ECON 5520 M ECONOMIC OF FINANCIAL MARKETS

PROF ANDREI SEMENOV

TERM PROJECT

"Exploring Financial Market Efficiency and Portfolio Diversification: An Analysis of Stochastic Dominance, Normality Tests, Random Walk and CAPM"

ALI MURTAZA HUSAIN

218410118

Table of Contents

1. Introduction 3

2. Data 4

3. Stochastic Dominance 5

   3.1 Overview 5

   3.2 Empirical Evidence 5

      3.2.1 Daily Frequency 6

      3.2.2 Monthly Frequency 8

   3.3 Inference 11

      3.3.1 Frequency Effect 11

      3.3.2 Size Effect 11

4. Tests of Normality and Random Walk 12

   4.1 Overview 12

      4.1.1 Test of Normality 12

      4.1.2 Test for Random Walk 1 13

      4.1.3 Test for Random Walk 3 14

   4.2 Empirical Evidence 15

      4.2.1 Daily Frequency 15

      4.2.2 Monthly Frequency 16

   4.3 Inference 16

      Size and Frequency Effects 16

5. CAPM 17

   5.1 Overview 17

   5.2 Empirical Evidence 18

   5.3 Inference 18

References 19

Appendix: Normality & Random Walk 20

1. Introduction

The world of finance is complex, dynamic, and constantly changing. The ability to accurately measure and evaluate financial risks and returns is critical for effective portfolio management and investment decision-making. This term project seeks to provide a comprehensive analysis of several statistical techniques that are commonly used to evaluate financial data and optimize investment portfolios. Specifically, this project will cover stochastic dominance, tests of normality and random walk, the Capital Asset Pricing Model (CAPM), and the Markowitz Portfolio Selection Model.

The project will begin with an overview of each of these techniques, followed by empirical evidence to support the practical usefulness of these techniques. For the purpose of applying these techniques, we collected data on historical prices of 30 American stocks, comprising of 15 large-cap and 15 small-cap stocks. The small-cap stocks were selected based on their market capitalization, which was less than $10 billion. We used a risk-free asset as a proxy for the market portfolio, namely SPY (SPDR S&P 500). The historical price data was collected from Yahoo Finance, and was used to calculate returns for each stock and the market portfolio over a specific time period. The analysis of these returns using the aforementioned statistical techniques provides insights into the performance of individual stocks, and helps in constructing optimized portfolios that balance risk and return.

Overall, this term project aims to provide a comprehensive understanding of the key statistical techniques used in finance, their practical applications, and their implications for portfolio management and investment decision-making.

2. Data

Data Legend

Large-cap Stocks

Serial No.	Ticker	Name	Name used for Daily Freq	Name used for Monthly Freq
01	AAPL	Apple Inc.	B01_AAPL_D	B01_AAPL_M
02	IBM	International Business Machines Corporation	B02_IBM_D	B02_IBM_M
03	MAR	Marriott International, Inc.	B03_MAR_D	B03_MAR_M
04	GOOG	Alphabet Inc. (Google)	B04_GOOG_D	B04_GOOG_M
05	TSLA	Tesla, Inc.	B05_TSLA_D	B05_TSLA_M
06	KO	The Coca-Cola Company	B06_KO_D	B06_KO_M
07	NKE	NIKE, Inc.	B07_NKE_D	B07_NKE_M
08	MA	Mastercard Incorporated	B08_MA_D	B08_MA_M
09	V	Visa Inc.	B09_V_D	B09_V_M
10	WMT	Walmart Inc.	B10_WMT_D	B10_WMT_M
11	META	Meta Platforms, Inc.	B11_META_D	B11_META_M
12	AMZN	Amazon.com, Inc.	B12_AMZN_D	B12_AMZN_M
13	MSFT	Microsoft Corporation	B13_MSFT_D	B13_MSFT_M
14	INTC	Intel Corporation	B14_INTC_D	B14_INTC_M
15	PEP	PepsiCo, Inc.	B15_PEP_D	B15_PEP_M

Small-cap Stocks

Serial No.	Ticker	Name	Name used for Daily Freq	Name used for Monthly Freq.
16	RL	Ralph Lauren Corporation	S16_RL_D	S16_RL_M
17	WWE	World Wrestling Entertainment, Inc.	S17_WWE_D	S17_WWE_M
18	DLB	Dolby Laboratories, Inc.	S18_DLB_D	S18_DLB_M
19	CROX	Crocs, Inc.	S19_CROX_D	S19_CROX_M
20	HAS	Hasbro, Inc.	S20_HAS_D	S20_HAS_M
21	MAT	Mattel, Inc	S21_MAT_D	S21_MAT_M
22	PII	Polaris Inc.	S22_PII_D	S22_PII_M
23	VAC	Marriott Vacations Worldwide Corporation	S23_VAC_D	S23_VAC_M
24	COLM	Columbia Sportswear Company	S24_COLM_D	S24_COLM_M
25	WEN	The Wendy's Company	S25_WEN_D	S25_WEN_M
26	LPX	Louisiana-Pacific Corporation	S26_LPX_D	S26_LPX_M
27	HE	Hawaiian Electric Industries, Inc.	S27_HE_D	S27_HE_M
28	DISH	DISH Network Corporation	S28_DISH_D	S28_DISH_M
29	MUSA	Murphy USA Inc.	S29_MUSA_D	S29_MUSA_M
30	SLAB	Silicon Laboratories Inc.	S30_SLAB_D	S30_SLAB_M

Risk-free Asset

Serial No.	Ticker	Name	Name used for Daily Freq	Name used for Monthly Freq.
31	SPY	SPDR S&P 500 ETF Trust	SPY_D	SPY_M

3. Stochastic Dominance

3.1 Overview

Stochastic dominance is a non-parametric statistical technique used to compare the probability distributions of the returns of two or more investment opportunities. It is a tool commonly used in finance to help investors make informed decisions about investment opportunities. The technique is based on the idea that one investment opportunity is said to dominate another if it has a higher probability of providing a greater return, regardless of the level of risk.

We need this non-parametric approach when the functional form of the utility function is unknown. Had it been known we could have just compared expected utilities from investment in every stock to determine which is better.

With the realistic assumption of investor preferring more to less or in other words, assuming that the first derivative of the utility function is positive, we can check for first-order stochastic dominance. First-order stochastic dominance occurs when the cumulative distribution function (CDF) of one investment opportunity is always below the other, indicating that it has a higher probability of providing a greater return. Thus, it is enough to look at the CDFs of two individual assets together to decide which one is a better investment.

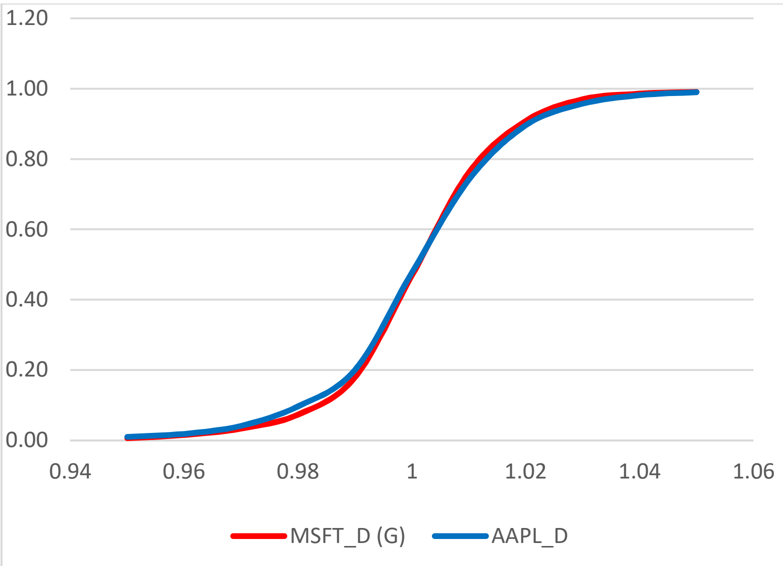
Had we have more information, for example knowing or realistically assuming that the investor is risk averse meaning the second derivative of his utility function is negative, we can also check for second-order stochastic dominance. Second-order stochastic dominance takes into account both the expected value and the degree of risk (as measured by standard deviation) of the return distributions. In order for one investment opportunity to be said to second-order stochastically dominate another, it must have both a higher expected return and a lower degree of risk than the other investment opportunity. This means that the CDF of the first investment opportunity must be to the right of the other investment opportunity at all points, except at the expected value where the CDFs cross. In other words, the probability of the first investment opportunity providing a greater return than the second is always greater than or equal to the probability of the second investment opportunity providing a greater return than the first. In simpler words, the area between the two CDFs should be more when the potential dominator is below than the other, than the part where it is above the other. In practice, Sum of differences S(t) between F(z) - G(z) should always be negative, where F and G are two CDFs for assets x and y.

3.2 Empirical Evidence

For the purpose of this exercise, we aim to apply the technique on 4 stocks: 2 large-cap and 2 small-cap. The stocks that we chose for this section are: AAPL, MSFT, WWE and RL. We will first compare each of the four stocks with the other three at the daily frequency and then at monthly frequency.

3.2.1 Daily Frequency

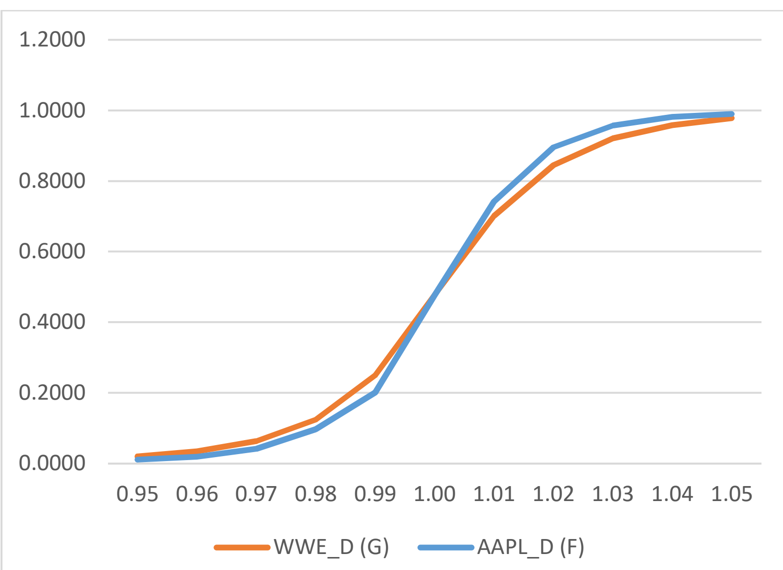
AAPL vs. MSFT



G-F	S(t)
-0.0036	
-0.0032	-0.006754
-0.0075	-0.014303
-0.0234	-0.037743
-0.0215	-0.059197
-0.0079	-0.067143
0.0203	-0.046881
0.0123	-0.034565
0.0127	-0.021851
0.0044	-0.017481
0.0004	-0.017084

SECOND ORDER STOCHASTIC DOMINANCE
MSFT_D Second Order Stochastic Dominates AAPL_D

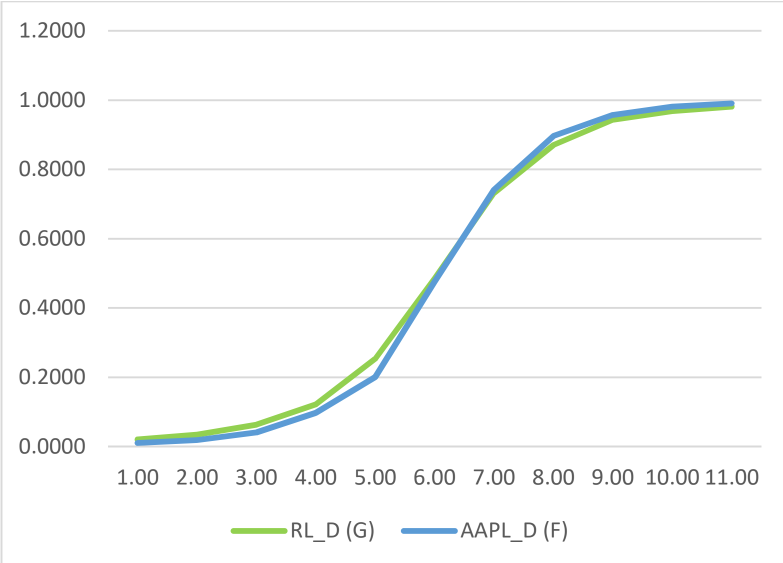
AAPL vs. WWE



F-G	S(t)
-0.0091	
-0.0151	-0.024235
-0.0215	-0.045689
-0.0266	-0.072308
-0.0489	-0.121176
-0.0012	-0.122368
0.0409	-0.081446
0.0509	-0.030592
0.0362	0.005562
0.0234	0.029003
0.0115	0.040524

NO STOCHASTIC DOMINANCE

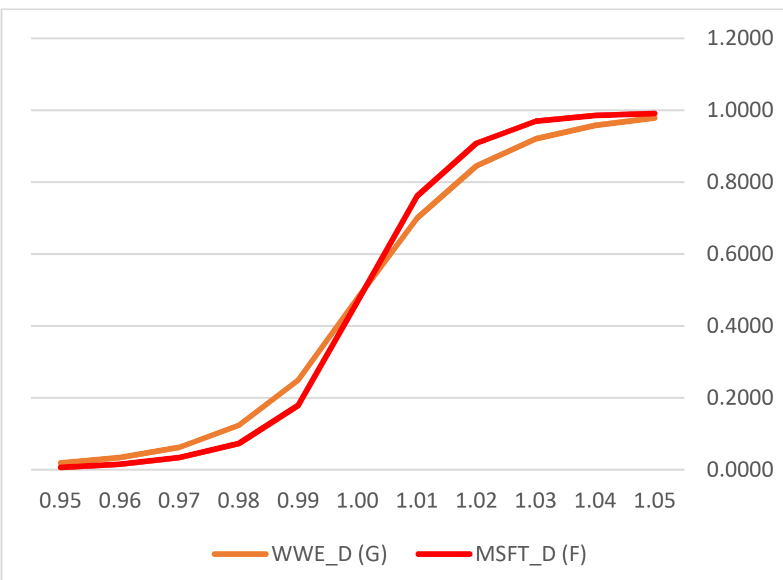
AAPL vs. RL



F-G	S(t)
-0.0103	
-0.0155	-0.025824
-0.0222	-0.048073
-0.0242	-0.072308
-0.0524	-0.124752
-0.0091	-0.133890
0.0091	-0.124752
0.0254	-0.099325
0.0135	-0.085816
0.0127	-0.073103
0.0091	-0.063965

SECOND ORDER STOCHASTIC DOMINANCE
AAPL_D Second Order Stochastic Dominates RL_D

MSFT vs. WWE



F-G	S(t)
-0.0127	
-0.0183	-0.030989
-0.0290	-0.059992
-0.0501	-0.110052
-0.0703	-0.180373
-0.0091	-0.189511
0.0612	-0.128327
0.0632	-0.065157
0.0489	-0.016289
0.0278	0.011522
0.0119	0.023441

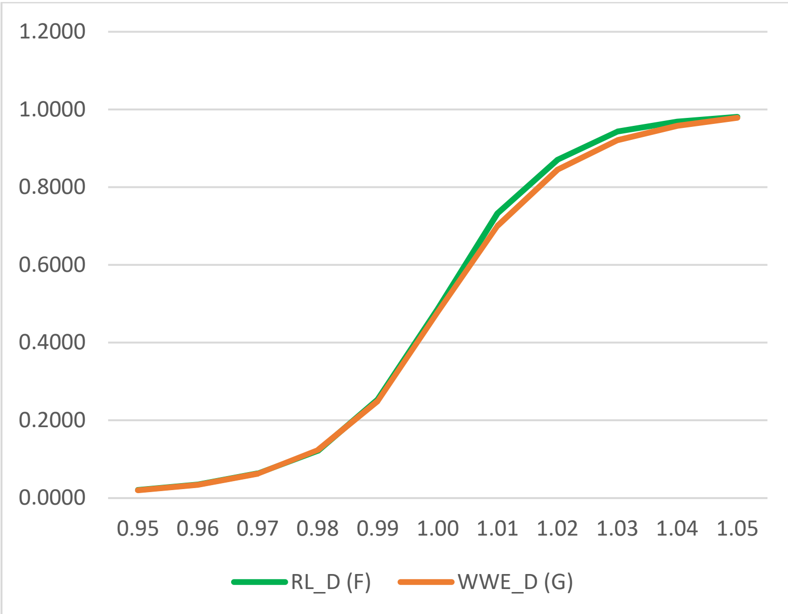
NO STOCHASTIC DOMINANCE

MSFT vs. RL

We don’t need to check here since MSFT second-order stochastic dominates AAPL and AAPL second-order stochastic dominates RL. Therefore MSFT also second-order stochastic dominates RL.

SECOND ORDER STOCHASTIC DOMINANCE
MSFT_D Second Order Stochastic Dominates RL_D

RL vs. WWE

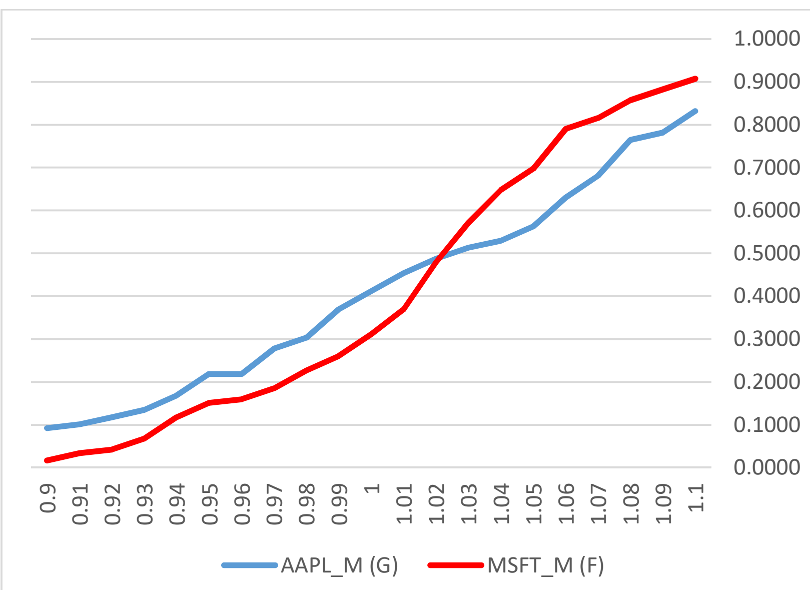


G-F	S(t)
-0.0012	
-0.0004	-0.001589
-0.0008	-0.002384
0.0024	0.000000
-0.0036	-0.003576
-0.0079	-0.011522
-0.0318	-0.043306
-0.0254	-0.068733
-0.0226	-0.091379
-0.0107	-0.102106
-0.0024	-0.104489

BOTH FIRST ORDER & SECOND ORDER STOCHASTIC DOMINANCE
WWE_D Stochastic Dominates RL_D (Both first and second order)

3.2.2 Monthly Frequency

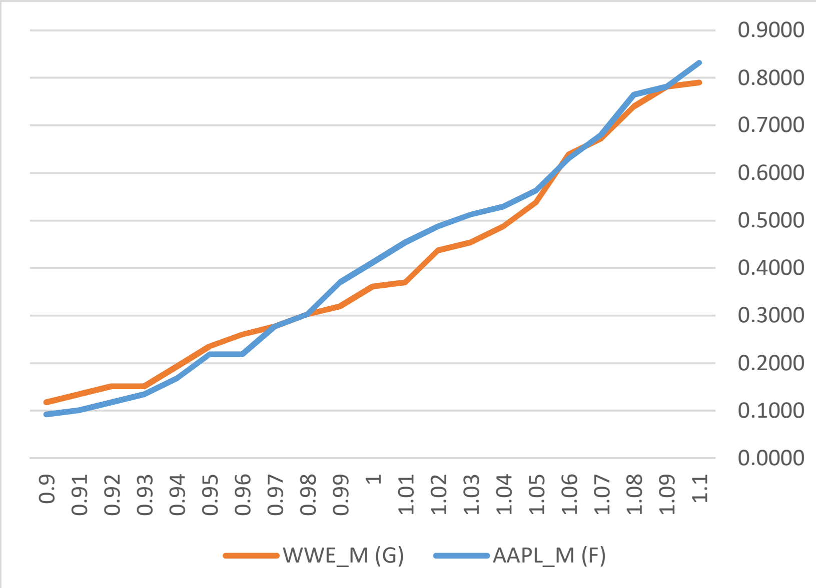
AAPL vs. MSFT



F-G	S(t)
-0.0756	
-0.0672	-0.142857
-0.0756	-0.218487
-0.0672	-0.285714
-0.0504	-0.336134
-0.0672	-0.403361
-0.0588	-0.462185
-0.0924	-0.554622
-0.0756	-0.630252
-0.1092	-0.739496
-0.1008	-0.840336
-0.0840	-0.924370
-0.0084	-0.932773
0.0588	-0.873950
0.1176	-0.756303
0.1345	-0.621849
0.1597	-0.462185
0.1345	-0.327731
0.0924	-0.235294
0.1008	-0.134454
0.0756	-0.058824

SECOND ORDER STOCHASTIC DOMINANCE
MSFT_M Second Order Stochastic Dominates AAPL_M

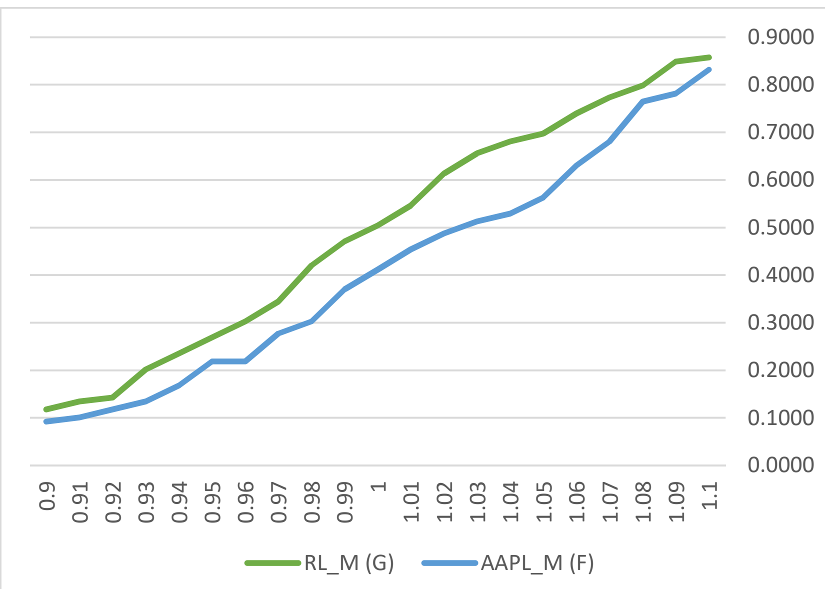
AAPL vs. WWE



F-G	S(t)
-0.0252	
-0.0336	-0.058824
-0.0336	-0.092437
-0.0168	-0.109244
-0.0252	-0.134454
-0.0168	-0.151261
-0.0420	-0.193277
0.0000	-0.193277
0.0000	-0.193277
0.0504	-0.142857
0.0504	-0.092437
0.0840	-0.008403
0.0504	0.042017
0.0588	0.100840
0.0420	0.142857
0.0252	0.168067
-0.0084	0.159664
0.0084	0.168067
0.0252	0.193277
0.0000	0.193277
0.0420	0.235294

NO STOCHASTIC DOMINANCE

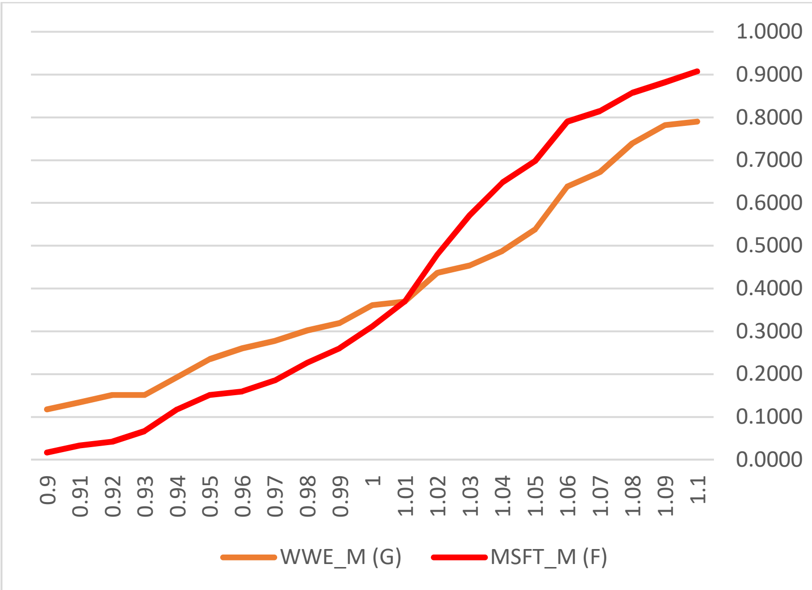
AAPL vs. RL



F-G	S(t)
-0.0252	
-0.0336	-0.058824
-0.0252	-0.084034
-0.0672	-0.151261
-0.0672	-0.218487
-0.0504	-0.268908
-0.0840	-0.352941
-0.0672	-0.420168
-0.1176	-0.537815
-0.1008	-0.638655
-0.0924	-0.731092
-0.0924	-0.823529
-0.1261	-0.949580
-0.1429	-1.092437
-0.1513	-1.243697
-0.1345	-1.378151
-0.1092	-1.487395
-0.0924	-1.579832
-0.0336	-1.613445
-0.0672	-1.680672
-0.0252	-1.705882

BOTH FIRST ORDER & SECOND ORDER STOCHASTIC DOMINANCE
AAPL_M Stochastic Dominates RL_D (Both first and second order)

MSFT vs. WWE



F-G	S(t)
-0.1008	
-0.1008	-0.201681
-0.1092	-0.310924
-0.0840	-0.394958
-0.0756	-0.470588
-0.0840	-0.554622
-0.1008	-0.655462
-0.0924	-0.747899
-0.0756	-0.823529
-0.0588	-0.882353
-0.0504	-0.932773
0.0000	-0.932773
0.0420	-0.890756
0.1176	-0.773109
0.1597	-0.613445
0.1597	-0.453782
0.1513	-0.302521
0.1429	-0.159664
0.1176	-0.042017
0.1008	0.058824
0.1176	0.176471

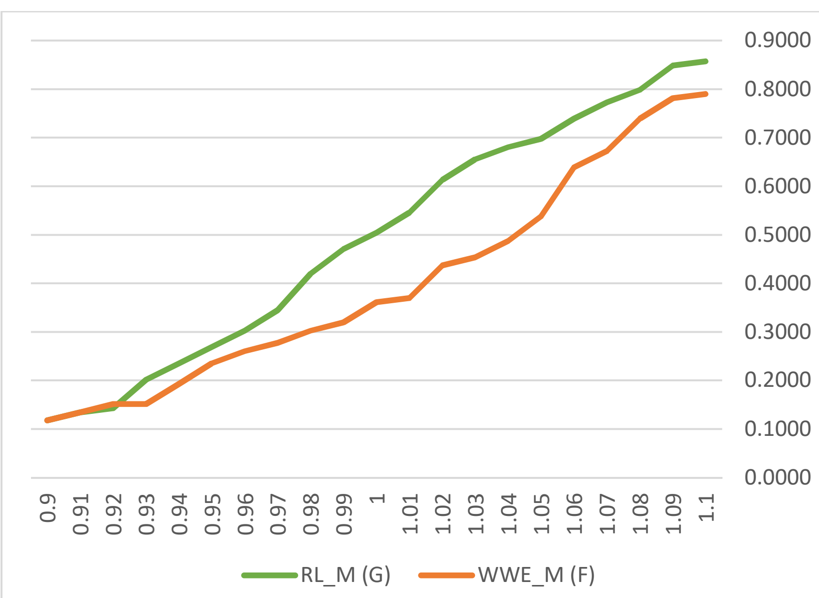
NO STOCHASTIC DOMINANCE

MSFT vs. RL

We don’t need to check here since MSFT second-order stochastic dominates AAPL and AAPL second-order stochastic dominates RL. Therefore MSFT also second-order stochastic dominates RL.

SECOND ORDER STOCHASTIC DOMINANCE
MSFT_D Second Order Stochastic Dominates RL_D

RL vs. WWE



G-F	S(t)
0.0000	
0.0000	0.000000
-0.0084	-0.008403
0.0504	0.042017
0.0420	0.084034
0.0336	0.117647
0.0420	0.159664
0.0672	0.226891
0.1176	0.344538
0.1513	0.495798
0.1429	0.638655
0.1765	0.815126
0.1765	0.991597
0.2017	1.193277
0.1933	1.386555
0.1597	1.546218
0.1008	1.647059
0.1008	1.747899
0.0588	1.806723
0.0672	1.873950
0.0672	1.941176

NO STOCHASTIC DOMINANCE

3.3 Inference

3.3.1 Frequency Effect

The frequency effect, or the impact of the frequency of data on the occurrence of stochastic dominance, is another factor that has been studied in the literature. The results of these studies have been mixed and depend on the specific data sets and methods used.

For our analysis, there doesn't seem to be any frequency effect as assets that stochastically dominate the other at daily frequency typically do so at the monthly frequency too. Small discrepancies can be seen but that can be due to the difference in number and interval of bins.

3.3.2 Size Effect

The literature on the size effect on stochastic dominance is mixed. Some studies have found that small-cap stocks have a higher probability of stochastic dominance over large-cap stocks, while others have found the opposite.

One explanation for the outperformance of small-cap stocks is that they are riskier investments, which would be consistent with second-order stochastic dominance. Small-cap stocks may have a higher degree of volatility and a wider distribution of returns, which could lead to higher expected returns and lower risk-adjusted returns.

However, other studies have found that the size effect is not statistically significant when using stochastic dominance as a method of comparison. These studies suggest that the relationship between stock size and stochastic dominance is more complex and cannot be easily generalized.

In our case, it's difficult to reach any conclusion based on the small number of sample we have from each of large-cap and small-cap stocks. However, what is evident in our sample is that twice large-cap stocks have stochastic dominated small-cap stocks while the converse hasn't happened. But again the stock dominated (RL) was dominated by the other small-cap stock (WWE) as well, so evidence is inconclusive.

4. Tests of Normality and Random Walk

4.1 Overview

4.1.1 Test of Normality

Testing the normality of returns is an important step in evaluating the suitability of various statistical models and investment strategies. One commonly used method for testing normality is the Jarque-Bera statistic.

The Jarque-Bera statistic is a goodness-of-fit test that measures the deviation of the distribution of returns from a normal distribution. It is based on the skewness and kurtosis of the data, which measure the degree of asymmetry and peakedness of the distribution, respectively.

To use the Jarque-Bera test to evaluate the normality of returns, one calculates the skewness and kurtosis of the returns data and computes the Jarque-Bera statistic using the formula:

$$JB = n\left[\frac{skewness^2}{6} + \frac{(kurtosis-3)^2}{24}\right]$$

where

$$skewness = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}}$$

$$kurtosis = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{2}}$$

If JB > $\chi^2_{(\alpha,2)}$, then the decision rejects the null hypothesis meant that data do not follow normal distribution.

If the returns data are normally distributed, the Jarque-Bera statistic should be close to zero. However, if the data are not normally distributed, the Jarque-Bera statistic will be larger, indicating a deviation from normality.

For testing the normality of returns for a portfolio of 30 stocks and one risk-free asset using the Jarque-Bera statistic, we calculated the returns for each individual stock and the risk-free asset, and then computed the Jarque-Bera statistic. The test statistics are then compared to the critical value from the Chi-square distribution χ (0.05, 2) = 5.991, as there were 2 restrictions in the null hypothesis i.e. skewness=0 and kurtosis=3.

Net returns should be used for this test but since at the daily frequency log returns are equal to net returns because of the property $\log(1+\varepsilon)=\varepsilon$ for small values of ε, we used log returns for daily frequency and net returns for monthly frequency.

4.1.2 Test for Random Walk 1

Random Walk is a theory that suggests that stock prices and other financial asset prices move randomly and are unpredictable. Testing whether a stock price series follows a random walk is important in determining the efficiency of financial markets and evaluating the effectiveness of investment strategies.

There are three types of random walk: Random Walk 1, Random Walk 2, and Random Walk 3. For the purpose of this piece in black and white we will limit ourselves to Random Walk 1 and 3. Random Walk 1, also known as the pure random walk, assumes that that returns are IID with a constant mean/drift and constant variance σ^2.

One commonly used method for testing the Random Walk 1 hypothesis is the autocorrelation test. The autocorrelation test is based on the principle that if a stock price series follows a random walk, then there should be no correlation between the returns at any given time and the returns at any other time. In other words, the returns should be uncorrelated over time.

To perform an autocorrelation test, one calculates the correlation coefficient between the returns at different time intervals. If the correlation coefficient is close to zero, then the returns are uncorrelated and the stock price series is consistent with the random walk hypothesis. However, if the correlation coefficient is significantly different from zero, then the returns are correlated and the random walk hypothesis is rejected.

The autocorrelation coefficients are tested for significance by calculating t-statistic and then comparing the statistic to the critical value of 1.96 (critical value at 0.05 significance level for large samples). If the stat is greater than the critical value the null hypothesis of insignificance is rejected.

The Portmanteau statistic test, also known as the Ljung-Box test, is another commonly used test for examining the randomness of a stock price series. This test is based on the principle that if a stock price series follows a random walk, then the autocorrelation coefficients should be small and statistically insignificant. The Portmanteau statistic test calculates a measure of the overall autocorrelation in the stock price series, and compares it to the expected autocorrelation under the random walk hypothesis. If the observed autocorrelation is significantly different from the expected autocorrelation, then the random walk hypothesis is rejected.

The Portmanteau Statistic, or Q-stat, is calculated using the following formula:

$$Q_m \equiv T \sum_{k=1}^{m} \rho^2(k).$$

The test stat is then compared to the critical value from the Chi-square distribution χ (0.05, m) where m is the number of autocorrelations we are jointly testing to be 0.

4.1.3 Test for Random Walk 3

Random Walk 3, also known as the random walk with a volatility change, relaxes the assumptions of independence and being identical. Under this assumption, the volatility of the stock price series is not constant and can be affected by external factors such as news events or changes in market conditions. Random Walk 3 is particularly useful in modeling financial crises and other periods of high volatility. It is a stronger measure of Random Walk then Random Walk since there are less assumptions and so nearer to the real world.

To test for Random Walk 3 we use the variance-ratio test statistic computed as follows:

$$VR(q) \stackrel{a}{=} 1 + \sum_{k=1}^{q-1} 2\left(1 - \frac{k}{q}\right) \rho(k)$$

The Standardized Test Statistic is as follows:

$$\psi^*(q) = \frac{\sqrt{nq}\left(\overline{VR}(q) - 1\right)}{\sqrt{\hat{\theta}}} \overset{a}{\sim} N(0,1)$$

$$\text{where } \hat{\delta}(k) = \frac{nq \sum_{j=k+1}^{nq} (r_j - \bar{r})^2 (r_{j-k} - \bar{r})^2}{\left[\sum_{j=1}^{nq} (r_j - \bar{r})^2\right]^2} \text{ and } \hat{\theta}(q) = \sum_{k=1}^{q-1} [2(1 - k/q)]^2 \hat{\delta}(k).$$

Therefore once the standardized test stat is calculated for each asset, it is compared to the critical value from the standard normal distribution at 5% significance level i.e. 1.96.

Important point to note for both Random Walks is that we use OLS to find coefficients of autocorrelation which needs the assumption of normality. While we do check for normality before testing for Random Walk, we assume returns are normally distributed irrespective of the result on the test for normality.

4.2 Empirical Evidence

While the autocorrelation coefficients from regressing each asset's log returns on its previous 10 lags, the Portmanteau test statistics and the variance ratio statistics can be found in Appendix A, summary of our findings can be found below

4.2.1 Daily Frequency

	Checking for Random Walk 1		Checking for Random Walk 3	Forecast Model
	Pearson's AutoCorrelation Coefficients	Portmanteau Statistics	Variance Ratio	Last.Lag.to.be.included
B01_AAPL_D	lag1 lag7 lag8 lag9	lag1 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
B02_IBM_D	lag4 lag5 lag6 lag7 lag8 lag9	lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
B03_MAR_D	lag1 lag3 lag6	lag1 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
B04_GOOG_D	lag1 lag6 lag7 lag8 lag9	lag1 lag2 lag6 lag7 lag8 lag9 lag10	6 7 8 9 10	lag9
B05_TSLA_D	lag7	Random Walk 1	Random Walk 3	0 (Random Walk)
B06_KO_D	lag4 lag5 lag6 lag7 lag9 lag10	lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
B07_NKE_D	lag4 lag6 lag7	lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
B08_MA_D	lag1 lag4 lag6 lag7 lag8	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	4 5 6 7 8 9 10	lag8
B09_V_D	lag1 lag4 lag6 lag7	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	1 2 3 4 5 6 7 8 9 10	lag7
B10_WMT_D	lag1 lag4 lag8 lag9	lag1 lag2 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
B11_META_D	lag8	lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
B12_AMZN_D	lag8	Random Walk 1	Random Walk 3	0 (Random Walk)
B13_MSFT_D	lag1 lag6 lag8 lag9	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	1 2 3 4 5 6 7 8 9 10	lag9
B14_INTC_D	lag1 lag2 lag8 lag9	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	1	lag9
B15_PEP_D	lag1 lag3 lag4 lag6 lag7 lag9	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
M_SPY_D	lag1 lag4 lag6 lag7 lag8 lag9	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S16_RL_D	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S17_WWE_D	lag10	Random Walk 1	Random Walk 3	0 (Random Walk)
S18_DLB_D	lag1 lag6 lag7 lag9	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S19_CROX_D	lag2 lag3	lag3 lag4 lag5 lag6	Random Walk 3	0 (Random Walk)
S20_HAS_D	lag8 lag9	lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S21_MAT_D	lag1 lag5	lag1	Random Walk 3	0 (Random Walk)
S22_PII_D	lag6 lag7 lag9	lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S23_VAC_D	lag2 lag3 lag4 lag5 lag6 lag7 lag9	lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S24_COLM_D	lag6	lag6 lag7 lag8 lag9 lag10	9 10	lag6
S25_WEN_D	lag1 lag5 lag6 lag7	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S26_LPX_D	lag6 lag7	lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S27_HE_D	lag2 lag3 lag4 lag6 lag7 lag9	lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S28_DISH_D	lag3 lag6 lag7	lag3 lag4 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S29_MUSA_D	lag1 lag3 lag4 lag5 lag7	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	Random Walk 3	0 (Random Walk)
S30_SLAB_D	lag1 lag7 lag8 lag9	lag1 lag2 lag3 lag4 lag5 lag6 lag7 lag8 lag9 lag10	1 2 3 4 5 6 7 8 9 10	lag9

4.2.2 Monthly Frequency

	Checking for Random Walk 1		Checking for Random Walk 3	Forecast Model
	Pearson's AutoCorrelation Coefficients	Portmanteau Statistics	Variance Ratio	Last Lag to be included
B01_AAPL_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
B02_IBM_M	lag1	lag1 lag2 lag3 lag4 lag5	1 2 3 4 5	lag1
B03_MAR_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
B04_GOOG_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
B05_TSLA_M	lag4	lag4 lag5	Random Walk 3	0 (Random Walk)
B06_KO_M	lag1 lag2	lag1 lag2 lag3 lag4 lag5	1 2 3 4 5	lag2
B07_NKE_M	lag2	Random Walk 1	Random Walk 3	0 (Random Walk)
B08_MA_M	lag2	lag2 lag3 lag4 lag5	2 3 4 5	lag2
B09_V_M	lag2	lag2 lag3 lag4 lag5	2 3 4 5	lag2
B10_WMT_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
B11_META_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
B12_AMZN_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
B13_MSFT_M	lag1	lag1	1	lag1
B14_INTC_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
B15_PEP_M	lag1 lag2	lag1 lag2 lag3 lag4 lag5	1 2 3 4 5	lag2
M_SPY_M	lag1	lag1 lag2 lag3	2 3 4	lag1
S16_RL_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S17_WWE_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S18_DLB_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S19_CROX_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S20_HAS_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S21_MAT_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S22_PII_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S23_VAC_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S24_COLM_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S25_WEN_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S26_LPX_M	lag1	lag1	1 2 3 4 5	lag1
S27_HE_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S28_DISH_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S29_MUSA_M	Random Walk 1	Random Walk 1	Random Walk 3	0 (Random Walk)
S30_SLAB_M	lag1	lag1 lag3	1 2 3 4 5	lag1

4.3 Inference

Size and Frequency Effects

	# Random Walk 1 Rejected		# Random Walk 3 Rejected		Total
	Daily	Monthly	Daily	Monthly	
Large-cap	13	7	5	6	31
Small-cap	13	2	2	2	19
Total	26	9	7	8	

As in literature, we reject Random Walk more often for large-cap stocks than small-cap stocks. Also, as in literature we reject Random Walk more often for Daily Frequency than Monthly Frequency.

5. CAPM

5.1 Overview

The Capital Asset Pricing Model (CAPM) is a financial model that attempts to explain the relationship between risk and return in a market. It provides a framework for understanding how the expected return of a portfolio or asset is related to the risk-free rate of return, the expected return of the market, and the potfolio/asset's sensitivity to market risk, as measured by its beta coefficient.

According to the CAPM, the expected return on a portfolio or asset can be calculated as the risk-free rate plus a premium that is proportional to the asset's beta coefficient multiplied by the difference between the expected return on the market and the risk-free rate. Mathematically, the formula is:

Expected return = Risk-free rate + Beta * (Expected market return - Risk-free rate)

The CAPM is widely used in finance and investment management to estimate the expected returns of stocks and other assets, and to calculate the cost of equity capital for businesses. It is also used as a benchmark for evaluating the performance of investment portfolios and asset managers. However, it has been subject to criticism for various assumptions and limitations, such as the assumption of efficient markets and the inability to account for all types of risk.

To empirically test the CAPM, we first estimate the following regression equation:-

$$Z_{it} = \alpha_{im} + \beta_{im} Z_{mi} + \epsilon_{it}$$

where

$Z_{it} = Excess\ return\ on\ asset\ i$

$Z_{mi} = Excess\ return\ on\ the\ market\ portfolio\ (risk - free\ asset\ serves\ as\ a\ proxy\ for\ this)$

We then test the following hypotheses:

$$H_0: \alpha = 0$$

$$H_1: \alpha \neq 0$$

To test that whether the vector of alphas is 0, we compute the Wald Test Statistic as follows:

$$J_1 = \frac{(T - N - 1)}{N} \left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}.$$

The test stat J_1, is then compared to the critical value from the F (N, T-N-1) distribution, where N is the number of assets and T is the number of observations in time series of each asset. We estimate and test the CAPM 4 times, for each size of stocks at each frequency level.

5.2 Empirical Evidence

Type	Large.Daily	Large.Monthly	Small.Daily	Small.Monthly
J1	1.31587656323146	1.1926209749143	0.839868897710858	1.16572479303767
F_cric	1.67038411336041	1.76454801841265	1.67038411336041	1.76454801841265
CAPM	Not Rejected	Not Rejected	Not Rejected	Not Rejected

5.3 Inference

Considering our small sample, we don't reject CAPM for any of the four groups of stocks. However we do get close to rejecting for Large Daily and Small Monthly which doesn't go well with the literature which says CAPM is more often rejected for Daily Frequency and Small-cap stocks but as mentioned our sample size isn't generalizable.

References

- *Yahoo! Finance*. https://ca.finance.yahoo.com/