**ORIGINAL ARTICLE**

**Journal Section**

# Probing the fear gauge

**Krishnan Raman1**∗  **| Anqi Wang1**† **| Jianing Li2**∗ **|** **Qianli Xu2**†

|  |
| --- |
| The Fear Gauge is the most watched metric on Wall Street. We provide a statistical methodology to probe the efficacy of the fear gauge vis-a-vis the three largest exchange traded funds on the market, SPY, DIA & QQQ.  **KEYWORDS**  ETF, VIX |

12nd year, Dept of Statistics, Purdue U

22nd year, Supply Chain and Operations

Management, Purdue U

**Correspondence**

Krishnan Raman

Email: raman23@purdue.edu

**1 | INTRODUCTION: ETFs & the VIX**

Exchange Traded Funds such as SPY (S&P 500), DIA (Dow Jones Industrials), QQQ (Nasdaq Tech Index), have accumulated over a trillion dollars of assets under management since they were first introduced. Their enormous popularity is driven by a bunch of investing trends - increased availability of passive, hands-off, low-fee funds from Blackrock, State Street & Vanguard, a stable risk profile historically averaging 7% annually over a long time-horizon & consumer trust in broad indices as opposed to individual stock picks. However, ETFs are NOT a safe haven. ETFs do lose money when traded speculatively, especially over shorter time duration. During times of recession, ETFs have lost over 20% at times! As such, investors are interested in keeping tabs on the average ETF’s price. However, measuring the volatility (standard deviation) of an ETF over a time horizon is a challenging proposition. An ETF is essentially a weighted sum of stocks. Eg. SPY, is a weighted sum of 500 stocks, DIA has 30 components & QQQ, 100. These component stocks rapidly change price every millisecond during the trading session. It is unclear how to capture these changing price deltas on 100s of timeseries & their standard deviations in a single number.

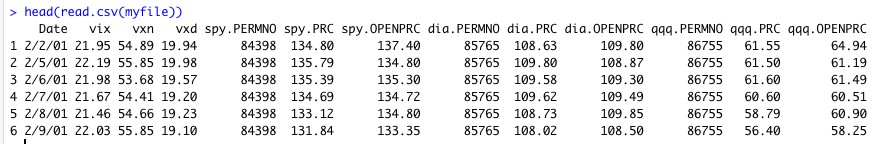
In 1993, Robert E. Whaley, currently a Management Professor at Vanderbilt U, wrote a seminal treatise in the Journal of Derivatives. Titled "***Derivatives on Market Volatility: Hedging Tools Long Overdue***", the paper outlines a "hedging tool" to systematically compute the volatility of an ETF into a single number. This tool went on to acquire outsize proportions and became the single most traded derivative in market history! Dr. Whaley called it the volatility index aka the VIX.

With the colossal success of the VIX came its critics. Numerous studies dispute the correlation of VIX with the ETF. In a widely cited paper *"****How Good is the VIX as a Predictor of Market Risk?"****,* Dr. Clemens Kownatzki, a Professor of Finance at Pepperdine U, alleges "VIX consistently over-estimates actual volatility in normal times...but underestimates volatility in times of market crashes...making it unsuitable for risk-management". Current literature either supports the VIX as a sound measure of the standard deviation of SPY, or fervently believes it is deficient in various statistical metrics.

We probe the efficacy of the VIX in order to answer a few simple queries:

1. Does the VIX statistically correlate with the ETF? Is it a genuine fear index?
2. Can probabilities of market decline be inferred from the movement of the VIX?
3. Does "Market Risk Contagion" exist i.e. are the ETFs themselves strongly correlated, even though the component holdings of each ETF are largely disjoint?
4. Suppose there is a major recession and the VIX peaks. Approximately how long(days/weeks/months) does it take for things to cool down?

**2 | METHODS: Contingency Tables**

To a statistician, an ETF on a given day is summarized by its closing price. Similarly, the VIX is a single number that summarizes the volatility of the ETF. We obtain voluminous data on the SPY, the QQQ and the DIA, for a period of eighteen years, starting February 2001 until December 2018. Each ETF is associated with its own VIX. Thus, we have six distinct time series for a total of 4500 days. A data snapshot is shown below. 

**Figure 1. Time Series Snapshot**

If there existed a contingency table that succinctly captured all that we know about a VIX & its ETF in a simple 2x2 matrix, any association would pop out right away! While there isn’t one, we propose a scheme to discretize a continuous variable, essentially moving from continuous price data to ordinal counts. Prices are not a suitable scale to do math in, since they vary in a very wide range.

However, time-windowed returns, such as the Daily Return, defined as

*Price*[*tomorrow*]− *Price*[*today*]

*Return* =

*Price*[*today*]

are nicely bounded. For instance, the Weekly VIX lies between [-0.5, 2.5], while the weekly ETF loses at-most 24% and gains 17%, lying in [-0.24, 0.17].

A 2x2 contingency table simply counts the number of occasions the VIX & the ETF move in tandem or in opposite directions.

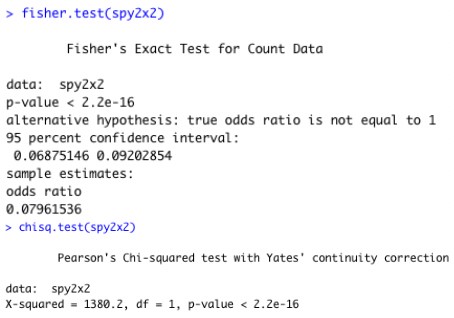
3



**Figure 2. Contingency Tables of VIX-SPY, VXD-DIA, VXN-QQQ**

Even before we conduct a formal test, it is obvious that the odds ratio is far from 1. We test for independence of rows & columns using Fisher’s Exact Test & Pearson’s Chi-square test.

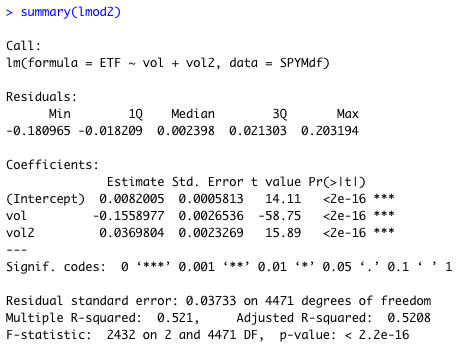
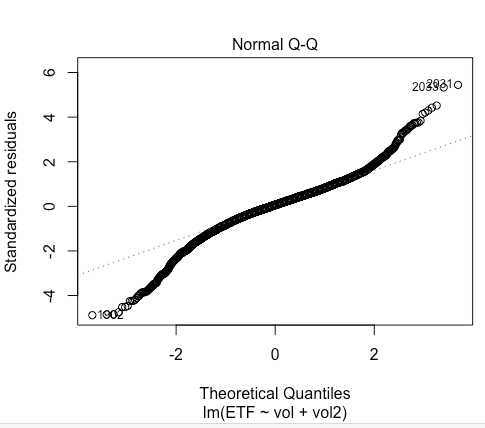
Both tests confirm our intuition:



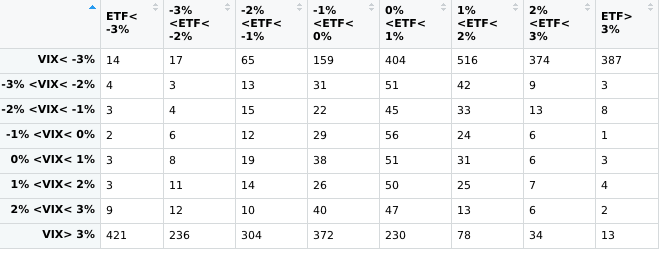
**FIGURE 3. Tests of Independence**

Thus, we have a rigorous confirmation over 18 years of data that the VIX & the ETF are strongly inversely correlated. Higher the VIX, lower the ETF & vice-versa.

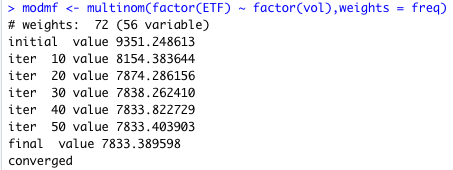
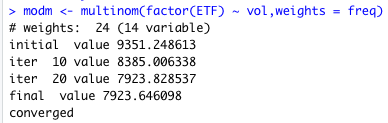
## 3 | METHODS: Linear, Multinomial logit & Proportional Odds Models

Inferring market decline probabilities via linear regression with quadratic predictors:i

However, it is clear that the residuals are far from normal. To fit multinomial models, we pick a suitable 8x8 contingency table with weekly returns.

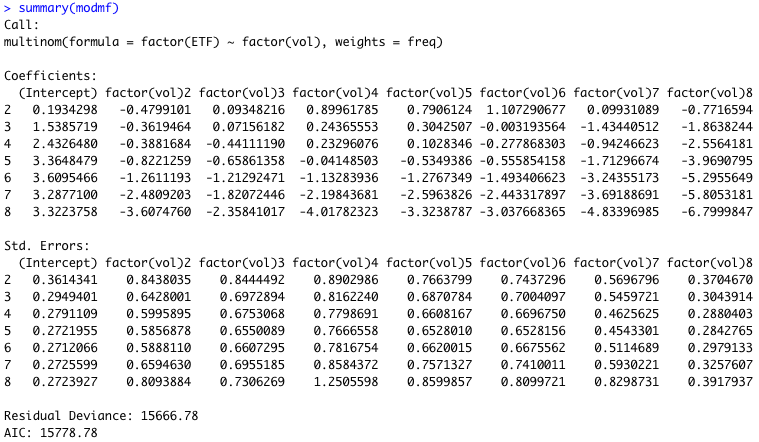


The 8x8 contingency table tells us that ETFs & VIXs move inversely on a more granular scale. E.g. there were 421 occasions when the VIX rose over 3% & the ETF had negative 3% or lower returns! On 387 occasions, opposite scenario played out. Response ETF is a categorical variable, but we have 2 choices: we can treat predictor VIX as continuous, or as a factor. Which model is preferable?

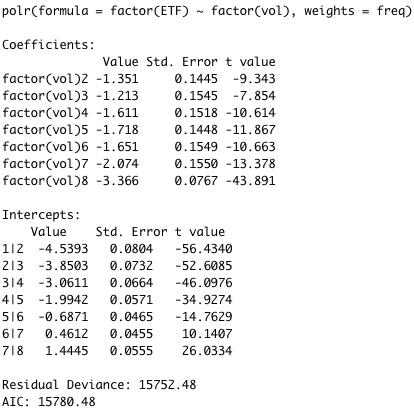




By far, the VIX as factor model is vastly better than VIX as continuous variate. We can infer probability of ETF decline as VIX rises, using multinomial model with parameters as below:



Rather than treat VIX & ETF counts as nominals, since there exists a natural ordering from -3% to 3%, a cumulative logit i.e. proportional odds model is the logical candidate:



The AICs are virtually identical, and *polr* parameters are parsimonious & simpler to interpret, as its easier to work with monotonically increasing cumulative probabilities.

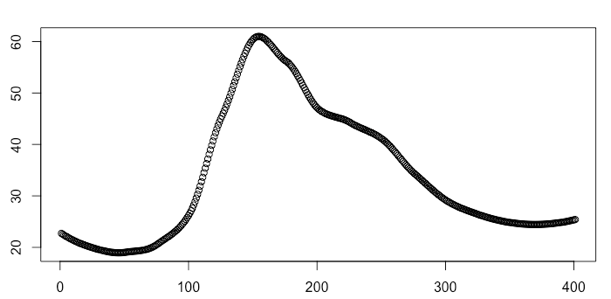
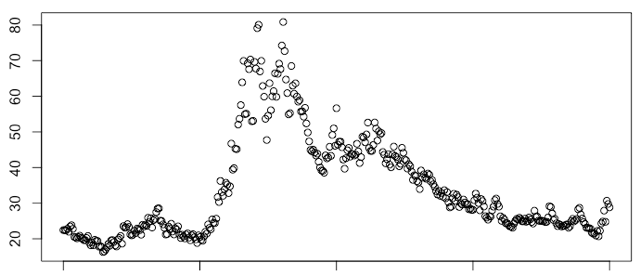
## 4 |METHODS: Estimating Market Risk Contagion with cross-correlated contingency tables

A popular misconception is to assume one is sufficiently hedged if funds are distributed among the Dow & the Nasdaq. Since DIA is primarily industrial equities, & QQQ is newfangled tech stocks, the two sets are disjoint. When tech does well, consumer staples lag. However, it is rather easy to show a so-called “regime change” that has created a clear correlation among indices. Indeed, not only are the DIA & QQQ & the SPY massively correlated, knowing the VIX for the SPY successfully predicts the DIA, something it was never intended to do! The 2 contingency tables below narrate this tale.

|  |  |  |
| --- | --- | --- |
|  | On the left, VIX, the SPY’s indicator, predicts DIA! On the right, we show a strong correlation between “disjoint” indices DIA & QQQ |  |

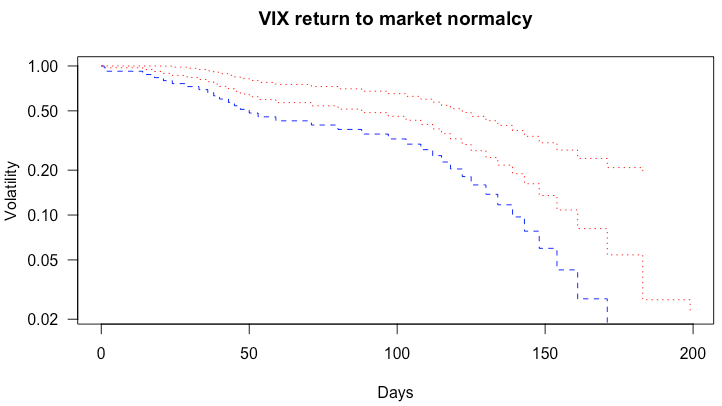
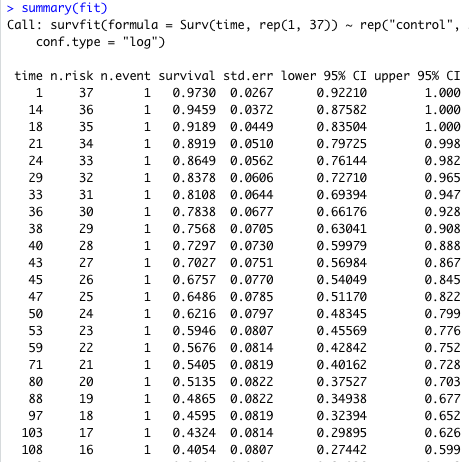
## 5 |METHODS: Estimating return to normality in a crisis via Survival Analysis, Loess.

When a recession is imminent or ongoing, the VIX spikes. These are rather sharp and prolonged, as the market experiences a dramatic & painful downturn. Often, traders have to estimate how long will the turbulence last. A popular technique is to treat the VIX as a sort of mortality curve, and as it “dies off”, one estimates how many more “patients” are at risk, until we return to a normal VIX level. We demonstrate with an example pertaining to the so-called Great Recession of 2008 triggered by the collapse in real estate markets. In the USA, it is estimated that the GDP downturn lasted from Q3 2008 to Q2 2009. However, the markets are a leading indicator, so they revert to normalcy in about 200 trading days.



We notice the VIX was treading water in the 20s, and suddenly spiked to 80. The market data is rather noisy, and we experiment with a loess to smooth out the variation.

Now that we have a monotonically decreasing function from the VIX peak, we fit a survival curve to estimate how long the recession lasts.



## 6 |RESULTS

We’ve eschewed traditionally intensive time-series techniques such as GARCH/ARIMA in favor of simpler cumulative logit models to convincingly demonstrate the negative correlation of market volatility, as exhibited by the VIX indicator, with popular market ETFs. Further, we show the risk of market contagion is real, as disjoint ETFs have gotten correlated over the years. Using a novel survival analysis approach, we estimate return to normalcy in events of market crises.

## 7 |DISCUSSION

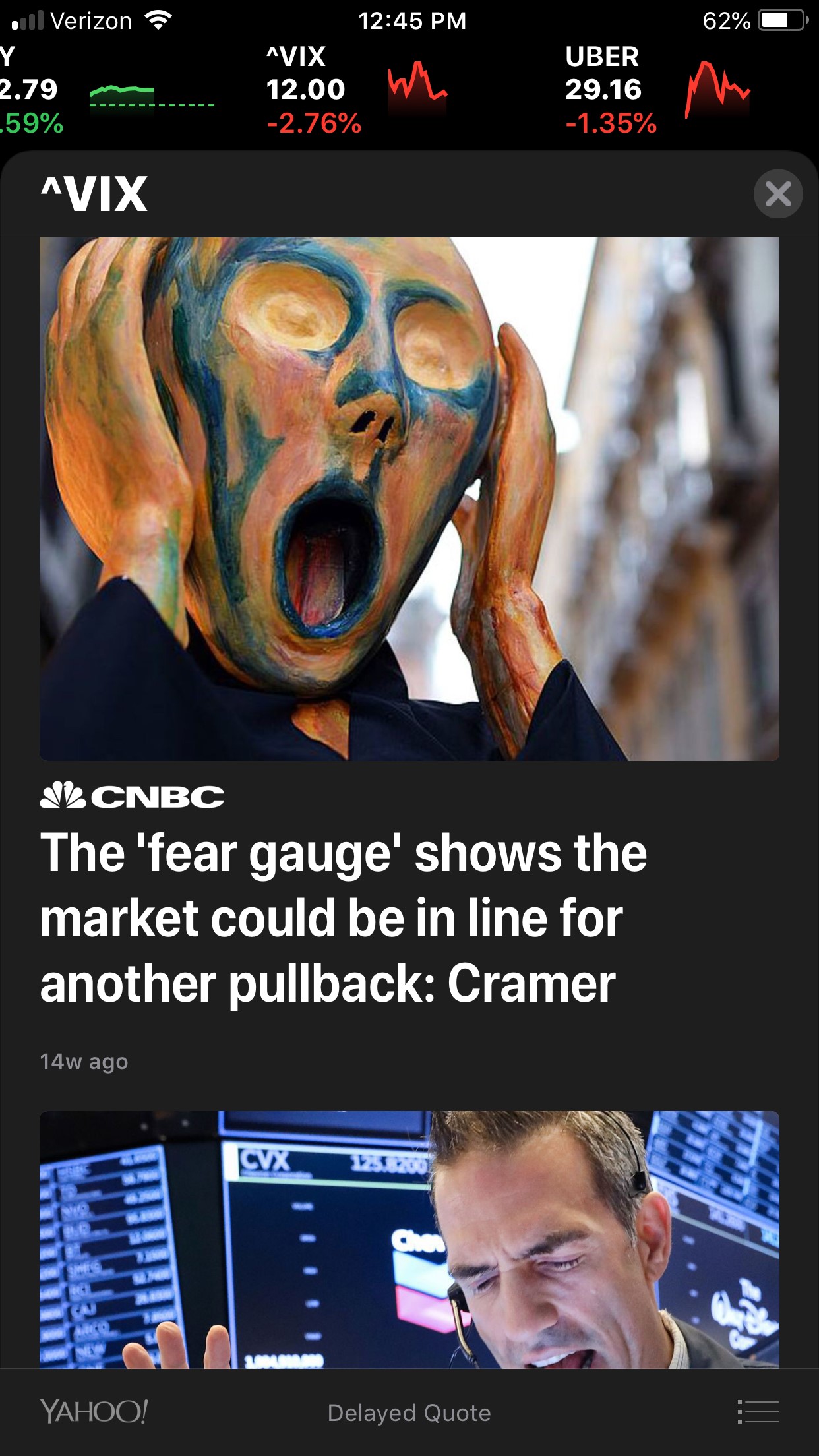
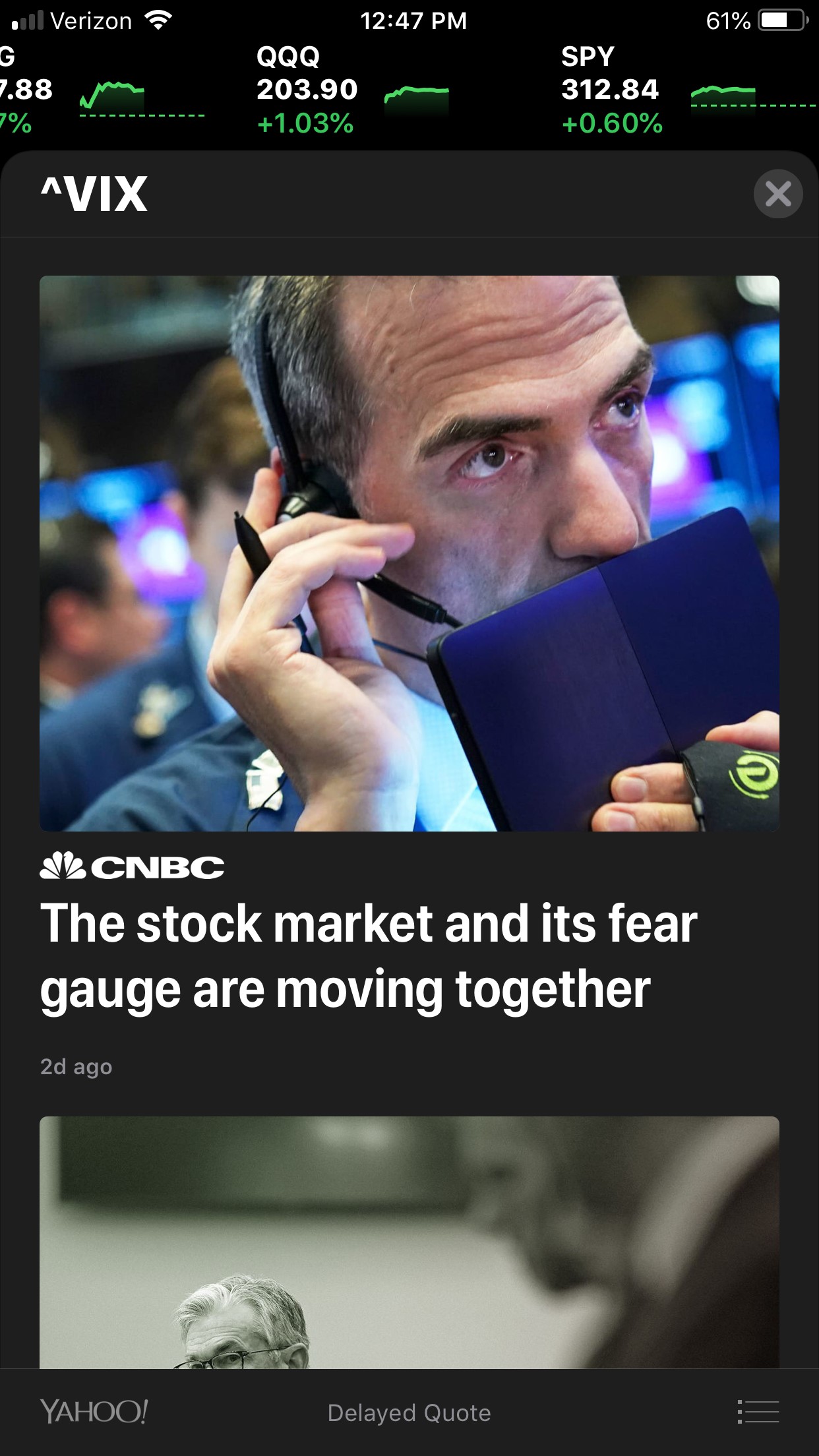
## Building a predictive VIX-ETF model turns out to be much harder than foreseen. Despite experimenting with varying time lags, we were unable to build a truly predictive model that would indicate market direction based on past week’s volatility. Concurrent models show much promise. Traders generally use GARCH to build predictive time-series models, since contingency tables are too simplistic in this scenario.

## 8 |REFERENCES

## 2001-2018 Market Data was obtained via subscription from the excellent Wharton Research Data Service. In addition to CNBC & Yahoo Finance, we consulted the following papers in our attempt to understand the strengths & deficiencies of the VIX.

* *How Good is the VIX as a Predictor of Market Risk? Clemens Kownatzki*
* *Derivatives on Market Volatility: Hedging Tools Long Overdue: Robert E Whaley*

## 9 |APPENDIX



rm(list=ls())

library(lubridate)

library(readr)

library(ggplot2)

library(nnet)

library(MASS)

# make an nx2 matrix of ETF-vol pair returns, bucketed by time Window

mkReturns<-function(filename, volIndex, underlyingIndex, timeWindow) {

mat<-data.matrix(read.csv(filename))

retmat<-matrix(NA,nrow=nrow(mat),ncol=2)

t = timeWindow

for (r in 1:(nrow(mat)-t)) {

c = volIndex

retmat[r,1] = unname((mat[r+t,c] - mat[r,c])/mat[r,c])

c = underlyingIndex

retmat[r,2] = unname((mat[r+t,c] - mat[r,c])/mat[r,c])

}

retmat<-na.omit(retmat)

colnames(retmat) <-c("VIX","ETF" )

return(retmat)

}

# make a predictive nx2 matrix of ETF-vol pair returns, bucketed by time Window

# the volatility will lag behind the ETF by past times the timeWindow

mkPredictiveReturns<-function(filename, volIndex, underlyingIndex, pastWindow, futureWindow) {

mat<-data.matrix(read.csv(filename))

retmat<-matrix(NA,nrow=nrow(mat),ncol=2)

t = pastWindow

t2 = pastWindow + futureWindow

for (r in 1:(nrow(mat)-t2)) {

c = volIndex

retmat[r,1] = unname((mat[r+t,c] - mat[r,c])/mat[r,c])

c = underlyingIndex

retmat[r,2] = unname((mat[r+t2,c] - mat[r+t,c])/mat[r+t,c])

}

retmat<-na.omit(retmat)

return(retmat)

}

# border = cutoff at border of the cont table, such as say 20%

# increment = stepping between columns in cont table, such as say 5%

# Makes a contingency table from -border to +border with bucket size = incrementpp

mkContTable<-function(mymatrix, border, increment) {

# equal number of rows & cols

rows = 2+ (2\*border/increment)

table<-matrix(0,nrow=rows, ncol=rows)

# walk thru return matrix, populate contingency table

for(r in 1:nrow(mymatrix)) {

e1 = mymatrix[r,1] #vix

e2 = mymatrix[r,2] #etf

i1 = ceiling(abs(e1)/increment)

i2 = ceiling(abs(e2)/increment)

if (e1 > 0) r1 = min(rows, (rows/2) + i1)

else r1 = max(1, (rows/2) - i1 + 1)

if (e2 > 0) c1 = min(rows, (rows/2) + i2)

else c1 = max(1, (rows/2) - i2 + 1)

table[r1,c1] = table[r1,c1] + 1

#print(paste(e1,",",e2,",",i1,",",i2,",",r1,",",c1))

}

return(table)

}

# Collapse a large contingency table into a 2x2 table

collapseTable<- function(bigtable) {

twobytwo = matrix(0,nrow=2,ncol=2)

r = nrow(bigtable)

r2 = r/2

sum = 0

for(i in 1:r2) {

for(j in 1:r2) {

sum = sum + bigtable[i,j]

#print(paste(i, ",", j,",", bigtable[i,j]))

}

}

twobytwo[1,1] = sum

sum = 0

for(i in 1:r2) {

for(j in (1+r2):r) {

sum = sum + bigtable[i,j]

}

}

twobytwo[1,2] = sum

sum = 0

for(i in (1+r2):r) {

for(j in 1:r2) {

sum = sum + bigtable[i,j]

}

}

twobytwo[2,1] = sum

twobytwo[2,2] = sum(bigtable) - sum(twobytwo)

colnames(twobytwo) <- c("ETF -ve", "ETF +ve")

rownames(twobytwo) <- c("VIX -ve", "VIX +ve")

return(twobytwo)

}

testTable<-function() {

mymat = matrix(c(-11,-11,-9,-9,9,9,11,11,9,-3), 5,2,byrow=TRUE)

table = mkContTable(mymat,10,10)

print(table)

}

# PLEASE CHANGE THIS PATH TO WHEREVER THE FILE IS LOCATED ON YOUR MACHINE

myfile = "~/Desktop/526/groud.csv"

# read file & convert to ETF-vol pair returns over 7-day & 30-day time windows

SPYWeekly = mkReturns(myfile, 2,6,7)

SPYMonthly = mkReturns(myfile, 2,6,30)

DIAWeekly = mkReturns(myfile, 4,9,7)

DIAMonthly = mkReturns(myfile, 4,9,30)

QQQWeekly = mkReturns(myfile, 3,12,7)

QQQMonthly = mkReturns(myfile, 3,12,30)

#mixed indices

# use vix go predict the dow

VIXDIAWeekly = mkReturns(myfile, 4,6,7)

# use vix to predict nasdaq

VIXQQQWeekly = mkReturns(myfile, 3,6,7)

#BEST Contigency table: USE THE 1% WEEKLY interval with 3% max

vixdia2x2 = collapseTable(mkContTable(VIXDIAWeekly, .03, .01))

fisher.test(vixdia2x2)

vixqqq2x2 = collapseTable(mkContTable(VIXQQQWeekly, .03, .01))

fisher.test(vixdia2x2)

#index vs index

QQQDIAWeekly = mkReturns(myfile, 3,4,7)

qqqdia2x2 = collapseTable(mkContTable(QQQDIAWeekly, .03, .01))

fisher.test(qqqdia2x2)

# Some contingency experiments -

#make contingency tables with 5% intervals, max 20%

#spy5wk = mkContTable(SPYWeekly, .20, .05)

#spy5mon = mkContTable(SPYMonthly, .20, .05)

#make contingency tables with 10% intervals, max 20%

#spy10wk = mkContTable(SPYWeekly, .20, .10)

#spy10mon = mkContTable(SPYMonthly, .20, .10)

#make contingency tables with 2% intervals, max 10%

#spy2wk = mkContTable(SPYWeekly, .10, .02)

#spy2mon = mkContTable(SPYMonthly, .10, .02)

#BEST Contigency table: USE THE 1% WEEKLY interval with 3% max

spyBest = mkContTable(SPYWeekly, .03, .01)

diaBest = mkContTable(DIAWeekly, .03, .01)

qqqBest = mkContTable(QQQWeekly, .03, .01)

#make 2x2 tables

spy2x2 = collapseTable(spyBest)

dia2x2 = collapseTable(diaBest)

qqq2x2 = collapseTable(qqqBest)

# CONCLUSION 1. VERY STRONG CONCURRENT SIGNAL

# Concurrent 2x2 tables show that

# if VIX decreases during a time window,

# SPY increases during SAME time window

# Conversely, if VIX increases during a time window,

# SPY decreases during SAME time window

# Very strong signal ( 5x ) in both cases

# But can we look into the future ? Make predictive returns ?

# Try to predict tomorrow index based on today's vix

SPYPred1 = mkPredictiveReturns(myfile, 2,6,1,1)

spyPred1Table = collapseTable(mkContTable(SPYPred1, .03, .01))

# Try to predict tomorrow index based on past week vix

SPYPred2 = mkPredictiveReturns(myfile, 2,6,7,1)

spyPred2Table = collapseTable(mkContTable(SPYPred2, .03, .01))

# Try to predict tomorrow index based on past two weeks vix

SPYPred3 = mkPredictiveReturns(myfile, 2,6,14,1)

spyPred3Table = collapseTable(mkContTable(SPYPred3, .03, .01))

# Try to predict weekly index return, based on past month vix

SPYPred4 = mkPredictiveReturns(myfile, 2,6,28,7)

spyPred4Table = collapseTable(mkContTable(SPYPred4, .03, .01))

# Analysis of SPY

# 1. Observe data

df <- read\_csv(myfile)

df$Date <- strptime(df$Date,'%m/%d/%Y')

df$weekday <- weekdays(df$Date)

df$month <- month(df$Date)

df$year <- year(df$Date)

SPYdf <- df[,c(1,2,6,14,15,16)]

names(SPYdf) <- c("time","vol","ETF","weekday","month","year")

ggplot(SPYdf, aes(x=as.POSIXct(time))) + geom\_line(aes(y=vol,col="vol")) + geom\_line(aes(y=ETF,col="ETF")) + xlab("Time") + ylab("")

# We can see an obvious negative correlation

# 1.2 Use spy2x2 table to test independence

# Use Fisher's exactly test

fisher.test(spy2x2)

# Use Pearson's Chi-squared test

chisq.test(spy2x2)

# CONCLUSION: we conclude that ETF and vol are not independent (both reject H0)

# 2. Try a linear model

SPYMdf <- data.frame(vol = as.numeric(SPYMonthly[,1]),

ETF = as.numeric(SPYMonthly[,2]))

lmod <- lm(ETF ~ vol,SPYMdf)

summary(lmod)

SPYMdf$vol2 <- SPYMdf$vol^2

lmod2 <- lm(ETF ~ vol + vol2 ,SPYMdf)

summary(lmod2)

anova(lmod,lmod2)

plot(lmod2)

# from the plot, we can find the normality assumption is not true

# 3. Try ordinal multinomial response, using spyBest table (8x8) to fit data

freq <- c(spyBest)

vol <- rep(c(1:8),8)

ETF <- rep(c(1:8),rep(8,8))

# 2.1 Try multinomial model and test whether should we use vol as continuous data

modm <- multinom(factor(ETF) ~ vol,weights = freq)

summary(modm)

c(modm$dev,modm$edf)

modmf <- multinom(factor(ETF) ~ factor(vol),weights = freq)

summary(modmf)

c(modmf$dev,modmf$edf)

pchisq(modm$dev-modmf$dev,modmf$edf-modm$edf,lower=F)

# CONCLUSION: we should not treat vol as continuous

# 2.2 Try proportional model and test whether it is better than multinomial model

modp <- polr(factor(ETF) ~ vol,weights = freq)

summary(modp)

c(modp$dev,modp$edf)

modpf <- polr(factor(ETF) ~ factor(vol),weights = freq)

summary(modpf)

c(modpf$dev,modpf$edf)

pchisq(modp$dev-modpf$dev,modpf$edf-modp$edf,lower=F)

# similar as before(2.1), should not treat vol as continous

pchisq(modpf$dev-modmf$dev,modmf$edf-modpf$edf,lower=F)

# CONCLUSION: multinomial model fits better than proportional model

# 2.3 Try monthly data and test whether it is better weekly data

spyBest2 = mkContTable(SPYMonthly, .03, .01)

freq2 <- c(spyBest2)

vol2 <- rep(c(1:8),8)

ETF2 <- rep(c(1:8),rep(8,8))

modmf2 <- multinom(factor(ETF2) ~ factor(vol2),weights = freq2)

summary(modmf2)

c(modmf2$dev,modmf2$edf)

modmf$dev-modmf2$dev

# using monthly data, we can get a lower deviance, so monthly data is better

# test the conclusion of (2.1) & (2.2) for monthly data

modm2 <- multinom(factor(ETF2) ~ vol2,weights = freq2)

summary(modm2)

c(modm2$dev,modm2$edf)

modm$dev-modm2$dev

pchisq(modm2$dev-modmf2$dev,modmf2$edf-modm2$edf,lower=F)

modp2 <- polr(factor(ETF2) ~ vol2,weights = freq2)

summary(modp2)

c(modp2$dev,modp2$edf)

modp$dev-modp2$dev

modpf2 <- polr(factor(ETF2) ~ factor(vol2),weights = freq2)

summary(modpf2)

c(modpf2$dev,modpf2$edf)

modpf$dev-modpf2$dev

pchisq(modp2$dev-modpf2$dev,modpf2$edf-modp2$edf,lower=F)

pchisq(modpf2$dev-modmf2$dev,modmf2$edf-modpf2$edf,lower=F)

# survival analysis

x=1800:2200

y=SPYdf$vol[1800:2200]

plot(y~x)

mod1 = loess(y~x,degree=2,span=0.4)

plot(mod1$fitted)

mod1$fitted

ceiling(mod1$fitted)

time = match(seq(61,25,-1),ceiling(mod1$fitted[149:400]))

fit = survfit(Surv(time, rep(1,37)) ~ rep("control", 37), conf.type="log")

summary(fit)

plot(fit, conf.int=TRUE, lty=3:2, col=c("red","blue"), log=T,

+ las=1, xlab="Days", ylab="Volatility", main="VIX return to market normalcy",

+ mark.tim