Darien Huang
NBA 6450 Advanced Investment Strategies, Fall 2016
Johnson Graduate School of Management, Cornell University

Problem Set 1: Time Series Predictability

This problem set is due at the start of class on **Monday, September 12th** at the start of class. Include a cover sheet with the name, section number, and Net ID of all group members. Hand in only one solution per group. Discussing the solution with other groups before the problem set has been handed in is not permitted. You can use any statistical package you want to answer the questions (state clearly what you use).

Question 1: Leverage Effect

(a) Download daily closing levels of the S&P500 Index and the VIX Index, for the period 1/2/1990 to 12/31/2015. Define "returns" of each process as $r_t = \log\left(\frac{X_t}{X_{t-1}}\right)$. We are ignoring dividends for the S&P500 and ignoring the fact that the VIX Index itself is not directly tradeable and use the word "returns" here for convenience. Make a scatterplot of S&P500 returns on the Y-axis against VIX returns on the X-axis.

(b) What are the standard deviations of S&P500 returns and VIX returns, respectively? What is the correlation between daily S&P500 returns and VIX returns?

(c) Run a linear regression of S&P500 returns on VIX returns. What are the estimated coefficients? Is the slope on VIX returns statistically significant?

(d) Given your estimate of the slope in part (c) and your results in part (b), what is the relationship between linear regression slope, correlation, and the standard deviations of the two variables?

(e) Skewness is a measure of the degree of asymmetry. A normally distributed random variable has skewness zero. Negative skewness represents longer left tail, positive skewness represents longer right tail. Market wisdom suggests that large declines in the stock market are more likely than large stock market rallies, and large increases in volatility are more likely than large drops in volatility. What is the skewness of S&P500 returns? What is the skewness of VIX returns? Is the evidence in line with market wisdom?

(f) Write no more than 1 paragraph summarizing the most salient results from your analysis of the relationship between aggregate stock returns and volatility.

Question 2: Price-Dividend Ratio

(a) Please use the file *vwret.csv* for this question. The variable $vwretd_t$ is the monthly return inclusive of dividends on the value-weighted composite AMEX-NYSE-NASDAQ index maintained by CRSP, which is the Center for Research in Security Prices at the Univer-

sity of Chicago. The variable $vwretx_t$ is the monthly return on the same index, excluding dividends. We will construct prices and dividends as follows. Initialize $P_0 = 1$. Recursively update $P_t = P_{t-1} \times (1 + vwretx_t)$. $D_t = P_{t-1} \times (vwretd_t - vwretx_t)$. Define the price-dividend ratio as the ratio of price over the sum of the past 12 months of dividends. Typically we work with the log of the price-dividend ratio. Make a plot of this series over time.

(b) Go to Robert Shiller's website (`www.econ.yale.edu/~shiller/data.htm`) and download the spreadsheet "U.S. Stock Markets 1871-Present and CAPE Ratio" from his book *Irrational Exuberance*. Take the log of the CAPE series, which is the cyclically adjusted price-earnings ratio. What is the correlation between the series you constructed in (a) and the log CAPE series?

(c) What is the persistence of log price-dividend ratio and log CAPE ratio in the data? How long are their half-lives?

(d) Is it econometrically sound to use the price-dividend ratio or CAPE in regressions? Either formally test for stationarity or make arguments based on their time series plots.

Question 3: Predictive Regressions

(a) Go to Kenneth French's website, select the U.S. Research Returns Benchmarks data, and download the Fama/French 3 Factors time series at a monthly frequency. The variable *MKT-RF* is the excess return on the aggregate stock market portfolio. Construct log returns as $r_{t+1} = \log(1 + R_{t+1})$. Create overlapping returns for horizons of 3 months, 1 year, 3 years, and 5 years. Use data from 1977 to 2013.

(b) Go to the FRED (Federal Reserve Bank of St. Louis Economic Database) website (`https://fred.stlouisfed.org/`) and download the Moody's Seasoned Aaa and Baa Bond Yields over the same period as in part (a) and create the default spread. Give descriptive statistics (mean, standard deviation, persistence) of the default spread.

(c) On the same website, download the 10-Year Treasury Constant Maturity Rate and 3-Month Treasury Bill: Secondary Market Rate over the same period as in part (a) and create the term spread. Give descriptive statistics of the term spread.

(d) The file *icc.csv* contains the time series of the implied cost of capital from Li, Ng, and Swaminathan (2013). Give descriptive statistics of the ICC. Predict future stock market excess returns with univariate regressions at horizons of 1 month, 3 months, 1-year, 3-years, and 5-years ahead using each of the two predictors in Question 2, and the three predictors you've constructed in Question 3. Report the slopes and R^2 from the regressions. Optional: Report the t-statistics using Newey-West standard errors to account for the overlapping returns.

(e) Run bivariate predictive regressions similar to (d), but now use ICC and each of the other predictors. Report the slopes and R^2 from the regressions. Optional: Report the t-statistics using Newey-West standard errors to account for the overlapping returns. Which predictor(s) seem to work best?