The Effects of Outliers on the AR(1)-GARCH(1,1) Process

Peter Verhoeven and Michael McAleer

Department of Economics, University of Western Australia, Nedlands, WA 6907

Abstract

This paper examines the effects of outliers on: (i) the estimated parameters of the AR(1)-GARCH(1,1) process; (ii) the second and fourth moment conditions of the process; and (iii) forecast errors. We find that outliers: (i) tend to dominate the maximum likelihood estimates, resulting in larger ARCH and smaller GARCH estimates; (ii) may give rise to spurious AR(1) and ARCH effects when the outliers are extremely large; (iii) significantly increase the frequency of violation of the fourth moment condition; (iv) significantly decrease the t-ratios of the GARCH estimates; (v) significantly increase the variability of the time series of the GARCH(1,1) volatility forecasts; and (vi) decrease the persistence measure of the volatility process. Removing outliers results in significantly improved daily volatility forecasts in terms of overprediction errors, but this does not translate into improved out-of-sample predictive power. For all the time series investigated, we find that the i.i.d. assumption of the standardised residuals cannot be rejected at the 5% significance level. In contrast, the assumption of conditional normality cannot be justified, even when all the outliers are removed. We find that the maximum excess kurtosis that can be captured by the GARCH(1,1) model under the assumption of conditional normality increases with the standard deviation (which is approximately 40).

1 Introduction

In finance, volatility is a fundamental measure of risk and of relative movements in the price of a security, such as stock, stock index or futures contract, over time. As the true underlying volatility of a security is unobservable, it must be estimated. While volatility can be expressed in different ways, the typical definition used in finance is the standard deviation of the returns of a security over a given period. Since volatility is an essential input to the optimisation of financial models describing the expected risk-return trade-off, it is of paramount importance to practitioners to use an adequate empirical model to measure the dynamics of volatility in financial securities.

Financial equity returns generally do not match the familiar bell-shaped normal distribution. Instead, financial markets frequently experience large and sudden price movements that predominantly consist of market crashes, not rallies. A recent example of extreme price movements is the October 1997 stock market crash originating in Asia. On 28 October 1997, the Hang Seng Stock Index (HSI) dropped by 14.7%, the German Stock Index (DAX) by 7.2%, the Standard & Poor's 500 Composite Index (S&P 500) by 5.0%, and the Japanese Stock Index (Nikkei 225) by 4.4%. A consequence of these outlying and extreme observations is the fat-tailed distribution of returns.

There is overwhelming evidence that the tail behaviour of equity returns evolves over time. In particular, absolute returns have significant positive serial correlation over long lags, implying that they have long term memory. This is known as volatility clustering: large (small) returns are more likely to be followed by large (small) returns than by small (large) returns, but of unpredictable sign. Hence, the returns are not independently and identically distributed (i.i.d.) over time. The implication for practitioners is that financial market volatility is highly predictable.

The autoregressive conditional heteroskedasticity (ARCH(p)) model of Engle [1982] and generalised ARCH (GARCH(p,q)) model of Bollerslev [1986] is currently the most popular volatility modelling and forecasting model for both market professionals and academic researchers. GARCH has been applied extensively to numerous financial time series, and its popularity is due to the fact that it: (i) captures the persistence of volatility; (ii) can account partly for the fat-tails of the returns distribution; and (iii) is simple, and also mathematically and computationally straightforward. In the majority of cases, it has been shown that the GARCH(1,1) model adequately represents the observed intertemporal dependencies in daily returns of most financial time series.

Of increasing importance in time series modelling and forecasting is the problem of outliers. Outlier analysis in time series is concerned with the detection and accommodation of abnormal observations in the data, and their influence on the mean and variance of financial time series. Despite the voluminous research that has estimated volatility using GARCH-type models, very little attention has been given to the effects of outlying observations in the data.

This paper examines the effects of outliers on the properties of the AR(1)-GARCH(1,1) model by applying the Chen and Liu [1993] outlier detection and removal procedure to the daily returns of a wide range of financial time series. The paper is organised as follows. Section 2 presents the AR(1)-GARCH(1,1) model. Section 3 describes outliers. Section 4 analyses the empirical results. Some concluding comments are given in Section 5.

2 The AR(1)-GARCH(1,1) model

Consider the AR(1)-GARCH(1,1) model, where the conditional mean (log returns) has the structure given by

$$y_t = \mu + \varphi y_{t-1} + \varepsilon_t \tag{1}$$

and the conditional variance of ε_t is generated by

$$\varepsilon_t = \eta_t \sqrt{h_t} \tag{2}$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{3}$$

where η_t is a sequence of normally, independently and identically distributed (n.i.i.d.) random variables (shocks) with zero mean and unit variance. Sufficient conditions for positivity of the conditional variance and the GARCH(1,1) process to exist are that $\omega > 0$, $\alpha > 0$ and $\beta \geq 0$.

Several statistical properties have been established for the GARCH(1,1) process in order to define the unconditional moments of $\{\varepsilon_t\}$ (Bollerslev [1986]). First, the second moment of $\{\varepsilon_t\}$ exists if $(\alpha+\beta) < 1$. This condition must be met in order for the GARCH(1,1) process to be strictly stationary and ergodic, and $\mathrm{E}(\varepsilon_t^2) < \infty$. Second, a necessary and sufficient condition for the existence of the fourth moment of $\{\varepsilon_t\}$ is $(k\alpha^2 + 2\alpha\beta + \beta^2) < 1$ (Bollerslev [1986]), where k is the conditional fourth moment of η_t. Under the assumption of conditional normality, $k \equiv E_t(\eta_t^4) = 3$, so that the condition becomes $(3\alpha^2 + 2\alpha\beta + \beta^2) < 1$.

3 Outliers

Outliers are considered as large observations (or a subset of observations) that lie far removed from the Gaussian (or heavy-tailed) conditional distribution assumed for the model generating the data. Although these observations could come from a Gaussian (or heavy-tailed) distribution, they should occur only rarely. For example, the Gaussian distribution predicts that, on average, observations will fall outside $\pm 4\sigma$ less than 0.003% of the time. However, in reality, such outlying observations are observed far too frequently to be consistent with normality in the data.

In autoregressive (integrated) moving average (AR(I)MA) models, it has been found that outliers can result in significantly negatively biased estimates of the AR(1) coefficient and positively biased estimates of the MA(1) coefficient, and may result in model misspecification (Ledolter [1989]). In the presence of outliers, the estimate of the ARCH effect in the GARCH(1,1) model can be severely biased upwards, and the estimate of the GARCH effect severely biased downwards (Engle and Lee [1993]; Franses and Ghijsels [1999]). Outliers also adversely affect both the size and power of the standard Lagrange Multiplier (LM) test used to detect ARCH effects, leading to model misspecification (Franses et al. [1998]; van Dijk et al. [1999]). In particular, a few outliers, either isolated or in patches, may result in spurious ARCH effects when none, in fact, is present.

Outliers may affect forecasts through the carry-over effect on the ARCH and GARCH terms, and may have a permanent effect on forecasts through their effects on the parameter estimates. In particular, the accuracy of the point forecast immediately following an outlier has been shown to be severely impaired (Ledolter [1989]). However, point forecasts are significantly less affected by additive outliers (AO) when they occur more than two periods away from the forecast origin. This result arises because the effects of past shocks on subsequent forecasts diminish exponentially with the number of periods away from the forecast origin.

We argue that returns consist of two types of observations: (i) ordinary observations arising from a conditional normal (or heavy-tailed) distribution, and which are realised in each period; and (ii) outlying (extraordinary) observations that are not generated by a normal distribution, and which are realised infrequently and irregularly. We further argue that outlying observations are independent of ordinary observations, and are not generated by a GARCH(1,1) process. Consequently, these outlying observations should not be used in estimating the AR(1)-GARCH(1,1) parameters. One way of proceeding is to identify outliers so that they can be down-weighted and their impacts removed. The principal idea behind such outlier adjustments is to apply the non-robust MLE procedure to the "cleansed" returns $\{y_t^*\}$, rather than to the original returns $\{y_t\}$.

We have applied a modified version of the Chen and Liu [1993] method to deal with outliers in the AR(1)-GARCH(1,1) model, as described in Franses and Ghijsels [1999]. This procedure jointly detects outliers and estimates the model parameters. It is a fully-automated iterative method, which consists of specification-estimation-detection-removal cycles to accommodate individually the most significant outliers. In each iteration, the maximum of the given test statistic is determined for a specific type of outlier, and the maximum across all types of outliers is examined. The maximum is then compared to a pre-specified critical value. If the test statistic exceeds the critical value, the outlier is effectively removed (e.g. down-weighted) and the parameters of the model are re-estimated. This cycle is continued until the test statistics for all data points lie below the critical value.

4 Empirical Results

4.1 Data

The effects of outliers on the AR(1)-GARCH(1,1) model were evaluated using 1000 trading days of five financial time series, comprising stock index, currency and commodity data: the daily close-to-close log index returns of the S&P 500, the Nikkei 225 and the HSI, the noon (Pacific time) British Pound-U.S. Dollar (GBP/USD) spot exchange rate, and the closing (London) Gold Bullion (GB) spot rate. The close-to-close index returns and the noon GB spot rate were obtained from Datastream, while the GBP/USD spot exchange rate was obtained from the Pacific Exchange Rate Service.

A rolling window of 500 trading days was used. The mean values for the parameter estimates, moment conditions and forecast errors were calculated using 500 one-day ahead volatility forecasts. For the outlier test statistic, we used critical values of 10, 8, 6 and 4. The forecasts were compared with the realised volatility according to the following definition:

$$\sigma_t = \mid y_t - \bar{y} \mid \tag{4}$$

where daily log returns are defined as $y_t = ln(\frac{P_t}{P_{t-1}})$, with P_t denoting the price in period t.

4.2 The effects of outliers on the unconditional returns distributions

None of the unadjusted time series is normally distributed, as indicated by the Jarque-Bera LM statistics. While the skewness of all the returns distributions is relatively small and negative (except for HSI), the kurtosis is large, implying that much of the departure from normality is due to leptokurtosis. The GB spot rates are the most volatile time series, having the largest standard deviation, followed by HSI, Nikkei 225, S&P 500 and GBD/USD.

The largest numbers of outliers and extreme observations are observed for the Nikkei 225 series, with 9 observations larger than 5σ, 20 observation larger than 4σ, and 47 observations larger than 3σ. This is followed by HSI (4, 13 and 41, respectively), S&P 500 (2, 12 and 26, respectively), GB (2, 11 and 23, respectively) and GBP/USD (0, 2 and 23, respectively), which are significantly greater than expected from a normal distribution. The Nikkei 225 series also has the highest kurtosis measure (51.06), and the largest (relative) positive (9.05σ) and negative (-13.55σ) outliers, with the large negative outlier corresponding to the 19 October 1987 stock market crash. Although both positive and negative outliers are observed, on average there appear to be more negative outliers. Moreover, the largest outlier is frequently negative, which is often subsequently followed by a large positive outlier. In addition, there appears to be a correspondence between the occurrence of outlying and extreme observations.

Removing outlying observations reduces the kurtosis measure, but the distribution remains non-normal. Similarly, the S&P 500, Nikkei 225, HSI and GB spot rates all become more symmetric, suggesting that the outliers are primarily responsible for the asymmetries in these series. In contrast, the negative skewness observed for the GBD/USD spot exchange rate does not approach zero, implying that the skewness in this series is caused by the ordinary returns.

4.3 The effects of outliers on the parameter estimates

Before examining the parameter estimates of the AR(1)-GARCH(1,1) model, we need to check whether the model is a valid description of the data. First, a minimum requirement is that the estimated η_t are i.i.d.. The Ljung-Box statistics do not indicate any violations of the i.i.d. assumption of η_t. Second, the assumption of conditional normality should not be rejected. However, the Jarque-Bera LM statistic indicates that for both the unadjusted and adjusted series, the assumption of conditional normality could be rejected. The degree of excess kurtosis induced by the GARCH(1,1) model with conditional normality increases with the standard deviation of the series (which is approximately 40).

For the HSI, Nikkei 225 and GBD/USD, the daily returns are positively correlated, while the daily returns are negatively correlated for GB. Removing outliers generally results in smaller and less significant AR(1) (φ) estimates. For example, when outliers are removed, the average φ estimate for the HSI and Nikkei 225 series drops by up to 40%. Extremely large (positive or negative) outliers result in substantially negatively biased φ estimates, leading to spurious AR(1) effects. These findings are consistent with those of Ledolter [1989].

For both the adjusted and unadjusted time series, the ARCH and GARCH estimates of the GARCH(1,1) model are significantly different from zero, with the β estimates having a substantially higher statistical significance. The average α estimates are also much smaller than the average β estimates. This indicates that, on average, there is a relatively weak reaction of conditional volatility to shocks (ARCH effects), but with a long-term memory (GARCH effects).

The results show that, in particular, the α estimates and, to a lesser extent, the β estimates vary considerably over time, even when outlying observations are removed. In particular, the α estimates are substantially larger (smaller) when volatility is high (low), while the β estimates are substantially smaller (larger). This implies that larger shocks have larger ARCH (short-run) effects but smaller GARCH (long-run) effects. For example, when outliers are removed for the HSI, the mean lag ($\frac{1}{1-\beta}$) of the GARCH process increases from 5.2 to 14.5.

The MLE is dominated by outlying observations, with the short-term memory of the outlying shocks overwhelming the long-term memory of the smaller shocks. For example, removing the seven largest outliers in the second least volatile index series, the S&P 500, causes the average α estimate to decrease by 0.064 (from 0.123 to 0.059), while the average β estimate increases by 0.085 (from 0.834 to 0.919). The most significant changes in the parameter estimates are observed for the Nikkei 225, which has the largest outliers. For this series, when the largest eight outliers are down-weighted and the maximum absolute outlier decreases from 13.55σ to 5.10σ, the average α estimate decreases by 0.279 (from 0.435 to 0.156) while the average β estimate increases by 0.398 (from 0.377 to 0.775).

For the individual time series, a sharp increase in the α estimate and a simultaneous and almost equivalent sharp decrease in the β estimate follows when an outlier enters the estimation period. The effects remain while the outlier remains in the estimation period, and its position in the estimation period appears to have little influence on the magnitude of its effect. When the outlier eventually drops out of the estimation period, the effect on the parameter estimates is reversed. The larger is the size of the outlier, the larger is its effect on the parameter estimates. This is particularly evident in the Nikkei 225 series, which include the 1987 stock market crash. When the extremely large outlier ($> 13\sigma$) of 19 October 1987 enters the estimation period, this causes a sharp and dramatic increase (0.571) in the α estimate and an almost equivalent sharp decrease (0.582) in the β estimate. Moreover, when the extremely large positive outlier ($> 9\sigma$) of 20 October 1987 enters the estimation period, no significant additional effect on the parameter estimates is observed. This illustrates the fact that the occurrence of just a single extreme outlier can yield spurious ARCH effects.

In addition to the dramatic negative effect of outliers on the β estimates, outliers have a similar effect on their t-ratios. In contrast, although outliers usually have a dramatic positive effect on the α estimates, no such effect is observed on their t-ratios. For example, for S&P 500, when the largest outlier (6.9σ) of the series enters the estimation period, this causes a sharp and dramatic decrease (from 23.58 to 9.82) in the t-ratio of the β estimate while the t-ratio of the α estimate drops only slightly (from 1.05 to 1.01). The effects on the t-ratios of the β estimates remain as long as the outlier remains in the estimation period.

4.4 The effects of outliers on the moment conditions

The average estimated value of $(\alpha + \beta)$ is usually very close to (but less than) unity, implying that the volatility process is highly persistent and close to a unit root. On average, outliers cause a slight decrease in the persistence measure $(\alpha + \beta)$ of the volatility process.

Only for the two series with the largest outliers, namely the Nikkei 225 and HSI, is the second moment condition violated. However, the largest estimated value is only 1.02. When the extreme outliers are removed, the second moment condition is satisfied. Violation of the second moment condition implies non-stationarity of the GARCH(1,1) process. However, if some weaker requirements (such as the log moment condition) are met, the volatility process may still be stationary even though $(\alpha + \beta)$ might exceed unity. For example, Nelson [1990] shows that when $\omega > 0$, $h_t < \infty$ and $\{\varepsilon_t, h_t\}$ is strictly stationary if and only if $E[ln(\beta + \alpha\eta^2)] < 0$.

The average $(3\alpha^2 + 2\alpha\beta + \beta^2)$ estimates of the fourth moment condition are also close to unity, but are usually slightly smaller than the average $(\alpha + \beta)$ estimates. Outliers have a dramatic effect on the $(3\alpha^2 + 2\alpha\beta + \beta^2)$ estimates. For example, when the extreme outlier of 19 October 1987 enters the estimation period for Nikkei 225, this causes a sharp doubling in its value (from 1.02 to 2.07). Clearly, the fourth moment regularity condition is more stringent than the second moment condition when outliers are present. For example, for Nikkei 225 and HSI, the fourth moment condition is violated more than half of the time. Removing outliers usually results in an increased stability of the $(3\alpha^2 + 2\alpha\beta + \beta^2)$ estimates over time and a decreased frequency of violation of the fourth moment condition.

Violation of the fourth moment regularity condition means that the assumption of asymptotic normality does not hold so that we cannot draw valid inferences. For example, the unconditional kurtosis measure for the Nikkei 225 may be meaningless. In addition, we cannot construct meaningful confidence intervals for forecasts.

4.5 The effects of outliers on the volatility forecasts

The empirical results show that the GARCH(1,1) forecasts are strongly positively biased (the mean error (ME) is positive) and, on average, overpredict volatility 70% of the time. Removing outliers results in a dramatic decrease in the rate of overprediction and ME. The higher rate of overprediction when outliers are present arises through higher α estimates. A larger proportion of the shocks is allowed to impact on the volatility process, thereby making the subsequent volatility forecasts inappropriately high.

Based on the mean absolute error (MAE), as well as the mean squared error (MSE) and root MSE (RMSE), a significant improvement (of up to 35%) in the mean forecast error is observed when outliers are removed.

A useful measure of the effects of outliers on forecasts is the MSE measure for positive forecast errors. This MSE measure is usually less than half the corresponding measure for negative forecast errors. Removing outliers results in a significant reduction (of up to 65%) in the MSE measure for positive forecast errors, but has no significant effect on the MSE measure for negative forecast errors. The explanation is that the GARCH(1,1) model usually underpredicts the largest outliers and overpredicts the first few subsequent observations.

Although removing outliers may result in significantly reduced MAE, MSE and RMSE measures, the corresponding median or relative forecast errors are not significantly affected. In fact, based on the adjusted R^2 measure, removing outliers results in a significant reduction in the predictive power of the GARCH(1,1) model for all data sets.

5 Conclusion

Using a variety of financial time series, we have shown that outliers significantly affect the parameter estimates of the AR(1)-GARCH(1,1) model. Outliers dominate the maximum likelihood estimates, causing a significant increase in the AR(1) and ARCH estimates and a significant decrease in the GARCH estimates. Furthermore, extremely large outliers may give rise to spurious AR(1) and ARCH effects. We also find that large and small shocks have different memory effects, and that both observations do not appear to be generated by the same GARCH process.

For all the time series investigated, the independent and identically distributed assumption of the conditional shocks cannot be rejected at the 5% significance level. In contrast, the assumption of conditional normality cannot be justified. We find that the maximum excess kurtosis that can be captured by the GARCH(1,1) model under the assumption of conditional normality increases with the standard deviation (which is approximately 40).

We also find that the regularity conditions, in particular the fourth moment condition, are harder to satisfy when outliers or extreme observations are present. Rejection of the fourth moment condition means that conditional shocks are not asymptotically normally distributed, so that the t-ratios of the parameter estimates are not asymptotically normal. Moreover, we cannot construct meaningful confidence intervals for the forecasts.

Removing outliers often results in significantly reduced forecast errors when measured in terms of MAE, MSE or RMSE. Most of this reduction is driven by a marked decrease in the positive forecast errors, arising from a substantially reduced ARCH estimate. However, failing to account for outlying observations significantly reduces the variation in the GARCH forecasts, thereby leading to significantly reduced out-of-sample predictive power.

6 Acknowledgements

The authors are most grateful to Shiqing Ling for helpful discussions, to Dick van Dijk for making available his GAUSS program, and the Australian Research Council for financial support.

7 References

Bollerslev, T., Generalized Autoregressive Conditional Heteroskedasticity, *J. Econometrics*, 31, 307-327, 1986.

Chen, C. and L. Liu, Joint Estimation of model parameters and outlier effects in time series, *J. Amer. Statist. Assoc.*, 88, 284-297, 1993.

van Dijk, D., P.H. Franses, and A. Lucas, Testing for ARCH in the presence of additive outliers, *J. Applied Econometrics*, 14, 539-562, 1999.

Engle, R.F., Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987-1007, 1982.

Engle, R.F. and G.G.J. Lee, A permanent and transitory component model of stock return volatility, Discussion Paper 92-44R, University of California, San Diego, 1993.

Franses, P.H., D. van Dijk, and A. Lucas, Short patches of outliers, ARCH and volatility modeling, Discussion Paper 98-057/4, Tinbergen Institute, Erasmus University Rotterdam, 1998.

Franses, P.H., and H. Ghijsels, Additive outliers, GARCH and forecasting volatility, *Int. J. Forecasting*, 15, 1-9, 1999.

Ledolter, J., The effects of additive outliers on the forecasts from ARIMA models, *Int. J. Forecasting*, 5, 231-240, 1989.

Nelson, D.B., Stationarity and persistence in the GARCH(1,1) model, *J. Econometrics*, 45, 7-38, 1990.