

# Homework 2

Erik Andersen

2024-04-23

## Problem 1

a) The data for this problem set are from Wharton. I use the CRSP value weighted return with and without dividends.

b) In this part, I construct the dividend price ratio  $\frac{D_t}{P_t}$  and the dividend growth ratio  $\frac{D_t}{D_{t-1}}$  using the returns. See the appendix for the code where this happens. Here I will derive how I construct these values using the given returns sequences. \ \ The formulas for the two return series I have access to are given below.

$$\text{returns with dividends} = \text{vwret}d = \frac{P_t + D_t}{P_{t-1}} - 1 \quad (1)$$

$$\text{returns no dividends} = \text{vwret}x = \frac{P_t}{P_{t-1}} - 1 \quad (2)$$

\ To construct the dividend price ratio, we will add one to (1) and (2) to remove the annoying constant, then divide (1) by (2). Finally, we subtract 1 from the ratio.

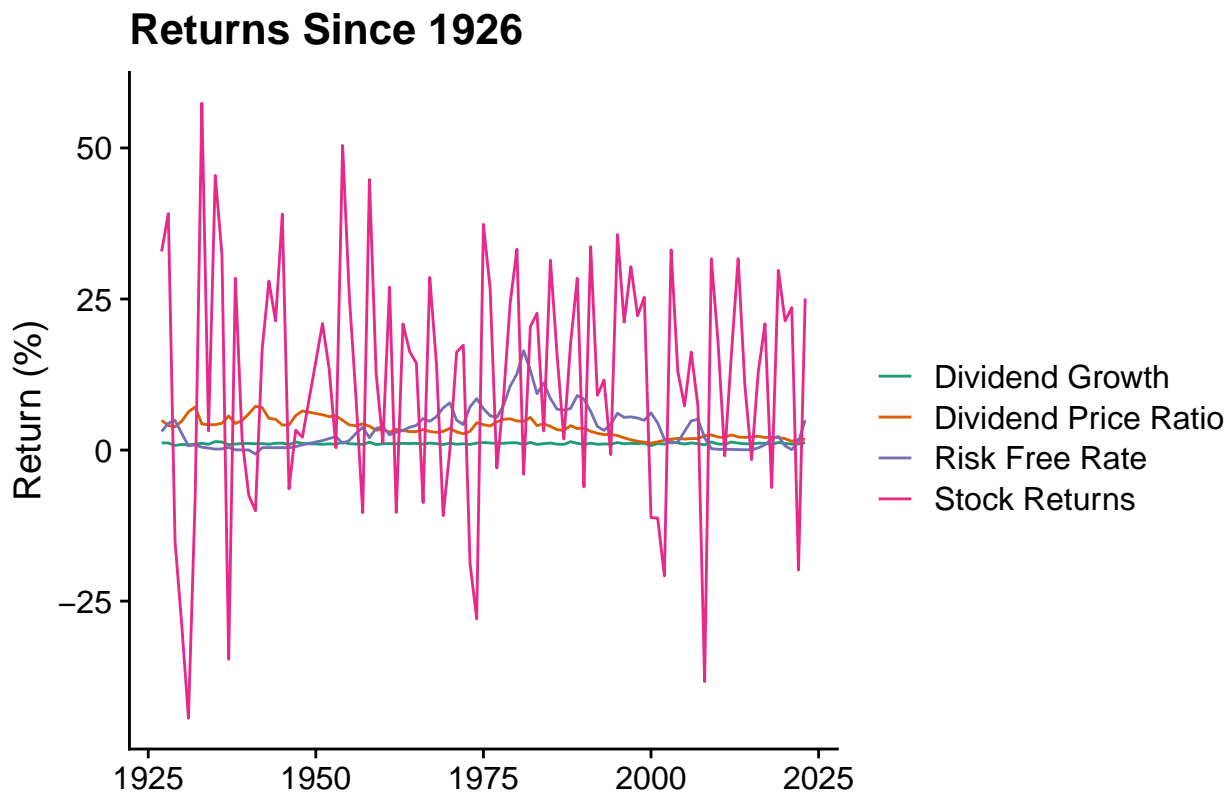
$$\begin{aligned} & \frac{\frac{P_t + D_t}{P_{t-1}}}{\frac{P_t}{P_{t-1}}} - 1 \\ &= \frac{P_t + D_t}{P_t} - 1 \\ &= \frac{D_t}{P_t} \end{aligned}$$

\ Using this we can now construct the dividend growth rate. To get that, we divide the dividend price ratio at time t+1 by the current period dividend price and multiply by returns with no dividends.

$$\begin{aligned} & \frac{\frac{D_{t+1}}{P_{t+1}}}{\frac{D_t}{P_t}} * \frac{P_{t+1}}{P_t} \\ &= \frac{D_{t+1} P_t}{D_t P_{t+1}} \frac{P_{t+1}}{P_t} \\ &= \frac{D_{t+1}}{D_t} \end{aligned}$$

c) Below is the graph of dividend growth, dividend price ratio, risk free return, and stock returns between 1926 and 2023. I only include the stock returns with dividends because including both is redundant. The

noise of the stock returns drowns out everything else, but if you squint we see the characteristic shape of the dividend price ratio we saw in class.



## Problem 2

In this problem, I will reproduce the following forecasting regressions and see how the standard errors differ under various different specifications.

$$R_{t,t+j} = a + b \left( \frac{D}{P} \right)_t + \epsilon_{t+1}$$

$$R_{t,t+j} - R_{t,t+j}^f = a + b \left( \frac{D}{P} \right)_t + \epsilon_{t+1}$$

$$\Delta D_{t,t+j} = a + b \left( \frac{D}{P} \right)_t + \epsilon_{t+1}$$

a) Below are the coefficients and  $R^2$  values for the regressions with time horizons varying from 1 to 10 years. First I'll look at the first regression on returns. Ignore the far too many digits I don't know how to reduce them. Notice that both  $R^2$  and  $\beta$  increase as the horizon increases. Both things indicate that over longer horizons the price dividend ratio explains more of the variation in returns, and has a larger impact on returns.

Table 1: Returns Regression

|           | Horizon Lengths    |                    |                    |                    |                    |                   |                   |                   |                   |                   |
|-----------|--------------------|--------------------|--------------------|--------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
|           | 1                  | 2                  | 3                  | 4                  | 5                  | 6                 | 7                 | 8                 | 9                 | 10                |
| Beta      | 2.44366653695284   | 4.76917987127556   | 7.50747127768786   | 10.4028238913409   | 13.2904255210478   | 17.3727535070467  | 22.6101265956181  | 28.1463465465393  | 35.4053953663635  | 43.3012843295402  |
| R-Squared | 0.0153989141080492 | 0.0375215959476639 | 0.0588451697676149 | 0.0791352848813604 | 0.0981157192571636 | 0.124398473075757 | 0.148624756711374 | 0.169458090530031 | 0.194238057626269 | 0.200288660509278 |

Next I'll look at the values for the second regression which looks at excess returns above the risk free rate.

The pattern is the same here. In the short run, the dividend price ratio does a bad job predicting returns, but as the horizon gets longer it improves.

Table 2: Excess Returns Regression

|           | Horizon Lengths     |                    |                    |                    |                    |                    |                    |                    |                    |                    |
|-----------|---------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|           | 1                   | 2                  | 3                  | 4                  | 5                  | 6                  | 7                  | 8                  | 9                  | 10                 |
| Beta      | 1.6203247961923     | 2.01419659773439   | 4.42056447809811   | 3.28157901338726   | 5.03164925663726   | 4.31013813047258   | 8.78758422709141   | 5.20312486122837   | 17.2137754786496   | 9.69566606868518   |
| R-Squared | 0.00572323243466721 | 0.0128319459924064 | 0.0211054068562909 | 0.0160186108987976 | 0.0169581260804565 | 0.0146928811691185 | 0.0242628120414151 | 0.0109641116592573 | 0.0478334972438223 | 0.0191706921224648 |

Finally, we'll look at the final regression which looks at dividend growth. This time the pattern is very different. The  $R^2$  decreases basically to 0 as the horizon increases. the coefficients remain around the same size and are very imprecisely estimated. We learn from this that at no horizon is the dividend price ratio predictive of dividend growth which contradicts the old style model of finance where all shifts in returns are driven by changes in dividends.

Table 3: Dividend Growth Regression

|           | Horizon Lengths    |                      |                      |                     |                    |                     |                    |                     |                     |                      |
|-----------|--------------------|----------------------|----------------------|---------------------|--------------------|---------------------|--------------------|---------------------|---------------------|----------------------|
|           | 1                  | 2                    | 3                    | 4                   | 5                  | 6                   | 7                  | 8                   | 9                   | 10                   |
| Beta      | 1.00785531020043   | 0.42194226644204     | 0.514971656622504    | 0.847141576007732   | 1.33228862052109   | 1.17966805352659    | 1.30496800379356   | 1.70805560366573    | 1.06698798046206    | -0.141622233030912   |
| R-Squared | 0.0064335245524883 | 0.000823805122841135 | 0.000801401124377455 | 0.00166192562998034 | 0.0035663820065183 | 0.00239098921350862 | 0.0024358712957306 | 0.00361275082365127 | 0.00121014404203085 | 1.64022629212497e-05 |

b) The coefficients and standard errors are biased by the overlapping data. It is entirely reasonable to expect that there is strong autocorrelation in the time series. OLS is not robust to autocorrelated errors, so both estimates will be biased.

c) Now we will compare three different types of standard errors for the first regression. I will compare the standard OLS standard errors to Hansen-Hodrick errors, and OLS errors if we estimate the regression by omitting the overlapping data. The first table shows the standard errors over the various time horizons. Notice that the standard OLS errors are too small as the horizon gets large because of the autocorrelation. The Hansen-Hodrick errors correct for this and are larger at every horizon and especially larger as the horizon gets large. The non-overlapping errors increase much faster than either of the other two. This is probably because they run out of data quickly. At the 10 year horizon for example, there are only 10 data points to use (1926-2016 by 10). This is probably not the best way to correct for the bias since it throws away so much data.

Table 4: Standard Errors

|                | Horizon Lengths  |                  |                  |                  |                  |                  |                  |                  |                  |                  |
|----------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
|                | 1                | 2                | 3                | 4                | 5                | 6                | 7                | 8                | 9                | 10               |
| OLS            | 2.0047731907849  | 2.4913492942349  | 3.11334560068052 | 3.69973113055147 | 4.22400382385755 | 4.85839849176364 | 5.7361819822464  | 6.64248092980196 | 7.7311844435551  | 9.33017948260699 |
| Hansen-Hodrick | 2.41356771085058 | 3.13810526171058 | 4.18319430435705 | 5.41789880288018 | 7.8882055166958  | 9.94027831581315 | 11.4950646915258 | 14.559625687556  | 15.9622316429848 | 18.2297535552623 |
| No Overlap     | 2.0047731907849  | 3.24752329859966 | 4.53744196937369 | 7.16944738781994 | 8.58044515783476 | 11.4476959005446 | 18.1002420185932 | 16.858141745642  | 28.1117205574785 | 41.4091991274208 |

Next, we'll look at the T-statistics. What we can notice here is that using the standard OLS errors leads to us failing to reject the null of no predictive effect at almost every time horizon. Starting at the three year horizon, the t-stats are above 2, so all of them are significant at the 5% level. However, using the standard errors robust to the autocorrelation, we never reject the null except barely at the longest horizons with the Hansen-Hodrick errors.

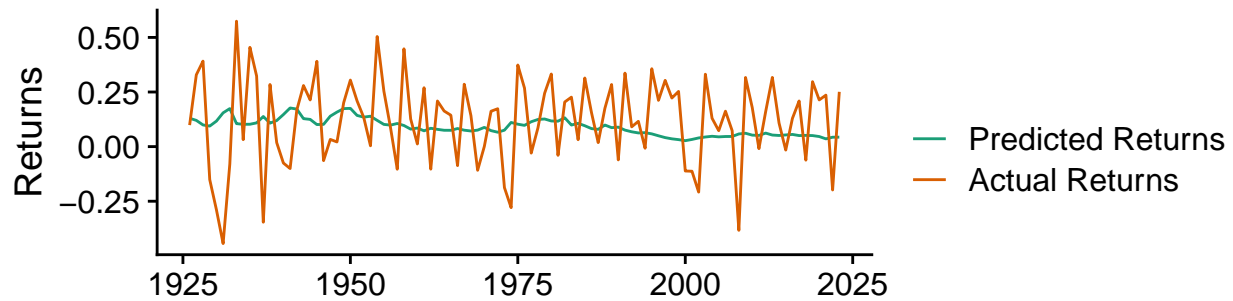
Table 5: T Statistics

|                | Horizon Lengths  |                  |                  |                  |                  |                  |                  |                  |                  |                  |
|----------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
|                | 1                | 2                | 3                | 4                | 5                | 6                | 7                | 8                | 9                | 10               |
| OLS            | 1.21892418962172 | 1.91429595292465 | 2.4113838425284  | 2.81177834936084 | 3.14640470872264 | 3.57581897337126 | 3.94166828486211 | 4.23732440393749 | 4.57955642176902 | 4.64099157044739 |
| Hansen-Hodrick | 1.01247067814462 | 1.51976414859834 | 1.79467429229103 | 1.92008456965101 | 1.68484777595081 | 1.74771298701062 | 1.96694209231257 | 1.93317789554134 | 2.21807302125726 | 2.37530826723879 |
| No Overlap     | 1.21892418962172 | 1.4685590934273  | 1.65456028492726 | 1.45099382541172 | 1.5489202805419  | 1.51757655133038 | 1.24916156217094 | 1.66959958999131 | 1.25945316274655 | 1.045692388213   |

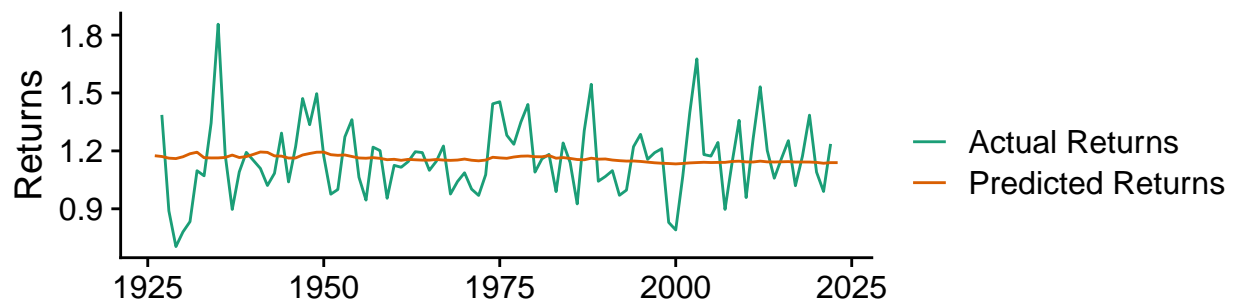
### Problem 3

Now that we've run the regressions to see how well the dividend price ratio can predict things and done some statistics, we can visually inspect how well our predictions do by graphing them. First we'll look at our predictions for returns and dividend growth using the one year time horizon. The graphs are below. We see that the dividend price ratio looks like it can do something to predict returns, but nothing at all to predict dividend growth. The line is basically flat.

#### One Year Horizon Predicted Versus Actual Returns

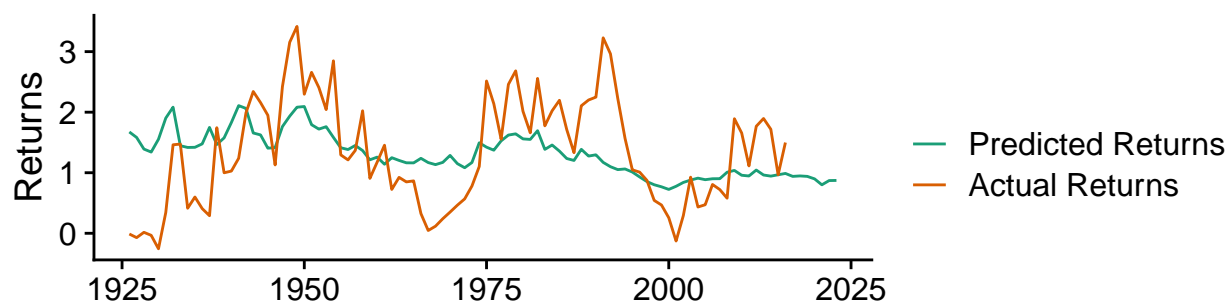


#### One Year Horizon Predicted Versus Actual Dividend Growth

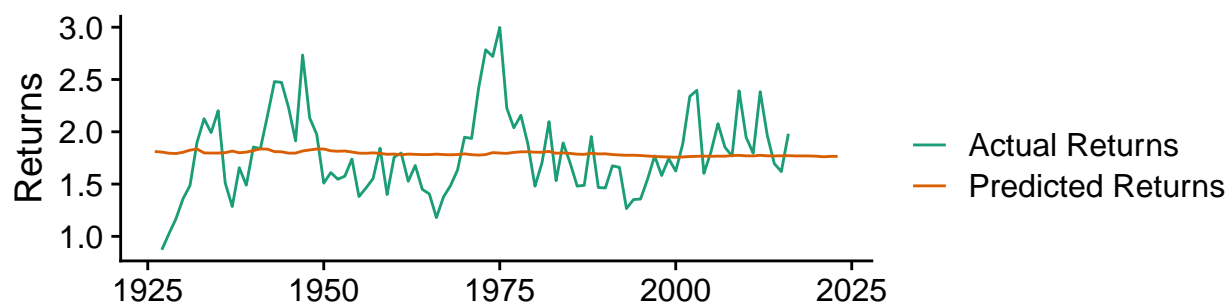


Now we'll look at the same thing but over a 7 year horizon. Now the returns look pretty well predicted by the dividend price ratio. It doesn't capture all of the noise, but it tracks the broad movements. Dividend growth still is not at all predictable and the line is still basically flat.

### Seven Year Horizon Predicted Versus Actual Returns



### Seven Year Horizon Predicted Versus Actual Dividend G



#### Problem 4

In this question, I will estimate the variance decomposition of the price dividend ratio. See the code appendix for the calculation. For the dividend part of the variation, I get a value of -16.4686 which matches fairly closely the value from table 20.3. It is slightly lower but that could be due to the changing macro environment since the table was created. I get a value of 143.0285 for the portion of the variation due to returns which matches the table almost exactly. Matching Cochrane's analysis, I find that variation in returns explains basically all the variation in the price dividend ratio.