FINM Data Analysis **Mark Hendricks**
August Review 2021

Homework #1
Basic Regressions

This homework uses the data posted on Canvas. The data file contains...

- Return rates, r_t^i, for 6 securities. The first security is an S&P 500 ETF, denoted r_t^M below.
- Return rates for one portfolio, r_t^p.
- Yields for 10-year treasuries.
- Dividend-price ratios for the S&P 500.

1 Regression

1. Estimate the regression of the portfolio return on the market return:

$$r_t^p = \alpha_t^{p,M} + \beta_t^{p,M} r_t^M + \epsilon_t^{p,M} \tag{1}$$

   Specifically, report alpha, beta, and the r-squared.

2. Estimate the regression of the portfolio return on the market return and on "CL1", the return on crude-oil futures, denoted as r_t^{oil}:

$$r_t^p = \alpha_t^{p,M,\text{oil}} + \beta_t^{p,M} r_t^M + \beta_t^{p,\text{oil}} r_t^{\text{oil}} + \epsilon_t^{p,M,\text{oil}} \tag{2}$$

   Specifically, report alpha, beta, and the r-squared.

3. Calculate the time-series of fitted regression values:

$$\hat{r}_t^p = \alpha_t^{p,M,\text{oil}} + \beta_t^{p,M} r_t^M + \beta_t^{p,\text{oil}} r_t^{\text{oil}}$$

   Your statistical package will output these fitted values for you, or you can construct them using the estimated parameters.

   What is the correlation between $\hat{r}_t^p$ and r_t^p? How does that compare to the r-squared of the regression in (2)?

4. How does $\beta_t^{p,M}$ differ in (1) and (2)? How does this relate to the correlation between r_t^M and r_t^{oil}?

5. Without doing any calculation, would you expect $\epsilon_t^{p,M,\text{oil}}$ or $\epsilon_t^{p,M}$ to have higher correlation to r_t^{oil}?

2 Decomposing and Replicating

1. Use linear regression on the portfolio to uncover the weights of the underlying assets that were used to build it: (NO INTERCEPT)

$$r_t^p = (\boldsymbol{\beta}_t^{p,\boldsymbol{r}})' \boldsymbol{r}_t + \epsilon_t^{p,\boldsymbol{r}} \tag{3}$$

where $\boldsymbol{r}$ denotes the vector of returns for the 6 individual securities.

Reminder: Many statistical packages include an intercept by default; you can use the formula for the OLS estimator if needed.

2. Suppose that we want to mimic AAPL using the other 5 securities. Run the following regression using data through the end of 2018. (NO INTERCEPT)

$$r_t^{AAPL} = (\boldsymbol{\beta}_t^{AAPL,\boldsymbol{r}})' \boldsymbol{r}_t + \epsilon_t^{AAPL,\boldsymbol{r}} \tag{4}$$

where $\boldsymbol{r}$ denotes the vector of returns for the 5 individual securities, excluding AAPL.

   (a) Report the r-squared and the estimate of the 5x1 vector, $\boldsymbol{\beta}_t^{AAPL,\boldsymbol{r}}$.

   (b) Report the t-stats of the explanatory returns. Which have absolute value greater than 2?

3. Using the estimates through 2018, calculate the fitted value of AAPL based on 2019 returns:

$$\hat{r}_{2019}^{AAPL} = \left(\hat{\boldsymbol{\beta}}_t^{AAPL,\boldsymbol{r}}\right)' \boldsymbol{r}_{2019}$$

What is the correlation between $\hat{r}_{2019}^{AAPL}$ and r_{2019}^{AAPL}? How does this compare to the $\mathcal{R}^2$ from the regression above based on data through 2018?