FOLIA OECONOMICA CRACOVIENSIA
Vol. LVIII (2017) PL ISSN 0071-674X

INVESTIGATION OF FORECASTING PERFORMANCE OF SELECTED VECM MODELS, FOR EUR/PLN EXCHANGE RATE[1]

ADRIAN BURDA[2]
e-mail: adrian.marek.burda@gmail.com

STRESZCZENIE

A. Burda. *Weryfikacja własności prognostycznych wybranych, proponowanych w literaturze modeli VECM dla kursu walutowego EUR/PLN*. Folia Oeconomica Cracoviensia 2017, 58: 59–84.

Celem niniejszego artykułu jest przetestowanie zdolności predykcyjnych wybranych wektorowych modeli korekty błędu (VECM) dla kursu walutowego EUR/PLN. Wspomniane modele inkorporują w sobie takie zjawiska i teorie ekonomiczne, jak parytet siły nabywczej (PPP), nieubezpieczony parytet stóp procentowych (UIP), efekt Harroda-Balassy-Samuelsona (HBS), terms of trade (TOT), jak i premię za ryzyko. Wyniki były porównywane pod względem dokładności prognoz punktowych oraz pod względem kierunkowej dokładności z takimi prognozami referencyjnymi jak: błądzenie przypadkowe bez dryfu, model AR1 oraz model VAR1 ze zmiennymi odzwierciedlającymi teorię PPP. Wyniki wskazują, że w odniesieniu do prognoz punktowych, żaden z rozważanych modeli VECM nie generuje istotnie bardziej dokładnych niż błądzenie przypadkowe. W odniesieniu do prognoz kierunku zmian, dla horyzontu jednego miesiąca większość modeli VECM wskazuje na istotną statystycznie zdolność do ich przewidzenia, a wyniki są odporne dla różnych podpróbek. W odniesieniu do dłuższych horyzontów, niektóre z analizowanych modeli VECM pozwalają na trafne przewidzenie kierunku zmian kursu walutowego, jednakże te wyniki nie są odporne dla różnych podpróbek.

ABSTRACT

The purpose of this paper is testing forecasting properties of selected VECM models for EUR/PLN. These models incorporate such theories as purchasing power parity (PPP), uncovered interest rate parity (UIP), Harrod-Balassa-Samuelson (HBS) effect, terms of trade (TOT), Net financial asset (NFA) theory and risk premium. Results had been compared in terms of point forecast accuracy

[1] This is part of research, financing from the research project "Bayesian inference in STVECM models — applications to exchange rate analysis", financing by National Science Centre in Poland in funding scheme "PRELUDIUM 7" (Contract UMO-2014/13/N/HS4/03593). Principal Investigator — Adrian Burda, Supervisor: Professor Mateusz Pipień.

[2] Ph.D candidate on the Cracow University of Economics.

and directional accuracy with benchmarks like random walk without drift and AR1 and VAR1 model for PPP variables. Results indicate that none of considered VECM model significantly beats random walk in terms of point accuracy. In terms of directional accuracy, all models except the most complex CHEER_BEER have ability to predict direction of change of EUR/PLN for 1 month horizon and the results are robust in subsamples. For other horizons some VECM models are found to have ability to predict direction of change EUR/PLN, while the results are not very robust.

SŁOWA KLUCZOWE — KEY WORDS

kurs walutowy, prognozy kursu walutowego, dokładność kierunkowa, modele VECM

Exchange rate, Exchange rate forecasts, directional accuracy, VECM models

1. INTRODUCTION

Since Meese and Rogof (1983) in the empirical literature it is found that exchange rate are very difficult to predict by economic model and simple a-theoretical model as random walk typically generate not worse and in some cases better forecast than economic models; see Rossi (2013), Meese and Rogoff (1983). Albeit many newer papers, in particular where error-correction models and Taylor rule is implemented indicate some positive results as in Ince and others (2016) or Galimberti and others (2013), they are sensible on the sample properties — Rossi (2013). In case of forecasting of PLN against to EUR or against to other currencies (e.g. USD, GBP), only few studies; see Galimberti and others (2013); Cuaresma and Hlouskova (2005) for some specifications found positive results. In others, analyzed specifications did not beat random walk, while in some cases random walk outmatched econometric models; e.g. Rubaszek and others (2010). Scarcity of this literature does not allow to drive general conclusion.

The objective of this article is testing predictive ability of popular in the empirical literature vector error correction models (VECM) to EUR/PLN exchange rate. The motivation of the research is twofold. Firstly, VECM models had been widely used in order to estimate equilibrium exchange rate and/or verify several economic hypothesis, e.g. Rubaszek and Serwa (2009), Kelm (2013), Kębłowski and Welfe (2012), Bęza-Bojanowska and MacDonald (2009). However in those articles predictive properties of developed VECM models are almost never investigated.

Secondly, overall empirical literature regarding forecasting exchange rate of Polish currency is scarce and not sufficient. While Cuaresma and Hlouskova (2005) tested predictive performance of VEC and Bayesian VEC (BVEC) models, due to sample specifics (1993–2000), containing mostly period with managed float regime, their conclusion are not relevant to recent period of floating regime. Much more recent paper of Galimberti and Moura (2013), containing

Taylor rule and UIP relation where models are estimated on sufficiently long sample is also not relevant to asses predictive properties of VECM models, as the used models are panel VECMs for sample including 15 countries. In the other proposed in the literature papers — Adric and others (2008), Mućk and Skrzypczyński (2012), Rubaszek and others (2010); VECM models had been not investigated in contrary to VAR models — Ardic and others (2008), Mućk and Skrzypczyński (2012); univariate time series models — Ardic and others (2008), Rubaszek and others (2010); survey forecasts — Naszodi (2011) or so called "structural models" — uncovered interest rate parity model and monetary models — Ardic and others (2008).

2. FORECASTING EXCHANGE RATE WITHIN VECM FRAMEWORK

2.1. VECM models and its properties

Vector Error Correction Models (VECM) could be described in the following way — Lütkepohl (2005),

$$\Delta y_t = [y_{t-1}:d_t^{co}]\begin{bmatrix}\beta \\ \eta\end{bmatrix}\alpha + d_t\xi + \sum_{i=1}^{p}\Delta y_{t-i}\Gamma_i + \sum_{i=1}^{q}x_{t-i}B_i + \varepsilon_t, \tag{1}$$

where $y_t = (y_{1t}, \ldots, y_{Kt})$ is vector of K endogenous variables,
$x_t = (x_{1t}, \ldots, y_{Mt})$ is vector of K endogenous variables,
d_t^{co} includes all deterministic terms in the cointegrated relations,
d_t contains all remaining deterministic terms (outside cointegrating relation).
The residual vector ε_t is assumed to be K-dimensional, zero mean white noise process with positive definite covariance matrix

$$E(\varepsilon_t\varepsilon_t') = \Sigma_\varepsilon.$$

The parameters matrix β contains cointegrating relations while α — loading coefficients. The Matrices β and α' have dimensions $(K * r)$ and have to have rank r, where r is the rank of cointegrating space.

The parameter matrix η has r columns and number of rows corresponding with dimension d_t^{co}.

This model could be rewritten as VAR$(p + 1)$ model. The reason is the fact, that K dimensional process VAR$(p + 1)$ process — for simplicity considered without deterministic term:

$$Y_t = \sum_{i=1}^{p+1} Y_i A_{t-i} + \varepsilon_t \tag{2}$$

could be written in VECM(p) form:

$$\Delta Y_t = Y_{t-1}\beta\alpha + \sum\nolimits_{i=1}^{p} \Delta Y_{t-i}\Gamma_i + \varepsilon_t, \tag{3}$$

or, alternatively, by

$$\Delta Y_t = Y_{t-1}\Pi + \sum\nolimits_{i=1}^{p} \Delta Y_{t-i}\Gamma_i + \varepsilon_t, \tag{4}$$

where $\Pi = \beta\alpha$.

Thus, it is easy to recover the corresponding VAR form (2), by nothing that:

$$A_i = \begin{cases} \Pi + I_K + \Gamma_1 & for\ i = 1 \\ \Gamma_i - \Gamma_{i-1} & for\ i = 2, \ldots, p \\ -\Gamma_p & for\ i = p+1. \end{cases} \tag{5}$$

It is important to note, that decomposition of the $(K\ x\ K)$ matrix $\prod$ as the product of two $(K\ x\ r)$ matrices, namely α and β is not unique, due to given any choice of α and β and any nonsingular $r\ x\ r$ matrix Q, where $\alpha^* = \alpha Q'$ $\beta^* = \beta Q^{-1}$ gives the same matrix $\prod = \beta^*\alpha^*$; see Johansen (2005). It means that cointegration relations are not unique. However, it is possible to impose restrictions on β and/or α, to get unique restrictions. They may be implied by subject matter consideration, or may be imposed for convenience, using the algebraic properties of associated matrix.

Forecasting variables incorporated into VECM model in levels is conventionally analyzed in the VAR representation of data generating process — see Lütkepohl (2005). Denoting the h-step forecast based on estimated coefficients by $\hat{Y}_t(h)$ and indicating estimators by hats gives following forecast:

$$\hat{Y}_t(h) = \hat{A}_1\hat{Y}_t(h-1) + \cdots + \hat{A}_p\hat{Y}_t(h-p), \text{ where } \hat{Y}_t(j) := \hat{Y}_{t+j} \text{ for } j \leq 0. \tag{6}$$

Selection of appropriate VECM model for empirical investigation includes such steps as: including adequate deterministic term, lag order selection, determination rank of cointegration, and eventual imposing restriction of coefficients. Deterministic terms should be included, taking into account properties of data generating process. However, it is possible in some procedures — as maximum likelihood (ML) to compare by likelihood ratio (LR) models with different deterministic terms. Impact of different variants of deterministic terms is derived below.

Define d_t as deterministic term. Let first $\Phi d_y = v_0 + v_1 t$, so that

$$Y_t = \Xi \sum_{i=0}^{t} (\eta_i + \upsilon_0 + \nu_1 i) + \Xi^* \sum_{i=0}^{\infty} (\eta_{t-i} + \nu_0 + \nu_1(t-i)) + A. \tag{7}$$

It shows that generally, a linear trend in the equation becomes a quadratic trend with coefficient $\frac{1}{2}\Xi t^2$ in the process. The process v_1 may be decomposed in the following way:

$\nu_i = \alpha\rho_i^{'} + \alpha_\perp\gamma_i^{'}$, where $\alpha'\nu_i = \rho_i^{'}$, and $\alpha_\perp^{'}\nu_i = \gamma_i^{'}$. It is then seen, that if $\gamma_1 = 0$, so that $\nu_i = \alpha\rho_\perp^{'}$, or $\alpha_\perp^{'}\nu_1 = 0$, then the quadratic term has coefficient zero, $\Xi\nu_1 = \Xi\alpha\gamma_\perp^i = 0$, so that only linear is present. Overall, five cases of deterministic case $\Phi d_y = \nu_0 + \nu_1 t$ may be distinguished. There are summarized in the table 1, below:

Table 1

The five models defined by restriction on the deterministic terms in equation 7

Model	Linear term	Restriction	Trend in Y_t	$E(\Delta Y_t)$	$E(\beta' Y_t)$
1	$\nu_0 + \nu_1 t$	none	quadratic	Linear	linear
2	$\nu_0 + \alpha\rho_1^{'} t$	$\xi_1 = 0$	linear	Constant	linear
3	ν_0	$\nu_1 = 0$	linear	Constant	constant
4	$\alpha\rho_1^{'}$	$\nu_1 = 0,\ \xi_0 = 0$	constant	Zero	constant
5	0	$\nu_1 = \nu_2 = 0$	zero	Zero	zero

Source: Johansen (2005).

Among those cases in the empirical research the most commonly used is model 3; see Lütkepohl (2005). Also in this research "case 3" is utilized, due to properties of EUR/PLN, similarly as in many other VECM studies for — e.g. Kelm (2013); Rubaszek and Serwa (2009).

Lag order selection is typically performed before as cointegration rank does not have to be known in order to choose appropriate lag, while many procedures for specifying the cointegration rank require selection of lag length. In order to select optimal lag information criteria — e.g. Akaike (1998); Schwarz (1978), residual autocorrelation — e.g. Ljung and Box (1978), or portmanteau tests could be used — Lütkepohl (2005). However, no single method of lags selection surpass others, in terms of detecting true data generating process, thus research should analyzed different criteria — Lütkepohl (2005).

In order to determine proper cointegration rank, several tests had been proposed, albeit also model selection criteria could be used to this purpose — e.g. Lütkepohl and Poskitt (1998). In this research, *trace test*, *maximum eigenvalue tests* — see Johansen (1988), (1995) — and *trace test with small-sample Bartlett correction* — Johansen (2002) — had been utilized. Thus, estimated models had been estimated, via maximum likelihood (ML) estimator, called Johansen procedure — see Johansen (1988), (1991), which takes into account the rank restrictions for $\Pi = \beta\alpha$.

Interestingly, assumptions about normality of the process is not necessary for the holding properties of the estimators $\tilde{\Gamma}$ and $\tilde{\Pi} = \tilde{\beta}\tilde{\alpha}$, as asymptotical efficiency and normal. Much of them are true under weaker conditions, if quasi ML estimators based on the Gaussian likelihood function are taken into consideration. Normality assumption is simply a convenience. Obviously, if ε_t is not Gaussian, $\tilde{\Sigma}_\varepsilon$ be compared with model with only normalization restriction, where $\beta^* = \begin{bmatrix} I_r \\ \beta^*_{(K^*-r)} \end{bmatrix}$ by likelihood ratio test.

2.2. Proposed specifications of VECM models

General specification VEC models investigated in the research for forecasting exchange rate is described below:

$$\Delta y_t = y_{t-1}\beta\alpha + d_t\xi + \sum_{i=1}^{p} \Delta y_{t-i}\Gamma_i + \varepsilon_t, \tag{8}$$

where
y_t — vector of endogenous variables, including nominal exchange rate (on the first place) and set of other variables influencing exchange rate (and vice versa),
β — cointegration coefficient matrix,
α — loading coefficient matrix,
d_t — deterministic terms, in this research it is only intercept,
p — number of lags — in this research for all models p = 2.

The number of cointegration relations had been assumed between 1 and 3 depending on the model described below, while numbers of lags — p, had been set as a compromise to selected number of lags in different procedure of lags selection, as: AIC, BIC, LM and investigating properties of residuals.

In the research several specifications of VECM models present had been considered. In the first specification, containing only PPP relation variable: it is assumed that

$$y_t = [s_t p_t p_t^*], \tag{9}$$

where
s_t — nominal exchange rate,
p_t — price index for Poland,
p_t^* — price index for foreign economy (Euro Area) and one cointegration relation (r = 1) representing PPP relation.

Furthermore, in the research the proportionality restriction in β is assumed, what could be described as: β = [1, –1,1]. This research specification is called further "PPP".

In the second specification beyond PPP relation, also uncovered interest rate (UIP) parities for short-term and long-term interest rates are included

$$y_t = [s_t p_t p_t^* i_{s_t} i_{s_t}^* i_{l_t} i_{l_t}^*], \tag{10}$$

where

i_{s_t} — short term domestic interest rate,

$i_{s_t}^*$ — short term foreign interest rate,

i_{l_t} — long term domestic interest rate,

$i_{l_t}^*$ — long term foreign interest rate.

There 3 cointegration relations are assumed, representing PPP and UIP for short-term and long-term interest rates with proportionality restriction for PPP. This specification is consistent with capital enhanced exchange rate equilibrium model (CHEER).

Third specification is CHEER model with risk premium indicator, where

$$y_t = [s_t p_t p_t^* i_{s_t} i_{s_t}^* i_{l_t} i_{l_t}^* rp_t], \tag{11}$$

where

rp_t — is risk premium indicator and which is called "CHEER_Risk_Premium".

All other properties are consistent with CHEER model without risk premium indicator.

Fourth and fifth specification is CHEER model without and with risk premium indicator and with Harrod-Balassa-Samuelson (HBS) effect indicator.

$$y_t = [s_t p_t p_t^* i_{s_t} i_{s_t}^* i_{l_t} i_{l_t}^* hbs_t], \tag{12}$$

$$y_t = [s_t p_t p_t^* i_{s_t} i_{s_t}^* i_{l_t} i_{l_t}^* rp_t hbs_t], \tag{13}$$

where hbs_t represents HBS effect indicator. These models are called in the research "CHEER_HBS" and "CHEER_HBS with risk premium" and has other properties (number of cointegration relations and PPP proportionality restriction) the same as CHEER models.

Last — sixth specification regards one the variant of hybrid CHEER-BEER (behavioral equilibrium exchange rate) model, in which several medium-term variables are added to the "CHEER_BEER" model. There:

$$y_t = [s_t p_t p_t^* i_{s_t} i_{s_t}^* i_{l_t} i_{l_t}^* hbs_t tot_t fdi_t ofl_t], \tag{14}$$

where
tot_t — terms of trade,
fdi_t — stock of foreign direct investment,
ofl_t — other liabilities (in net international investment position) than direct investment.

Primarily also risk premium had been intended to be included in the model, but the "CHEER_BEER" model with this variable failed to be estimated. Also in this model cointegration rank is equal 3 and PPP proportionality restrictions are imposed.

2.3. Criteria of forecast evaluation

In this research forecast are evaluated by the measures of ex-post point accuracy measures, Diebold-Mariano test; see Diebold and Mariano (1995) and by measure and test of directional accuracy; see Leitch and Tanner (1991). In order to present and discuss their properties, forecast error, for period t and horizon h, e_{ht} is defined as follow:

$$e_{ht} = y_t - \hat{y}_{ht}, \forall t = 1 \ldots T, \tag{15}$$

where
T — is length of out of sample period used in the forecast evaluation.

Then, considered in the researched *ex-post* forecast accuracy measures are defined as follows — e.g. Makridakis and others (1998); Zeliaś and others (2008); Papież and Śmiech (2015):

1) Mean absolute error (MAE):

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |e_{ht}|. \tag{16}$$

2) Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} e_{ht}^2}. \tag{17}$$

Both measures have slightly different properties, as MAE is optimal criterion for linear loss function, while RMSE for quadratic loss function. RMSE gives higher penalty for single large errors than MAE. As in the research forecast performance is compared for the same variable (exchange rate), expressed in logs and for the same sample, there is no need to calculate these measures in relative terms as mean average percentage error (ang. MAPE).

In the research not only MAE and RMSE for different model are compared, but also it is tested, if the particular model generate "better" forecast than benchmark — random walk without drift (RW), which is found the toughest benchmark to beat in case of exchange rate; Rossi (2013). The most common test used to this purpose is Diebold-Mariano (DM) test; Diebold and Mariano (1995). In this research, due to relative small sample, its corrected version proposed by Harve, Leybourne and Newbold (1997) (HLN) is utilized.

In DM test two series of forecast errors — generated from two different methods are compared, which are defined as

$$e_{1,ht} = y_t - \hat{y}_{1,ht}, \forall t = 1 \ldots T,$$

$$e_{2,ht} = y_t - \hat{y}_{2,ht}, \forall t = 1 \ldots T,$$

— or more strictly — losses related to them. Loss function — $L(e_{i,ht})$ — had following properties: takes the value zero when no error is made, is never negative and increasing in size as the errors become larger magnitude. Typically $L(e_{i,ht})$ is square (squared-error-loss) or the absolute value (absolute error loss), so respectively:

$$L(e_{i,ht}) = (e_{i,ht})^2, \tag{18}$$

$$L(e_{i,ht}) = |e_{i,ht}|. \tag{19}$$

Thus, loss differential between the two forecast for the horizon h is defined as:

$$d_{ht} = L(e_{1,ht}) - L(e_{2,ht}). \tag{20}$$

Thus, two forecast have equal accuracy if, and only if the loss differential is equal zero. Therefore, null hypothesis for DM test, that two forecasts has equal accuracy is written:

$$H_0 : E(d_{ht}) = 0. \tag{21}$$

Alternative hypothesis could be two-sided: $H_1 : E(d_{ht}) \neq 0$ or be one sided:

$$E(L(e_{1,ht}))E(L(e_{2,ht})) \text{ or } E(L(e_{1,ht})) < E(L(e_{2,ht})). \tag{22}$$

Assuming that $\{d_{ht}\}$ is covariance stationarity — Diebold (2012) — DM test statistics asymptotically follows normal distribution:

$$DM = \frac{\sqrt{T}\,\bar{d}}{\sqrt{Var_{\bar{d}}}}, \tag{23}$$

where

$\bar{\bar{d}} = \frac{1}{T}\sum_{t=1}^{T} d_t$ and $Var_{\bar{d}} = 2\pi \hat{f}_d(0)$ is consistent estimator of long-term variance of d_t.

However, in the simulation experiments in Diebold and Mariano (1995) it is shown that normal distribution could be very poor approximation of the DM test's finite-sample null distribution — rejecting the null hypothesis too often — depending on the degree of serial correlation among the forecast errors and sample size; see Thus, Harve, Leybourne and Newbold (1997), (HLN) suggested improvement of small sample properties of DM test statistics and comparing corrected statistics with a Student-t distribution with $(T - 1)$ degrees of freedom:

$$HLN - DM = \sqrt{\frac{T + 1 - 2h + h(h-1)}{T}}\,DM. \tag{24}$$

It should be noted that *HLN-DM* could be used for nested models only under some regularity conditions — Giacomini and White (2006) — such as models estimated on rather rolling window data, than expanding one.

As in the research HLN-DM test is utilized to compare forecast accuracy between VEC models and RW, which are not nested models this research, this issue is not relevant.

Directional accuracy or direction of change statistics could be defined as particular loss function, against the random prediction of direction; see Cheung and others (2005):

$$d_t = \begin{cases} 1\ if\ \ forecast\ predict\ correctly\ direction\ of\ change, \\ 0\ otherwise. \end{cases} \tag{25}$$

If $\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t > 0.5$, the forecasting method has ability to predict direction of change correctly.

The test statistics s:

$$s = \frac{\bar{\bar{d}} - 0.5}{\sqrt{\frac{0.25}{T}}} \tag{26}$$

is distributed as Students-t distribution with $T - 1$ degrees of freedom, while in large samples — as standard Normal. Direction of change statistics is particular useful from investors perspective; see Melvin and others (2013). Following it allows to achieve profits more frequently than using other measures of forecast accuracy; Leitch and Tanner (1991).

2.4. Strategy of the research

In this research, forecasting performance of VEC models described in the chapter 2.2 is conducted for 1, 3, 6 and 12 months ahead horizon, for recursive scheme. These models at the beginning are estimated for the sample from January 1999 to January 2006 and respectively out of sample forecasts of logarithm of EUR/PLN for 1, 3, 6 and 12 months ahead are generated — respectively till February 2006, April 2006, July 2006, January 2007. In the further step, the estimation sample is extended by one observation, by adding February 2006 observation and the procedure is repeated. This scheme ends, when for out of sample forecasts for 12 months reaches end of sample, which is December 2015. This recursive updating scheme give for each forecast horizon for each model 108 forecasts. The motivation to use recursive scheme instead of popular in exchange rate forecasting rolling scheme; e.g., Engel and others (2007) is relative shortness of sample. Analogous rolling scheme, containing the same number of generated forecasts could have only 7 year of monthly observations (72 observations) what could be insufficient to estimate more complex VEC models.

Mentioned VEC models had been compared in terms of forecasting performance with naïve methods: random walk without drift, AR1 model for logarithm exchange rate and VAR1 model for variables which are included in the PPP model — in levels. In order to assess accuracy of all methods, MAE, RMSE and direction of change statistics (with testing ability to forecast direction of change) had been computed. Furthermore, HLN-DW test for each of mentioned models, except RW (which is benchmark) for absolute and quadratic loss function had been conducted. These forecast performance measures had been compared for the whole period of described rolling scheme and for rolling window of 60 time series of forecast errors. It allowed to assess forecast performance of mentioned models both in the period of crisis 2008–2009, when Polish zloty depreciated sharply after "appreciation anomaly" — see Kelm (2013) — and in less volatile post-crisis period.

3. EMPIRICAL RESULTS

3.1. Data

The models are estimated using monthly data from January 1999 till December 2015 for Poland and Euro Area. Albeit Euro Area throughout the research sample expanded, from 12 countries in 1999 to 19 in 2015, for data availability and clearness of calculation most of data for Euro Area had been calculated for 19 countries. As all countries which joined Euro Area after 2002 are relatively small economies, impact of this simplification should be low, as time series for 19 Euro Area countries (EA19) are very highly correlated with series for first 12 Euro Area members (EA12)[3]. All data are revised series (state as June 2016), calculated according to newest methodologies (e.g. EA 2010 in case of National Account or BPS6 in case of Balance of payment statistics). Back-extending of data and other adjustments which had been made are described below. To estimation purposes all series expect interest rates and risk premium indicator are transformed to indexes, where values for 2000=1 and after it — logarithmized by natural logarithm.

Data regards EUR/PLN exchange rate had been obtained from Eurostat and they are average monthly exchange rate calculated from daily exchange rate by National Bank of Poland.

Price indexes are producer price indexes (PPI) for manufacturing and they are provided by OECD.

Short-term interest rate are three-month WIBOR and EURIBOR rates, and they are provided by OECD. It had been transformed to monthly interest rate according to formula

$$i_{m,t} = (1 + i_{y,t})^{1/12} - 1,$$

where
$i_{y,t}$ — yearly interest rate,
$i_{m,t}$ — monthly interest rate.

Long-term interest rate are average rates of 10-year government bonds, and they are provided by OECD. As OECDs data do not include observations for the period from January 1999 to December 2000, this observation had been back-extended, utilising Kelm (2013) dataset, which included analogous time series.

Indicator of Harrod-Balassa-Samuelson effect had been calculated by the author, according to formula:

$$HBS_t = \frac{\frac{EMP_{MA,pl,t}}{GVA_{MA,pl,t}}}{\frac{EMP_{NMA,pl,t}}{GVA_{NMA,pl,t}}} * \frac{\frac{EMP_{NMA,ea,t}}{GVA_{NMA,ea,t}}}{\frac{EMP_{MA,ea,t}}{GVA_{MA,ea,t}}},$$

where
EMP — regards to total employment in the sector,
GVA — refers to the Gross Value Added in the Sector,

[3] Detailed calculation could be send by author by request.

MA — manufacturing sector,
NMA — aggregate of other sectors than manufacturing,
pl — Poland,
ea — Euro Area.

All of single factors needed to calculate HBS are gathered from Eurostat. They had been disaggregated to monthly from quarterly time series, by the method described by Pipień and Roszkowska (2015), with own modifications. These modification regards to use specific explanatory variables — monthly indicators of economic activity and employment in sector of interests and by balancing procedure. Furthermore, due to data availability, HBS indicator for the period 1999:01–1999:12 had been back-extended using Kelm's data (2013).

Relative Terms of trade (TOT) indicator data had been taken directly from Kelm (2013) — till June 2011, and extended till December 2015, using TOT indicator calculated by Polish Statistical Office and exports and imports price data for EA19 provided by Eurostat.

Net Foreign Direct Investment (FDI) is analysed in relative terms, in relation to GDP:

$$fdi_t = (-1) * \frac{net\ direct\ investment\ (asset - liabilities)_t}{\sum_{i=t-3}^{t} GDP_i},$$

where
net direct investment are quarterly time series, denominated in PLN gathered from International Investment Position statistics from National Bank of Poland and GDP is quarterly GDP in current prices gathered from Eurostat.

The quarterly time series had been disaggregated to monthly time series, by usage automatic interpolation procedure in Gretl Software. Due to change of Balance of Payment methodology, data consistent with new methodology (BMP6) are available only since 2004. Thus earlier data had been back-extended using Kelm's data (2013).

Other liabilities than direct investment (OFL) had been defined as

$$ofl_t = (-1) * \frac{net\ \ international\ investment\ \ position_t - net\ direct\ investment_t}{\sum_{i=t-3}^{t} GDP_i}.$$

Data sources and adjustment procedures had been analogous as in case of FDI.

As a measure of risk premium in this research CBOE Volatility Index (VIX) had been utilised. This daily time series had been aggregated to monthly by unweighted average. VIX is based on the S&P 500 Index (SPX), the core index for U.S. equities, and estimates expected volatility by averaging the weighted prices of SPX puts and calls over a wide range of strike prices. Despite the fact, that in reality it could not reflect truly forward realized volatility; see e.g. Adhikari and Hillard (2014) it is commonly used by market participants to predict future volatility and could reflect global uncertainty.

3.2. Estimation results for the whole sample

After estimation of models described in the chapter 2 for recursive scheme generated via procedures described in the chapter 2.4, the correlation between forecast error from RW and others forecasting models had been calculated. Results, summarized in the Table 1, indicate high correlation between forecast errors generated by RW and other models. Furthermore, the all models experienced more similarity in scale and direction forecast errors in times of crisis, than in other periods, what is displayed on the Figures 1 and 2.

Table 2

Correlation of forecast errors from analyzed models with forecast errors from RW

Forecast horizon	AR1	VAR1	PPP	CHEER	CHEER_Risk Premium	CHEER_HBS	CHEER_HBS3_ riskPrem
1m	0.997	0.981	0.830	0.833	0.824	0.818	0.810
3m	0.995	0.962	0.777	0.869	0.870	0.863	0.847
6m	0.992	0.935	0.861	0.902	0.913	0.907	0.898
12m	0.983	0.849	0.938	0.854	0.907	0.872	0.931

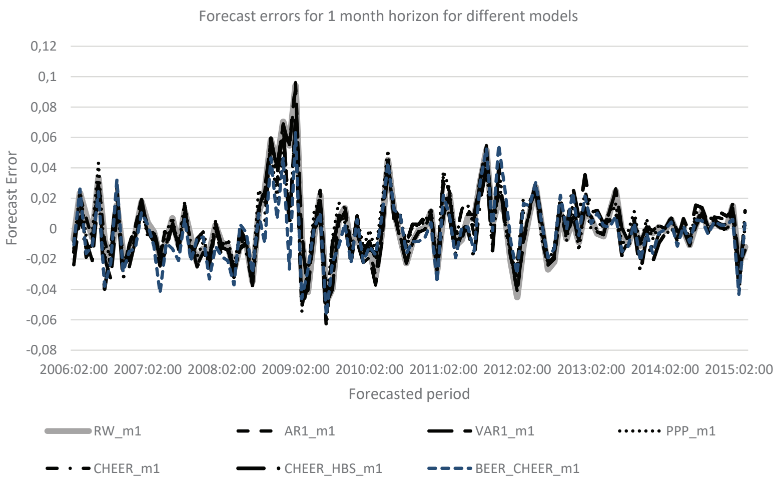

Figure 1. Forecast errors for 1 month horizon for analysed forecasting methods

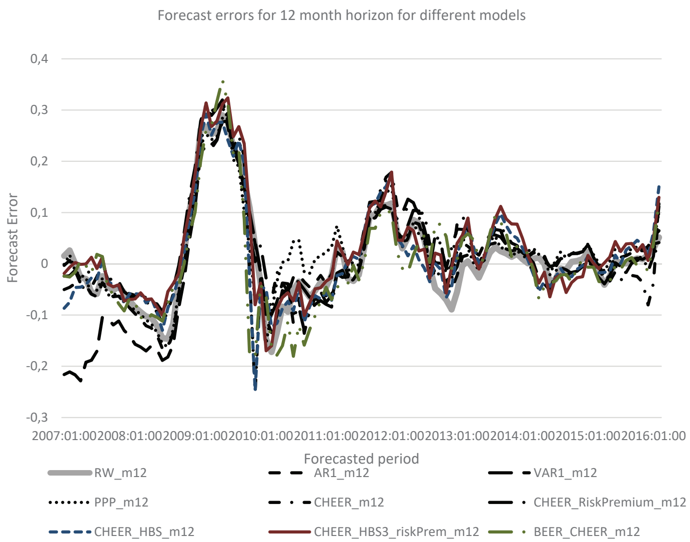

Figure 2. Forecast errors for 12 month horizon for analysed forecasting methods

MAE for forecast generated from all models, expect AR1 model and PPP model is higher than for forecast generated from RW. Slightly more promising results could be seen for RMSE generated from all models — being lower than RMSE not only for AR1 model and PPP (for 1 month and 12 month horizon), but also from BEER-CHEER models for 1 and 3 month horizon. Interestingly relative RMSE (where the reference is RW) for analyzed VECM models are slightly more favorable than relative MSE. Details are presented in Tables 3–4.

Table 3

RMSE statistics for the full out of sample, all models and horizons

Model	forecast horizon	RW	AR1	VAR1	PPP	CHEER	CHEER_Risk Premium	CHEER_HBS	CHEER_HBS_ RiskPremium	BEER_CHEER
RMSE	1	0.0222	0.0221	0.0227	0.0200	0.0223	0.0236	0.0219	0.0243	0.0218
	3	0.0489	0.0480	0.0508	0.0563	0.0533	0.0583	0.0497	0.0607	0.0465
	6	0.0746	0.0731	0.0815	0.0789	0.0829	0.0854	0.0759	0.0859	0.0779
	12	0.0936	0.0891	0.1091	0.0897	0.1038	0.1002	0.0970	0.1018	0.1016
Model	forecast horizon	RW	AR1	VAR1	PPP	CHEER	CHEER_Risk Premium	CHEER_HBS	CHEER_HBS_ RiskPremium	BEER_CHEER
Relative RMSE (RW = 1)	1	1.00	0.99	1.02	0.90	1.00	1.06	0.99	1.09	0.98
	3	1.00	0.98	1.04	1.15	1.09	1.19	1.02	1.24	0.95
	6	1.00	0.98	1.09	1.06	1.11	1.14	1.02	1.15	1.04
	12	1.00	0.95	1.17	0.96	1.11	1.07	1.04	1.09	1.09

Table 4

MAE statistics for the full out of sample, all models and horizons

Model	forecast horizon	RW	AR1	VAR1	PPP	CHEER	CHEER_Risk Premium	CHEER_HBS	CHEER_HBS_ RiskPremium	BEER_CHEER
MAE	1	0.0156	0.0156	0.0164	0.0151	0.0161	0.0166	0.0161	0.0169	0.0166
	3	0.0323	0.0320	0.0350	0.0397	0.0386	0.0366	0.0360	0.0389	0.0362
	6	0.0462	0.0448	0.0548	0.0501	0.0539	0.0521	0.0489	0.0528	0.0535
	12	0.0642	0.0621	0.0799	0.0616	0.0743	0.0678	0.0692	0.0702	0.0707

Model	forecast horizon	RW	AR1	VAR1	PPP	CHEER	CHEER_Risk Premium	CHEER_HBS	CHEER_HBS_ RiskPremium	BEER_CHEER
Relative MAE (RW = 1)	1	1.00	1.00	1.05	0.97	1.03	1.06	1.03	1.08	1.06
	3	1.00	0.99	1.08	1.23	1.19	1.13	1.11	1.20	1.12
	6	1.00	0.97	1.18	1.08	1.17	1.13	1.06	1.14	1.16
	12	1.00	0.97	1.24	0.96	1.16	1.05	1.08	1.09	1.10

However, none of mentioned results above is statistical significant according to HLN-DM test. Furthermore, for some horizons, lower forecast accuracy than forecasts generated by RW is significantly less accurate according to HLN-DM test. It regards e.g. CHEER model for 3–12 month horizon and all other models marked dark grey in the Tables 5–6.

Table 5

HLM-DM test results (p value) for quadratic loss function, alternative hypothesis: model-based forecast more accurate than RW

	1M	3M	6M	12M
AR1	0.506	0.325	0.345	0.334
VAR1	0.971	0.959	0.819	0.807
PPP	0.336	0.992	0.729	0.341
CHEER	0.678	0.996	0.919	0.966
CHEER_RiskPremium	0.813	0.959	0.885	0.769
CHEER_HBS	0.666	0.906	0.974	0.791
CHEER_HBS_RiskPremium	0.861	0.984	0.972	0.792
BEER_CHEER	0.798	0.794	0.981	0.779

Table 6

HLM-DM test results (p value) for linear loss function, alternative hypothesis: model-based forecast more accurate than RW

	1M	3M	6M	12M
AR1	0.258	0.129	0.213	0.187
VAR1	0.901	0.933	0.895	0.782
PPP	0.113	0.908	0.862	0.256
CHEER	0.526	0.955	0.956	0.967
CHEER_RiskPremium	0.743	0.913	0.933	0.967
CHEER_HBS	0.418	0.625	0.969	0.768
CHEER_HBS_RiskPremium	0.800	0.925	0.908	0.885
BEER_CHEER	0.399	0.394	0.964	0.850

Direction of change statistic for most of models and horizons shows that forecasts generated from VECM more frequently predict direction of change of exchange rate. Furthermore, in most cases these results are statistically significant, what is displayed in Table 7 in light grey. It means that VECM models could be useful for investors, in particular those which have availability to predict direction of change in 1month horizon, as: PPP, CHEER, CHEER_RiskPremium, CHERR_HBS, CHEER_HBS_RiskPremium.

Table 7

Direction of change statistics for the full out of sample, all models and horizons

Model	forecast horizon	AR1	VAR1	PPP	CHEER	CHEER_Risk Premium	CHEER_HBS	CHEER_HBS_ RiskPremium	BEER_CHEER
Directional change	1	0.505	0.450	0.615	0.606	0.596	0.624	0.596	0.550
	3	0.514	0.440	0.468	0.431	0.550	0.514	0.495	0.459
	6	0.569	0.440	0.569	0.505	0.532	0.606	0.569	0.523
	12	0.541	0.495	0.495	0.587	0.569	0.578	0.550	0.596
Model	**forecast horizon**	**AR1**	**VAR1**	**PPP**	**CHEER**	**CHEER_Risk Premium**	**CHEER_HBS**	**CHEER_HBS_ RiskPremium**	**BEER_CHEER**
p-value of test statustcs for	1	0.462	0.853	0.009	0.014	0.023	0.005	0.023	0.147
	3	0.387	0.892	0.748	0.924	0.147	0.387	0.538	0.805
	6	0.076	0.892	0.076	0.462	0.252	0.014	0.076	0.317
	12	0.195	0.538	0.538	0.035	0.076	0.053	0.147	0.023

3.3. Estimation results for rolling-window

Comparison of forecasts accuracy and direction measures for rolling windows allows to investigate robustness of forecasting performance of analyzed methods in time. Consistently with the chapter 2.4, rolling-windows containing 6 year period.

For RMSE not surprisingly all methods generates higher RMSE in samples including pre-crisis and crisis period, than later samples including only post-crisis period, when volatility of EUR/PLN became milder. This difference is bigger for longer horizons. Relative RMSE against RW for VECM models in most cases, in particular for shorter horizon in most cases is lower for. However almost all analyzed VECM models for at least one subsample generated higher RMSE statistics than RW. Interestingly forecasts from AR1 models, generated lower RMSE statistics for all subsamples, for horizon longer than 3 months. This analysis is presented only for RMSE, as MSE for the whole out of sample period for most of VECM models is larger than for RW.

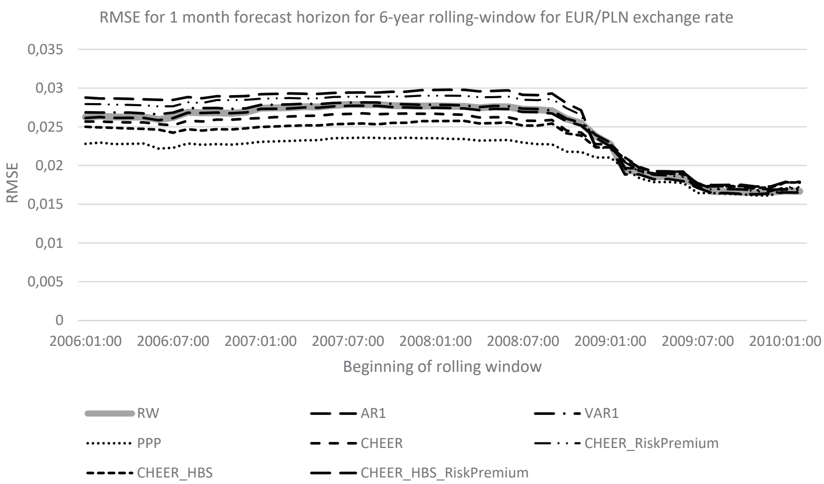

Figure 3. RMSE for 1 month horizon for 6-year rolling-window for analysed forecasting methods

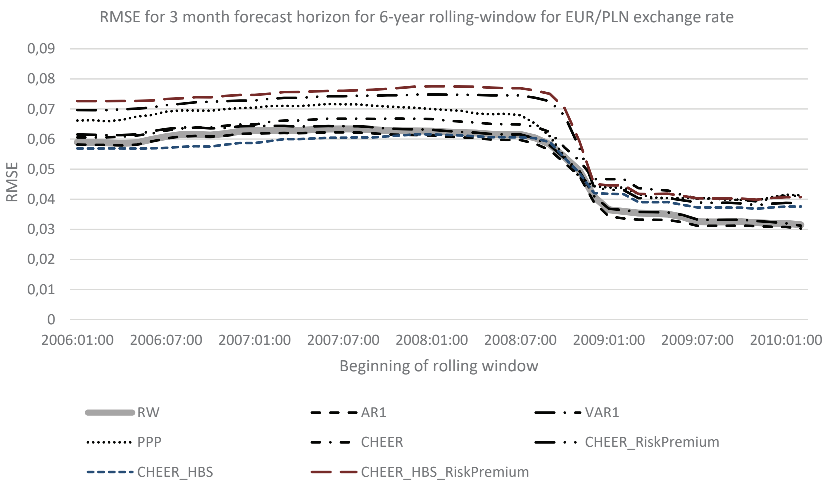

Figure 4. RMSE for 3 month horizon for 6-year rolling-window for analysed forecasting methods

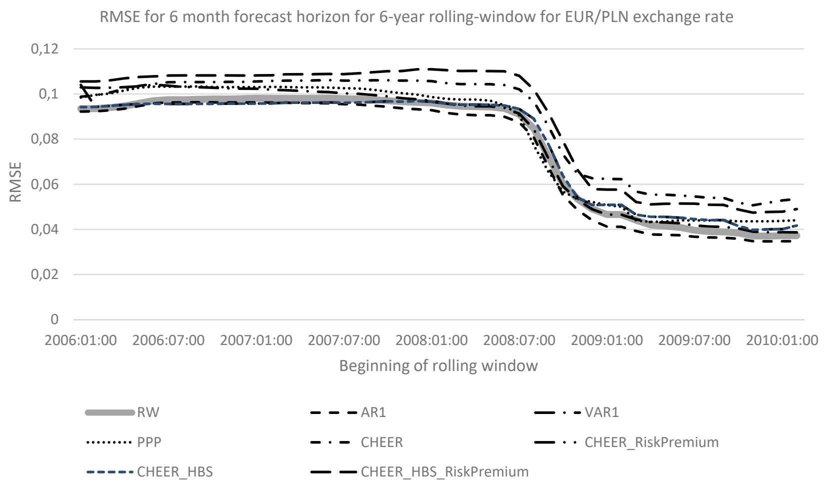

Figure 5. RMSE for 6 month horizon for 6-year rolling-window for analysed forecasting methods

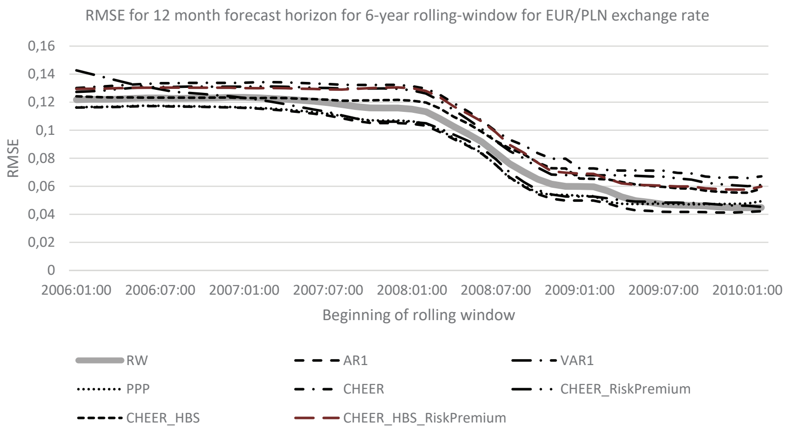

Figure 6. RMSE for 12 month horizon for 6-year rolling-window for analysed forecasting methods

Direction of change statistics for rolling-windows for forecasts generated from VECM models show that in particular for longer horizons ability of those models to predict direction of change of exchange rate EUR/PLN varies in time. However, for 1 month horizon: *PPP*, *CHEER*, *CHEER_RiskPremium*, *CHEER_HBS*, *CHEER_HBS_RiskPremium* for all analyzed subsamples have statistically significant ability to predict direction of change of exchange rate. AR1 and VAR1 models have no statistically significant ability in many subsamples to predict direction of change statistics.

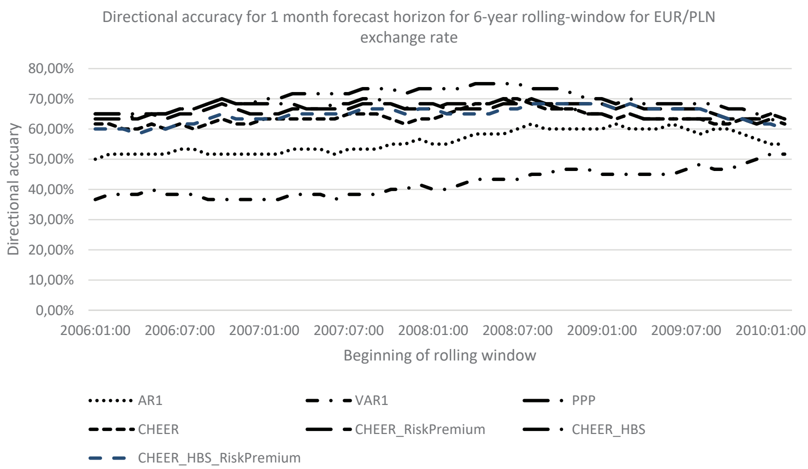

Figure 7. Direction of change statistics for 1 month horizon for 6-year rolling-window for analysed forecasting

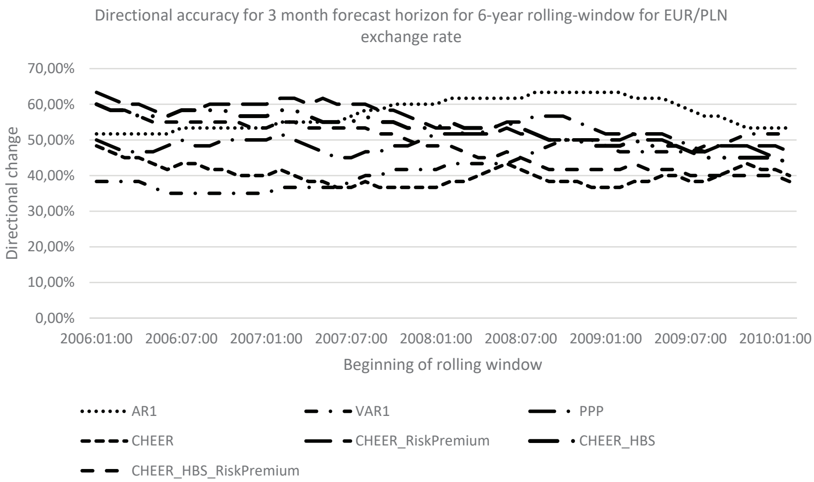

Figure 8. Direction of change statistics for 3 month horizon for 6-year rolling-window for analysed forecasting

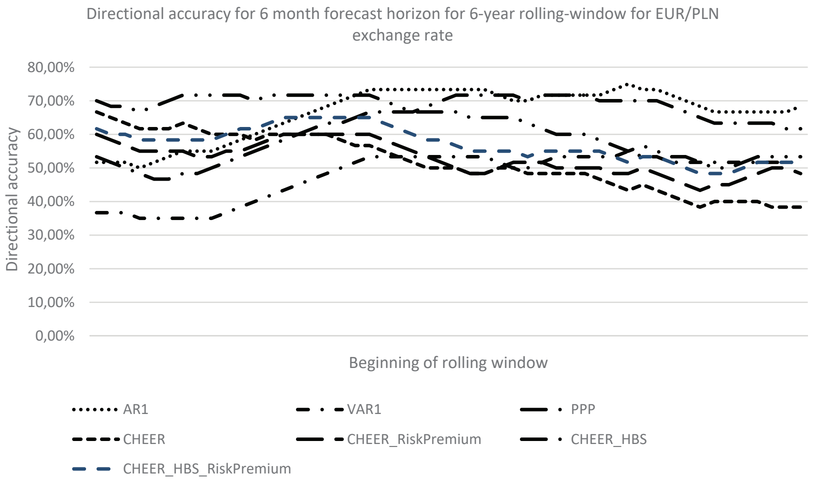

Figure 9. Direction of change statistics for 6 month horizon for 6-year rolling-window for analysed forecasting

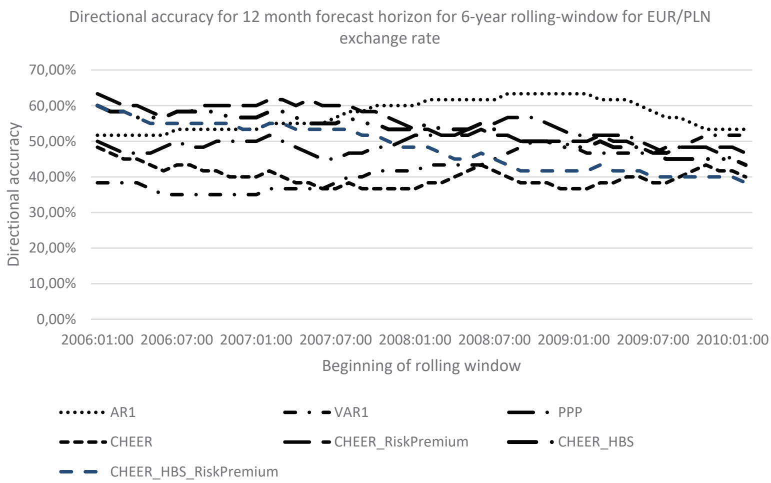

Figure 10. Direction of change statistics for 12 month horizon for 6-year rolling-window for analysed forecasting

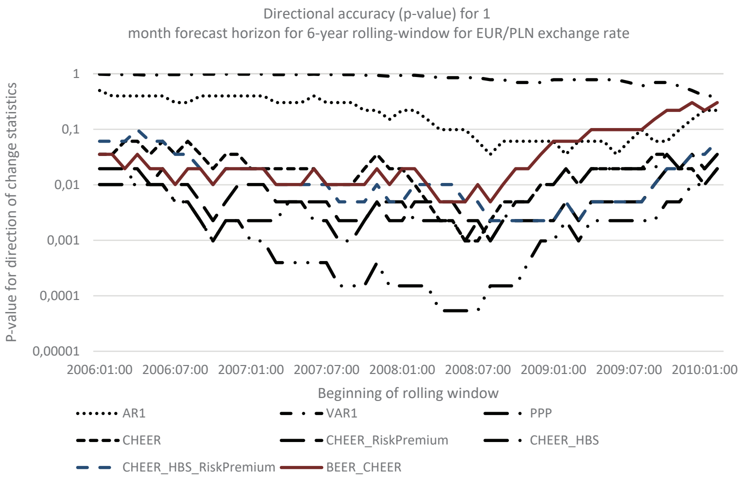

Figure 11. P-value for direction of change statistics for 1 month horizon for 6-year rolling-window for analysed forecasting

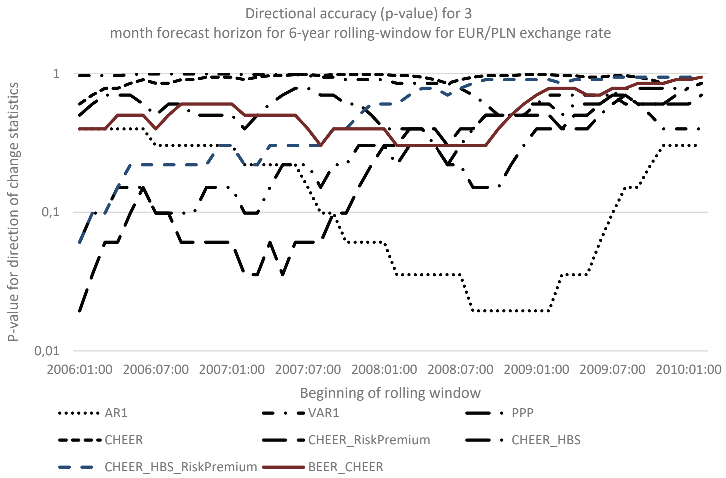

Figure 12. P-value for direction of change statistics for 3 month horizon for 6-year rolling-window for analysed forecasting

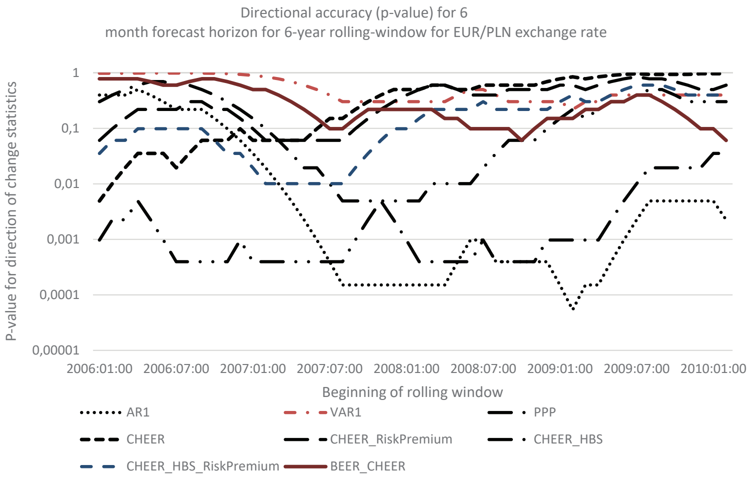

Figure 13. P-value for direction of change statistics for 6 month horizon for 6-year rolling-window for analysed forecasting

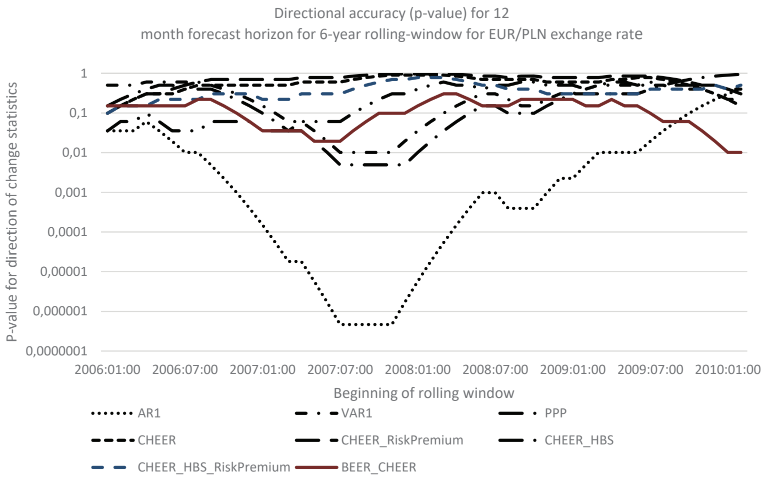

Figure 14. P-value for direction of change statistics for 12 month horizon for 6-year rolling-window for analysed forecasting

4. CONCLUSION

In the research forecasting properties of various VECM models — incorporating different exchange rate theories had been investigated for EUR/PLN exchange rate. In particular point forecast accuracy and directional accuracy had been assessed for recursively estimated models for out of sample containing period 2006:02–2015:12. RW without drift was a benchmark, while for comparison purposes forecasts from AR1 for exchange rate and VAR1 for PPP variables had been generated.

Not surprisingly, none of VECM model significantly beats random walk, neither for absolute nor for quadratic loss function. Furthermore, forecasts from some VECM models for certain horizons are significantly beaten by RW, as CHEER model for all horizons expect 1 month. However, many of VECM models has ability to predict correctly direction of change of EUR/PLN. In contrast to Cheung and others (2005) the strongest and the most robust (tested for rolling windows) results had been found for 1 month horizon forecasts. For other horizons, expect 3 month horizon for some VECM models ability to predict correctly direct of change was found as for *CHEER*, *CHEER_RiskPremium*, *CHEER_HBS* and *CHEER_BEER* for 12 month horizon. However these results are not robust in time.

Looking into particular VECM models it is hard to select the one which robustly surpass others in time and forecast accuracy measure. For 1-month horizon PPP model in terms of RMSE beats other VECM models also for rolling-window out of sample. However, in terms of directional accuracy, for 1-month horizon *CHEER_HBS* model slightly beats PPP model. Interestingly for AR1 model for some subsamples generate slightly more accurate forecasts (measured by RMSE) than any investigated VECM model and also than VAR1 model. All in all the research do not give support for estimation more complex VECM models which contains more than one cointegration relation to forecasting exchange rate purposes, despite the fact that those models could be useful for economic policy purposes; e.g. Kelm (2013).

BIBLIOGRAPHY

Adhikari B.K., Hillard J.E. (2014), *The VIX, VXO and realised volatility: a test of lagged and contemporaneous relationships,* International Journal of Financial Markets and Derivatives vol. 3, issue 3.

Akaike H. (1998), I*nformation Theory and an Extension of the Maximum Likelihood Principle*, in: E. Parzen, K. Tanabe, G. Kitagawa (eds), *Selected Papers of Hirotugu Akaike*, Springer Series in Statistics, 199–213.

Ardic N., Ergin O., Senol G. (2008), *Exchange rate forecasting: Evidence from the emerging central and eastern european economies*, MPRA Paper, (7505), https://mpra.ub.uni-muenchen.de/7505/1/MPRApaper505.pdf.

Bęza-Bojanowska J., MacDonald M.R. (2009), *The behavioral zloty/euro equilibrium exchange rate.,* NBP Working Paper, (55), http://www.nbp.pl/publikacje/materialyistudia/55en.pdf.

Cheung Yin-Wong, Chinn M.D., Pascual A.G. (2005), *Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?*, Journal of International Money and Finance, 24 (7): 1150–75.

Cuaresma J., Hlouskova J. (2005), *Beating the random walk in central and Eastern Europe*, Journal of Forecasting, (24), 189–201, http://onlinelibrary.wiley.com.wiley.han.uek.krakow.pl/doi/10.1002/for.952/epdf.

Diebold F.X. (2012), *Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests* (no. w18391), National Bureau of Economic Research.

Diebold F.X., Mariano R.S. (1995), *Comparing Predictive Accuracy*, Journal of Business & Economic Statistics, 13(3), 253–263.

Engel C., Mark N.C., West K.D. (2007), *Exchange Rate Models Are Not as Bad as You Think*, NBER Macroeconomics Annual 2007, D. Acemoglu, K.S. Rogoff, M. Woodford (eds), 381–441, Chicago and London: University of Chicago Press.

Galimberti J.K., Moura M.L. (2013), *Taylor rules and exchange rate predictability in emerging economies*, Journal of International Money and Finance, 32, 1008–1031, http://www.sciencedirect.com.sciencedirect.han.uek.krakow.pl/science/article/pii/S02615 60612001581.

Giacomini R., White H. (2006), *Tests of Conditional Predictive Ability*, Econometrica, 74, issue 6, 1545–1578, http://EconPapers.repec.org/RePEc:ecm:emetrp:v:74:y:2006:i:6:p:1545-1578.

Harve D., Leybourne S., Newbold P. (1997), *Testing the equality of prediction mean square errors*, International Journal of Forecasting, 13(2), 281–291.

Ince O., Molodtsova T., Pappel D.H. (2016), *Taylor rule deviations and out-of-sample exchange rate predictability*, Journal of International Money and Finance, 69, 22–44.

Johansen S. (1988), *Statistical analysis of cointegration vectors*, Journal of Economic Dynamics and Control, 12, 70–86.

Johansen S. (1991), *Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models*, Econometrica, 59(6), 1551–1580.

Johansen S. (1995), *Likelihood-Based Inference in Cointegrated Vector Auto-Regressive Models*, Oxford University Press, New York.

Johansen S. (2002), *A small sample correction for the test of cointegration rank in the vector autoregressive model*, Econometrica, Vol.70, No. 5, 1929–1961.

Johansen S. (2005), *The interpretation of cointegrating coefficients in the cointegrated vector autoregressive model*, Oxford Bulletin of Economics and Statistics, 67, 93–104.

Kębłowski P., Welfe A. (2012), *A risk-driven approach to exchange rate modelling*, Economic Modelling, 29, 1473–1482, http://www.sciencedirect.com /science/article/pii/S0264999312000375.

Kelm R. (2013), *Kurs złoty/euro: teoria i empiria*, Wyd. UŁ, Łódź.

Leitch G., Tanner J.E. (1991), *Economic forecast evaluation: profits versus the conventional error measures*, American Economic Review, 81, 580–590.

Ljung G.M., Box G.E.P. (1978), *On a Measure of Lack of Fit in Time Series Models*, Biometrika, 65 (2), 297–303, doi:10.1093/biomet/65.2.297.

Lütkepohl H., Poskitt D.S. (1998), *Consistent estimation of the number of cointegration relations in a vector autoregressive model*, in: R. Galata, H. Küchenhoff (eds), *Econometrics in Theory and Practice. Festschrift for Hans Schneeweiß*, Physica, Heidelberg, 87–100.

Lütkepohl H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin, Heidelberg, New York.

Melvin M., Prins J., Shand D. (2013), *Forecasting Exchange Rates: An Investor Perspectiv*e, in: Handbook of Economic Forecasting, Volume 2B, E. Graham, A. Timmermann (eds), 721–50, Amsterdam and Oxford: Elsevier, North-Holland.

Meese R.A., Rogoff K. (1983), *Empirical models of the exchange rate: do they fit out of sample*, Journal of International Economics, (14), 3–24.

Makridakis S., Wheelwright S., Hyndman R.J. (1998), *Forecasting: Methods and applications*, eds. 3, John Wiley and Sons, New York.

Mućk J., Skrzypczyński P. (2012), *Can we beat the random walk in forecasting CEE exchange rates?*, NBP Working Paper, (127), http://pki.nbp.pl/publikacje/materialy_i_studia/127_en.pdf

Naszodi A. (2011), *Beating the random walk in central and eastern europe by survey forecasts,* MNB Working Papers, (3), https://www.mnb.hu/letoltes/wp-2011-03.pdf.

Papież M., Śmiech S. (2015), *Modelowanie i prognozowanie cen surowców energetycznych*, C.H. Beck, Warszawa.

Pipień M., Roszkowska S. (2015), *Quarterly estimates of regional GDP in Poland — application of statistical inference of function of parameters,* NBP working paper, (219), http://www.nbp.pl/publikacje/materialy_i_studia/219_en.pdf.

Rossi B. (2013), *Exchange rate predictability*, Journal of Economic Literature, 51 (4), 1063–1119.

Rubaszek M., Serwa D. (2009), *Analiza kursu walutowego*, C.H. Beck, Warszawa.

Rubaszek M., Skrzypczyński P., Koloch G. (2010), *Forecasting the Polish zloty with nonlinear models,* Central European Journal of Economic Modelling and Econometrics, 2(2), 151–167, http://cejeme.org/publishedarticles/2013-12-15-634938559231250000-2549.pdf.

Schwarz G. (1978), *Estimating the Dimension of a Model*, The Annals of Statistics 6, (2), 461–464, doi:10.1214/aos/1176344136.

Zeliaś A., Pawełek B., Wanat S. (2008), *Prognozowanie ekonomiczne: teoria, przykłady, zadania*, PWN, Warszawa.