Financial Econometrics assignment 2

Ethan Wright

November 2025

1

The goal of Clark's paper is to explain why asset prices exhibit non gaussian behavior. in the introduction he presents a plot of returns that is much different than a classic normal distribution. Clark argues that a random walk/brownian motion itself is unable to recreate what is seen in markets. Clark starts by defining the prices of cotton X as a function of time he defines this as

$$X_t = X_{t-1} + \epsilon_t$$

he also defines how time will move. Instead of just having time we also have this idea of "trading time". When new information is randomly revealed investors will arbitrage this meaning that the price of a stock can evolve very quickly or very slowly depending on these shocks. He defines the time increment or the directing process $T(t)$. This is non decreasing, has a finite mean and is independent with $X(t)$. Due to this finite mean $X(t)$ has some nice properties

$$var(x(T(t))|\Delta T(t) = v) = v\sigma^2$$

and the unconditional variance is

$$E_{\Delta(T(t))}(v\sigma^2) = \alpha\sigma^2$$

. Clark then shows how the kurtosis will be a function of the variance of the directing process. assuming that $\Delta T(t) = v$

$$E(\delta X(T(t))^4|\delta T(t) = v) = 3v^2\sigma^4$$

the unconditional expectation will be

$$E_{\Delta X(T(t))}(3v^2\sigma^4) = 3\sigma^4(\alpha^2 + var(v))$$

and we get the function of kurtosis being

$$k_{\Delta X(T(t))} = 3\left[\frac{\alpha^2 + var(v)}{\alpha^2}\right]$$

This also means that as long as change in T(t) (which is strictly increasing) is positive then we will have kurtosis greater than 3. This explains the fat tails we observe in commodity prices as well as returns of stocks.

Clark then provides a model for the distribution of a normal function subordinated by a lognormal trading time. lognormal time has distribution

$$f(x;\mu,\sigma_1^2) = \frac{1}{2\pi\sigma_1^2 x} exp\left(-\frac{\log(x-\mu)^2}{2\sigma_1^2}\right)$$

this leads X(T(t)) to have the distribution

$$f_{LNN}(y) = \frac{1}{2\pi\sigma_1^2 x\cdot\sigma_2^2}\int_0^\infty v^{-\frac{3}{2}}\cdot exp\left(\frac{-\log(v-\mu)^2}{2\sigma_1^2}\right)\cdot exp(\frac{-y^2}{2v\sigma_2^2})dv$$

. This is clearly no longer normal due to our subordinate process. This shows what is the most important part of this paper is that since trading time is not constant we should not expect the distribution of returns to be normal. Clark then goes on to test the distribution of trading time and the conditional distribution of X(T(t)) given T(t) but this is more econometrics rather than the mathematical theory of the paper.

2 a

In the previous assignment I used data from 2004 to better compare the VIX to the VXO. In this assignment I used data from 1990 since this gives more observations. Looking at the ACF, I observe a stationary, slow decaying ACF. I supply the dickey fuller test in the appendix which shows that we do in fact have a stationary process. This looks to me as though it is a AR(1) process. I also supply the PACF in the appendix which shows a very large value on the first lag with small values in all other lags. This further corroborates my decision to go with an AR(1) process.

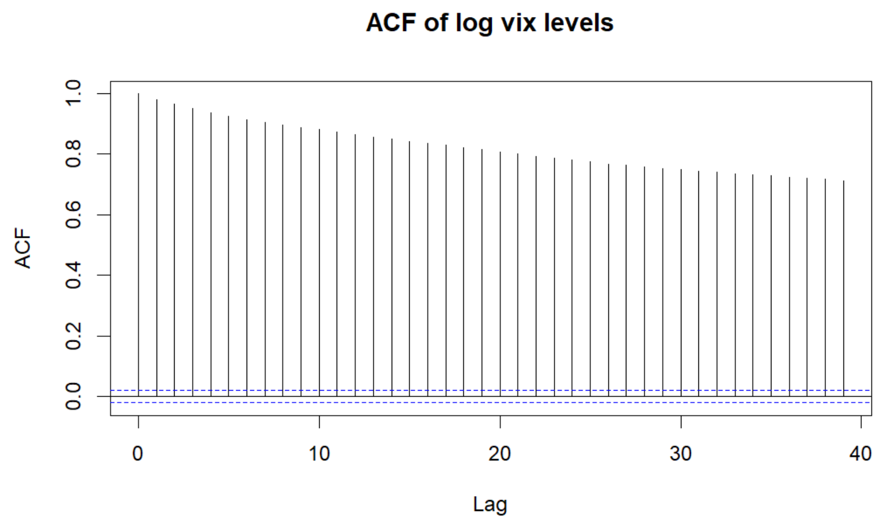

Figure 1:

Before I present my AR(1) process I would also like to discuss the "Akaike information criterion" (aic) this is a test to aid in model selection. When I use aic I get a AR specification on 12. This can help with model choice but also can overfit the data. Since this is so clear from ACF/PACF I will use the AR(1) specification.

My AR(1) model is specified as:

$$y_t = 2.90471 + 0.9803572 y_{t-1} + \epsilon_t$$

as we can see the coefficient on our lag is near 1. This means that we have very strong autocorrelation with the previous lag or that shocks are persistent in the VIX. This makes intuitive sense as volatility levels in markets yesterday should affect volatilty levels in markets today.

2.1 returns

we now repeat the analysis which we did on the VIX log levels but on the log daily returns on the VIX. We define the returns as

$$\Delta \log(VIX_t) = \log(VIX_t) - \log(VIX_{t-1})$$

Looking at the ACF we now see a very different picture. While previously we saw very strong positive autocorrelation the returns show weak negative autocorrelation.

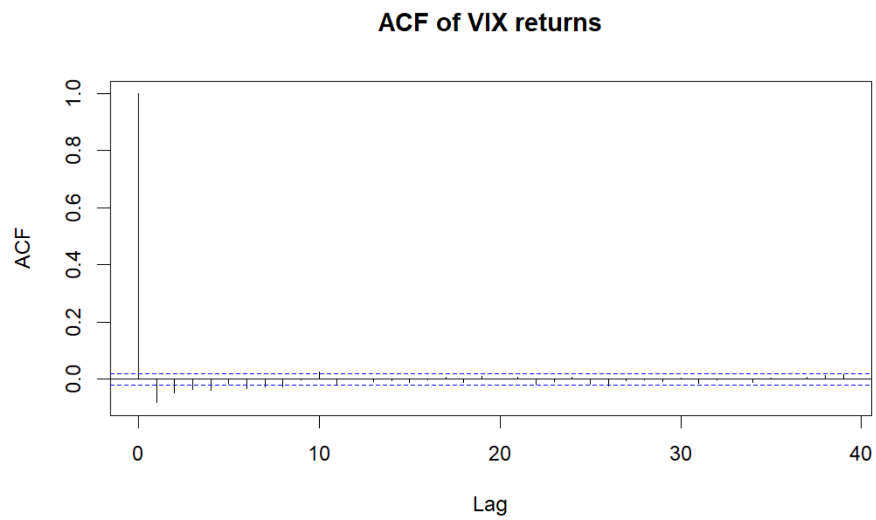

Figure 2:

We see up until the fourth lag our spikes are outside of the confidence interval but these are just slightly outside. I once again provide the PACF in the appendix to improve my analysis. When I look at this it does not seem that there is much of any autocorrelation here. It seems that returns are not serially auto-correlated. This makes sense as returns are not auto-correlated. For my AR(1) specification I will use AIC to aid my selection in this case. AIC chooses a 11 lag model with very small negative values at each lag (all less than .1)

Table 1: AR(11) Model Coefficients

Term	Coefficient
AR(1)	-0.0984
AR(2)	-0.0693
AR(3)	-0.0588
AR(4)	-0.0608
AR(5)	-0.0400
AR(6)	-0.0531
AR(7)	-0.0473
AR(8)	-0.0443
AR(9)	-0.0184
AR(10)	0.0111
AR(11)	-0.0248

Table 2: *
Estimated $\sigma^2 = 0.004516$

When I compare this to the VIX levels we see much less autocorrelation in returns when compared to levels. AR models are also harder to specify as they are generally used for longer term trends. It may be better to use a moving average model or for returns which seem to have much smaller return correlation.

2.2 regression

Table 3:

	Dependent variable:
	VIX_returns
L(VIX_returns, 1)	-0.063^{***}
	(0.011)
I(L(VIX_returns, 1)ˆ2)	-0.296^{***}
	(0.055)
Constant	0.001^{*}
	(0.001)
Observations	9,055
R^2	0.010
Adjusted R^2	0.010
Residual Std. Error	0.068 (df = 9052)
F Statistic	45.777^{***} (df = 2; 9052)
Note:	*p<0.1; **p<0.05; ***p<0.01

When regressing on lagged returns and squared lagged returns of VIX we see that both are statistically significant at the 99% level of significance. If we consider what the semi-strong Efficient Market Hypothesis (EMH) says, assets reflect all public information about it at the time. This also means that one should not be able to profitable trade on past information of an asset. What the EMH does not do is discuss volatility. Since we are regressing on the returns of the VIX which measures the volatility of stocks. So, having significance at either the lag or the square lag does violate the EMH here since it does not really apply. Additionally the VIX cannot be traded as we discussed in the previous assignment it is derived from options on the s&p 500. This means that having predictable returns does not provide an opportunity for arbitrage.

2.3 residuals

For this I will provide the histogram in the paper and the ACFs in the appendix. Almost all lags are not significant, this implies that the residuals are not auto-correlated. The square of the residuals have a somewhat stronger auto-correlation when compared to the non squared residuals. We can see somewhat significant autocorrelation until the seventh lag. We should note that this auto-correlation is quite weak only peaking at about .2 at the third lag. When looking at the histogram of the residuals it seems to be about normal. Of course it is somewhat hard to discern if tails are normal just from the distribution itself so I also provided a Q-Q plot in the appendix which allows me to get a better idea of the distribution. We can see from this that the distribution seems to have fat tails as the distribution is above the normal line on the high end and below the normal line on the low end.

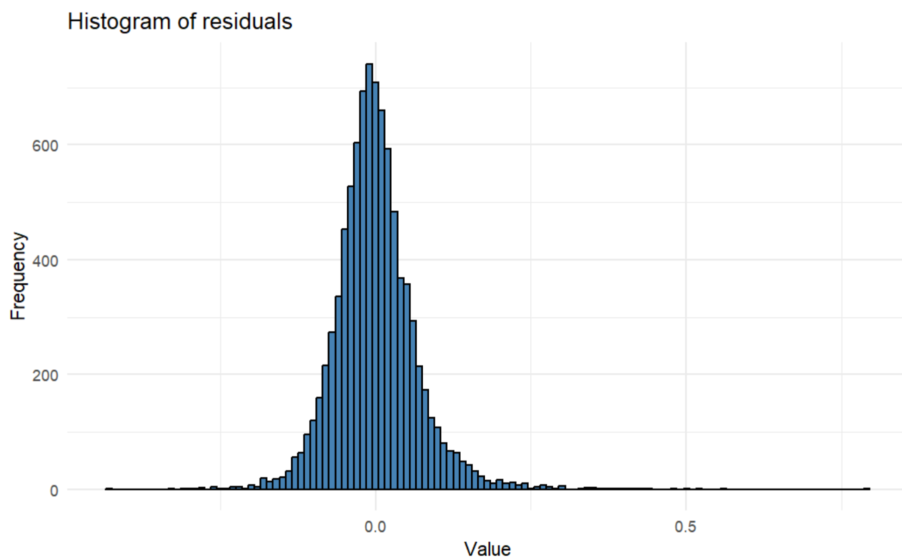

Figure 3: Enter Caption

3 Linear Regression

I perform a rolling calculation to create returns from h = 1,...,100. I then run a regression and plot this in R. As h increases we can see that R_h^2 has a slight upward trend while $\hat{\beta}_h$ becomes increasingly negative.

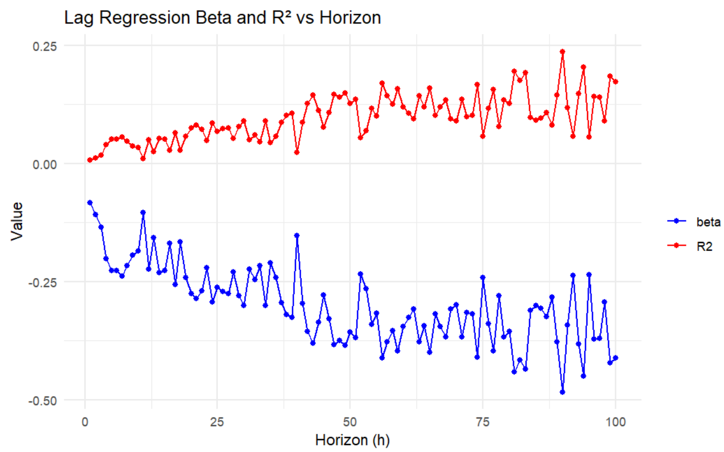

Figure 4: Enter Caption

3.1

As we have seen that R^2 increases with horizon h. This is similar to what Fama and French observed with the returns of stocks. They hypothesize that the reason this happens is that there is two parts that make up stock returns. A random walk part as well as a decaying stationary part. This means that in the short run we can observe predictable variance and return of stocks as we have this natural mean reverting tendency. On the other hand we have a white noise that wants to push the prediction to zero. As time increases the white noise variance eventually dominates our stationary components and after about 3-5 years we have no predictive power for returns. If we know about this mean reverting component we should be able to predict how volatility will act during these periods. This shows why we can explain some of the variance in the VIX with these lagged values! Fama and French have another paper which explains that the movement of stocks and bonds are correlated, this is due to the buisness cycle. With both stocks and bonds as the economy gets stronger, default spreads fall and stock yields fall. When business conditions are weak, default spreads rise and stock yields rise. Applying this same idea to volatility we can assume that stock volatility will change over the business cycle. This once again explains why we see a non-zero β and positive R^2 in the graph. Overall we can explain both a mean return tendency as well as the business cycle to explain why stocks see predictable returns and volatility in the stock market.

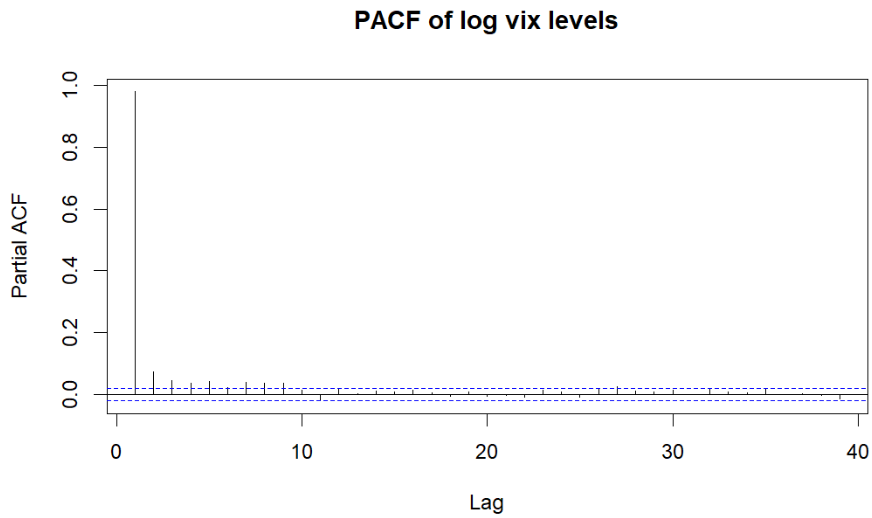

Figure 6: PACF of VIX levels

4 Appendix

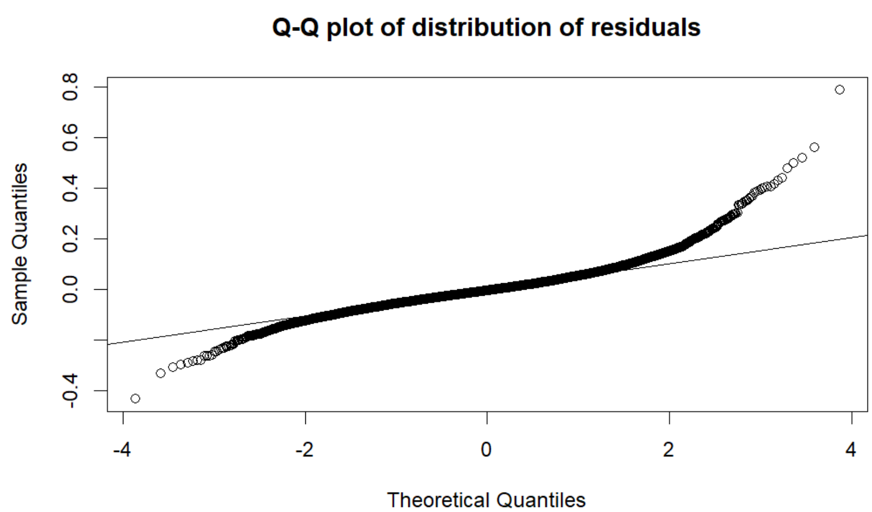

Figure 5: Quantiles of returns distribution

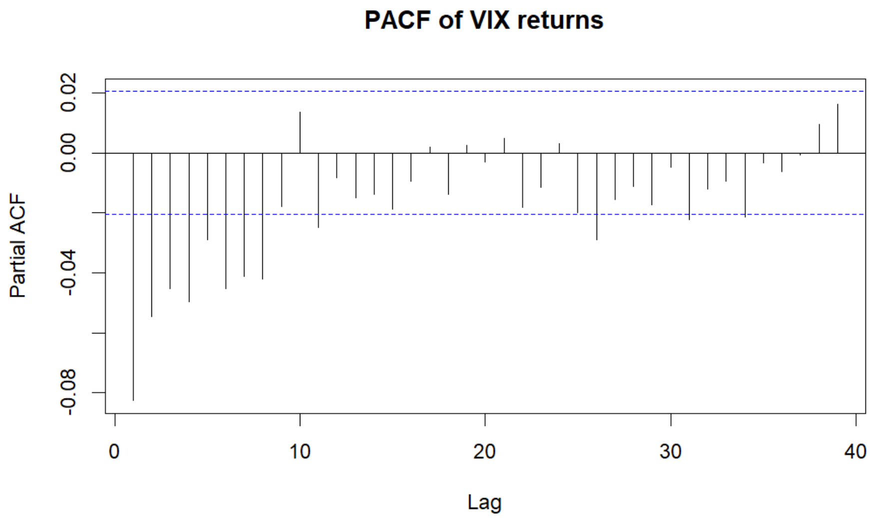

Figure 7: PACF of VIX returns

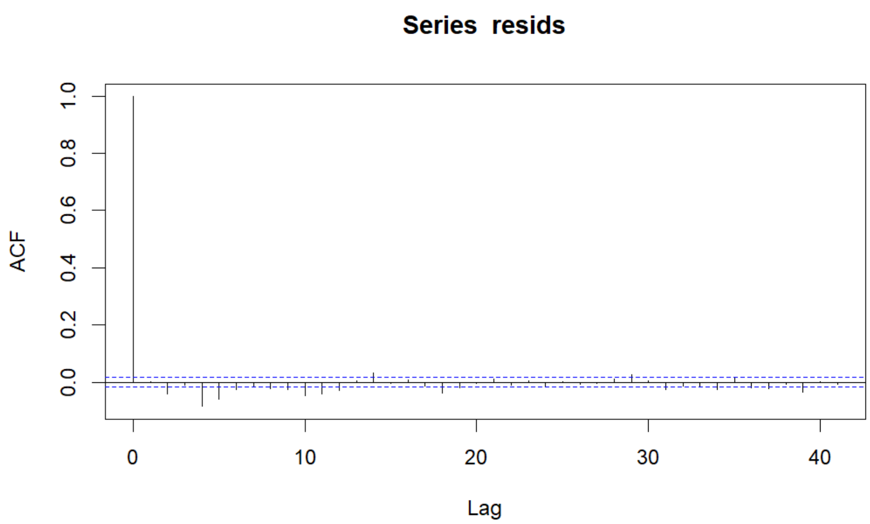

Figure 8: ACF of regression residuals

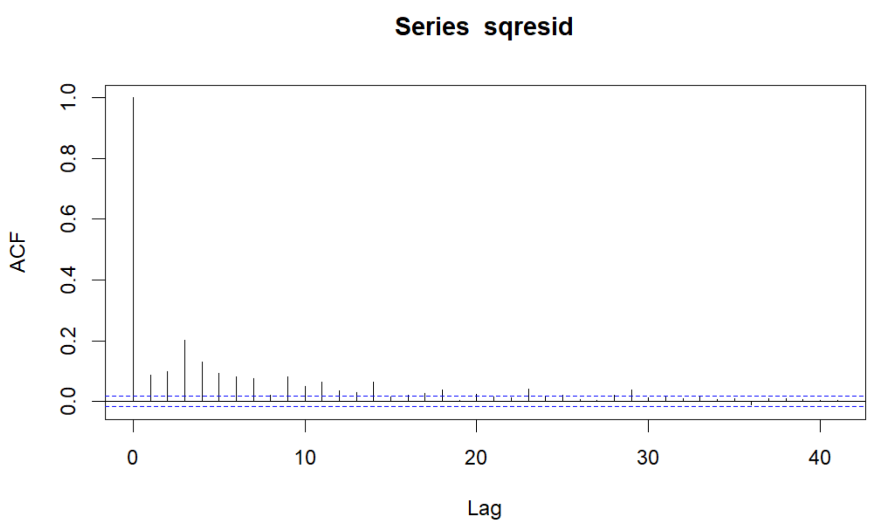

Figure 9: ACF regression square residuals