heteroskedasticity. Here's the formula for the simple regression. Usually we set $p = T^{1/4}$.

$$
\begin{aligned}
\text{var}\left(\hat{\beta}\right) &= \left(\frac{1}{\sum x_t^2}\right)^2 S_T \\
S_T &= S_0 + \frac{1}{T}\sum_{j=1}^{p} w(j) \sum_{t=j+1}^{T} 2\hat{\epsilon}_t\hat{\epsilon}_{t-j}\left(x_t x_{t-j}\right) \\
S_0 &= \sum \hat{\epsilon}_t^2 x_t^2 \\
w\left(j\right) &= 1 - \frac{j}{p+1}
\end{aligned}
$$

8. Don't worry. There is an option in Eviews to do this computation and automatically get Newey-West t-ratios. But here's the rule: **In time-series regression always do Newey-West.**

   Go back to lh_dy.wf1

7.3 Dividend yield as predictor of future return

A toy model as motivation. Let P be the stock price (not log price), d be the dividend (not log). $\beta = \frac{1}{1+\rho}$ is the subjective discount factor, ρ is the discount rate. Present value model,

$$
P_t = E_t \sum_{j=0}^{\infty} \beta^j d_{t+j}
$$

Let's say we expect dividends to grow at rate δ each period, so that

$$
\begin{aligned}
E_t d_{t+1} &= \left(1+\delta\right) d_t \\
E_t d_{t+j} &= \left(1+\delta\right)^j d_t
\end{aligned}
$$

and $\rho > \delta$. Then

$$\begin{aligned}
P_t = E_t \sum_{j=0}^{\infty} \beta^j d_{t+j} = \sum_{j=0}^{\infty} \left(\frac{1+\delta}{1+\rho}\right)^j d_t &= \left(\frac{1+\rho}{\rho-\delta}\right) d_t \\
E_t P_{t+1} &= \left(\frac{1+\rho}{\rho-\delta}\right)(1+\delta)\, d_t \\
E_t P_{t+2} &= \left(\frac{1+\rho}{\rho-\delta}\right)(1+\delta)^2\, d_t \\
E_t \frac{P_{t+1}}{P_t} &= \left(\frac{1+\rho}{\rho-\delta}\right)(1+\delta)\, \frac{d_t}{P_t} \\
E_t \frac{P_{t+2}}{P_t} &= \left(\frac{1+\rho}{\rho-\delta}\right)(1+\delta)^2\, \frac{d_t}{P_t}
\end{aligned}$$

Must be the case that

$$E_t \frac{P_{t+k}}{P_t} = \left(\frac{1+\rho}{\rho-\delta}\right)(1+\delta)^k\, \frac{d_t}{P_t}$$

Let $R^e_{t,t+h}$ be the future $h-$horizon excess return.

$$R^e_{t,t+h} = \alpha_h + \beta_h \left(\frac{d_t}{P_t}\right) + \epsilon_{t,t+h}$$

Think of fitted part of regression as conditional expectation. The expected excess return is a risk premium, so the regression estimates the risk premium.

In a plot of the dividend yield against the 7 year return. If you bought stocks in 1980, great! Bought stocks in 2000–not so good.

The lesson is that $E_t R^e_{t+k,t}$ varies a lot. Expected excess returns are large and **variable**. Why?

The 1970s economy was terrible. Since 2008, the economy has been terrible. In the bad state, people have low tolerance for risk.

You might think P is high because people forecast high dividend growth. Empirical says that's wrong. Prices are high relative to dividends because people expect low future returns. It's prices that move around dividends. Dividends don't move around prices.

Prices are high (dividend yield low) when people are willing to take the risk (expansion/boom). Prices are low (dividend yield high) when people are unwilling to bear risk (recession).

8 Some Necessary Matrix Algebra

Text: pp. 28-35.

8.1 Definitions

- **Scalar**: a single number

- **Matrix:** a two-dimensional array. $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{23} \end{pmatrix}$ is a (3×2) matrix–that is, 3 rows and 2 columns. We say the number of rows then columns. a_{11} is the (1,1) element of A, and is a scalar. The **subscripts** of the elements tell us which row and column they are from.

- **Vector**: a one-dimensional array. If we take the first column of A and call it $A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}$, it is a (3×1) **column vector**. If we take the second row of A and call it $A_2 = \begin{pmatrix} a_{21} & a_{22} \end{pmatrix}$, it is a (1×2) **row vector**.

- **Square matrix:** An $m \times n$ matrix is square if $m = n$.
  $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ is a square matrix. The **diagonal** of the matrix A are the elements a_{11}, a_{22}, a_{33}.

  It only makes sense to talk about the diagonal of square matrices.

- **Symmetric matrix**. For a square matrix, if the elements $a_{ij} = a_{ji}$, for $i \neq j$, then the matrix is symmetric. (notice the correspondence of the bold entries).
  $$A = \begin{pmatrix} 2 & \mathbf{3} & \mathbf{4} \\ \mathbf{3} & 10 & \mathbf{6} \\ \mathbf{4} & \mathbf{6} & 11 \end{pmatrix}$$

- **Transpose of a matrix.** The $i - th$ row becomes the $i - th$ column. The transpose of an $(m \times n)$ matrix is $(n \times \times m)$.

  $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix}$, then $A' = A^T = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix}$.

  $A = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix}$, then $A' = A^T = \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix}$.

$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$, then $A' = A^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \end{pmatrix}$

- **Zero matrix:** All the entries are 0. $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ is a zero matrix.
- **Identity matrix:** A square matrix with 1s on the diagonal elements and 0 on the off-diagonal elements is called the identity matrix. $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is a (3×1) identity matrix. We always call an identity matrix I.

8.2 Matrix Operations

- To add two matrices or to subtract one from the other, they must have the same dimensions. We do element by element addition or subtraction.
  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$,
- **Addition**
  $$C = A + B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \underbrace{\begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}}_{C}$$
- **Subtraction**
  $$C = A - B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} \\ a_{21} - b_{21} & a_{22} - b_{22} \end{pmatrix}$$
- **Scalar multiplication**. Multiplying a matrix by a scalar means you multipliy every element by that scaler.
  Let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, and c be a scalar.
  $$cA = c\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \end{pmatrix} = Ac$$
- **Matrix multiplication.** If A is $(m \times n)$ and B is $(n \times k)$, they can be multiplied as AB, because the columns of A matches the rows of B. But you cannot multiply BA, because the columns of B doesn't match the rows of A. The result of multiplying an $(m \times n)$ matrix to an $(n \times k)$ matrix is $(m \times k)$.