STAT 133 FINAL PROJECT

GROUP NAME: EAT AND PLAY TEAM

An Analysis of the Chinese Stocks Listed in the US Exchanges

Authors:

Yi LI (SID: 25132875)
Zhang QIN (SID: 25190195)
Zhaoqin YE (SID: 25940732)

May 2016

1 Introduction

As the second largest economy in the world, China has been attracting much attention globally. In recent years, a lot of Chinese companies are seeking capital from the larger investor pool in the global markets. The liquidity, transparency and lack of government manipulation in foreign exchanges are very attractive for Chinese companies. Successfully launching their IPOs in foreign exchanges, the companies are occupying headlines of both foreign and domestic media. However unfortunately, there are few studies devoted to this unique group of companies. Thus, we started this project to serve as a basic analysis of the stocks of Chinese companies listed in the US exchanges.

Firstly, to get some common traits of the companies and general trends, we analyzed the general information of the companies, such as industry, IPO year and market cap. We focused mainly on answering the following questions:

- What is the industry decomposition of these stocks, and what's the most popular area of all?
- How are these stocks distributed among different exchanges and are there any general trends in industry, exchanges and IPO year?

After these analysis, the technology sector stands out to us because it takes a large part of all the companies, both in terms of number of companies and market capitalization. Therefore, we focused on this sector as a representative for our a performance analysis. By examining the historical prices and comparing them to the performance of both Chinese domestic and US markets, we focused our analysis on the following questions:

- How are these stocks performing compared to the industry average in the same industry in the US exchanges?
- How are these stocks performing compared to similar stocks listed domestically in Chinese exchanges and the domestic industry average?
- With which market- US or domestic- are the companies more closely correlated to?

Note that in this project, Chinese companies listed in the US are limited to companies registered and located in mainland China but have their stocks listed in one of the exchanges in the US; Chinese domestic exchanges refer to Shanghai Stock Exchange and Shenzhen Stock Exchange, and we focus only on the US exchanges with Chinese companies listed, namely NYSE, NASDAQ and AMEX.

2 Data Collection and Processing

We obtained data by filtering and downloading the data into csv files from the websites and then did further processing and cleaning up using R script.We implemented methods learned in class such as using XML library to scrape data table in HTML page and processing and mutating date variables in R. We also solved some new problems by exploring features of R by ourselves. Some examples are data scraping for non-html tables using Scraper plugin in Chrome, dealing with non-ASCII encoded data table information by changing encoding to utf-8 and cleaning HTML page for link tag information as data input using XPathApply.

As the main subject of our study, the list of Chinese companies listed in the US is obtained from the NASDAQ website. We also obtained lists of Chinese stocks listed domestically, US Tech sector ETFs, China US Exchange ETF, and China tech sector index because we want to compare their performance with the US listed Chinese companies. Then we used functions to iterate through the list and downloaded data of the historical prices of the stocks.

The data sources:

- List of Chinese Stocks listed in the US:

- – http://www.nasdaq.com/screening/companies-by-region.aspx?region=Asia&country=China
- Lists of Chinese stocks listed domestically, US Tech sector ETFs, China US Exchange ETF, and China tech sector index:
  - – http://quote.eastmoney.com/
  - – http://etfdb.com/etfdb-category/technology-equities/#overview&page=1
  - – http://www.etf.com/sections/features/22712-3-etfs-for-hot-chinese-tech-sector.html?nopaging=1
  - – http://finance.yahoo.com/q?s=000915.SS&ql=0
- individual stock historical price data, both domestically listed and US listed:
  - – http://finance.yahoo.com/q/hp?s=SYMBOL

We then selected the variables necessary for our analysis. For the stock name lists data table, we primarily used stock name and symbol, the country, exchange, IPO year, market capitalization and subsector. For individual stock's historical price, we used its everyday open and close prices and dates to analyze the time series.

For some of the data we need to combine several csv documents into one data frame to see the complete time series or calculate the average. We did so by functions iterating through a list of documents and combining them the way we wanted. The detailed functions and methods are included in the Appendix section.

3 Analysis and Data Visualization

3.1 Analysis of General Information of Chinese Companies in US Stock Market

What kind of companies want their stocks listed in the US? Are there any patterns or trends? We looked at the sector, exchange distribution, market cap, and IPO years of the companies To summarize the patterns.

3.1.1 Data Preparation

In this section, we get the list of Chinese Company US Stocks `ChinaStockInUS.csv` from NASDAQ official website listing at (http://www.nasdaq.com/screening/companies-by-region.aspx?region=Asia&country=China). The website also contains the listing of stocks in major US exchange markets `NASDAQ.csv`, `NYSE.csv`, `AMEX.csv`, downloaded from http://www.nasdaq.com/screening/company-list.aspx

3.1.2 Industry decomposition

As we can see from Figure 1, the Technology sector dominates all other sectors by the number of companies. Among all the 144 companies, more than 40 are in the field of technology. Finance and consumer services follow by 22 companies and 17 companies respectively. Companies in those fields want their stocks listed in the US the most, probably to take advantage of the favorable environment in the sectors in the US. This is especially true for the tech companies, which are more growth oriented and have higher investor acceptence in the US exchanges.

However, when we take the market capitalization into account, the miscellaneous sector and energy sector surpassed the technology by taking 29.02% and 27.3% of the total market cap respectively, as is shown in Figure 2. This is mainly due to the few large cap companies in the sectors, such as Alibaba, PetroChina and China Petroleum & Chemical. The technology sector closely followed, with some leading Chinese tech companies like Baidu and Qihoo 360 listed.

The Technology sector ranked at the top 3 both in terms of number of companies and market capitalization. Therefore, we decided to focus our performance analysis on this sector later in this project.
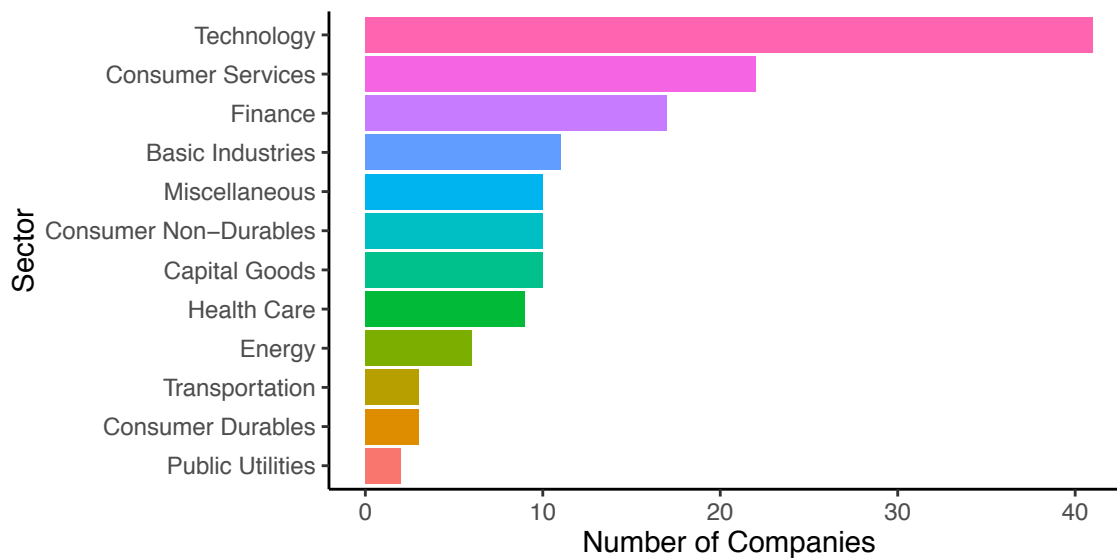
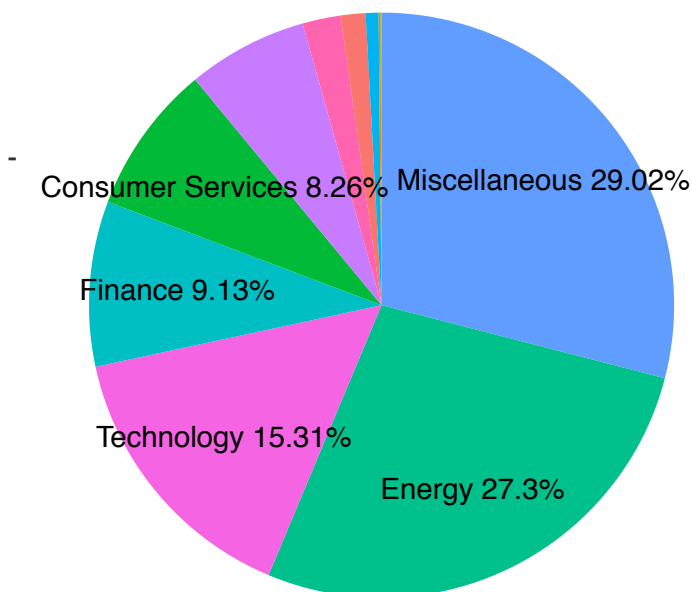

Figure 1: Number of Companies in Each Sector

Figure 2: Percentage of Market Cap of Each Sector

3.1.3 Exchange Distribution

Exchange	Count	Percentage
NASDAQ	91	63.19%
NYSE	49	34.03%
AMEX	4	2.78%

From this table we can see that about two-thirds of all the companies are listed in the NASDAQ exchange while about one-third of them are in NYSE. There are only 4 companies listed in AMEX. Companies trading on the NYSE are often larger, bottom-line focused companies that pay dividends and grow their top lines slowly but steadily, while companies listed on the NASDAQ are seen as more growth-oriented, which means more volatility to go along with increased upside potential.The listing cost per share in NASDAQ is also lower. This result aligns with the fact that most of the Chinese companies listed in the US are growth-oriented tech-related companies.

3.1.4 Market Cap

The average market cap for all the companies is $ 5.47 Billion, which is not very large. Most companies are categorized as small cap(market cap <$ 2 billion) in both exchanges. There are more large cap (>$ 10 billion) and mid caps listed in NYSE, but more small caps listed in NASDAQ. Overall the companies are mostly small cap companies in NASDAQ exchange.
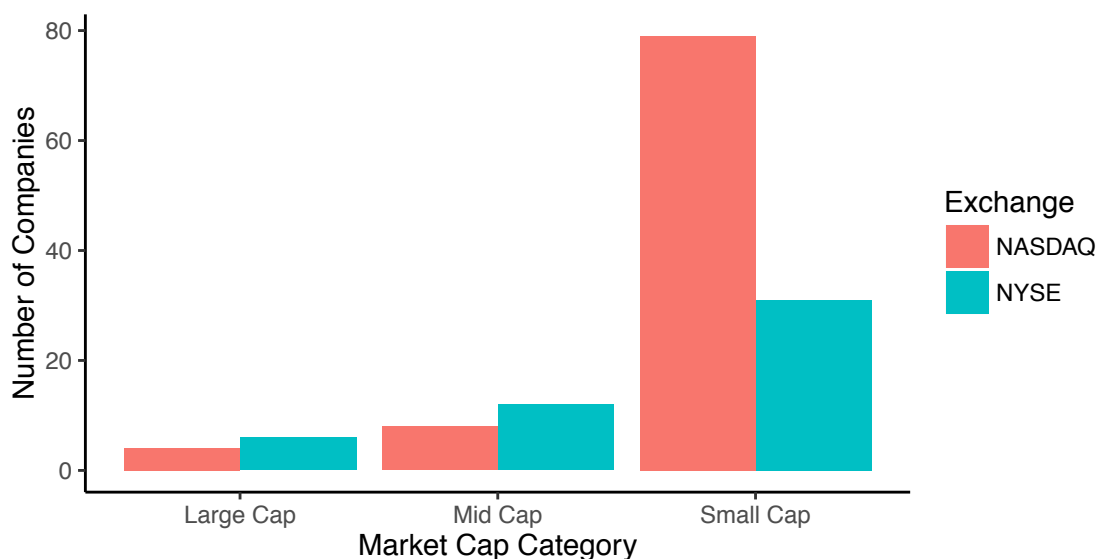

Figure 3: Market Capitaliztion Categories

3.1.5 IPO Year

From the stacked bar plot Figure 4, we can see that the peak of IPOs occured back in 2010, when the global market was recovering from the great recession. There is also an upward trend of IPOs in recent years, with more and more companies listed in the US every year in the past 4 years. Thus, it is also reasonable to expect more Chinese companies in the US exchange in the near future.

When we look at the number decomposed by the exchanges, we would find that NYSE used to dominate but in recent years more companies choose to have their stocks listed in NASDAQ. The result is consistent

with that in 3.1.3, which indicates that an increasing number of growth-oriented companies are seeking opportunities in the US.
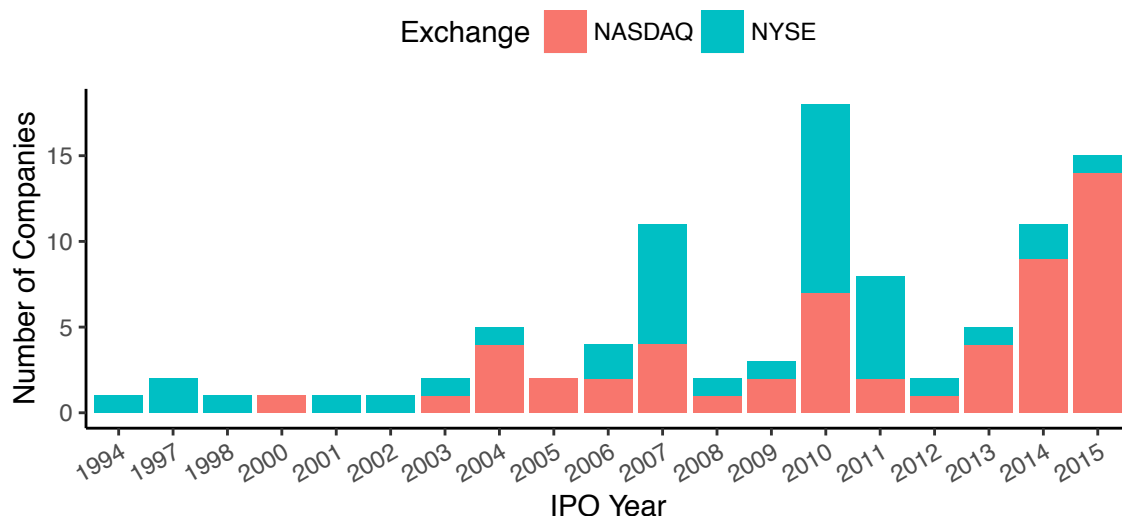

Figure 4: Number of Companies Each year

3.2 Analysis of Chinese Companies' US Stock Performance against US Stock Market

In 3.1 we found that the Tech sector is not only a major part of all the companies but are also in an upward trend. Thus, for the following analysis in 3.2, 3.3 and 3.4, we would only concentrate on the Tech sector as a representative to do the performance analysis. We would first like to find out how these stocks performed compared to the US companies in the same sector.

3.2.1 Data preparation

We downloaded the full history of Chinese technology company US stocks at Yahoo Finance. We also found lists of technology ETFs holding only China tech company US stocks at http://www.etf.com/sections/features/22712-3-etfs-for-hot-chinese-tech-sector.html?nopaging=1, and top US Tech ETFs at http://etfdb.com/etfdb-category/technology-equities/#overview&page=1, obtaining the history in the same way.

Because all history data are obtained in the same format, we can append the all stocks history of China tech Company US stock performance over 2005 to present, group by date and get their average.

3.2.2 Performance compared to US Tech Sector ETFs

We used the Tech sector ETFs as indicators of the performance of the US stocks in the Tech sector. ETFs, or Exchange Traded Funds, are a marketable security that tracks an index, a commodity, bonds, or a basket of assets like an index fund. We chose to use ETFs because first, it takes out the influence of the outliers and extremities compared to directly taking the average of all the stocks. Also, it has diversifed risk which we deem better to compare with the average of the Chinese cross-border listings that we calculated.
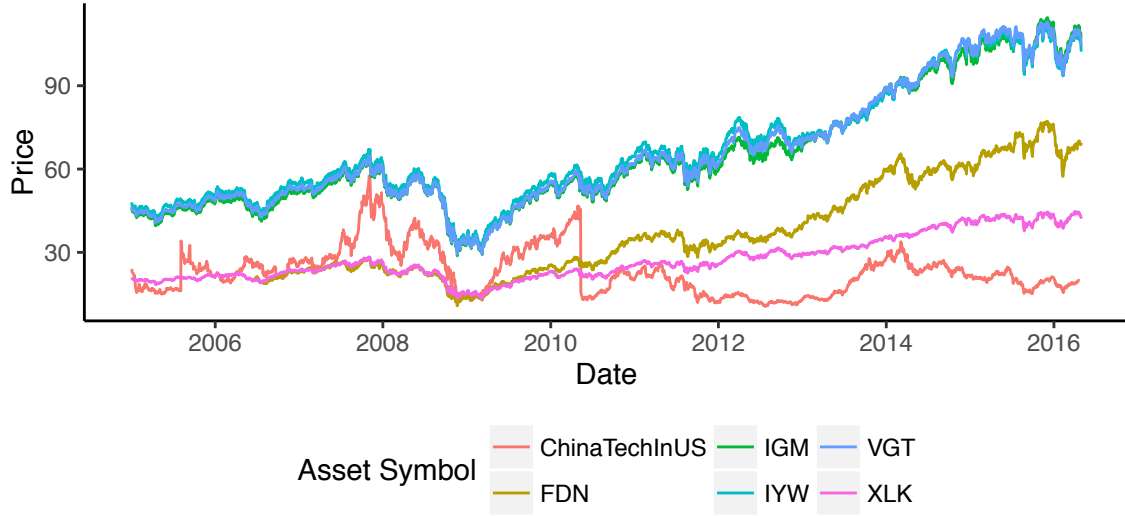
Figure 5: Performance against US top 5 Tech ETF

From Figure 5 we can find the historical price of the top 5 US tech sector ETFs and the average price of all the Chinese tech stocks listed in the US exchanges. In spite of the ups and downs, the red line, which represents the average of all the Chinese tech companies listed in the US, is almost uniformly below the other lines after 2010. This indicates that on average all US tech ETFs outperform Chinese tech companies. Some reasons for the underperformance of Chinese companies might include investor acceptance, as well as the influence of domestic market performance.

3.2.3 Performance compard to US ETFs holding only Chinese stocks listed in the US
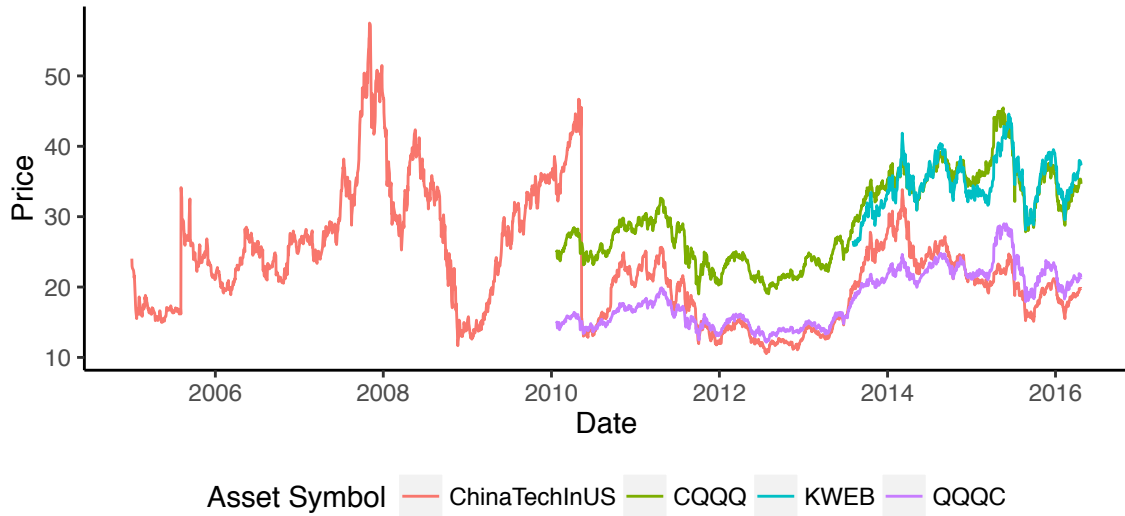

Figure 6: Performance against ETFs holding Chinese Cross Border Listed Stocks

Besides ETFs holding US tech stocks, there are ETFs traded in the US exchanges holding only the cross border listed Chinese Tech companies as well. We did a comparison between the performance of those ETFs and the average of the Chinese stocks.

Figure 6 shows the trend of the three ETFs (CQQQ,KWEB,QQQC) holding only Chinese cross-border listed tech stocks. As indicated on Figure 6, the Chinese stocks average (red line) and the ETFs follow very similar patterns in their historical prices. We can see that the three ETFs are good representation of general trend of the average performance of the Chinese companies stocks.

3.3 Analysis of Chinese Companies' US Stock Performance against Chinese Domestic Stock Market

After looking at the performance against US Tech stocks, we are also interested in learning about how the stocks performed compared to domestically listed stocks in the same sector, in the hope that we can make sense of the companies' decision to list their stock in the global market.

3.3.1 Data preparation

We found the listing of Chinese technology sector stocks in two subsectors, software and hardware, at http://quote.eastmoney.com/center/list.html#28002737_0_2. We then obtained the historical prices of these stocks in Yahoo Finance in the same way as before.

With the same idea as in the previous section, we got the average of Software subsector, hardware subsector, and tech sector as a whole.

Because these stocks are quoted and traded in Chinese yuan (RMB), we converted the price to US dollar (USD).The exchange rate doesn't have much influence on the result of our analysis since the exchange rate has been relatively stable in recent years, so we used a fixed exchange rate of 1 USD= 6.3 RMB in our calculation for simplicity.

3.3.2 Performance against Domestically Listed Tech Stocks

Figure 7 and Figure 8 are both graphs showing the trend of Chinese domestically listed stocks compared to US-listed stocks. In Figure 7 the prices of domestically listed stocks are converted to USD to compare the asset value while in Figure 8 the prices remain in RMB to take out the influence of exchange rate and see the trend more clearly.

We can see the average of Chinese Tech companies in US stock market (the purple line) is always above the other lines on Figure 7, which indicates that the cross border listed companies have absolute advantage over the domestically listed ones in total asset value, which is a factor contributing to the companies' decision to list their stock globally. However, we are not able to see the trend of the domestically listed stocks clearly because the curves are flattened by the exchange rate.
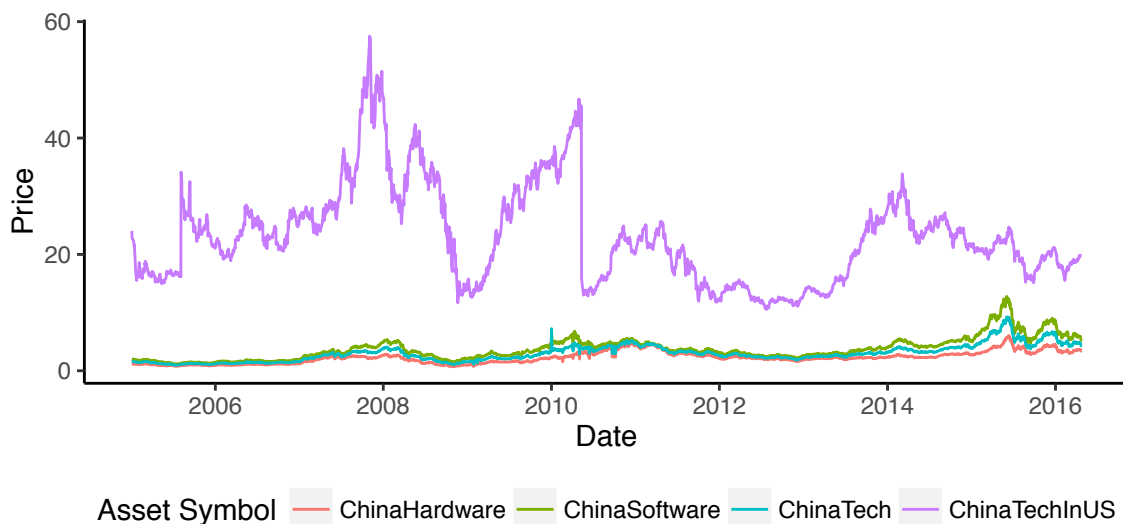
Figure 7: Performance against Domestically Listed Stocks (USD)

Thus we looked at Figure 8, which has the domestically listed stocks quoted in RMB. As we are only concerned with the trend, the absolute value of the stocks are not relevant so we can take out the exchange rate. We can see that comparing the trend actually domestically listed and US listed stocks have very similar average trends before 2015, implying the China's domestic price is a relatively great driven force for US listed Chinese company stock price. However, during 2015 domestic stock tends to have a stronger upward trend compared to the stocks listed in foreign market(the purple line) because of China's stock boom.
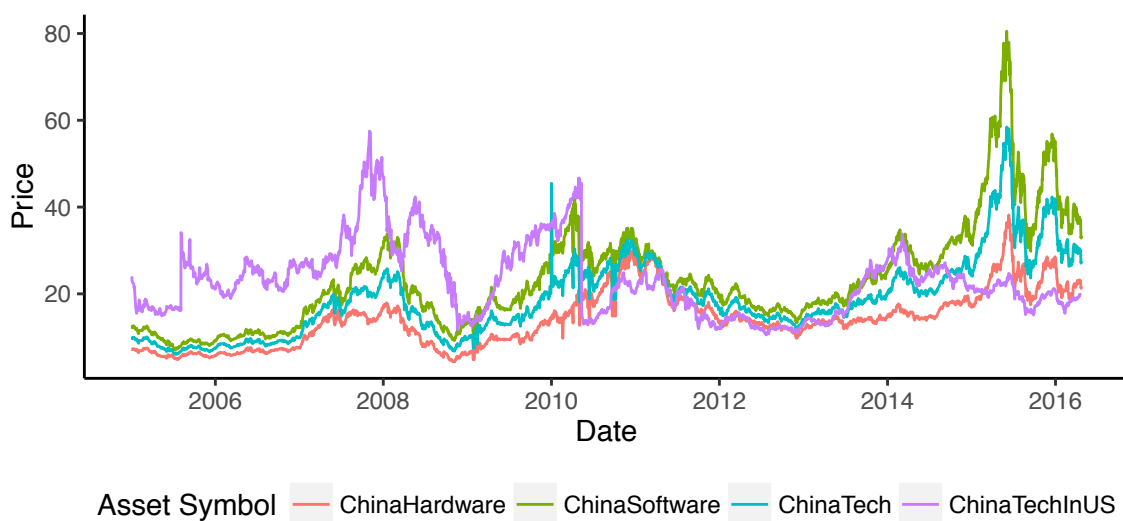
Figure 8: Performance against Domestically Listed Stocks (RMB)

3.4 Correlation Analysis

Because the cross border listed stocks are somewhat related to both the Chinese and the US markets, we would expect both markets to have influence on their performance. To see if this is true, we ran multiple linear regressions on the monthly change in the average price of the US Tech sector ETFs, the Chinese domestic Tech sector stocks, and the Chinese tech stocks listed in the US.

8

3.4.1 Data preparation

We reused the previously calculated averages, and then calculated the change in their monthly average price respectively.
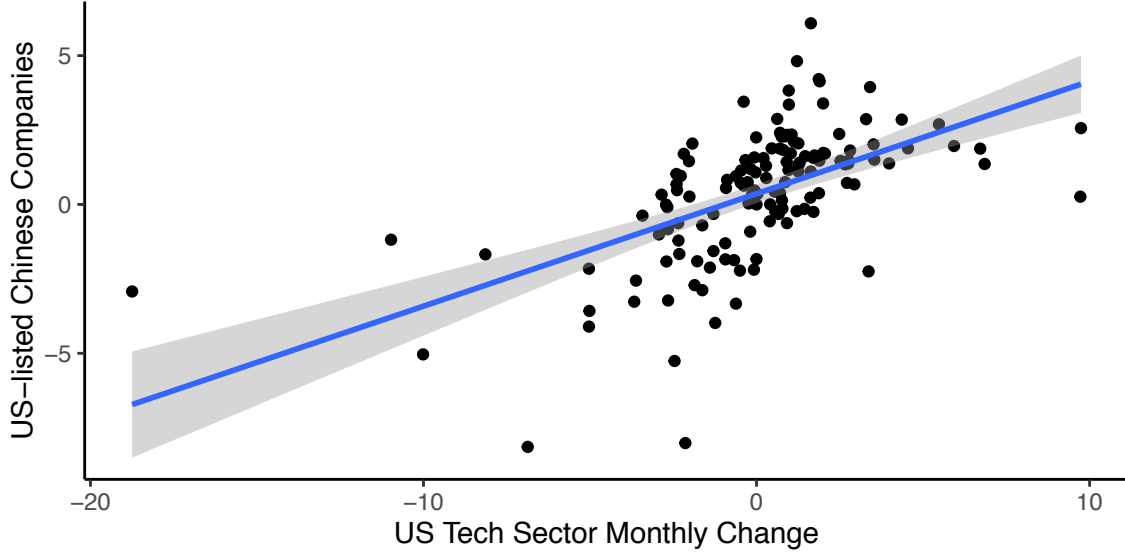
3.4.2 Linear Regressions

Figure 9: Correlation against US market

Figure 9 shows the relationship between the monthly change of the US Tech sector ETFs and the monthly change of the Chinese companies listed in the US. The changes are positively correlated, with an R^2 value of 0.32, ρ of 0.56 (relatively strongly correlated) and a p value of < 0.01 (highly significant). The result indicates that the Chinese cross border listed stocks follows the trend of the US stock markets relatively closely.

Similarly we analyzed the relationship against Chinese tech stocks, as is shown in Figure 10. The correlation is still positve, but not as strong as the correlation with that of the US markets, with an R^2 value of 0.056, ρ of 0.24 (weakly correlated) and a p value of < 0.01 (highly significant). The result shows that the Chinese cross border listed stocks have some positive relationship with the trend of the domestic stock market, but the relationship is relatively weak.
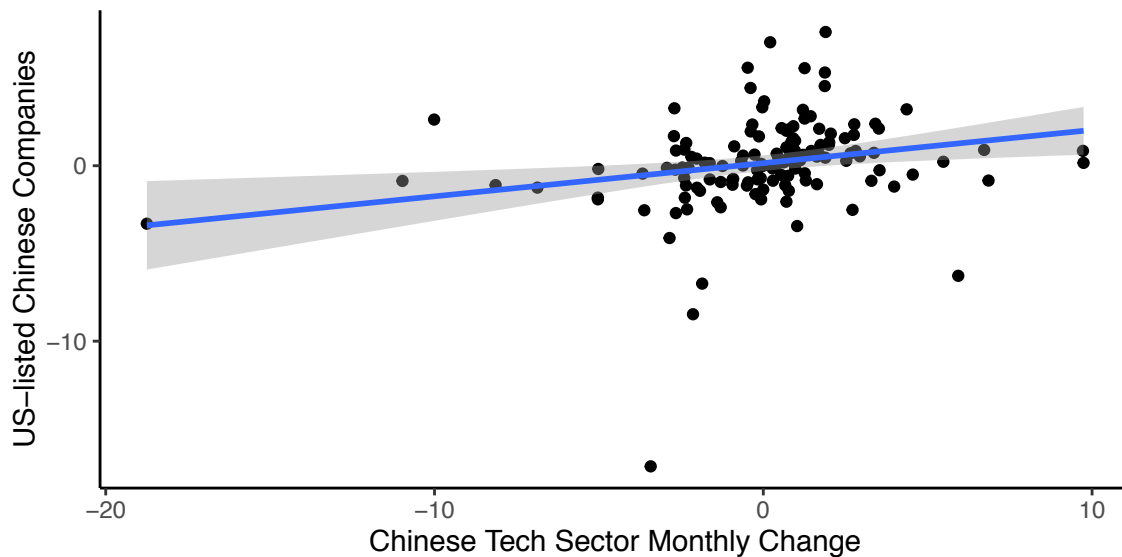
Figure 10: Correlation against Chinese Domestic market

Although they are located and registered in China, the companies are more influenced by the trend in the US stock market and changes together with the US stocks in the same sector. If we look at the multifactor regression against both the Chinese and US markets at the same time, we would find an R^2 value of 0.33, ρ of 0.58, which is a bigger value than the single factor correlations. Both the Chinese and US markets have influences on the special group of companies.

4 Conclusion

In this project, we used the company information and the historical prices from various websites to analyze the Chinese companies that have their stocks listed in the US stock exchanges.

We found that most of the companies are in the technology sector. About two- thirds of the companies are listed in NASDAQ, and most of them are small cap stocks. The results consistently picture the companies to be a growth oriented group seeking opportunity in the global markets. The number of IPOs each year is in an upward trend right now, so we can expect the group of companies to be larger and larger in the near future.

Then we compared the performance of the companies with both US tech sector stocks and the domestically listed stocks. We found that the Chinese companies tend to underperform the US companies in the same sector, probably due to lack of scale and investor acceptance. The US listed companies on average have larger asset value than domestically listed companies, but with a similar trend in performance.

We also looked at the correlation between the performance of the companies against both US and Chinese markets. The stocks are positively correlated with both markets, but the US market has a greater influence. In analyzing and predicting the performance of this special group of companies, we need to take both the domestic markets and the markets of listing in to account.

Cross-border listings play an increasingly important and valuable role for companies and investors in an ever-more-global economy-and they do promote the mobility of capital, competition between exchanges, and greater strategic flexibility for companies. We believe that this group of companies is well worth more in depth research beyond this project. For example, how are the volatilities of the stocks compared to domestic or foreign markets? What are some factors influencing the performance of the stocks? How are the stocks performing compared to stocks listed in other countries? These questions can serve as starting points of future research projects.

5 Appendix

The R code we used are listed as follows:

- Loading packages and setting data path

```
library(DataComputing)
library(printr)
library(XML)
dataPath <- "~/Desktop/proj/data/"
```

- Data Preparation

```
##################################################
# Data preparation for all analysis in project#
##################################################

library(DataComputing)
library(XML)
workPath <- "~/Desktop/proj/"
dataPath <- "~/Desktop/proj/data/"

downloadSymbols <- function(symbolList, pathID) {
  dir.create(paste(dataPath, pathID, sep=""));
  for (s in symbolList) {
    url <- paste("http://finance.yahoo.com/q/hp?s=", s, sep="");
    doc <- htmlParse(url);
    if (is.null(doc) == FALSE) {
      links <- xpathSApply(doc, "//a/@href");
      links <- links[grepl("csv", links)];
      if (length(links) > 0) {
        download.file(links[[1]], destfile=paste(
          dataPath, pathID, "/", "history_", s, ".csv",sep=""));
      }
    }
  }
}

########################################################################
# download history prices of each stock according to ChinaStockInUS#
########################################################################

# part 1.1: download all China tech company US Stocks history prices
ChinaStockInUS <- read.csv(paste(dataPath, "ChinaStockInUS.csv", sep=""))
ChinaTechInUS <- ChinaStockInUS %>%
  filter(Sector=="Technology")
downloadSymbols(ChinaTechInUS$Symbol, "ChinaTechInUS")

# part 1.2: download all China tech US Stock ETF history prices
downloadSymbols(c("CQQQ", "QQQC", "KWEB"), "ChinaTechUSETF")

# part 2: download top 10 US tech ETF history prices
```

11

```r
url <- "http://etfdb.com/etfdb-category/technology-equities/"
doc <- htmlParse(url)
tables <- readHTMLTable(doc)
Top10ETF <- head(tables[[1]]$Symbol, 10)
downloadSymbols(Top10ETF, "USTechETF")

# part 3.1: download China tech stocks history prices in informaition software subsector
dir.create(paste(dataPath, "ChinaTech", sep=""));
for (i in 1:5) {
  StockList <- read.csv(paste(dataPath, "ChinaTech", i, ".csv", sep=""))
  Symbols <- vector('character')
  for (s in StockList$??????) {
    symbol <- toString(s)
    len <- nchar(symbol)
    if (len < 6) {
      for (j in 1:(6-len)) {
        symbol <- paste(0, symbol, sep="")
      }
    }
    if (substring(symbol,1,1) == "6") {
      Symbols <- append(Symbols, paste(symbol,".SS", sep=""))
    } else {
      Symbols <- append(Symbols, paste(symbol,".SZ", sep=""))
    }
    downloadSymbols(Symbols, "ChinaTech/Software")
  }
}

# part 3.2: download China tech stocks history prices in electrical hardware subsector
for (i in 6:13) {
  StockList <- read.csv(paste(dataPath, "ChinaTech", i, ".csv", sep=""))
  Symbols <- vector('character')
  for (s in StockList$??????) {
    symbol <- toString(s)
    len <- nchar(symbol)
    if (len < 6) {
      for (j in 1:(6-len)) {
        symbol <- paste(0, symbol, sep="")
      }
    }
    if (substring(symbol,1,1) == "6") {
      Symbols <- append(Symbols, paste(symbol,".SS", sep=""))
    } else {
      Symbols <- append(Symbols, paste(symbol,".SZ", sep=""))
    }
    downloadSymbols(Symbols, "ChinaTech/Hardware")
  }
}
```

- 3.1.2

```r
ChinainUS <- read.csv(file=paste(dataPath, "ChinaStockInUS.csv", sep="")) %>%
  filter(Sector!="n/a")
```

```r
SectorCount <- ChinainUS %>%
  group_by(Sector) %>%
  summarise(Count=n()) %>%
  arrange(desc(Count)) %>%
  mutate(Percentage=paste(round(100*Count/nrow(ChinainUS), 2), "%", sep=""))
summkt <- sum(ChinainUS$MarketCap)
SectorCap <- ChinainUS %>%
  group_by(Sector) %>%
  summarise(Cap=sum(MarketCap)) %>%
  arrange(desc(Cap)) %>%
  mutate(Percentage=paste(round(100*Cap/summkt, 2), "%", sep=""))
```

```r
# Fig. 1
SectorCount %>%
  transform(Sector = reorder(Sector, Count)) %>%
  ggplot(aes(x=Sector, y=Count, fill=Sector)) +
    geom_bar(stat="identity") +
    coord_flip() +
    ylab("Number of Companies") +
    theme(
      legend.position="none",
      panel.background=element_blank(),
      axis.line=element_line())
```

```r
#Fig. 2
at <- cumsum(SectorCap$Cap) - SectorCap$Cap/ 2
ggplot(data=SectorCap, aes(x = factor(1), y=Cap, fill =Sector)) +
 geom_bar(width = 1, stat="identity")+ coord_polar(theta = "y")+ theme(
   axis.text=element_blank(),
   axis.title=element_blank(),
   panel.background=element_blank(),
   legend.position="none" )+
  annotate(geom = "text", x = 1.2, y=at[1:5],
           label = paste(SectorCap$Sector[1:5],
                         SectorCap$Percentage[1:5]))
```

- 3.1.3

```r
NYSE <- read.csv(file=paste(dataPath, "NYSE.csv", sep=""))
NASDAQ <- read.csv(file=paste(dataPath, "NASDAQ.csv", sep=""))
AMEX <- read.csv(file=paste(dataPath, "AMEX.csv", sep=""))
for (i in 1:nrow(ChinainUS)){
  if (ChinainUS$Symbol[i] %in% NYSE$Symbol){
    ChinainUS$Exchange[i]="NYSE"
  }else if(ChinainUS$Symbol[i] %in% NASDAQ$Symbol){
    ChinainUS$Exchange[i]="NASDAQ"
  }else if(ChinainUS$Symbol[i] %in% AMEX$Symbol){
    ChinainUS$Exchange[i]="AMEX"
  }else{
    ChinainUS$Exchange[i]="Other"
  }
}
```

```r
ExchangeCount <- ChinainUS%>%
  group_by(Exchange)%>%
  summarise(Count=n())%>%
  arrange(desc(Count))%>%
  mutate(Percentage=paste(round(100*Count/nrow(ChinainUS), 2), "%", sep=""))
kable(ExchangeCount)
```

- 3.1.4

```r
# Calculating mean market cap
mean(ChinainUS$MarketCap)
```

```
## [1] 5555280662
```

```r
for (i in 1:nrow(ChinainUS)){
  if (ChinainUS$MarketCap[i]<2000000000){
    ChinainUS$Mkt[i]="Small Cap"
  }else if(ChinainUS$MarketCap[i]>10000000000){
    ChinainUS$Mkt[i]="Large Cap"
  }else {
    ChinainUS$Mkt[i]="Mid Cap"
  }
}
```

```r
# Fig. 3
ChinainUS%>%
  filter(Exchange!="AMEX")%>%
  ggplot(aes(Mkt))+geom_bar(aes(fill=Exchange), position="dodge")+
  theme(
    panel.background=element_blank(),
    axis.line=element_line())+
  xlab("Market Cap Category")+ylab("Number of Companies")
```

- 3.1.4

```r
# Fig. 4
ChinainUS%>%
  filter(IPOyear!="n/a")%>%
  ggplot(aes(IPOyear))+geom_bar(aes(fill=Exchange),position="stack")+
    theme(
      legend.position="top",
      axis.text.x = element_text(angle =30,hjust = 1),
      axis.line=element_line(),
      panel.background=element_blank())+
  xlab("IPO Year")+ylab("Number of Companies")
```

- 3.2

Because the process of getting the average of several stocks and plotting the time series are the same, we wrote the following functions:

```
getStockTimeSeries <- function(filePath, title) {
  stock <- read.csv(filePath) %>%
    mutate(date=as.Date(Date), price=Close) %>%
    filter(year(date) >= 2005) %>%
    select(date, price)
  return(mutate(stock, title=rep(title, nrow(stock))))
}

getStocksAverageTimeSeries <- function(folderPath, title) {
  files <- list.files(folderPath)
  stock <- read.csv(paste(folderPath, files[1], sep=""))
  for (i in 2:length(files)) {
    stock <- rbind(stock, read.csv(paste(folderPath, files[i], sep="")))
  }
  stock <- stock %>%
    mutate(date=as.Date(Date)) %>%
    group_by(date) %>%
    summarise(price=mean(Close)) %>%
    filter(year(date) >= 2005) %>%
    select(date, price)
  return(mutate(stock, title=rep(title, nrow(stock))))
}

plotSingleTimeSeries <- function(timeSeries) {
  ggplot(timeSeries, aes(x=date, y=price)) +
    geom_line(color="blue")+
    theme(
      panel.background=element_blank(),
      axis.line=element_line())
}

plotMultipleTimeSeries <- function(timeSeriesList) {
  tbl <- timeSeriesList[[1]]
  for (i in 2:length(timeSeriesList)) {
    tbl <- rbind(tbl, timeSeriesList[[i]])
  }
  ggplot(tbl, aes(x=date, y=price, color=title)) +
    geom_line()+
    theme(
      panel.background=element_blank(),
      legend.position="bottom",
      axis.line=element_line())+
    ylab("Price")+xlab("Date")+
    guides(color=guide_legend(title="Asset Symbol"))
}
```

- Time series for average price of Chinese companies listed in the US

```
ChinaTechInUS <- getStocksAverageTimeSeries(
  paste(dataPath, "ChinaTechInUS/",sep=""),"ChinaTechInUS")
```

- 3.2.2

```
timeSeriesList <- list(ChinaTechInUS)
USETFPath <- paste(dataPath, "USTechETF/", sep="")
total <- 1
for (etf in list.files(USETFPath)) {
  total <- total + 1
  newTimeSeries <- getStockTimeSeries(
    paste(USETFPath, etf, sep=""), gsub("(history_|.csv)","",etf))
  timeSeriesList[[total]] <- newTimeSeries
}

# Fig. 5
plotMultipleTimeSeries(timeSeriesList)
```

3.2.3

```
timeSeriesList <- list(ChinaTechInUS)
USETFPath <- paste(dataPath, "ChinaTechUSETF/", sep="")
total <- 1
for (etf in list.files(USETFPath)) {
  total <- total + 1
  newTimeSeries <- getStockTimeSeries(
    paste(USETFPath, etf, sep=""), gsub("(history_|.csv)","",etf))
  timeSeriesList[[total]] <- newTimeSeries
}

#Fig. 6
plotMultipleTimeSeries(timeSeriesList)
```

- 3.3

```
# Getting data for domestically listed stocks
ChinaSoftwareRMB <- getStocksAverageTimeSeries(
  paste(dataPath, "ChinaTech/Software/",sep=""),"ChinaSoftware")
ChinaHardwareRMB <- getStocksAverageTimeSeries(
  paste(dataPath, "ChinaTech/Hardware/",sep=""),"ChinaHardware")
ChinaTechRMB <- rbind(ChinaSoftwareRMB, ChinaHardwareRMB) %>%
  group_by(date) %>%
  summarise(price=mean(price))
ChinaTechRMB <- mutate(ChinaTechRMB, title=rep("ChinaTech", nrow(ChinaTechRMB)))

# Converting RMB to USD
ChinaSoftwareUSD <- mutate(ChinaSoftwareRMB, price=price/6.3)
ChinaHardwareUSD <- mutate(ChinaHardwareRMB, price=price/6.3)
ChinaTechUSD <- mutate(ChinaTechRMB, price=price/6.3)
```

- 3.3.2

```
timeSeriesList <- list(ChinaTechInUS, ChinaTechUSD, ChinaSoftwareUSD, ChinaHardwareUSD)
# Fig.7
plotMultipleTimeSeries(timeSeriesList)
```

```
timeSeriesList <- list(ChinaTechInUS, ChinaTechRMB, ChinaSoftwareRMB, ChinaHardwareRMB)
#Fig.8
plotMultipleTimeSeries(timeSeriesList)
```

- 3.4

```
# function for getting monthly change
getchange <- function(x){
  change <- 0
  for (i in 2:nrow(x)){
    change[i] <- x$price[i]-x$price[i-1]
  }
  return (change)
}

ChinaTechDomesticMonthly <- ChinaTechRMB %>%
  mutate(month=year(date)+month(date)*0.01) %>%
  group_by(month) %>%
  summarise(price=mean(price))

USETFAverage <- getStocksAverageTimeSeries(
  paste(dataPath, "USTechETF/",sep=""),"USETFAverage") %>%
  mutate(month=year(date)+month(date)*0.01) %>%
  group_by(month) %>%
  summarise(price=mean(price))

ChinaTechInUSMonthly <- ChinaTechInUS %>%
  mutate(month=year(date)+month(date)*0.01) %>%
  group_by(month) %>%
  summarise(price=mean(price))
USChange <- getchange(USETFAverage)
DomesticChange <- getchange(ChinaTechDomesticMonthly)
ChinainUSChange <- getchange(ChinaTechInUSMonthly)
ForeignVSDomestic <- data.frame(
  US=USChange, Domestic=DomesticChange, ChinainUS=ChinainUSChange)
```

```
#Fig. 9
ggplot(ForeignVSDomestic, aes(x=ChinainUS, y=US)) + geom_jitter() +
  geom_smooth(method="lm")+
  theme(
    panel.background=element_blank(),
    legend.position="bottom",
    axis.line=element_line())+
  xlab("US Tech Sector Monthly Change")+ylab("US-listed Chinese Companies")
```

```
#Fig. 10
ggplot(ForeignVSDomestic, aes(x=ChinainUS, y=Domestic)) + geom_jitter() +
  geom_smooth(method="lm")+
  theme(
    panel.background=element_blank(),
    legend.position="bottom",
    axis.line=element_line())+
  xlab("Chinese Tech Sector Monthly Change")+ylab("US-listed Chinese Companies")
```

- 3.4.2 Calculating correlation, R^2, and p-value

```
fit1 <- lm(ChinainUS~ US, data=ForeignVSDomestic)
summary(fit1)
```

```
##
## Call:
## lm(formula = ChinainUS ~ US, data = ForeignVSDomestic)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.9490  -1.1286   0.0374   1.2114   9.8073
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.3082     0.2421  -1.273    0.205
## US            0.8521     0.1068   7.976 5.93e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.788 on 134 degrees of freedom
## Multiple R-squared:  0.3219, Adjusted R-squared:  0.3169
## F-statistic: 63.62 on 1 and 134 DF,  p-value: 5.931e-13
```

```
fit2 <- lm(ChinainUS~ Domestic, data=ForeignVSDomestic)
summary(fit2)
```

```
##
## Call:
## lm(formula = ChinainUS ~ Domestic, data = ForeignVSDomestic)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.7198  -1.3760   0.2885   1.4218   9.7372
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.04475    0.28243  -0.158  0.87435
## Domestic     0.29754    0.10526   2.827  0.00542 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.289 on 134 degrees of freedom
## Multiple R-squared:  0.05627,    Adjusted R-squared:  0.04923
## F-statistic:  7.99 on 1 and 134 DF,  p-value: 0.005424
```

```
fit3 <- lm(ChinainUS~ Domestic+US, data=ForeignVSDomestic)
summary(fit3)
```

```
##
## Call:
## lm(formula = ChinainUS ~ Domestic + US, data = ForeignVSDomestic)
```

18

```
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -15.7807  -1.1937   0.1263   1.3084   9.7506
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.30879    0.24243  -1.274    0.205
## Domestic     0.07686    0.09434   0.815    0.417
## US           0.82258    0.11296   7.282 2.59e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.791 on 133 degrees of freedom
## Multiple R-squared:  0.3253, Adjusted R-squared:  0.3151
## F-statistic: 32.06 on 2 and 133 DF,  p-value: 4.33e-12
```