Empirical Asset Pricing
Problem Set 1

Yu Zhou, HKUST

September 22, 2020

Data sources

- Annual value-weighted market return with and without dividends (all markets: nyse, amex and nasdaq)
  - Source: WRDS $\rightarrow$ CRSP $\rightarrow$ Index/Stock File Indexes $\rightarrow$ Market Cap. - Annual $\rightarrow$ vwretd and vwretx
  - Time range: 1925–2019
- Risk-free rate
  - Source: Amit Goyal's website
  - Time range: 1871–2019
- cay
  - $cay_t = c_t - \omega a_t - (1-\omega) y_t$ (see Lettau and Ludvigson (2001, JF))
  - Source: Martin Lettau's website
  - Time range: 1952–2019

Q1: Assembling data

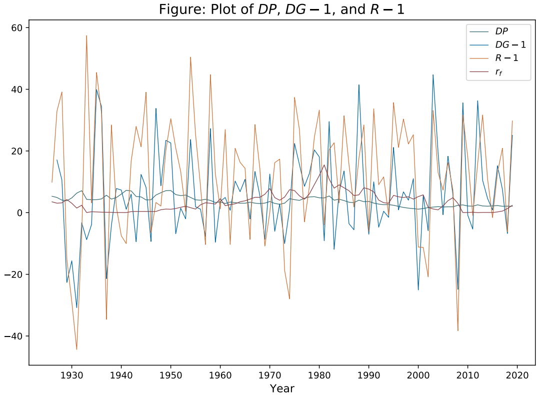

Figure: From the plot, we can see that the dividend growth rates and stock returns are quite volatile. In constrast, the dividend yield and risk-free rates moves slowly.

Q1: Assembling data (cont'd)

- mean and standard deviation of returns and dividend growth rate:

	$R - 1$	$DG - 1$
mean	11.78%	5.89%
std	0.2	0.15

- Sharpe ratio is 0.42.

Q2: Forecasting regressions

	1	2	3	4	5	6	7	8	9	10
Return										
b	2.97	5.91	8.86	12.45	14.65	18.44	23.75	29.60	36.00	44.33
R^2	0.05	0.09	0.13	0.15	0.15	0.19	0.23	0.25	0.27	0.29
Excess return										
b	3.11	6.17	9.23	12.93	15.29	19.28	24.84	31.03	37.96	46.86
R^2	0.05	0.09	0.13	0.15	0.16	0.20	0.25	0.28	0.31	0.34
Dividend growth										
b	-0.27	-0.96	-1.41	-1.63	-2.04	-2.62	-2.69	-2.41	-2.84	-3.12
R^2	0	0	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table: Both coefficients and R^2 of returns on dividend yield increases with horizons. This fact implies that dividend yield can predict one-period returns. And one-period return predictability persists. In constrast, the dividend growth rate can not be predicted using dividend yield.

Q2: Forecasting regressions (cont'd)

- The OLS coefficients and R^2 are not biased. But the standard errors and t-statistics are biased because there is autocorrelation between resisual terms for overlapping data.

Q2: Hansen-Hodrick Errors

- When run forecasting regressions of overlapping long-horizon returns on variables, heteroskedasticity and autocorrelation in residual terms arise:
  - $R_{t\to t+k} = \Pi_{i=1}^{k} R_{t+i}$, $R_{t+1\to t+1+k} = \Pi_{i=1}^{k} R_{t+1+i}$, $Cov(R_{t\to t+k}, R_{t+1\to t+1+k}) \neq 0$.
  - It's natrual to consider the covariance of residual terms as following form:

$$S = \frac{1}{n} E\left(\sum_{i,j} x_i x_j^T \varepsilon_i \varepsilon_j \right) = \frac{1}{n} E\left(\sum_{|i-j|<k} x_i x_j^T \varepsilon_i \varepsilon_j \right)$$

- I Realized it using Python.

Q2: Forecasting regressions (cont'd)

	1	2	3	4	5	6	7	8	9	10
t_{OLS}	2.32	3.45	3.66	3.85	4.52	5.08	5.88	6.73	6.94	6.97
t_{HH}	2.32	2.70	2.92	3.06	3.09	2.71	2.93	3.15	3.08	3.25
t_{NO}	2.32	2.27	2.39	2.16	2.02	2.35	2	3.53	2.54	1.66

Table: The standard errors for Hansen-Hodrick estimator and non-overlapping estimator are smaller than the OLS estimator. The one-period return predictability is less strong and persistent than before.

Q2: Forecasting regressions (cont'd)

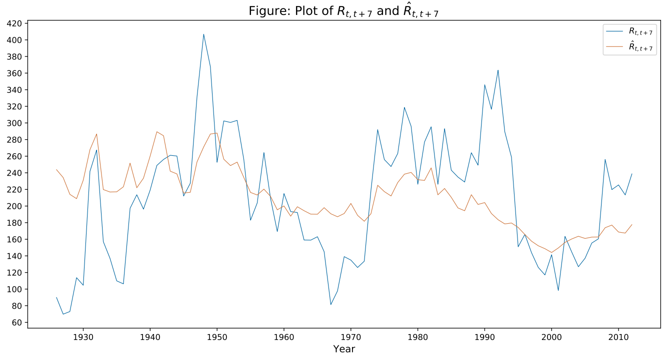

Figure: This figure shows realized and predicted returns. It suggests that return seems to be predictable.

Q2: Forecasting regressions (cont'd)

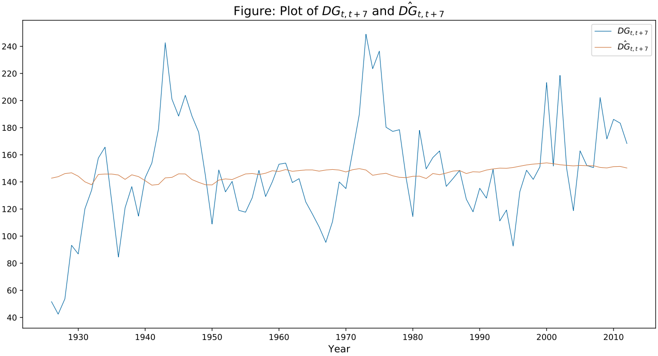

Figure: This figure shows realized and predicted dividend growth. It suggests that dividend growth is not predictable.

Q3: Multivariate forecast

	1	2	3	4	5	6	7	8	9	10
DP	3.27	5.9	6.98	8.27	12.5	18.62	24.51	30.83	39.77	48.37
t	1.63	1.49	1.41	1.67	2.37	2.74	3.58	4.31	5.84	6.41
R^2	0.03	0.05	0.05	0.04	0.07	0.13	0.18	0.20	0.25	0.28
cay	1.88	4.51	7.07	9.1	10.54	12.17	12.47	14.61	18.73	22.17
t	1.6	2.15	2.73	2.63	2.55	2.23	1.7	1.73	2.17	2.55
R^2	0.03	0.09	0.17	0.18	0.15	0.15	0.12	0.11	0.14	0.16
DP	3.19	5.83	7.11	8.63	13.26	19.88	26.04	32.92	42.33	52
t_{DP}	1.58	1.52	1.43	1.54	2.05	2.5	3.24	3.74	4.57	5.77
cay	1.83	4.48	7.12	9.22	10.89	12.9	13.55	16.18	20.65	24.91
t_{cay}	1.59	2.19	2.69	2.70	2.87	3.31	2.71	3.36	4.83	5.43
R^2	0.05	0.14	0.22	0.24	0.23	0.3	0.32	0.35	0.43	0.5

Q3: Multivariate forecast (cont'd)

- Observations:
  - In this shorter sample (bacause *cay* starts from 1952), the regression result are quite similar.
  - The coefficients and R^2 for *cay* and DP are increasing through horizon $1-3$. However, after horizon 3, the coefficients and R^2 for *cay* becomes stable which implies that *cay* is not persistent.
  - *cay* and DP seem to complement with each other instead of substitute for each other. This relationship exists in all horizons.

Q3: Multivariate forecast (cont'd)

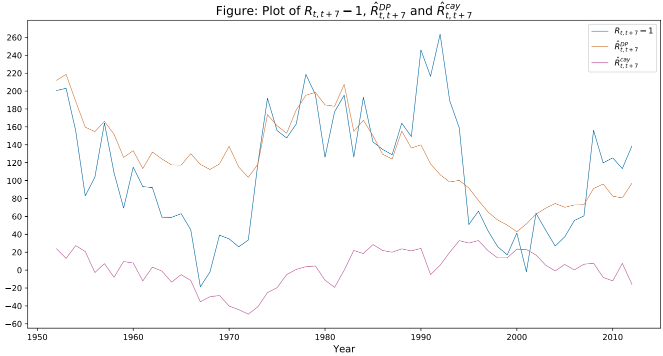

Figure: *DP* and *cay* are complement in this way: *DP* captures the slow-moving variation of returns, and *cay* captures the quick-moving variations.

Q4: Variance decomposition

- The log-linear approximation of $R_{t+1} = (P_{t+1} + D_{t+1})/P_t$ (omitting the constant term):

$$\begin{aligned} & r_{t+1} \approx \rho pd_{t+1} - pd_t + \Delta d_{t+1} \\ \Leftrightarrow \quad & \Delta d_{t+1} \approx r_{t+1} + \rho dp_{t+1} - dp_t, \text{ where } \rho = \exp(\bar{pd})/(1+\exp(\bar{pd})) \end{aligned}$$

- Using sample average ($\hat{\rho} = 0.97$), this identity becomes to be

$$\Delta \hat{d}_{t+1} \approx r_{t+1} + 0.97 dp_{t+1} - dp_t$$

Q4: Variance decomposition

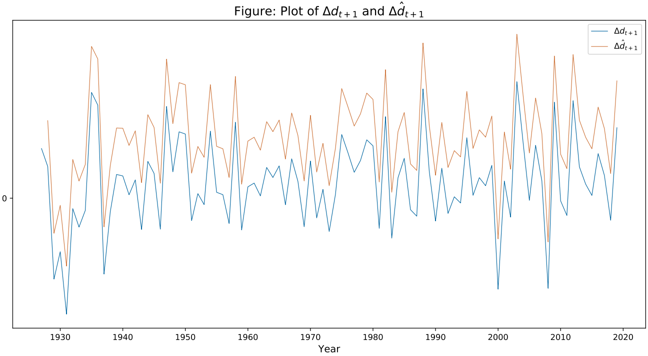

Figure: This figure shows the dividend growth rate in data and dividend growth rate implied by log-return identity are the same.

Q4: Variance decomposition (Cont'd)

- Suppose the state variables at t are $X_t = (r_t, dp_t)^T$ which proceed following $VAR(1)$.

	r_t	dp_t	$\bar{R}^2$
r_{t+1}	0.05	0.08	0.01
dp_{t+1}	-0.21	0.94	0.89

- R^2 for the return is low. But it is not a problem. Because from the plot, we can see that the realized return is quite volatile. It's plausible that the variation of expected return compared to the variation of the realized returns are low.
- $\sigma(E_t(r_{t+1}))$: 0.04, $E(r_t)$: 0.09. The magnitude of the variation in expected returns is nearly one-half of the unconditional expected returns and is economically significant.

Q4: Variance decomposition (Cont'd)

	r_t	dp_t	$\frac{\sigma^2()}{\sigma^2(dp_t)}$
$e_1 G$	-0.12	0.8	0.65
$e_d G$	-0.12	-0.2	0.04

- Observations:
  - $\frac{\sigma^2(pd_t^r)}{\sigma^2(pd_t)} = 0.65$ and $\frac{\sigma^2(pd_t^d)}{\sigma^2(pd_t)} = 0.04$ imply that most of variation in dp associates the variation of long-run expected return instead of the long-run dividend growth.
  - Long-run expected return is predictable.

Q4: Variance decomposition (Cont'd)

- $Cov(dp_t^r, dp_t^d)/Var(dp_t) = -0.16$. Long-run expected returns and dividend growth are negatively correlated. This implies that when investors anticipate a high dividend growth in the future, they require a less expected return.
- $Var(dp_t^r) + Var(dp_t^d) - 2Cov(dp_t^r, dp_t^d) = Var(dp_t)$.
- $X_t \subset \mathcal{F}_t$ does not matter. Because adding more variable as regressors will increase the regression's R^2, and thus reinforce the return predictability.

Q5: Ito's lemma

- Diffusion model of y_t:

$$\mathrm{d}y_t = \mu \, \mathrm{d}t + \sigma \, \mathrm{d}w_t$$

- Using Ito's lemma, we can derive the diffusion model of $x_t = \exp(y_t)$:

$$\begin{aligned}
\mathrm{d}x_t &= \big(\exp(y_t)\mu + \frac{1}{2} \exp(y_t)\sigma^2 \big) \, \mathrm{d}t + \exp(y_t)\sigma \, \mathrm{d}w_t \\
&= \big(x_t\mu + \frac{1}{2} x_t\sigma^2 \big) \, \mathrm{d}t + x_t\sigma \, \mathrm{d}w_t \\
&= \big(\mu + \frac{1}{2}\sigma^2 \big) x_t \, \mathrm{d}t + \sigma x_t \, \mathrm{d}w_t
\end{aligned}$$

Q5: Brownian motion

- From discrete-time to continuous-time stochatic processes:

$$
\begin{aligned}
& w_{t+k} - w_t = \Sigma_{i=1}^{k} \varepsilon_{t+i} \sim N(0, k) \\
\Rightarrow & w_{t+\Delta t} - w_t \sim N(0, \Delta t) \\
\Rightarrow & \, \mathrm{d}w_t / \, \mathrm{d}t \to +\infty
\end{aligned}
$$

- Hence,

$$
\begin{aligned}
& \frac{(w_{t+\Delta t} - w_t)^2}{\Delta t} \sim \chi_1^2 \\
\Rightarrow & E_t (w_{t+\Delta t} - w_t)^2 = \Delta t \\
& Var_t (w_{t+\Delta t} - w_t)^2 = 2\Delta^2 t \\
\Rightarrow & (\mathrm{d}w_t)^2 / \, \mathrm{d}t \sim \chi_1^2
\end{aligned}
$$

- The standard deviation is in the $\mathcal{O}(\Delta t)$ order, which will vanish as $\Delta t \to 0$.

Q5: Brownian motion (cont'd)

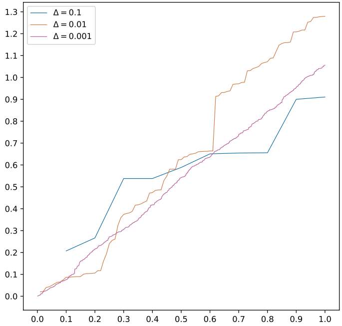

Figure: This figure shows that as Δt becomes smaller, the cumulative sum of $(z_{t+\Delta} - z_t)^2$ converges to the straight line.