The modeling of earnings per share of Polish companies for post-financial crisis period using random walk and ARIMA models*

Wojciech Kuryłek Ph.D. – University of Warsaw,
Faculty of Management

***This project is XXX funded**

I. Aims

Project

- Is focused on the **forecasting of EPS firms listed on the Warsaw Stock Exchange**
- **Only a small fraction (20%) of companies is covered by financial analysts in Poland,** contrary to the situation in US, **so the impotence of time series forecasting matters**
- **The most recent data** coming from a period of relative earning stability i.e. ranging **from the last financial crisis 2008-2009 to the pandemia shock of 2020**
- **Instead of mean absolute percentage error (MAPE), a modification of this measure is used (MAAPE)**

II. The modeling of EPS in literature

The history of research

- 1960's – the beginning of EPS forecasting literature and comparion of mechanical forecasts with security analysts' predictions [Cragg and Malkiel (1968)]
- 1970 – development of ARIMA time series models [Box and Jenkins]
- 1972-1977 – development the premier models of ARIMA type for EPS [Ball and Watts (1972), Watts (1975), Griffin (1977), Foster (1977), Brown and Rozeff (1977)]
- 1979 – 1984 - Building a consensus that ARIMA-type models performed the best [Lorek (1979), Bathke and Lorek (1984)]
- 1987 - the groundbreaking work that forecasts provided by financial analysts were better than those made by time series models [Brown et al. (1987)]
- 2020 – questioning the superiority of analysts over time series [Pagach and Warr (2020)]

I II. The modeling of EPS in literature

The problems with existing research

- The literature is **mostly focused on the US** with few exceptions [Bao (1996), Grigaliūnienė (2013)]
- All the existing research is **limited to the time period ending prior to the 2009 year**
- The most popular **MAPE** error metric that is related to the **explosion** of this measure when its denominator is very small i.e. when **actual earnings are close to zero**, which is often a case

III. Hypothesis and results

Hypothesis

- “ are technically **complex ARIMA models** more appropriate for EPS forecasting of WSE companies, than **naive random walk models**? ”

Research results

- **The best model**, is the **seasonal random walk** (SRW) model across all examined quarters, which describes quite well the behavior of the Polish market compared to other models. Hence, **conclusions drawn for the US might not hold for emerging economies** because of the **much simpler behavior** of these markets.
- Medians of errors of the firm-specific (**BJ**) model are **statistically not different from** the best seasonal random walk (**SRW**) model for the most of analyzed periods

IV. Data

- Data source is **EquityRT,** which is a product of the Turkish company **RASYONET**
- **267** companies listed on **Warsaw Stock Exchange**
- **Excluded** companies with **splits/reverse splits** because such operations influence substantially EPS behavior
- **Quarterly** data
- **Q1 2010 - Q4 2018** (36 quarters) are used for the **estimation** of various models
- **Q1 2019 - Q4 2019** (4 quarters) are used as **hold-out validation** sample for testing forecast accuracy

V. Methodology

TESTED MODELS

- The **naive** models:
  - **The random walk model (RW)**:

$$E_{t-1}(Q_t) = Q_{t-1}$$

  - **The random walk model with drift (RWD)**:

$$E_{t-1}(Q_t) = \delta + Q_{t-1}$$

  - **The seasonal random walk model (SRW)**:

$$E_{t-1}(Q_t) = Q_{t-4}$$

  - **The seasonal random walk model with drift (SRWD)**:

$$E_{t-1}(Q_t) = \delta + Q_{t-4}$$

V. Methodology

TESTED MODELS

- The **SARIMA** models:

$$\varphi(B)(1-B)^d\Phi(B^S)(1-B)^D Q_t = \theta(B)\Theta(B^S)\varepsilon_t + \theta_0$$

where $BQ_t = Q_{t-1}$ and $B^S Q_t = Q_{t-4}$ and $\varphi(B)$ and $\Phi(B^S)$ are polynomials

- **The Griffin-Watts (GW) model** is the SARIMA of order (0,1,1)×(0,1,1):

$$E_{t-1}(Q_t) = Q_{t-1} + (Q_{t-4} - Q_{t-5}) - \theta_1\varepsilon_{t-1} - \Theta_1\varepsilon_{t-4} - \theta_1\Theta_1\varepsilon_{t-5}$$

- **The Foster (F) model** is the SARIMA of order (1,0,0) × (0,1,0):

$$E_{t-1}(Q_t) = Q_{t-4} + \varphi_1(Q_{t-1} - Q_{t-5}) + \theta_0$$

- **The Griffin-Watts (GW) model** is the SARIMA of order (0,1,1)×(0,1,1):

$$E_{t-1}(Q_t) = Q_{t-4} + \varphi_1(Q_{t-1} - Q_{t-5}) - \Theta_1\varepsilon_{t-4}$$

- **The firm-specific Box-Jenkins (BJ) model**:

parameters (p,d,q)×(P,D,Q), as well as the constant term θ_0, are chosen individually for every company

V. Methodology

ERROR METRIC

- **An absolute percentage error (APE)** of the forecasts for an i-th individual company in the j-th quarter of 2019 is defined as:

$$APE_j^i = \left|\frac{A_j^i - F_j^i}{A_j^i}\right|$$

but APE has a significant disadvantage: it produces infinite or undefined values when the actual values are zero or close to zero, which is a common occurrence in the forecasting of earnings

- **An arctangent absolute percentage error**, which is a novel approach in the literature:

$$AAPE_j^i = arctan\left(\left|\frac{A_j^i - F_j^i}{A_j^i}\right|\right)$$

- **The mean arctangent absolute percentage error**: $MAAPE^i = \frac{1}{4}\sum_{j=1}^{4} AAPE_j^i$

V. Methodology

STASTICAL TESTS

- **The Kruskal-Wallis one-way H-test** [Corder and Foreman (2009)] - a nonparametric test that avoids difficulties concerning the potential normality of the errors:

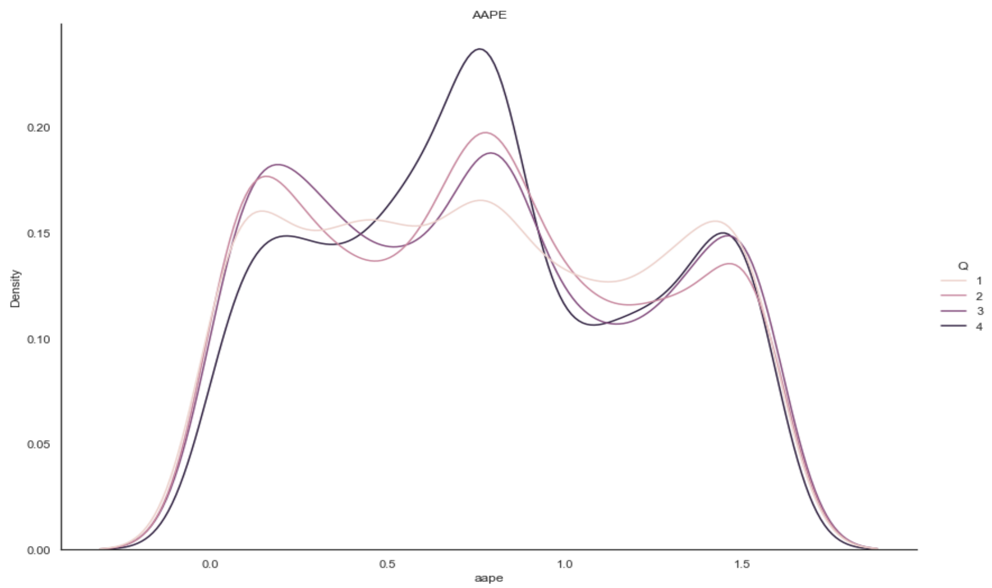
$$H_o\text{: } medians\ of\ AAPEs\ of\ all\ 8\ models\ are\ the\ same$$

- **The Wilcoxon two-sided test** [Wilcoxon (1945)] – a nonparametric test for all model paires:

$$H_o\text{: } medians\ of\ AAPEs\ of\ a\ pair\ of\ models\ are\ the\ same$$

VI. Empirical Results

The kernel density estimators of arctangent absolute percentage errors for forecasted quarters



- Surprisingly, **forecast errors don't increase with forecast horizons**

VI. Empirical Results

Summary statistics on forecast terrors and Kruskal-Wallis test:

		Quarters								All Quarters	
		Q1 MAAPE	Q1 Rank	Q2 MAAPE	Q2 Rank	Q3 MAAPE	Q3 Rank	Q4 MAAPE	Q4 Rank	MAAPE	Rank
model	RW	0,89	5,21	0,80	4,68	0,83	5,01	0,74	3,97	0,81	4,72
	RWD	0,92	5,81	0,84	5,26	0,88	5,59	0,79	4,96	0,85	5,40
	SRW	0,66	3,69	0,70	3,98	0,65	3,74	0,74	3,97	0,69	3,85
	SRWD	0,70	4,03	0,73	4,35	0,73	4,25	0,80	4,67	0,74	4,33
	GW	0,78	4,51	0,80	4,81	0,77	4,52	0,82	4,84	0,79	4,67
	F	0,77	4,38	0,75	4,49	0,75	4,35	0,80	4,75	0,77	4,49
	BR	0,75	4,16	0,74	4,24	0,71	4,14	0,80	4,62	0,75	4,29
	BJ	0,71	4,20	0,69	4,19	0,74	4,40	0,73	4,23	0,72	4,25
	H statistics		63,92		19,79		38,18		10,79		36,56
	H pvalue		0,00		0,01		0,00		0,15		0,00

- **SRW** model performs **the best** having **the lowest rank** in respective quarters as well as for all quarters
- The **Kruskal-Wallis test** shows **that null hypothesis,** that median of arctangent absolute percentage errors (AAPEs) of all 8 models are statically the same*, **can be rejected in all cases except 4th quarter**

* - at 0.05 statistically significance level

VI. Empirical Results

P-values of paired Wilcoxon test for forecast errors in Q1 2019

model	RWD	SRW	SRWD	GW	F	BR	BJ
RW	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
RWD		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
SRW			0,0218	0,0001	0,0000	0,0005	0,0487
SRWD				0,0052	0,0129	0,0887	0,5389
GW					0,4606	0,0609	0,0240
F						0,7939	0,1090
BR							0,1573

Only **BJ** model is not statistically significantly different than **SRW** model

P-values of paired Wilcoxon test for forecast errors in Q2 2019

model	RWD	SRW	SRWD	GW	F	BR	BJ
RW	0,0000	0,0004	0,0210	0,3844	0,0412	0,0066	0,0004
RWD		0,0000	0,0003	0,0541	0,0012	0,0001	0,0000
SRW			0,0036	0,0002	0,0001	0,5705	0,9455
SRWD				0,0215	0,2108	0,9248	0,2197
GW					0,0763	0,0010	0,0007
F						0,4492	0,0856
BR							0,4630

Only **BR**, **BJ** models are not statistically significantly different than **SRW** model

* - at 0.05 statistically significance level

VI. Empirical Results

P-values of paired Wilcoxon test for forecast errors in Q3 2019

model	RWD	SRW	SRWD	GW	F	BR	BJ
RW	0,0000	0,0000	0,0001	0,0028	0,0001	0,0000	0,0003
RWD		0,0000	0,0000	0,0001	0,0000	0,0000	0,0000
SRW			0,0001	0,0020	0,0000	0,1113	0,0005
SRWD				0,1770	0,2032	0,2569	0,5654
GW					0,1947	0,0441	0,6852
F						0,1285	0,9419
BR							0,2883

Only **BR** model is not statistically significantly different than **SRW** model

P-values of paired Wilcoxon test for forecast errors in Q4 2019

model	RWD	SRW	SRWD	GW	F	BR	BJ
RW	0,0000	0,0000	0,0000	0,0011	0,0000	0,0213	0,7377
RWD		0,0000	0,4202	0,1339	0,8936	0,4939	0,0785
SRW			0,0000	0,0011	0,0000	0,0213	0,7377
SRWD				0,1578	0,6280	0,8037	0,0281
GW					0,2343	0,0502	0,0045
F						0,8973	0,0196
BR							0,0547

Only **BJ** model is not statistically significantly different than **SRW** model

* - at 0.05 statistically significance level

VI. Empirical Results

P-values of paired Wilcoxon test for forecast errors for all quarters

model	RWD	SRW	SRWD	GW	F	BR	BJ
RW	0,0000	0,0000	0,0000	0,0163	0,0003	0,0000	0,0000
RWD		0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
SRW			0,0001	0,0000	0,0000	0,0042	0,0930
SRWD				0,0066	0,0183	0,7726	0,2826
GW					0,0984	0,0008	0,0007
F						0,0282	0,0131
BR							0,3392

Only **BJ** model is not statistically significantly different than **SRW** model

* - at 0.05 statistically significance level

VI. Roboustness check

- The models are estimated using **expanding window approach** i.e. the sample Q1 2010 – Q4 2017 is used for their estimation and Q1 2018 – Q4 2018 for their testing. Then, the same procedure is applied taking the year 2017 to validate the results.

		2017		2018		2019	
		MAAPE	Rank	MAAPE	Rank	MAAPE	Rank
model	RW	0,83	4,78	0,86	4,97	0,81	4,72
	RWD	0,85	5,42	0,88	5,60	0,85	5,40
	SRW	0,69	3,86	0,71	3,81	0,69	3,85
	SRWD	0,72	4,29	0,76	4,27	0,74	4,33
	GW	0,79	4,75	0,80	4,62	0,79	4,67
	F	0,75	4,45	0,78	4,41	0,77	4,49
	BR	0,74	4,24	0,75	4,19	0,75	4,29
	BJ	0,72	4,21	0,73	4,14	0,72	4,25
	H statistics		32,07		40,28		36,56
	H pvalue		0,00		0,00		0,00

SRW model is characterized by the lowest rank i.e. gives **the best results** not only in **2019**, but also in **2018** and **2017**

- P-values of paired Wilcoxon test of forecast errors for all quarters 2017-2019 and SRW model

year	model	RWD	SRWD	GW	F	BR	BJ
2017	SRW	0,0000	0,0013	0,0000	0,0000	0,0007	0,0233
2018	SRW	0,0000	0,0000	0,0000	0,0000	0,0488	0,1686
2019	SRW	0,0000	0,0001	0,0000	0,0000	0,0042	0,0930

Only the errors of **SRW** and **BJ** model are statistically different in 2018 and 2019

* - at 0.05 statistically significance level

VII. Conclusions

- **Forecast errors don’t increase with forecast horizons**, as one would expect
- **The best model**, with the lowest rank, is **the seasonal random walk** (SRW) model across all quarters, which describes quite well the behavior of the Polish market compared to other models
- The **medians of errors** of the analyzed models **differ statistically** significantly **in almost all quarters**
- Medians of errors of the firm-specific (**BJ**) model are **statistically not different** from the best seasonal random walk **(SRW**) model for the most of analyzed periods

VII. Conclusions

- The superiority of the seasonal random walk model (SRW) implies that the underlying **EPS generating process exhibits neither autoregressive nor moving average parts ant there is no drift.**
- **The horizontal performance of** the stock market index **WIG** during the analyzed period **implies the absence of a trend**.
- In the context of emerging markets, **the absence of moving average part is consistent with the fact that a lower fraction of companies publishes the forecasts of their earnings** compared to developed markets, and hence not for so many companie past forecast errors result in the correction of the performance of future earnings.
- **The non-existence of autoregressive part may** in turn **be related to the dominance of seasonal component relative to past EPS behavior**, which might imply that the emerging market companies are more seasonal than those operating on the developed markets.