Determinants of Short-term Volatility at the Warsaw Stock Exchange: In-sample vs. Out-of-sample Forecasts from Factor and Predictive GARCH Models*

Janusz Brzeszczynski‡

Aleksander Welfe§

October 2004
Discussion Paper 2004/08

Abstract

The paper presents factor and predictive GARCH(1,1) models of the Warsaw Stock Exchange (WSE) main index WIG. An approach where the mean equation of the GARCH model includes a deterministic part is applied. The models incorporate such explanatory variables as volume of trade and major international stock market indices. The paper exploits the direction quality measures that can be used as alternative measures to evaluate model goodness of fit. Finally, the in-sample versus the out-of-sample forecasts from the estimated models are compared and model forecasting performance is discussed.

Keywords: stock market; factor GARCH, predictive GARCH; in-sample versus out-of sample forecasts; direction quality measures; emerging markets

JEL Classification: G15, C51

* We are grateful to Michael Melvin and Richard Lyons for helpful comments on the earlier versions of the paper. Janusz Brzeszczynski wishes to thank the Fulbright Commission for a research grant at the Arizona State University. The usual disclaimer applies.

‡ CERT, Heriot-Watt University. Contact address: Centre for Economic Reform and Transformation, School of Management and Languages, Heriot-Watt University, Edinburgh EH14 4AS, UK. Email: J.Brzeszczynski@hw.ac.uk.

§ Chair of Econometric Models and Forecasts, University of Łódź, Poland.
1. Introduction

The volatility of financial markets, when measured on low frequency data, is usually determined by macroeconomic factors. In case of high frequency observations, however, most of the macro-variables cannot be used due to the data unavailability. On the other hand the volatility of financial market prices is in the short-run influenced by the inflow of new information. It determines decisions of the investors whether to shift demand or supply of a given financial instrument. This results in price changes. Due to the random nature of market information arrival (and different reactions of investors to the same piece of information caused by the heterogeneity of opinions and views), the problem of modeling financial market variability is obviously difficult to quantify.

The efforts to describe volatility of financial instruments prices have led to the discovery of some causal relationships existing in financial markets. From this point of view volatility can be perceived as dependent on other factors – both external and internal to the given system (market).

The paper presents GARCH models of the Warsaw Stock Exchange (WSE) main index WIG. The following hypotheses were put forth and verified.

Firstly, the volatility in a short-run depends on the volume of traded stocks as well as its realizations in the past.

Secondly, the volatility of the WIG index is influenced by the situation on the international financial markets. As the indicator of the short-term trends (and the primary cause for the volatility on other markets worldwide) the following indices were chosen for the proposed model specifications: Dow Jones Industrial Average (DJIA), NASDAQ Composite as well as two major European indices DAX and FTSE.
2. Causal relationships on financial markets

The search of causal dependencies existing on any financial market worldwide usually focuses on two directions: the price-volume relationship present on the local market and the interdependency between the given market and other international stock exchanges.

The first relationship exists due to the globalization of financial market and the increasing role of short-term funds circulating around the globe and transmitting the “shocks” among major international exchanges and – as a side-effect - to emerging markets. It is widely believed that the behavior of the two main indices in the USA: DJIA and NASDAQ Composite is in the short-run the primary source of international capital movements and resulting stock price changes. Very often their volatility has even stronger impact on the local stock exchanges than the current (macro)economic situation in a given country.

The price-volume relationship means the correlation between stock prices (indices) and volume of trading. It is one of the most extensively explored causal dependency in the field of high frequency data financial econometrics because of the data availability and congruence of different frequencies of observations for both variables.

The existence of price-volume relationship has been widely documented in the economic literature of over four decades. The first attempts to describe this dependency were made by Osborne (1959), Granger, Morgenstern (1963), (1970) and Godfrey et al. (1964). Since then many further studies, most of them empirical in nature, appeared (see: Lyons, 1996, Campbell et al., 1992 and Blume et al.,
The review of the research on price-volume relationship can be found in Karpoff (1987) and Hsu (1998).

In spite of a strong empirical evidence there is little theoretical justification for the presence of this phenomenon. A few possible explanations of price-volume dependency on different financial markets were, however, proposed. The first one includes the hypothesis that the trading volume is positively correlated with price volatility because it is related to the number of intra daily transactions which in turn depend on the variance of daily price changes (see: Clark, 1973). Secondly, it is postulated that the price-volume relationship exists due to the extent of disagreement in the traders’ beliefs, which is positively correlated with the volume itself (see: Epps, Epps, 1976). The third hypothesis assumes that the daily price change and daily volume depend upon the average daily rate of information arrival on the market, the magnitude of traders disagreement in reaction to new information and market depth represented by the number of active traders present in the market (see: Tauchen, Pitts, 1983). Assuming the constancy of the first two factors, price-volume relationship can be explained by the changes in the market activity generated by the market participants.

Parallel to that, several interesting properties of volume in relation to price volatility have been discovered. Campbell et al. (1992) and Morse (1980) observe the dependency between volume of trading and the serial correlation of stock returns, although their results contradict each other. On the other hand, Lamoreux, Lastrapes (1990) find that the introduction of volume into the conditional variance equation of an ARCH class model leads to the vanishing of the ARCH effect. This may suggest that in case of financial time series data the clustering in trading volumes is responsible for the existence of the ARCH effects. Lyons (1996)
argues that on the FX market the low-intensity trades are more informative than when the trading intensity is high.

The intuition suggests that price-volume dependency is strongest on the thinly traded emerging markets, less efficient in the Fama (1970) sense, than on the well established markets.

3. Stock market in Poland

The Polish stock market was re-established in 1991. Since then a dynamic development of its infrastructure has taken place. It includes the creation of the institutions organizing and regulating the market, such as the Warsaw Stock Exchange (WSE) and Polish Securities Commission, as well as many of its active participants, i.e. the brokerage houses, investments funds etc. Furthermore, about one million individual investors have emerged.

Initially, the increase of the stock supply was the result of privatization of the previously state-owned companies. Later on also private sector started to compete for capital through new stock issues. Also the government started to be active in the area of fixed-income securities by issuing bonds and treasury bills. Recently, new derivative instruments have been introduced which marks the beginning of a new stage in market development.

The history of the market can be divided into two sub-periods. The first one embraces speculative growth and decline in the years 1993/94. The second period refers to more recent years, starting from 1995, when the stock prices behaviour has been much less volatile.

The pace of stock market growth and its relatively short history determine its following specific features. Firstly, there is still a low number of shares listed on
the WSE (in 2001 it hardly exceeded 200 companies). Secondly, the volume and the capitalization of the market are also very thin. The market capitalization / GDP ratio for Poland reaches the only the 10% level whereas in many developed countries it exceeds 100% (against the European average of 54%). Thirdly, low volume and liquidity lead to relatively high variance and frequent clustering of volatility.

The development of the stock market in Poland led to the introduction of a continuous trading system replaced the system of one price fixing sessions (the first companies started to be quoted in continuous trading in 1996).

The stock market in Poland has been very sensitive to both domestic (macro)economic news as well as developments on international financial markets.

4. Methodology

4a. Modeling approach

Volatility clustering is a well-known feature of any financial time series. In case of traditional econometric models, it causes obvious problems as it usually transmits to the variability of an error term and results in its heteroscedasticity. Therefore, the application of standard estimation techniques leads to the over- or underestimated structural parameters of the investigated model meaning that the results obtained are far from the true ones. Due to this fact, GARCH models were applied as the method of estimation (cf. Engle, 1982, Bollerslev, 1986 and for the summary: Bollerslev et al., 1992 and Bollerslev et al., 1994).
In current study, a standard linear generalized autoregressive conditional heteroskedastic model GARCH(S,Q):

\[ r_t = x_t \alpha + \xi_t , \quad (1a) \]

\[ \xi_t = \vartheta_t \sqrt{h_t} , \quad (1b) \]

\[ h_t = \gamma_0 + \sum_{s=1}^{S} \gamma_s \xi_{t-s}^2 + \sum_{q=1}^{Q} \phi_q h_{t-q} \quad (1c) \]

where \( \vartheta_t \) : IID(0,1) , \( \xi_t \) : IID(0,\( \sigma_\xi^2 \)) ,

\[ E(r_t) = E(r_t \mid r_{t-1}) = x_t \alpha \quad (2a) \]

\[ D^2(r_t \mid r_{t-1}) = h_t \quad (2b) \]

\( x_t \) stands for the vector of k explanatory variables and \( \alpha \) for the vector of the structural parameters, is applied. The dependent variable in (1a) is the rate of return, \( r_t = \ln y_t - \ln y_{t-1} \), where \( y_t \) is the corresponding financial market price (i.e. stock or currency price).

The model (1a)-(1c) can be extended in two different directions. Firstly, the conditional variance function \( h_t \) in (1c) can be further expanded to a more sophisticated form aiming at better description of the volatility of \( \xi_t \). Most of the empirical studies conducted for different markets and instruments indicate, however, that the best results are given by the standard GARCH(1,1) model with the following conditional variance function:

\[ h_t = \gamma_0 + \gamma_1 \xi_{t-1}^2 + \phi_1 h_{t-1} \ . \quad (3) \]

Secondly, the deterministic part in basic regression equation (1a) can incorporate the explanatory variables which have causal effects on the dependent variable. This approach ensures that the model is not deprived of the causal dependencies.
which (if properly captured) may themselves explain a large portion of the volatility and, on the other hand, it guarantees the desired properties of the estimators.

The modeling approach adopted in this investigation is thus focused on expanding the deterministic part in the mean equation while the conditional variance function does not change and in all the models estimated has the GARCH(1,1) specification. Such a strategy is based on the assumption that, for modeling the returns of the financial instruments, it is more important to capture the factors which represent the capital flows within different segments of the financial markets (represented by the respective financial instruments) rather than striving to discover a better conditional variance function in the GARCH model (or more general: in the ARCH-class model). In this approach GARCH plays the role of correcting the estimator (and the estimates themselves), which leads to obtaining better estimates of the parameters in the mean equation.

4b. Factor and predictive GARCH models

In the financial literature the factor models usually do not posses a dynamic structure (what is more – in many of their formulations there is not even any time subscript indicated). It means that the implicit assumption behind their theoretical background is that the relationships which the model must capture are simultaneous or – in the best cases - they are defined as expectations in time \( t \) for future periods from \( t + 1 \) to some \( t + s \) (as for example in the portfolio theory or the APT theory). Furthermore, the classical factor models are based on standard econometric assumptions, including the constancy of the variance of the error term. Similarly, the factor GARCH models do not have any lagged variables nor incorporate any intertemporal relationships in their structure.
On the contrary the predictive GARCH models should possess only the lagged variables (if the specification is mixed and contains both lagged and current values of the explanatory variable, the model is considered from such a perspective as eclectic).

From the point of view of the estimation of any GARCH-type model, the specification of the mean equation (dynamic or non-dynamic structure) does not make any significant difference as long as the ARCH effects are present. The relevant distinction arises when the model is used for prediction purposes. Obviously, factor GARCH models cannot be applied in such cases.

The question is however whether the predictive GARCH models containing only lagged explanatory variables posses the specification complete enough to generate accurate forecasts. The answer is an empirical issue.

4c. In-sample versus out-of-sample forecasts

The performance of all the models should be measured both in-sample and out-of-sample. Satisfactory results for the out-of-sample forecasts not only decide about the models’ quality, but also allow for verification if the relationships found in-sample are stable over time and if the model’s predictive properties are maintained in the longer run.

The modeling experience based on the out-of-sample comparisons indicates, however, that in many cases the out-of-sample forecasts are not of the same quality as those computed in-sample. Furthermore, it is often argued that they are not even better than the forecasts generated by the simple random walk model. The results for different currency models leading to such conclusions are presented in: Meese, Rogoff (1983).
The key question here, however, is what criteria should be used to evaluate the out-of-sample forecasts' performance. The approach applied in Meese, Rogoff (1983) focuses on the comparison of the three basic forecast error measures, such as: mean error (ME), mean absolute error (MAE) and root mean square error (RMSE). In most cases, reported values were higher and thus inferior to those obtained by the random walk model.

An alternative possibility to compare the forecasts' results with their empirical values is to employ direction quality measures (more in section 5.b). It cannot be excluded that although the models are characterized by the high values of the above-mentioned forecast errors, they still can produce good forecasts when measured by the compliance of the direction of change of the predicted and real values. In such a case the results in terms of the rate of return generated over the whole forecast horizon can prove to be better than those from the random walk model – although not necessarily superior to those from the in-sample analysis.

The methodology for generating the out-of-sample forecasts applied in this investigation is based on the standard rolling regression procedure.

If the in-sample period starts in time $t$ and ends in time $T$, then the first out-of-sample forecast generated in time $T$ for the time $T + 1$ for the model of the returns $r_t$:

$$r_t = \alpha_0 + \sum_{j=1}^{J} \alpha_j r_{t-j} + \xi_t,$$

where $r_{t-j}$ is the value of the explanatory variable $k$ in time $t - 1$, and $k$ stands for the names of the individual financial instruments, is the following:

$$r_{t+1}^{*} = \hat{\alpha}_0 + \sum_{j=1}^{J} \hat{\alpha}_j r_{t-j}^{*} + \tilde{\xi}_t,$$
where $\hat{\alpha}_0$ and $\hat{\alpha}_j$ are the estimates of the structural parameters. The out-of-sample forecasts are then generated by iterating:

$$\sum_{s=1}^S r_{T+s}$$

(6)

where $S$ is the end of the out-of-sample period.

### 5. Goodness of fit measures for GARCH models

#### 5a. Traditional goodness of fit measures

If $\Psi_{t-1}$ is the set of information available at time $t - 1$ then for the properly specified model the following relation holds:

$$E(t^2 | \Psi_{t-1}) = E(\xi^2 | \Psi_{t-1}) = h_t$$

(7)

which leads to the conclusion (see: Andersen, Bollerslev, 1998) that the goodness of fit measure for the GARCH model is the determination coefficient, $R^2$, from the equation:

$$r_t^2 = \varphi_0 + \varphi_1 h_t + \zeta_t$$

(8)

For the intradaily data the rate of return in the $1/b$ interval ($b$ stands for the number of observations within one day) equals $r_t = \ln y_t - \ln y_{t-1/b}$ for $t = 1/b, 2/b, ..., b/b$. The corresponding goodness of fit measure for the model based on such data frequency is the $R^2$ from the following model:

$$r_{t+1/b}^2 = \varphi_0 + \varphi_1 h_{t+1/b} + \zeta_{t+1/b}$$

(9)

called $R^2_{ARCH}$ in further references.

According to (7) $E(r_{t+1/b}^2 | \Psi_t) = h_{t+1/b}$, so $\hat{\varphi}_0$ should take on the value close to zero while both $\hat{\varphi}_1$ and $R^2_{ARCH}$ - close to unity. Empirical investigations have proved
that $\bar{R}^2$ for the model (9) reaches surprisingly low values, usually smaller than 0.05. Thus, Andersen and Bollerslev (1998) suggested to re-define the explanatory variable in (8) and to use the sum of squared returns characterized by higher frequency: $\sum_{a=1}^{b} r_{t+al/b}^2$, where $t = 0, 1/b, 2/b, \ldots$. In consequence (9) can be written as:

$$\sum_{a=1}^{b} r_{t+al/b}^2 = \phi_0 + \phi_1 h_{t+1} + \zeta_{t+1}$$

(10)

where $t = 0, 1, 2, \ldots$.

Replacement of $r_{t+al/b}^2$ by $r_{t+1}^2 = \sum_{a=1}^{b} r_{t+al/b}^2$ means an aggregation of the data with $1/b$ frequency to the $t$ frequency. Therefore the variables at both sides of the equation (10) have the same frequency of observations.

In practice the higher the frequency of the data used for the aggregation of the dependent variable, the higher the value of $\bar{R}^2$ in the model (10) reaching even 0.5 for the data of 5 minutes frequency ($b = 288$) (see: Andersen, Bollerslev, 1998).

The improvement in the goodness of fit values as the data frequency increases results from the fact that usually:

$$\sum_{a=1}^{b} r_{t+al/b}^2 \neq r_{t+1}^2$$

(11)

The $\bar{R}^2$ from (10), called $\bar{R}_{AB}^2$, is a much improved gauge of the GARCH model performance.
5b. Direction quality measures for GARCH models

The idea of the direction quality measures was introduced by Pesaran and Timmermann (1992) and developed by the Olsen and Associates group; for summary see: Dacorogna et al. (1998) and Dacorogna et al. (2001). The direction quality measures are based on the belief that in many cases it is more important that the model properly captures direction of changes rather than features a good fit. Such measures should therefore be especially useful in case of models based on the rates of return.

There are a few basic direction quality measures (see: Welfe, Brzeszczynski, 1999). The simplest one is:

\[
Q^1 = \frac{N\{r_t \hat{r}_t > 0\}}{N\{r_t \hat{r}_t \neq 0\}},
\]

where:

\(\hat{r}_t\) - theoretical values of the dependent variable,

\(N\{r_t \hat{r}_t > 0\}\) - number of observations for which \(r_t \hat{r}_t > 0\),

\(N\{r_t \hat{r}_t \neq 0\}\) - number of observations for which \(r_t \hat{r}_t \neq 0\).

The ability of the model to predict turning points can be measured by:

\[
Q^2 = \frac{N\{r_t \hat{r}_t > 0\ r_{t-1} r_t < 0\}}{N\{r_t \hat{r}_t \neq 0\ r_{t-1} r_t < 0\}},
\]

where \(N\{r_t \hat{r}_t > 0\ r_{t-1} r_t < 0\}\) is the number of observations for which \(r_t \hat{r}_t > 0\) under the condition that \(r_{t-1} r_t < 0\). The above is a ratio of predicted turning points to the number of all turning points in the sample.

Another measure capturing turning points takes into account an additional condition \(\hat{r}_{t-1} \hat{r}_t > 0\) in the numerator:
\[
W_1 = \frac{N\{r_t \hat{r}_t > 0 \mid r_{t-1} < 0 \wedge \hat{r}_{t-1} \hat{r}_t > 0\}}{N\{r_t \hat{r}_t \neq 0 \mid r_{t-1} < 0\}} .
\] (14)

It is a ratio of predicted turning points to the number of all turning points in the sample provided that the model properly reflected two consecutive observations in time \( t - 1 \) and \( t \).

Existence of the transaction costs (usually between 0.5% and 1% per trade) implies, however, that those measures should be „filtered” to make the evaluation realistic. For a, say, 1%-filter \( Q_1 \) takes the form:

\[
Q_{1(1\%)} = \frac{N\{r_t \hat{r}_t > 0 \mid r_t > 1\%\}}{N\{r_t \hat{r}_t \neq 0 \mid r_t > 1\%\}} \tag{15}
\]

The measures capturing the magnitude of reflected changes, which is very important from the point of view of investment decisions and strategies, can be constructed as follows:

\[
Q_3 = \frac{\sum_{i=1}^{T} |r_i|}{\sum_{i=1}^{T} |r_i^*|} \tag{16}
\]

\[
Q_4 = \frac{\left(\sum_{i=1}^{T} |r_i|\right) / N\{r_t \hat{r}_t > 0\}}{\left(\sum_{i=1}^{T} |r_i^*|\right) / N\{r_t \hat{r}_t > 0\}} \tag{17}
\]

where:

\[
r_t = \begin{cases} 
0 & \text{for } r_t \hat{r}_t \leq 0 \\
\hat{r}_t & \text{for } r_t \hat{r}_t > 0
\end{cases}
\]

\[
r_t^* = \begin{cases} 
0 & \text{for } r_t \hat{r}_t > 0 \\
r_t & \text{for } r_t \hat{r}_t \leq 0
\end{cases}
\]

The measure \( Q_3 \) is the ratio of the sum of absolute rates of return in case when model properly reflects the direction of change to the sum of absolute returns
when it fails to do so. By analogy, $Q^4$ is the ratio of average returns in these two cases. For the well specified model both measures $Q^3$ and $Q^4$ should be greater than unity.

6. Empirical results

6a. Models specifications and estimation

The database used in the investigation contains daily quotations and covers the period of 5 years: from 02.01.1998 to 31.12.2002. It includes the Warsaw Stock Exchange main index WIG, the volume of traded stocks at WSE, two American indices: Dow Jones Industrial Average (DJIA), NASDAQ Composite, as well as German main index Deutscher Aktienindex DAX and British Financial Times Stock Exchange (FTSE100) index.

The data for WIG and WSE volume of trade comes from the Warsaw Stock Exchange, whereas the source for the international indices is Reuters (Reuters Serwis Polski - RSP and Reuters X-tra). The database was adjusted to incorporate non-trading days (such as national holidays) in Poland, USA, Germany and UK to make all the time series comparable.

In the first step of the empirical investigation factor GARCH models of the WIG index were estimated. In all cases very strong ARCH effects were detected (the values of the $TR^2$ and $\chi^2$ test statistic for all the samples and sub-samples were significant at the 0.001 significance level). As mentioned before, the GARCH(1,1) specification has been adopted for all the models proposed.

The mean equation in the factor models include only those international indices which do not overlap in time with the WIG index. Due to this assumption,
only the two European indices were chosen as the ones which come from the stock exchanges located in the same geographical region as the Warsaw Stock Exchange and have corresponding trading-day hours. The working hours of the stock exchanges in Frankfurt and London fully cover those in Warsaw. On the other hand in the predictive models only the lagged variables, which values are known before the opening of the trading day at the Warsaw Stock Exchange, are considered for the models specifications.

The factor GARCH model takes the following general form:

\[ r_t^{WIG} = \alpha_0 + \alpha_{20} r_t^{VOL} + \sum_{j=3}^{5} \alpha_j r_t^k + \xi_t \]  

(18)

where \( r_t^{VOL} \) denotes daily returns (daily percentage changes) of the volume of trade and \( j \) refers to the daily returns of the respective international index, that is:

\( j = 3 \) stands for DAX (\( k = DAX \)) and \( j = 4 \) for FTSE (\( k = FTSE \)).

The predictive GARCH model is:

\[ r_t^{WIG} = \alpha_0 + \alpha_{11} r_{t-1}^{WIG} + \alpha_{21} r_{t-1}^{VOL} + \sum_{j=3}^{6} \alpha_j r_{t-1}^k + \xi_t \]  

(19)

where additionally the autoregression of \( r_t^{WIG} \) has been added and \( j = 5 \) stands for DJIA (\( k = DJIA \)) and \( j = 6 \) for NASDAQ (\( k = NASDAQ \)).

The results are reported for the predictive models with only one lag of the explanatory variables. The reason for this is that more lags proved to be not significant in almost all the models tested.

In search of the best form of the mean equation the following specifications of (1a) were proposed (see also: Brzeszczynski, 2002).

The starting model was based on the price-volume relationship exploiting the returns of trading volume \( r_t^{VOL} \) (model: Factor1). Different lengths of the rate of
return, from one to five days, were tested. The best results are given by the three
days long return: \( r_{t}^{\text{VOL}} = vol_{t} - vol_{t-3} \), where \( vol_{t} \) stands for the volume of trade in
day \( t \). Such definition of this variable is used in all the models investigated. A
possible explanation why the three days long return is statistically significant (and
one day return is not) is connected with the technical analysis and corresponds
with the commonly accepted principle stating that only the change of the trading
volume for a few consecutive days in a row constitutes relevant information in
relation to (future) price changes (it is believed that the changes in volume
“confirm” or “support” trends and different technical analysis formations) and thus
only as such is informative.

In case of all other variables only one day long returns are calculated and
used in the subsequent models’ specifications.

The assumption about the dependency between WIG and major international
indices led to a further extension of the model by adding two European indices:
DAX and FTSE (\( r_{t}^{\text{DAX}} \) and \( r_{t}^{\text{FTSE}} \)) chosen as the indicators from the biggest stock
exchanges in Europe as measured by their capitalization (models: Factor2 and
Factor3). Finally the models were estimated without the volume of trade in order to
find the impact of the international indices only without any control variable
(models: Factor4 and Factor5). Due to the relatively high correlation between
\( r_{t}^{\text{DAX}} \) and \( r_{t}^{\text{FTSE}} \), as high as 0.7 in the investigated sample, all the proposed
specifications include only one foreign index (either DAX or FTSE) to avoid the
problem of multicollinearity and the obvious negative consequences of it.

The parameters’ estimates for the models Factor1 through Factor5 are
presented in Table 1. The sample of 3 years (36 months) embraces the following
period: 02.01.1998-29.12.2000 and counts 760 daily observations. The estimation results allow for drawing the following conclusions:

1. The volume of trade proves to be an important explanatory variable in the factor GARCH models analyzed. The estimate of its parameter is highly significant in all the models, and it is stable (at the 0.014 level) regardless of the two other (control) variables added to the equation.

2. There is a significant influence of the two European indices DAX and FTSE on the variability of WIG index. Their parameters are highly significant and positive in all the cases investigated. This result allows for the conclusion that there exist contemporaneous co-movements effects of the WIG index and the two main European stock markets over the same trading day. The parameter’s value for FTSE is, however, roughly by half higher in case of FTSE comparing to DAX.

3. In all the models the constant is not significant.

4. The GARCH(1,1) parameters are strongly significant in all the models investigated.

The specifications for the predictive GARCH models were built according to the following two criteria. Firstly, all the explanatory variables have to be lagged by one day – which means that the information about their values from time $t-1$ is known before the trading starts at the Warsaw Stock Exchange. Secondly, the explanatory variables cannot be too much correlated with each other which would lead to the above-mentioned problem of multicollinearity. The correlation coefficients for the DJIA and NASDAQ is equally, and thus unacceptably, high as in the case of DAX and FTSE, reaching the level of 0.6. However, for other pairs of variables - one European and one American index - it takes on much lower values: between 0.3 and 0.4.
Therefore, all the models tested include one lagged European index (either DAX or FTSE) and one lagged American index (either DJIA or NASDAQ). The estimation results are presented in Table 2. They include the estimates for 8 models. The first four possess different combinations of the international indices: $r_{t-1}^{DAX}$, $r_{t-1}^{FTSE}$, $r_{t-1}^{DJIA}$ and $r_{t-1}^{NASDAQ}$ (models: Pred1 to Pred4). The next four have the same specification, extended by adding lagged explained variable $r_{t-1}^{WIG}$ in order to improve their the dynamic structure (models: Pred5 to Pred8).

The data sample is the same as in the case of the factor GARCH models and it covers the period: 02.01.1998-29.12.2000. The analysis of the results leads to the following conclusions:

1. All the lagged explanatory variables representing the returns of the four international indices are significant. The signs of their estimates are in all cases positive which means that the direction of the movements of the American and European stock exchange indices is transmitted from time $t-1$ and provides signals for the direction of change of the WIG index in time $t$.

2. The highest value of the parameter is reached in the case of the DJIA (over 0.6) indicating that the relationship between that index and the WIG is the strongest. Much lower values possess the parameters related to NASDAQ (about 0.32), FTSE (from 0.21 to 0.24) and DAX (from 0.15 to 0.17).

3. Volume of trade is significant in most of the models (except for the two with autoregression of the WIG index, $r_{t-1}^{WIG}$: Pred6 and Pred8). The value of the respective parameters takes on the values from 0.002 to 0.004.

4. Lagged explained variable $r_{t-1}^{WIG}$ as well as the constant are not significant in any of the four models tested.
5. As in the case of the factor models, the GARCH(1,1) parameters are strongly significant in all the predictive models analyzed. The results are in line with the expectations concerning the transmission of signals from the largest and the most important American market to other smaller markets which trade after its closing. The only result against the anticipations is lack of the significance of the lagged dependent variable. This means that the “external” signals coming from foreign markets are stronger and more important than any “internal” relationships, such as autoregressive dependencies etc.

**6b. In-sample and out-of-sample forecasts**

To evaluate the models’ performance, direction quality measures were applied. The analysis was conducted by comparing the in-sample forecasts from the predictive GARCH models, computed for the whole estimation period of 3 years: 1998-2000 (02.01.1998-29.12.2000), with the out-of-sample forecasts generated from the same models in the period of the next 2 years: 2001-2002 (02.01.2001-31.12.2002).

For this investigation the best 4 models: Pred1 to Pred4 were used. The out-of-sample forecasts were obtained on the basis of the rolling regression procedure. Therefore, the forecasts for time \( t+1 \) are calculated based only on the values of the explanatory variables in time \( t \) with iterations rolling until the end of the sample.

The values of the direction quality measures for the in-sample forecasts are presented in Table 3 and for the out-of-sample forecasts in Table 4.

For the in-sample forecasts the measure \( Q! \) indicates that the analyzed models explain about 0.663 – 0.674 of the WIG index value direction of change.
The turning points, as measured by $Q^2$, are captured at the 0.569 – 0.599 level. The measures adjusted by 1% and 0.5% filters reach the values of 0.520 – 0.545 and 0.377 – 0.394 for $Q_{1(0.5\%)}$ and $Q_{1(1\%)}$ respectively. For the measures $Q_{2(0.5\%)}$ and $Q_{2(1\%)}$, they take on the values from 0.591 to 0.618. The turning points measured by $W1$ are reflected at the 0.196 – 0.207 level.

Interesting results are delivered by the analysis of the measures $Q^3$ and $Q^4$. They prove that sum of absolute rates of return when the model successfully reflected the direction of change is about 2.734 – 3.062 higher than in cases when the model was missing the right direction. By analogy, the average rate of return is about 1.386 – 1.480 times higher than otherwise.

The values of the above-reported measures do not differ a lot among all the four models investigated. However, it can be noticed that systematically the highest values are achieved by the model Pred2 (with returns of the trading volume, DAX and NASDAQ) and the lowest – by the model Pred1 (with returns of the trading volume, DAX and DJIA).

The values of the measures calculated for the out-of-sample forecasts are lower - but satisfactory in most cases.

The measure $Q1$ takes on the values from 0.545 to 0.552; however, $Q2$ only from 0.432 to 0.461. Their modifications adjusted by 1% and 0.5% filters, $Q_{1(0.5\%)}$ and $Q_{1(1\%)}$, achieve the results of 0.364 – 0.372 and 0.242 – 0.245. The decrease of values in the out-of-sample case is relatively the smallest in the case of $Q_{2(0.5\%)}$ and $Q_{2(1\%)}$ as their values are 0.461 – 0.513 and 0.506 – 0.529 respectively. The measure $W1$ reaches 0.177 – 0.222.
The most satisfactory result is obtained for the measure $Q^3$. Its value ranges from 1.393 to 1.426, which means that in the whole sample of the years 2001 - 2002 the sum of the correctly predicted returns is about 1.4 times higher than the sum of the mispredicted returns. The ratio of average returns $Q^4$ is also greater than unity, and it reaches the values: 1.155 – 1.198. It has to be emphasized that good results of $Q^3$ are especially valuable from the practical point of view when the models analyzed are used as a tool for the investment decisions and trading strategies.

Similar to the in-sample forecasts, slightly better results also achieved the model Pred2 with returns of volume, DAX and NASDAQ.

In case of both in-sample and out-of-sample forecasts, the goodness of fit measure $R^2_{ARCH}$ does not exceed the value of 0.1 in all the models analyzed while $R^2_{AB}$ (for the daily data aggregated to monthly) is much higher and approaches 0.4.

7. Conclusions

Empirical results prove the existence of a very strong dependency between WIG index and the indices from international markets - especially DJIA and NASDAQ Composite. The European indices DAX and FTSE also prove significant. It is characteristic, however, that the most important factors influencing the situation at the Warsaw Stock Exchange come in the form of signals from the American stock market – which is in line with the general perception of this phenomenon by market participants (investors).

Comparison of the in-sample and the out-of-sample forecasts indicates that the latter yield worse and less accurate results in terms of the index direction of movement. Still those results can be perceived as satisfactory because the most
important direction quality measures reach the values above their respective benchmarks. Also the measure which seems to be the most important from the practical point of view \( Q^3 \) shows that the models are capable of generating positive rates of return - even when the forecasts are made out-of-sample.

The results obtained in the current study raise a few new questions of the fundamental importance.

The first one concerns the incorporation of the real transaction costs to the presented analysis. The main concern is whether they would significantly influence the forecasts and yielded rates of return (and if so – to what degree).

The second question refers to the analysis of the volatility dynamics within the trading day, which leads to the extension of the presented investigation by taking into account the stop-loss and/or take-profit orders. Depending on their level they may either improve or worsen the out-of-sample forecasting performance.

The third question is how useful and efficient will be those out-of-sample forecasts in terms of real trading strategies basing on the day-trading rules.

The above-outlined issues sketch some important directions for further investigations and open up a space for future research.
References


Table 1. Estimates of the GARCH(1,1) factor models with the returns of volume of trade and the returns of the indices DAX and FTSE:

\[ r^\text{WIG}_t = \alpha_0 + \alpha_{20} r^\text{VOL}_t + \alpha_{30} r^\text{DAX}_t + \alpha_{40} r^\text{FTSE}_t + \xi_t \] (sample: 02.01.1998-29.12.2000).

<table>
<thead>
<tr>
<th>Model</th>
<th>( \alpha_0 )</th>
<th>( \alpha_{20} )</th>
<th>( \alpha_{30} )</th>
<th>( \alpha_{40} )</th>
<th>( \gamma_0 )</th>
<th>( \gamma_1 )</th>
<th>( \phi_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fact1</td>
<td>0.062 (1.032)</td>
<td>0.014 (7.184)</td>
<td>-</td>
<td>-</td>
<td>0.383 (4.404)</td>
<td>0.193 (4.822)</td>
<td>0.703 (13.885)</td>
</tr>
<tr>
<td>Fact2</td>
<td>0.045 (0.764)</td>
<td>0.014 (8.220)</td>
<td>0.281 (8.539)</td>
<td>-</td>
<td>0.477 (4.391)</td>
<td>0.203 (5.174)</td>
<td>0.653 (10.826)</td>
</tr>
<tr>
<td>Fact3</td>
<td>0.050 (0.872)</td>
<td>0.014 (8.597)</td>
<td>-</td>
<td>0.427 (10.310)</td>
<td>0.644 (4.397)</td>
<td>0.246 (5.163)</td>
<td>0.554 (7.048)</td>
</tr>
<tr>
<td>Fact4</td>
<td>0.030 (0.502)</td>
<td>-</td>
<td>0.278 (7.926)</td>
<td>-</td>
<td>0.325 (4.165)</td>
<td>0.157 (5.110)</td>
<td>0.753 (18.160)</td>
</tr>
<tr>
<td>Fact5</td>
<td>0.024 (0.397)</td>
<td>-</td>
<td>-</td>
<td>0.431 (9.817)</td>
<td>0.551 (3.768)</td>
<td>0.194 (5.143)</td>
<td>0.644 (9.782)</td>
</tr>
</tbody>
</table>

(\( t \)-statistics in parentheses)

Note: The table reports exclusively the results for the model with zero-restrictions. In order to avoid multicollinearity, all specifications possess only one European index (either DAX or FTSE).
Table 2. Estimates of the GARCH(1,1) predictive model with lagged returns of volume of trade and the returns of the indices DJIA, NASDAQ, DAX and FTSE: 

\[ r_t^{\text{WG}} = \alpha_0 + \alpha_{11} r_{t-1}^{\text{WG}} + \alpha_{21} r_{t-1}^{\text{VOL}} + \alpha_{31} r_{t-1}^{\text{DAX}} + \alpha_{41} r_{t-1}^{\text{FTSE}} + \alpha_{51} r_{t-1}^{\text{DJIA}} + \alpha_{61} r_{t-1}^{\text{NASDAQ}} + \xi_t \]  


<table>
<thead>
<tr>
<th>Model</th>
<th>$\alpha_0$</th>
<th>$\alpha_{11}$</th>
<th>$\alpha_{21}$</th>
<th>$\alpha_{31}$</th>
<th>$\alpha_{41}$</th>
<th>$\alpha_{51}$</th>
<th>$\alpha_{61}$</th>
<th>$\gamma_0$</th>
<th>$\gamma_1$</th>
<th>$\phi_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred1</td>
<td>-0.016</td>
<td>-</td>
<td>0.004</td>
<td>0.160</td>
<td>-</td>
<td>0.622</td>
<td>-</td>
<td>0.328</td>
<td>0.151</td>
<td>0.732</td>
</tr>
<tr>
<td></td>
<td>(-0.281)</td>
<td></td>
<td>(2.438)</td>
<td>(4.196)</td>
<td></td>
<td>(11.666)</td>
<td></td>
<td>(3.800)</td>
<td>(4.223)</td>
<td>(13.070)</td>
</tr>
<tr>
<td>Pred2</td>
<td>0.1x10^{-4}</td>
<td>-</td>
<td>0.002</td>
<td>0.171</td>
<td>-</td>
<td>-</td>
<td>0.324</td>
<td>0.292</td>
<td>0.192</td>
<td>0.716</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td>(1.664)</td>
<td>(4.346)</td>
<td></td>
<td></td>
<td>(12.769)</td>
<td>(4.048)</td>
<td>(4.912)</td>
<td>(14.401)</td>
</tr>
<tr>
<td>Pred3</td>
<td>-0.017</td>
<td>-</td>
<td>0.003</td>
<td>-</td>
<td>0.207</td>
<td>0.617</td>
<td>-</td>
<td>0.339</td>
<td>0.165</td>
<td>0.715</td>
</tr>
<tr>
<td></td>
<td>(-0.316)</td>
<td></td>
<td>(2.391)</td>
<td></td>
<td>(3.979)</td>
<td>(11.941)</td>
<td></td>
<td>(4.105)</td>
<td>(4.432)</td>
<td>(12.840)</td>
</tr>
<tr>
<td>Pred4</td>
<td>0.001</td>
<td>-</td>
<td>0.002</td>
<td>-</td>
<td>0.237</td>
<td>-</td>
<td>0.321</td>
<td>0.307</td>
<td>0.208</td>
<td>0.696</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td></td>
<td>(1.646)</td>
<td></td>
<td>(4.658)</td>
<td></td>
<td>(13.063)</td>
<td>(4.368)</td>
<td>(5.096)</td>
<td>(13.889)</td>
</tr>
<tr>
<td>Pred5</td>
<td>-0.017</td>
<td>0.018</td>
<td>0.004</td>
<td>0.154</td>
<td>-</td>
<td>0.621</td>
<td>-</td>
<td>0.324</td>
<td>0.151</td>
<td>0.733</td>
</tr>
<tr>
<td></td>
<td>(-0.305)</td>
<td>(0.527)</td>
<td>(2.200)</td>
<td>(3.965)</td>
<td></td>
<td>(11.651)</td>
<td></td>
<td>(3.879)</td>
<td>(4.245)</td>
<td>(13.339)</td>
</tr>
<tr>
<td>Pred6</td>
<td>-0.001</td>
<td>0.022</td>
<td>0.002</td>
<td>0.166</td>
<td>-</td>
<td>-</td>
<td>0.323</td>
<td>0.287</td>
<td>0.190</td>
<td>0.719</td>
</tr>
<tr>
<td></td>
<td>(-0.026)</td>
<td>(0.604)</td>
<td>(1.387)</td>
<td>(4.145)</td>
<td></td>
<td></td>
<td>(12.694)</td>
<td>(4.136)</td>
<td>(4.897)</td>
<td>(14.815)</td>
</tr>
<tr>
<td>Pred7</td>
<td>-0.018</td>
<td>0.014</td>
<td>0.004</td>
<td>-</td>
<td>0.202</td>
<td>0.617</td>
<td>-</td>
<td>0.336</td>
<td>0.165</td>
<td>0.717</td>
</tr>
<tr>
<td></td>
<td>(-0.331)</td>
<td>(0.396)</td>
<td>(2.195)</td>
<td></td>
<td>(3.734)</td>
<td>(11.931)</td>
<td></td>
<td>(4.164)</td>
<td>(4.414)</td>
<td>(13.067)</td>
</tr>
<tr>
<td>Pred8</td>
<td>0.1x10^{-3}</td>
<td>0.010</td>
<td>0.002</td>
<td>-</td>
<td>0.234</td>
<td>-</td>
<td>0.321</td>
<td>0.305</td>
<td>0.206</td>
<td>0.698</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.276)</td>
<td>(1.462)</td>
<td></td>
<td>(4.478)</td>
<td></td>
<td>(13.034)</td>
<td>(4.407)</td>
<td>(5.053)</td>
<td>(14.052)</td>
</tr>
</tbody>
</table>

(*t*-statistics in parentheses)

Note: The table reports exclusively the results for the model with zero-restrictions. In order to avoid multicollinearity, all specifications possess only one American index (either DJIA or NASDAQ) and only one European index (either DAX or FTSE).
Table 3. Direction quality measures for the *in-sample* forecasts from the GARCH(1,1) predictive models for WIG index.

<table>
<thead>
<tr>
<th></th>
<th>Pred1</th>
<th>Pred2</th>
<th>Pred3</th>
<th>Pred4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Q1</strong></td>
<td>0.663</td>
<td>0.674</td>
<td>0.670</td>
<td>0.664</td>
</tr>
<tr>
<td><strong>Q2</strong></td>
<td>0.599</td>
<td>0.586</td>
<td>0.591</td>
<td>0.569</td>
</tr>
<tr>
<td><strong>W1</strong></td>
<td>0.196</td>
<td>0.196</td>
<td>0.196</td>
<td>0.207</td>
</tr>
</tbody>
</table>

*Total number of turning points*: 362

<table>
<thead>
<tr>
<th></th>
<th>Pred1</th>
<th>Pred2</th>
<th>Pred3</th>
<th>Pred4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Q1 (1%)</strong></td>
<td>0.377</td>
<td>0.394</td>
<td>0.389</td>
<td>0.388</td>
</tr>
<tr>
<td><strong>Q2 (1%)</strong></td>
<td>0.618</td>
<td>0.647</td>
<td>0.623</td>
<td>0.623</td>
</tr>
</tbody>
</table>

*Number of turning points (for $r_i > 1\%$)*: 170

<table>
<thead>
<tr>
<th></th>
<th>Pred1</th>
<th>Pred2</th>
<th>Pred3</th>
<th>Pred4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Q1 (0.5%)</strong></td>
<td>0.520</td>
<td>0.545</td>
<td>0.531</td>
<td>0.533</td>
</tr>
<tr>
<td><strong>Q2 (0.5%)</strong></td>
<td>0.606</td>
<td>0.618</td>
<td>0.602</td>
<td>0.591</td>
</tr>
</tbody>
</table>

*Number of turning points (for $r_i > 0.5\%$)*: 254

| **Q3** | 2.734 | 3.062 | 2.990 | 2.809 |
| **Q4** | 1.386 | 1.480 | 1.472 | 1.424 |
Table 4. Direction quality measures for the *out-of-sample* forecasts from the GARCH(1,1) predictive models for WIG index.

<table>
<thead>
<tr>
<th>Model</th>
<th>Pred1</th>
<th>Pred2</th>
<th>Pred3</th>
<th>Pred4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>0.545</td>
<td>0.552</td>
<td>0.547</td>
<td>0.547</td>
</tr>
<tr>
<td>Q2</td>
<td>0.436</td>
<td>0.461</td>
<td>0.440</td>
<td>0.432</td>
</tr>
<tr>
<td>W1</td>
<td>0.177</td>
<td>0.222</td>
<td>0.189</td>
<td>0.206</td>
</tr>
</tbody>
</table>

**Total number of turning points**: 243

<table>
<thead>
<tr>
<th>Model</th>
<th>Pred1</th>
<th>Pred2</th>
<th>Pred3</th>
<th>Pred4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1 (1%)</td>
<td>0.242</td>
<td>0.242</td>
<td>0.242</td>
<td>0.246</td>
</tr>
<tr>
<td>Q2 (1%)</td>
<td>0.506</td>
<td>0.529</td>
<td>0.518</td>
<td>0.529</td>
</tr>
</tbody>
</table>

**Number of turning points (for $r > 1\%$)**: 85

<table>
<thead>
<tr>
<th>Model</th>
<th>Pred1</th>
<th>Pred2</th>
<th>Pred3</th>
<th>Pred4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1 (0.5%)</td>
<td>0.366</td>
<td>0.372</td>
<td>0.364</td>
<td>0.370</td>
</tr>
<tr>
<td>Q2 (0.5%)</td>
<td>0.461</td>
<td>0.513</td>
<td>0.468</td>
<td>0.481</td>
</tr>
</tbody>
</table>

**Number of turning points (for $r > 0.5\%$)**: 154

<table>
<thead>
<tr>
<th>Model</th>
<th>Pred1</th>
<th>Pred2</th>
<th>Pred3</th>
<th>Pred4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3</td>
<td>1.397</td>
<td>1.426</td>
<td>1.393</td>
<td>1.444</td>
</tr>
<tr>
<td>Q4</td>
<td>1.168</td>
<td>1.155</td>
<td>1.156</td>
<td>1.198</td>
</tr>
</tbody>
</table>