Risk Analysis Using GARCH and ARIMA Models

Spencer Lyon

April 2, 2013

Macroeconomics and Computational Laboratory (MCL)

BYU Economics

Abstract

Put an awesome abstract here.

CONTENTS

1 Introduction **1**

2 Estimation **2**

2.1 The Data 2

2.2 ARIMA Models 3

2.3 (G)ARCH Models 5

3 Results **5**

4 Conclusion **5**

Risk Analysis Using GARCH and ARIMA Models

1 Introduction

One longstanding phenomenon economists see in financial data is known as the equity premium puzzle (Mehra and Prescott, 1985). The heart of the equity premium puzzle can be summarized one of two equivalent questions: why is demand for risk free government bonds so high when returns on those assets are so low? The second question could be stated as follows: why is demand for stocks, and hence stock prices, so low despite the relatively high expected returns

Over the years economists have posed many possible explanations for why the equity puzzle premium (EPP) exists, or even if it exists. Modern economics is based largely on rational expectations and risk aversion. An appeal to these theories to explain the EPP seems natural and intuitive; however, this type of explanation falls very short. To put this disparity in perspective Mankiw and Zeldes (1991) pointed out that in order for risk aversion to explain the EPP, investors would have to be indifferent between taking a sure $51,209 or having 50% at winning either $50,000 and $100,000. Other common explanations have generally been rooted in tax policy, market failures, liquidity constraints, and the implied volatility of the underlying equities.

The focus of this paper will be analyzing the ability of ARIMA and (G)ARCH models for identifying and explaining the equity premium puzzle. ARMIA methods have been around for some time and have held a significant foothold in the finance literature. An ARIMA model allows the economist to model the effect of historical moving averages on current values. This flexibility is particular appealing to someone trying to model financial time series because things like seasonality, momentum, and market lags and/or inefficiencies lead to these moving averages storing a lot of relevant data.

ARCH and GARCH models are, in a sense, very similar to ARIMA models. Instead of trying to model the effect of changing moving averages, (G)ARCH models attempt to capture changes in the variance of underlying parameters. It is immediately apparent the trained financial eye that focusing on the changes in expected variance might actually be more instructive than moving averages when applied to forecasting models.

The remainder of this paper will be structured as follows: Section 2 will discuss ARIMA and (G)ARCH models as well as the data I am using, Section 3 will show the results of my analysis, and Section 4 will have a few concluding thoughts and ideas for further research.

2 Estimation

2.1 The Data

The data for this paper is quite simple. To mimic a risk free asset I have gathered daily data for the price of both 3-month and 10-year US treasury bills. I obtained this data from the St. Louis Federal reserve under the data set names DGS3MO and DGS10 for the 3-month and 10-year maturities, respectively. The units on these data sets are percent returns, the data is not seasonally adjusted, and the data was sampled at a daily frequency.

To model risky assets I chose to use the daily returns on the S&P 500 stock market index. The units on this data set is US dollars ($) and I retrieved the data from Yahoo! Finance. The S&P 500 was a natural choice because it is a very common metric used in finance to measure the overall movement of the stock market. Because some explaination of the EPP refer to the volatility of stocks I also chose to use the VIX, also from Yahoo! Finance. [1]

For the most part, these four time series had observations on the same days: all of them are reported for week-days only. However, the big difference is that the data from Yahoo! Finance is only reported on trading days, while the data from FRED is reported every Monday through Friday, including most holidays. To overcome this I simply dropped observations for days where FRED reported data, but the stock markets were closed. In Table 2.1 I have included a summary of this data and a simple plot of the data in Figure 1.

[1] The VIX is constructed using the implied volatilities on call and put options for the stocks making up the S&P 500. It is the generally accepted measure of overall stock market volatility.

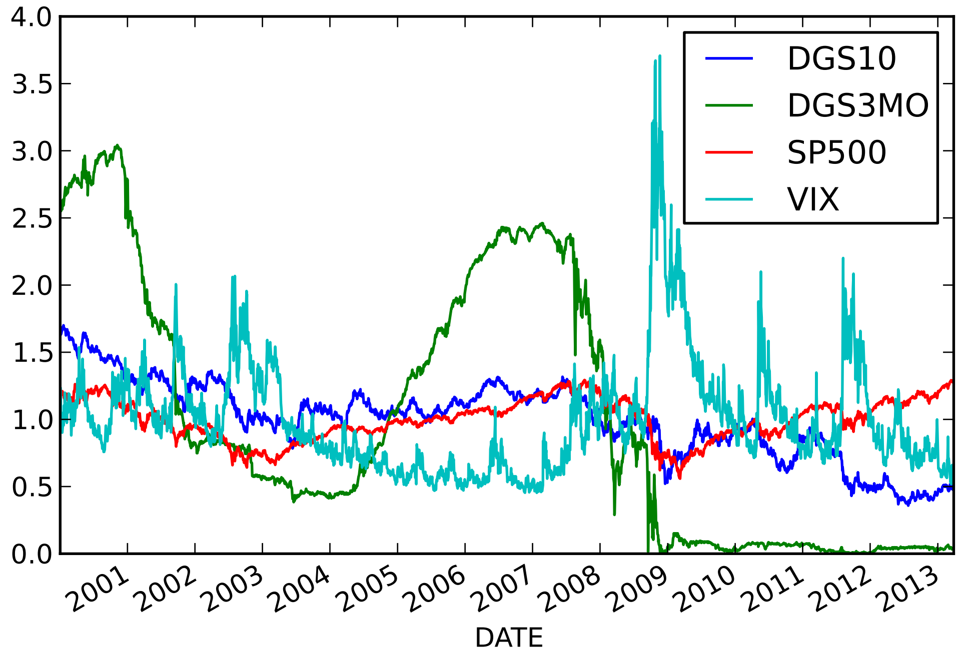

Figure 1: Plot of the data used in this paper. Note that because the units are different and the value of the S&P 500 is much larger than the percents on the T-bill data sets, I have divided each data set by its mean before plotting.

	DGS10 (%)	DGS3MO (%)	SP500 ($)	VIX ($)
obs	3305	3305	3305	3305
mean	3.99311	2.10994	1210.33	21.8055
std	1.14046	2.01772	187.527	9.16702
min	1.43	0	676.53	9.89
25%	3.38	0.15	1095.45	15.54
50%	4.12	1.59	1214.91	20.03
75%	4.74	3.85	1357.98	25.4
max	6.79	6.42	1569.19	80.86

Table 1: Summary of the data . All the data was sampled from 1/1/2000 to 3/28/2013. Note the rows with percent labels correspond to percentiles

2.2 ARIMA Models

Time series modeling is all about identifying time-varying trends in data and extrapolating those trends into the future with as little error as possible. One extremely general family of time series models is known as the ARIMA model, or autoregressive integrated moving average model. There are at least three roles an ARIMA model (or any forecasting model) can play:

1. By themselves to predict future values of a dependent variable
2. To predict the future value of an independent variable in an economic model
3. To estimate future values of systematic residuals in an economic model

In this paper I will be using ARIMA models within the first and second categories listed above. In order to understand how I will be using them I will first explain what the model is in a general form. An ARIMA model is defined in terms of three parameters $p, q,$ and $, d$ and is defined as follows:

$$Y_t^* - \phi_1 Y_{t-1}^* - \cdots - \phi_p Y_{t-p}^* = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \tag{2.1}$$

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right) Y_t^* = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t \tag{2.2}$$

Where L is the lag operator and $Y_t^* \equiv (1-L)^d Y_t$.

The first major roadblock when trying to build an ARIMA model is identifying the parameters $p, q,$ and d. I will start by trying to identify p and q. To do this I need to look at the autocorrelation and partial autocorrelation coefficients (see McDonald notes section 8). In Figure 2 I have included a plot of the autocorrelation and partial autocorrelation coefficients for each data set.

As can be clearly seen, for each of my time series we have an initial spike in the partial autocorrelation coefficient for 1 lag and then nothing. Also, in each picture we see that the autocorrelation coefficients slowly decline for all 4 time series. Together, these two pieces of information suggest that we should approximate an AR(1) model by setting $p = 1$and $q = 0$.

The last identification step is to determine a suitable value for d. This is a bit more subjective, but can still be formalized. The easiest way to find a value for d is to plot successive differences of the time series and stop increasing the number of differences when you see that the time series is fairly constant in mean and variance. In Table 2.2 I have included a table that reports the mean and variance of the first 4 differences for all 4 data sets.

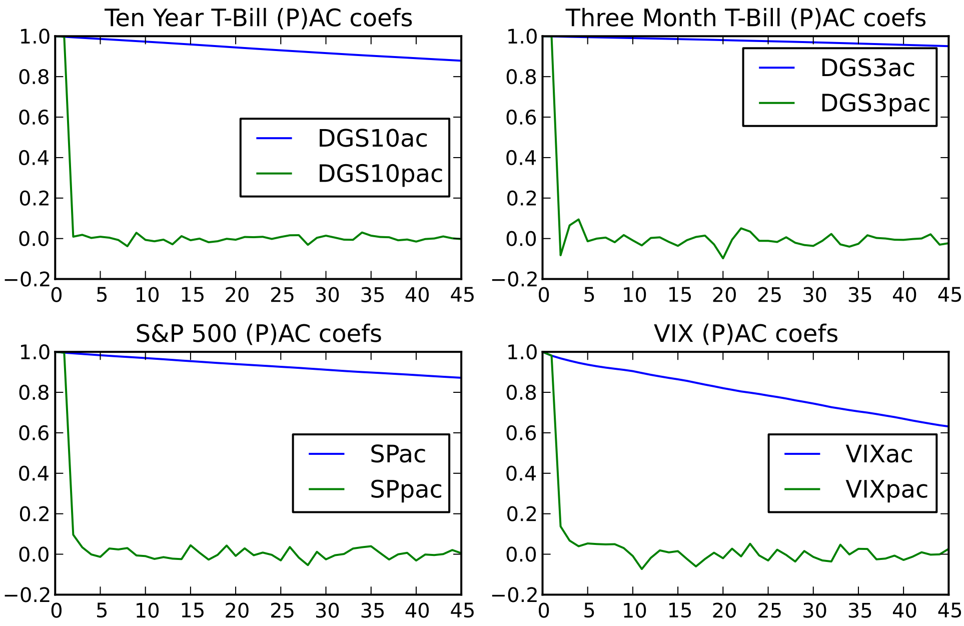

Figure 2: Autocorrelation and partial autocorrelation coefficients for each data set. Used to find the parameters q and p in an ARIMA model.

2.3 (G)ARCH Models

3 Results

4 Conclusion

DataSet	Lags	Mean	Variance	DataSet	Lags	Mean	Variance
DGS10	1	-0.001425545	0.003990852	SP500	1	0.03449455	224.2908
	2	-0.002824705	0.007977331		2	0.08398426	406.3557
	3	-0.00424894	0.01156277		3	0.1329679	574.528
	4	-0.005655862	0.01518609		4	0.1779158	748.1064
DGS3MO	1	-0.001637409	0.003366346	VIX	1	-0.003483656	3.13206
	2	-0.003254617	0.007774232		2	-0.007680896	5.355797
	3	-0.00488189	0.01147704		3	-0.01181405	7.261939
	4	-0.006498031	0.01407682		4	-0.01544986	9.018458

Table 2: TODO: Write this caption

References

Mankiw, N. Gregory and Stephen P. Zeldes, "The Consumption of Stockholders and Nonstockholders," *Journal of Financial Economics*, March 1991, *29* (1), 97–112.

Mehra, Rajnish and Edward C. Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145 – 161.