

Volatility Estimation through ARCH/GARCH Modeling

Aric LaBarr, Institute for Advanced Analytics
North Carolina State University

ABSTRACT

Volatility estimation plays an important role in the fields of statistics and finance. Many different techniques address the problem of estimating volatility of financial assets. Autoregressive conditional heteroscedasticity (ARCH) models and the related generalized ARCH models are popular models for volatility. This paper will introduce the need for volatility modeling as well as introduce the framework of ARCH and GARCH models. A brief discussion about the structure of ARCH and GARCH models will then be compared to other volatility modeling techniques.

INTRODUCTION

Volatility modeling plays a critical role in mathematical finance and statistical applications. The ability to estimate and forecast volatilities for different assets and groups of assets leads to a better understanding of current and future financial risk. There are many applications in finance that require accurate estimation and forecasting the volatilities of financial assets:

- Estimating the value at risk of a portfolio
- Optimizing allocations of assets in a portfolio
- Hedging the risk of portfolios
- Pricing multiple assets in an option

Many different methods of volatility estimation have been developed over the past few decades. The following notion will be used throughout the rest of the paper. Let r_t be the daily return of a financial asset, modeled by $r_t = \sqrt{\sigma_t^2} \varepsilon_t$, and ε_t is a random error term with a mean of zero and variance of 1. This means that h_t is the variance of the returns and this variance is what will be modeled. These daily returns are defined as the logarithm of the relative prices $\log\left(\frac{P_t}{P_{t-1}}\right)$ with P_t as the current price in dollars.

SIMPLE VOLATILITY ESTIMATION

One of the worst things to do is to assume that volatility stays constant across time. Understanding the characteristics of financial assets helps develop estimation methods for volatilities. Figure 1 displays the returns of the S&P 500 Index.

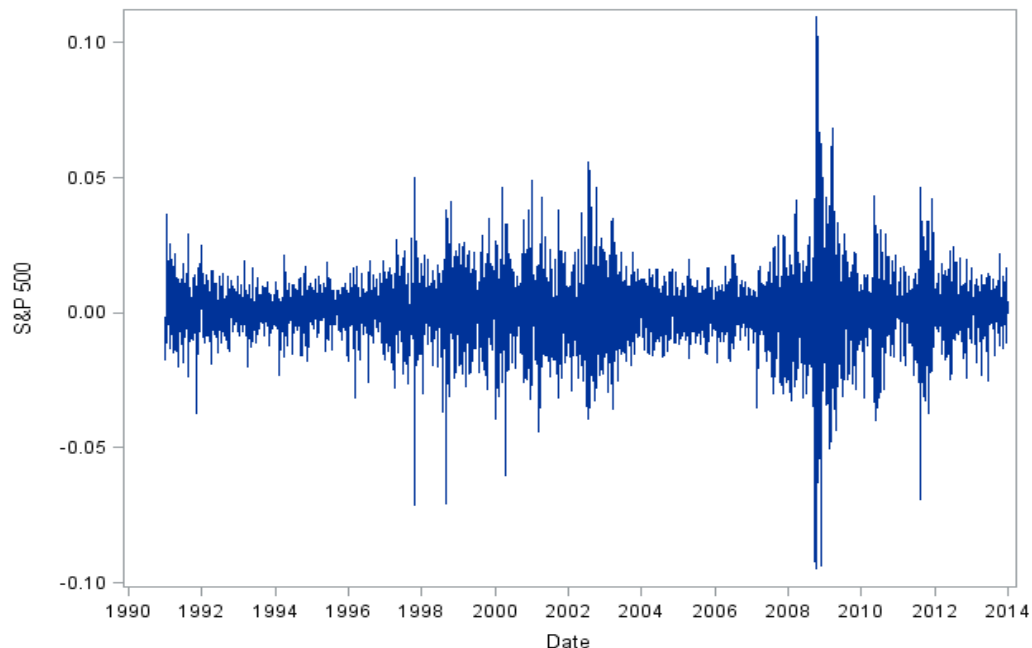


Figure 1. Returns for S&P 500 from 1/1/1990 – 12/31/2013

From this figure we can see that volatility seems to vary over time instead of remaining constant. The volatilities also exhibit some persistence, or dependence over time. This can appear as a clustering effect where large (or small) values of returns are followed by large (or small) returns of either sign.

One simple way to try and account for changes in the variances over time is through a basic rolling window calculation. In this calculation, the variance (volatility) is just an estimate of the m most recent squared observations:

$$\sigma_t^2 = \frac{1}{m} \sum_{i=1}^m r_{t-i}^2$$

One of the problems with this approach is when you are shifting from periods of low (or high) volatility to periods of high (or low volatility) as seen in Figure 2.

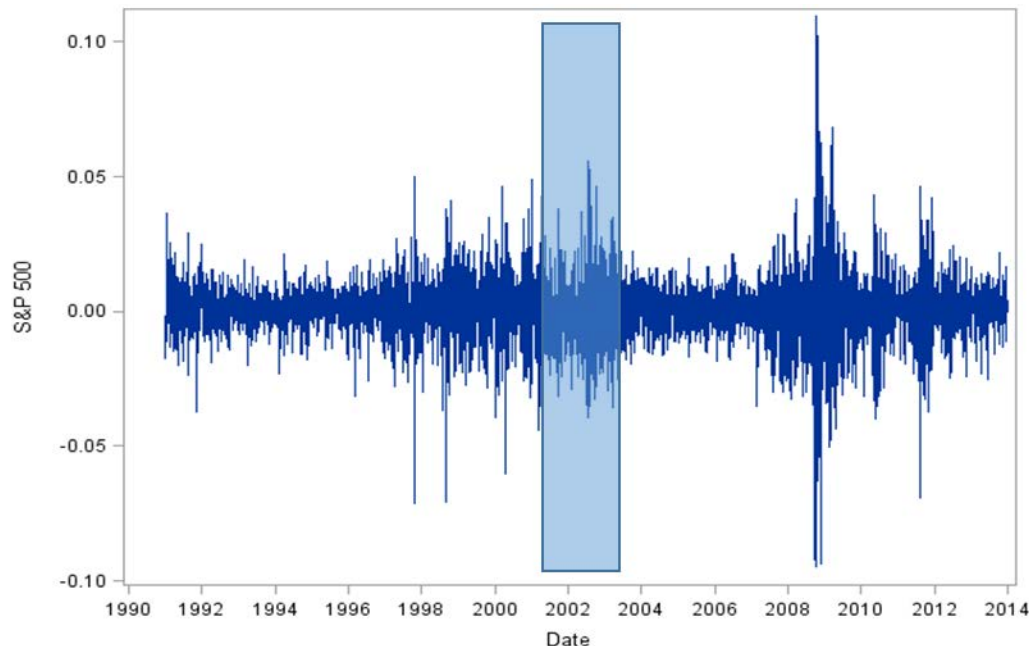


Figure 2. Rolling Window on Returns for S&P 500

The shaded region is a region of relatively high volatility compared to the region that follows it. Although a few recent observations single a shift in the volatility pattern, these values get averaged out in the rolling window calculation. This could lead to inflated estimations of volatility until the window adequately covers the period of low volatility.

One approach to getting around this problem would be a weighting scheme applied to the squared observations where more recent observations have a higher weight than past observations. The formula would be the following:

$$\sigma_t^2 = \sum_{i=1}^n \omega_i r_{t-i}^2$$

The weights (ω_i) will be larger for more recent observations and should all sum to one. Although many different weighting schemes have been proposed over the years, an exponential weighting scheme is probably the most popular. This is basically a simple exponential smoothing approach:

$$\sigma_t^2 = \omega r_{t-1}^2 + (1 - \omega) \hat{\sigma}_{t-1}^2$$

This was the basic approach taken by JP Morgan when they released their RiskMetrics™ data in 1990s. The downside to this approach is that the weight, ω , stays the same. This is true even if you have a portfolio of assets where potentially different weights should exist for different factors in the model.

ARCH/GARCH MODELS

In the 1980's a different family of approaches was born that has become one of the most popular fields and volatility estimation techniques used today. It all started with Engle's (1982) autoregressive conditional heteroscedasticity (ARCH) model and later adapted into Bollerslev's (1986) generalized ARCH model. This family of models has been studied and adapted countless times over the past few decades.

AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY (ARCH)

Instead of a weighting scheme, Engle uses an autoregressive time series approach to account for persistence in volatility estimation. He also assumes that the volatilities are conditional on each other over time. Engle's model is defined as the following:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2$$

In this model, the intercept α_0 is positive, and the parameter coefficient α_1 is nonnegative so that the volatility estimation is positive. The returns here in Engle's model are also assumed to be Normally distributed with a mean of zero and variance of σ_t^2 , conditional on all information up to the current time. Sometimes this model is referred to as the Normal ARCH model for this reason. Of course this model can easily be extended to include as many lags of returns as you like since it would be unreasonable to assume that today's volatility only depends on yesterday's returns. This is called the ARCH(q) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_q r_{t-q}^2 = \alpha_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2$$

However, in practice, we often need large values of q to accurately model real world data.

GENERALIZED ARCH (GARCH)

Bollerslev avoids the problem of large values of q in Engle's ARCH model by generalizing the ARCH(q) model in the generalized autoregressive conditional heteroscedasticity (GARCH) model, in much the same way as an autoregressive (AR) model extends to the autoregressive moving average (ARMA) model. The GARCH model allows for a longer memory process with more flexibility. The GARCH model stills assumes Normality in much the same as Engle's model did, but now includes previous estimates of volatility in the calculation:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Just like the ARCH model can be extended to the ARCH(q) model, the GARCH model can be extended to the GARCH(q, p) model as follows:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2$$

Just like with Engle's model, the intercept must be positive, while the coefficients must be nonnegative to ensure that the volatility estimate is positive. In many cases, $q = p = 1$, is an adequate model to fit real world data.

The GARCH framework of models looks very similar to typical ARMA models that are used often in time series modeling. This is because that the GARCH model is actually a special case of an ARMA model. Through some simple algebra, one can convert the GARCH(1,1) model into an ARMA(1,1) model.

S&P500 Index Example

Let's use the GARCH(1,1) model to estimate of the volatility for December 3, 2007 (the official beginning of the recession). This model should quickly adjust to the raising volatilities of the stock market at that time compared to the simpler approaches described earlier.

The following SAS® code will produce estimates for a GARCH(1,1) model:

```
proc autoreg data=SP500;
    model SP500_Vol = / garch=(q=1, p=1);
run;
```

The estimated coefficients for this model can be seen in Output 1 below.

Parameter Estimates					
Variable	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	0.000518	0.000122	4.24	<.0001
ARCH0	1	6.1063E-6	9.6829E-8	6.31	<.0001
ARCH1	1	0.0591	0.004529	13.04	<.0001
GARCH1	1	0.9354	0.004978	187.92	<.0001

Output 1. Output from PROC AUTOREG

From Output 1 we get estimates for $\alpha_0 = 0.0000006$, $\alpha_1 = 0.0793$, and $\beta_1 = 0.9129$. The intercept listed in the output is the intercept on the returns themselves, not the GARCH volatility model.

ARCH/GARCH FAMILY

Countless adaptations have been made to the ARCH/GARCH modeling framework over the last few decades. I will not list all of the different branches of ARCH/GARCH models that have been created, but there are three important ones that I will quickly mention here:

1. t -GARCH
2. IGARCH
3. EGARCH

t -GARCH Model

Bollerslev (1987) proposes one of the first extensions of the ARCH/GARCH modeling system by making an adjustment to the conditional distribution of the error term. Bollerslev notes the usefulness of the ARCH and GARCH models in portraying the clustering of volatilities in financial data. However, he also notes that financial data are conditionally leptokurtic (fatter tails than the Normal distribution). Therefore, Bollerslev proposes to switch the conditional error distribution to a t -distribution instead of a Normal distribution. The degrees of freedom of the t -distribution would be estimated along with the other parameters in the model. SAS can calculate the t -GARCH model using the following distribution option in the model statement:

```
model SP500 = / garch=(q=1, p=1) dist=t;
```

When run, SAS will produce the usual estimates along with an estimate for TDFI, the degrees of freedom for the t -distribution.

IGARCH Model

Bollerslev and Engle (1986) also noted that there is a possibility of a unit root in the GARCH framework that they created. As mentioned earlier, the GARCH(q, p) model can be manipulated into an ARMA time series process through simple algebra. The resulting ARMA process is the following:

$$r_t^2 = \sum_{i=1}^p (\alpha_i + \beta_i) r_{t-i}^2 - \sum_{j=1}^q \beta_j v_{t-j} + v_t$$

For those with experience in ARMA time series models, the above model is only stationary if the sum of the above α_i 's and β_i 's is less than one. However, if the sum is equal to one, then a unit root process exists and the model must become an integrated GARCH (IGARCH) model instead. SAS can calculate the IGARCH model using the following type option in the model statement:

```
model SP500 = / garch=(q=1, p=1, type=integ);
```

This process forces SAS to have the coefficients mentioned above sum to one.

EGARCH Model

All of the models we have discussed so far have had symmetric underlying distributions. However, not all instances would have this characteristic. Nelson (1991) noted that some financial data sets contain the leverage effect. The leverage effect is when variance (volatility) increases more when a return is negative compared to when a return is positive. This adheres to the typically observed market condition that people are risk adverse and will tend to react differently to negative returns than positive ones. To get around this, Nelson created the exponential GARCH model (EGARCH):

$$\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^q \alpha_i g\left(\frac{r_t}{\sigma_t}\right) + \sum_{i=1}^p \beta_i \log(\sigma_{t-i}^2)$$

SAS can calculate the EGARCH model using the following type option in the model statement:

```
model SP500 = / garch=(q=1, p=1, type=exp);
```

Not only is this model developed using the natural logarithm of volatility instead of volatility itself, but it also uses a more general distribution than the Normal distribution proposed by Engle. Nelson uses the Generalized Error Distribution instead.

CONCLUSION

Volatility estimation is key to the financial world and analysts within it. The ARCH/GARCH family of modeling techniques is one of the most popular and fundamental modeling approaches for volatility forecasting and estimation. From the accurate estimation of future volatility in financial assets, one can better hedge against risk, optimize a portfolio of assets, and even price options on assets. Although this was only a brief introduction to the framework, it hopefully opens the eyes of financial modelers to important techniques.

REFERENCES

Engle, Robert. Autoregressive Conditional Heteroscedasticity with Estimates of the Variances of United Kingdom Inflation. *Econometrica*, 987-1007, 1982.

Bollerslev, Tim. Generalized autoregressive heteroscedasticity. *Journal of Econometrics*, 31:307-327, 1986.

Bollerslev, Tim. A conditionally heteroscedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 69:542-547, 1987.

Bollerslev, Tim & Engle, Robert. Modeling the persistence of conditional variances. *Econometric Reviews*, 5:1-50, 1986.

Nelson, Daniel. Conditional heteroscedasticity in asset returns: A new approach. *Econometrica*, 59:347-370, 1991.

Alexander, Carol. 2008. Practical Financial Econometrics: Market Risk Analysis II. Pp. 89-199. West Sussex, England: John Wiley & Sons Ltd.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Aric LaBarr
Organization: Institute for Advanced Analytics
Address: 920 Main Campus Dr, Suite 530
City, State ZIP: Raleigh, NC 27606
Work Phone: 919-513-2957
Email: aric_labarr@ncsu.edu
Web: <http://www4.ncsu.edu/~adlabarr/>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.