регрессоров.

Ключевые слова: панельные данные, объединенная модель, модель со случайным эффектом, модель с фиксированным эффектом, обобщенный тест Дарбина – Уотсона, эндогенность, метод инструментальных переменных

EVALUATION OF GOLD-RESERVE RESERVES UNDER FRAMEWORK DYNAMIC MODEL FOR PANEL DATA

Babeshko L.O., Orlova I.V., Bronnikov E.V.

The Financial University under the Government of the Russian Federation, Moscow, e-mail: LBabeshko@fa.ru

The purpose of this study is to analyze and model the gold and foreign exchange reserves of eight countries: Germany, France, the USA, China, Russia, the Netherlands, Switzerland and Japan. As an econometric tool, the model apparatus for panel data is chosen, which allows to take into account the individual dynamics of objects, and increases the accuracy of estimates. Aprrobation of the models was carried out according to the panel data of gold reserves in eight countries for seventeen years from 2000 to 2016 in the software environment R: the replacement of missed observations with the Multivariate Imputation by Chained Equations package, the estimation of model parameters and testing of the character of the effects taking into account the hierarchical structures of econometric models for panel data – using the plm package. The lag values of the GNP of each country and the gold reserve are used as regressors of the model. In dynamic models for panel data, the problem of endogenous regressors is present irrespective of the nature of the individual effect. Therefore, to solve this problem, a method of instrumental variables that does not require external tools (the Hausman-Taylor approach) is applied in this paper. As a result of empirical testing, the hypothesis of a positive correlation of gold reserves from the selected regressors is confirmed.

Keywords: panel data, pooled model, random effect model, fixed effect model, generalized Darbin-Watson test, endogenous, instrumental variables estimator

Целью исследования является анализ и моделирование золотовалютных резервов с привлечением аппарата моделей для панельных данных. Статистическая выборка панельных данных, представляющих динамику запасов золота восьми стран: Германии, Франции, США, Китая, России, Нидерландов, Швейцарии и Японии – за семнадцать лет с 2000 по 2016 г. была сформирована из открытых баз данных Всемирного банка [1] и Всемирного золотого совета [2]. В качестве регрессоров модели использован ВВП каждой страны и запас золота (лаговые переменные). При формировании панельных данных была решена задача замены пропущенных значений, которые требовали либо удаления, что существенно сужает исследовательские

возможности, либо их замещения. Замена пропущенных наблюдений выполнена с помощью пакета Multivariate Imputation by Chained Equations, предоставляемым языком программирования R.

Страны, которые мы рассматриваем в нашем исследовании, можно разделить на две группы. В первую группу входят технически развитые страны. В этих странах золото широко используется в различных областях техники и промышленных отраслях, а также для изготовления ювелирных изделий. К этим странам относятся Япония, США и Германия. В этих странах золото выступает как индикатор развития высоких технологий. За рассматриваемый период в этих странах не наблюдается тенденций к росту запасов золота. Вторая группа – это страны, интен-

сивно наращивающие запасы золота. Это Россия и Китай (рис. 1–2). Россия активно скупает золото. Причиной этого могут быть санкции США и Евросоюза, которые вынуждают российское руководство относиться к резервам в долларах и евро с большой осторожностью. Китай также стремится увеличивать свои золотые резервы.

Многие аналитики считают, что Китай и Россия, скорее всего, будут увеличивать запасы золота для защиты от экономической неопределённости в будущем.

В настоящее время наблюдается ситуация, когда финансовые власти России продают долговые обязательства правительства США ускоренными темпами. При этом Центробанк России продолжает скупать золото у российских золотодобытчиков максимальными за последние 12 лет объёмами. Сейчас Россия находится на пятом месте среди стран – держателей золота, обогнав недавно Китай. Впереди находятся США, Германия, Италия и Франция [3].

Выбор эконометрического инструментария

Выборочные данные состоят из наблюдений однотипных объектов (8 стран) в течение небольшого периода времени (17 лет). Поэтому в качестве инструмента выбран аппарат моделей для панельных данных, позволяющий выполнять оценивание при небольшом объеме выборки для отдельных экономических единиц [4–6]. Одним из преимуществ аппарата панельных данных является способность учитывать индивидуальную динамику (динамику на «индивидуальном уровне» [7, 8]), что особенно важно при моделировании зависимости объёма золотых резервов от ВВП страны, в которой текущий объем существенно зависит

от объема и ВВП предыдущего года. Основными регрессионными моделями, применяемыми к панельным данным, являются:

- объединённая модель (*Pooled model, Pool*): спецификация не учитывает индивидуальные особенности объектов

$$y_{it} = \mu_i + x_{it} \cdot \beta + \varepsilon_{it}, \mu_i = \text{const} = \mu, \quad (1)$$

где y_{it} – зависимая переменная, x_{it} – вектор строка регрессоров (размерностью k), ε_{it} – случайное возмущение; $i = 1, \dots, n$, $t = 1, \dots, T$, n – число объектов (панелей), T – число наблюдений в рамках одной панели, $\mu = \text{const}$ – постоянное для всех объектов значение свободного члена – параметр «местоположения», $\beta = \beta_{\text{pool}}$ – вектор параметров «влияния», постоянный для всех объектов (панелей);

- модель с фиксированными эффектами (*fixed effect model, FE*): гетерогенность объектов учитывается индивидуальными параметрами местоположения μ_i , $i = 1, \dots, n$

$$y_{it} = \mu_i + x_{it} \cdot \beta + \varepsilon_{it}, \mu_i \neq \text{const}, \quad (2)$$

- модель со случайными эффектами (*random effect model, RE*): гетерогенность объектов учитывается независимой от времени специфической составляющей ошибки m_i :

$$y_{it} = \mu_i + x_{it} \cdot \beta + \varepsilon_{it}, \mu_i \neq \text{const}, \mu_i = \mu + m_i,$$

$$y_{it} = \mu + x_{it} \cdot \beta + m_i + \varepsilon_{it} = \mu + x_{it} \cdot \beta + v_{it}. \quad (3)$$

В статических моделях (1)–(3) предполагается, что $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$, и оценки параметров получаются состоятельными и эффективными, при обоснованном выборе характера эффектов.

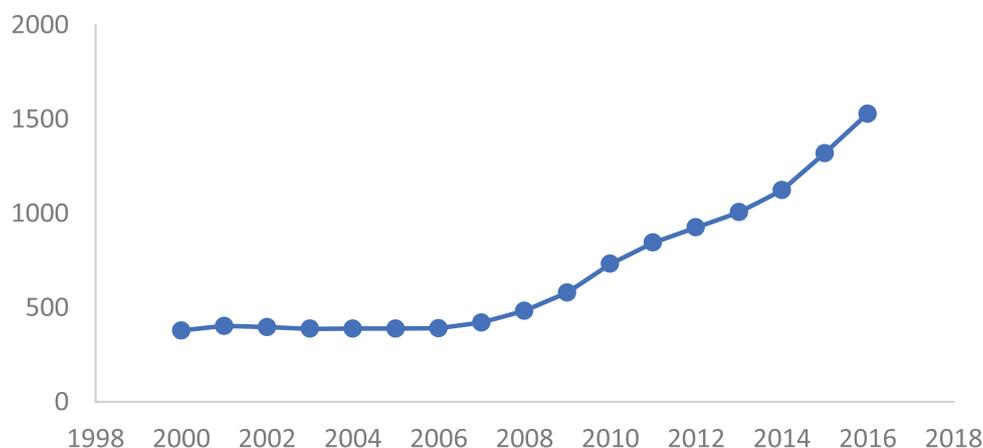


Рис. 1. Динамика запасов золота в России

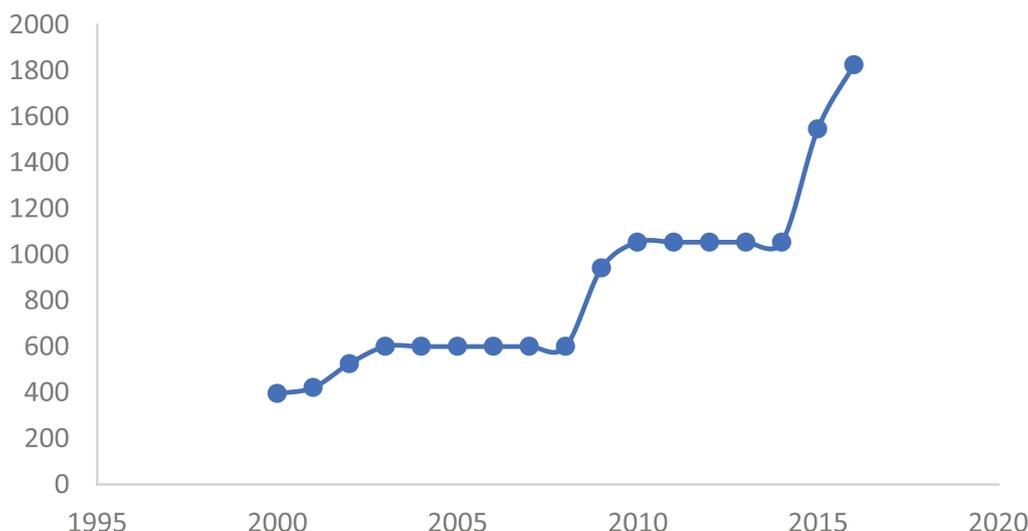


Рис. 2. Динамика запасов золота в Китае

Результаты оценивания

Формальное применение аппарата моделей для панельных данных приводит к следующему результату:

Объединенная модель:

$$\hat{y}_t = 57,806 + 0,999 \cdot y_{t-1} + 0,002 \cdot x_{t-1},$$

(15,171) (0,003) (0,0004)

$$R^2 = 0,9989, \quad (4)$$

$RSS = 900500$, $n = 8$ – число панелей, $T = 16$ – число наблюдений в рамках одной панели, $nT = 128$ – объем панельных данных. В скобках под оценками параметров приводятся их стандартные ошибки, RSS (Residuals Sum of Squares) – сумма квадратов остатков.

Модель с фиксированными эффектами:

$$\hat{y}_t = 1,029 \cdot y_{t-1} + 0,004 \cdot x_{t-1}, \quad R^2 = 0,9330,$$

(0,027) (0,0008)

$$RSS = 530350. \quad (5)$$

В оцененной модели (5) приводятся оценки параметров влияния (общие для всех панелей). Для получения их состоятельных оценок, в модели с фиксированными эффектами, используется процедура раздельного оценивания параметров влияния и местоположения. Для спецификации в матричном виде [9, 10]

$$Y^* = X^* \cdot \beta_w + \varepsilon^*, \quad (6)$$

где $Y^* = (I - P_w)Y$ – вектор значений эндогенных переменных, центрированных по индивидуальным средним (средним по каждой панели), $X^* = (I - P_w)X$ – матрица ре-

грессоров, центрированных по индивидуальным средним, $I - P_w = I - DD'/T$ – оператор внутригруппового преобразования (центрирования по индивидуальным средним),

$D = I_{n,n} \otimes I_T$ – матрица фиктивных переменных, учитывающая гетерогенность панельных данных в моделях с фиксированным эффектом,

$$\hat{\beta}_w = (X^* X^*)^{-1} X^* Y^*, \quad (7)$$

– вектор МНК-оценок параметров влияния. Оценки параметров местоположения вычисляются по оценкам параметров влияния:

$$\hat{\mu} = A_D Y - A_D X \hat{\beta}_w = \bar{Y} - \bar{X} \cdot \hat{\beta}_w, \quad (8)$$

где

$$A_D = (D'D)^{-1} D' = \frac{1}{T} I_{n,n} \cdot D' = \frac{1}{T} D'$$

– оператор формирования вектора индивидуальных средних. В табл. 1 приводятся оценки параметров местоположения, полученные в рамках раздельной процедуры оценивания (5)–(8).

Модель со случайными эффектами:

$$\hat{y}_t = 45,897 + 0,998 \cdot y_{t-1} - 0,001 \cdot x_{t-1},$$

(17,543) (0,004) (0,0004)

$$R^2 = 0,998, \quad RSS = 864160. \quad (9)$$

Минимальную сумму квадратов остатков обеспечивает модель с фиксированными эффектами. Адекватность этой модели подтверждается и тестированием характера гетерогенных эффектов. F -тест

($H_0 : \mu_i = \mu_j$, функция в программной среде R $pFtest()$) [11]:

$$F = \frac{(RSS_{pool} - RSS_{FE}) / (n-1)}{(RSS_{FE}) / (nT - n - k)} = 11,765,$$

$$p\text{-value} = 2,708e-11,$$

где k – число параметров влияния, RSS_{pool} – сумма квадратов остатков объединённой модели, RSS_{FE} – сумма квадратов остатков модели с фиксированными эффектами. Тест множителей Лагранжа (тестирование *Pooled*-модели против модели *RE*, $H_0 : \sigma_m^2 = 0$, (функция в программной среде R $plmtest()$):

$$LM = \frac{nT}{2(T-1)} \left(\frac{e'DD'e}{e'e} - 1 \right)^2 = 2,115,$$

$$p\text{-value} = 0,146,$$

где e – вектор-столбец остатков в объединённой модели, D – матрица фиктивных переменных, не выявил значимых случайных эффектов для различных стран. Тест Хаусмана (тестирование модели *RE* против модели *FE*, $H_0 : Cov\{\mu_i, x_{jt}\} = 0$, (функция в программной среде R $phtest()$):

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' \hat{\Sigma}^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}) = 84,733 > H_{\alpha=0,05}(2) = 5,99,$$

где $\hat{\Sigma}$ – оценка автоковариационной матрицы вектора $b = \hat{\beta}_{FE} - \hat{\beta}_{RE}$, отклоняет нулевую гипотезу об адекватности модели со случайными эффектами.

Тестирование наличия автокорреляции в моделях для панельных данных может быть выполнено при помощи обобщенной статистики Дарбина – Уотсона в рамках модели с фиксированными эффектами, так как оценки параметров модели с FE состоятельны и в случае адекватности модели со случайными эффектами [7, 12]:

$$DW_p = \frac{\sum_{i=1}^n \sum_{t=2}^T (e_{it} - e_{i,t-1})^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} = 1,203,$$

$$p\text{-value} = 2.212e-06. \quad (10)$$

Так как области неопределенности обобщенного теста являются малыми и изменения критических значений, в зависимости от параметров (n , T , K) незначительные, то полученный результат показывает наличие автокорреляции случайного возму-

щения в модели *FE*. Одной из причин автокорреляции является ошибка спецификации модели. К сожалению, в моделях (4), (5), (6) присутствует проблема эндогенности регрессоров, и применение «статического» аппарата моделей для панельных данных приводит к смещённым и несостоятельным оценкам, поскольку лаговое значение эндогенной переменной $y_{i,t-1}$ коррелирует с индивидуальным (фиксированным или случайным) эффектом μ_i , т.е. ни внутригрупповое (6), ни межгрупповое преобразования переменных не исключают проблему эндогенности.

Спецификация динамической модели, с учетом индивидуального эффекта, может быть записана следующим образом:

$$y_{it} = \mu_i + \beta_1 \cdot y_{i,t-1} + \beta_2 \cdot x_{it} + \varepsilon_{it}, \quad (11)$$

где y_{it} – текущее значение объёмов золота в золотовалютных резервах i -й страны, $y_{i,t-1}$ – лаговое значение объёмов золота в золотовалютных резервах i -й страны, x_{it} – текущее значение ВВП страны. Для корректировки моделей с эндогенными регрессорами используется метод инструментальных переменных МИП [13, 14]. Для выбора инструмента для лагового регрессора $y_{i,t-1}$ выполним следующие преобразования модели (11). Избавимся от индивидуальных эффектов путем перехода к первым разностям переменных

$$y_{it} - y_{i,t-1} = \beta_1 (y_{i,t-1} - y_{i,t-2}) + \beta_2 (x_{it} - x_{i,t-1}) + (\varepsilon_{it} - \varepsilon_{i,t-1}),$$

$$t = 3, \dots, T. \quad (12)$$

Из спецификации (12) следует, что

$$Cov\{y_{i,t-2}; y_{i,t-1} - y_{i,t-2}\} \neq 0,$$

$$Cov\{y_{i,t-2}; \varepsilon_{it} - \varepsilon_{i,t-1}\} = 0,$$

т.е. лаговая переменная $y_{i,t-2}$ может быть выбрана в качестве инструментальной переменной для регрессора $y_{i,t-1} - y_{i,t-2}$. Регрессор $x_{it} - x_{i,t-1}$ – экзогенный, поэтому может быть использован инструментом для самого себя. МИП-оценка параметров модели (12) в матричной форме имеет вид

$$\tilde{\beta} = (Z'X)^{-1} Z'Y, \quad (13)$$

где Z – матрица инструментальных переменных, X – матрица регрессоров, Y – вектор значений эндогенной переменной. Структура матриц МИП-оценки (13) для модели (12) представлена в табл. 2.

Таблица 1

Оценки параметров местоположения модели FE

Номер страны i	Страна	$\hat{\mu}_i$	$\hat{\sigma}_i$	t -статистика
1	Китай	51,658	28,472	1,814
2	Германия	-264,348	11,694	-2,367
3	США	-431,583	241,356	-1,788
4	Россия	21,402	26,746	0,800
5	Япония	-166,763	35,976	-4,635
6	Франция	-267,497	91,898	-2,911
7	Нидерланды	-222,088	49,561	-4,481
8	Швейцария	-408,438	78,278	-5,141

Таблица 2

Структура матриц МИП-оценки

$Z - (nT \times 2)$ – матрица инструментальных переменных		$X - (nT \times 2)$ – матрица регрессоров		Y – вектор эндогенных переменных
Z1	Z2	X1	X2	Y
$y_{i,t-2}$	$x_{it} - x_{i,t-1}$	$y_{i,t-1} - y_{i,t-2}$	$x_{it} - x_{i,t-1}$	$y_{it} - y_{i,t-1}$
$i = 1, \dots, n, t = 3, \dots, T$				

Получена следующая МИП-оценка (13) модели (12):

$$\hat{y}_{it} = y_{i,t-1} + \hat{\beta}_1 (y_{i,t-1} - y_{i,t-2}) + \hat{\beta}_2 (x_{it} - x_{i,t-1}), \quad t = 3, \dots, T, \quad i = 1, \dots, n, \quad (14)$$

$$\hat{y}_{it} = y_{i,t-1} + \underset{(0,0019)}{0,7515} (y_{i,t-1} - y_{i,t-2}) - \underset{(0,0024)}{0,00076} (x_{it} - x_{i,t-1}), \quad \hat{\sigma}_{IV}^2 = 4429,8. \quad (15)$$

Обобщенная статистика Дарбина – Уотсона (10): $DW_p = 2,15$, показывает отсутствие автокорреляции остатков в модели (14) с МИП-оценкой (15).

Использование инструментов Z в оценке (13) оправдано только в случае корреляции регрессоров с возмущениями модели. Если корреляция отсутствует ($H_0 : Cov\{x_{it}, \varepsilon_{is}\} = 0$), то МНК-оценки точнее МИП-оценок. Формальный тест, позволяющий определить, какой метод оценивания следует применять, это тест Хаусмана, основанный на сравнении оценок параметров в основной и альтернативной моделях. Статистика теста имеет вид квадратичной формы [15]:

$$H = (\hat{\beta}_{OLS} - \hat{\beta}_{IV})' \hat{\Sigma}^{-1} (\hat{\beta}_{OLS} - \hat{\beta}_{IV}) \sim \chi^2(m), \quad (16)$$

где m – число инструментов, $\hat{\Sigma}$ – оценка автоковариационной матрицы вектора $b = \hat{\beta}_{OLS} - \hat{\beta}_{IV}$, которая в рамках нулевой гипотезы принимает вид

$$\begin{aligned} \Sigma = C_{bb} &= Cov\{\hat{\beta}_{OLS} - \hat{\beta}_{IV}, b\} = Cov\{\hat{\beta}_{OLS}, b\} - Cov\{\hat{\beta}_{IV}, b\} = \\ &= C_{OLS} - C_{IV} = \sigma_{OLS}^2 (X'X)^{-1} - \sigma_{OLS}^2 (Z'Z)^{-1}, \end{aligned}$$

так как

$$Cov\{\hat{\beta}_{OLS}, b\} = Cov\{\hat{\beta}_{OLS}, \hat{\beta}_{OLS} - \hat{\beta}_{IV}\} = Cov\{\hat{\beta}_{OLS}, \hat{\beta}_{OLS}\} - Cov\{\hat{\beta}_{OLS}, \hat{\beta}_{IV}\} \underset{H_0}{=} C_{OLS} - C_{OLS} = 0,$$

$$Cov\{\hat{\beta}_{IV}, b\} = Cov\{\hat{\beta}_{IV}, \hat{\beta}_{OLS} - \hat{\beta}_{IV}\} = Cov\{\hat{\beta}_{OLS}, \hat{\beta}_{IV}\} - Cov\{\hat{\beta}_{IV}, \hat{\beta}_{IV}\} \underset{H_0}{=} C_{OLS} - C_{IV}.$$

Оценка автоковариационных матриц статистики Хаусмана

МНК-оценка		МИП-оценка	
$\hat{C}_{OLS} = \hat{\sigma}_{OLS}^2 (X'X)^{-1}$		$\hat{C}_{IV} = \hat{\sigma}_{OLS}^2 (Z'Z)^{-1}$	
0,0054	0,0000606	0,00000372	-0,00000157
0,0000606	0,00000588	-0,00000157	0,00000587

При справедливости $H_0 : Cov\{x_{it}, \varepsilon_{is}\} = 0$,

$$Cov\{\hat{\beta}_{IV}, \hat{\beta}_{OLS}\} = Cov\{(Z'X)^{-1} Z'Y, (X'X)^{-1} X'Y\} = \\ = (Z'X)^{-1} Z' Cov\{Y, Y\} X (X'X)^{-1} = \sigma_{\varepsilon}^2 (X'X)^{-1} = C_{OLS}.$$

Используя в качестве оценки дисперсии возмущения σ_{ε}^2 оценку, вычисленную по остаткам обычного МНК – $\hat{\sigma}_{OLS}^2$ (МНК-оценка более устойчива к слабым инструментам), вычислим оценки автоковариационных матриц \hat{C}_{OLS} , \hat{C}_{IV} . Результаты оценивания приведены в табл. 3.

Результат тестирования: $H = 0,43 < \chi^2(1) = 3,84$, следовательно, для исследуемых панельных данных можно использовать обычный МНК для спецификации (14), который дает более точные результаты, чем МИП:

$$\hat{y}_{it} = y_{i,t-1} + \underset{(0,074)}{0,7029}(y_{i,t-1} - y_{i,t-2}) - \underset{(0,0024)}{0,0013}(x_{it} - x_{i,t-1}), \hat{\sigma}_{OLS}^2 = 4413,5. \quad (17)$$

Построенная модель может быть улучшена посредством увеличения рассматриваемого временного периода и расширения как перечня наблюдаемых объектов (стран), так и перечня влияющих факторов. Периодическое тестирование на обновленных панельных данных решает задачи актуализации построенных моделей и возможности построения более точных прогнозов.

Список литературы

1. Официальный сайт Базы данных Всемирного банка. URL: <https://data.worldbank.org/> (дата обращения: 10.03.2018).
2. Официальный сайт Всемирного золотого совета. URL: <https://data.worldbank.org/> (дата обращения: 10.03.2018).
3. Россия продаёт трежерис, но покупает золото. URL: <http://gold.ru/news/rossija-prodajot-trezheris-2018-no-pokupaet-zoloto-dlja-rezervov.html> (дата обращения: 10.03.2018).
4. Baltagi B.H. *Econometric Analysis of Panel Data* (4th ed.) / B.H. Baltagi. – Chichester: John Wiley & Sons, 2008. – 314 p.
5. Green W.H. *Econometric Analysis* (7th ed.) / W.H. Green. – N.Y., 2012. – 1024 p.

6. Бабешко Л.О. Эконометрическое моделирование влияния инвестиций в сферу туризма на объем туристического потока в программной среде R // Международный журнал прикладных и фундаментальных исследований. – 2018. – № 1. – С. 167–171.

7. Вербик М. Путеводитель по современной эконометрике. Пер. с англ. В.А. Банникова. Науч. ред. С.А. Айвазян / М. Вербик. – М.: Научная книга, 2008. – 616 с.

8. Носко В.П. Эконометрика. Кн. 2: учебник / В.П. Носко. – М.: Издательский дом «Дело» РАНХиГС, 2011. – 576 с.

9. Бабешко Л.О. Прогнозирование финансово-экономических показателей по разнородным данным / Л.О. Бабешко, В.А. Бывшев. – М.: РУСАЙНС, 2017. – 356 с.

10. Бабешко Л.О. Эконометрическое прогнозирование по разнородной информации / Л.О. Бабешко. – М.: Вега-Инфо, 2016. – 232 с.

11. Kleiber C., Zeileis A. *Applied Econometrics with R*. Springer-Verlag. – N.Y., 2008. – 222 p.

12. Wooldridge J.M. *Econometric Analysis of Cross Section and Panel Data*. – Cambridge, MIT Press, 2007. – 776 p.

13. Эконометрика и эконометрическое моделирование: учебник / Л.О. Бабешко, М.Г. Бич, И.В. Орлова. – М.: Вузский учебник: ИНФРА-М, 2018. – 385 с.

14. Эббес П. Инструментальные переменные и эндогенность: нетехнический обзор / П. Эббес // Квантиль. – 2007. – № 2. – С. 3–20.

15. Цыплаков А. Эскурсе в мир инструментов / А. Цыплаков // Квантиль. – 2007. – № 2. – С. 21–49.