FIN 625 Empirical Asset Pricing
Hanno Lustig
Spring 2022

Steven Zheng
steven_zheng@berkeley.edu
April 11, 2022

Problem Set 1

1 Paper and Pencil Question

Part a

Consider a world in which a slow-moving state variable μ_t drives expected dividend growth:

$$r_{t+1} = \varepsilon^r_{t+1} \tag{1}$$
$$\Delta d_{t+1} = \mu_t + \varepsilon^d_{t+1} \tag{2}$$
$$\mu_{t+1} = b\mu_t + v_{t+1}. \tag{3}$$

All variables are demeaned logs. $\varepsilon^i_{t+1}, i \in \{r, d\}$ are white noise.

No question, great!

Part b

Use the Campbell-Shiller present value identity to derive an expression for the log dividend yield.

Start with the result for dividend yield in slide 62 and substituting in the above dynamics,

$$dp_t = -\frac{k}{1-\rho} + \sum_{j=0}^{\infty} \rho^j r_{t+j+1} - \sum_{j=0}^{\infty} \Delta d_{t+j+1} \tag{4}$$
$$= -\frac{k}{1-\rho} + \sum_{j=0}^{\infty} \rho^j \varepsilon^r_{t+1+j} - \sum_{j=0}^{\infty} \mu_{t+j} - \sum_{j=0}^{\infty} \varepsilon^d_{t+j+1}. \tag{5}$$

Since the identity holds ex-post and ex-ante, take conditional expectations and note that we have white noise terms

$$dp_t = -\frac{k}{1-\rho} - \mathbb{E}_t \sum_{j=0}^{\infty} \rho^j \mu_{t+j} \tag{6}$$
$$= -\frac{k}{1-\rho} - \frac{\mu_t}{1-\rho b}. \tag{7}$$

The (log) dividend yield depends on a constant term and a time-varying component, where the latter depends on expected dividend growth μ_t.

Part c

Compute the AR(1) coefficient for the log dividend yield dp_t

$$dp_{t+1} = \phi_0 + \phi_1 dp_t + \nu_{t+1}. \tag{8}$$

Notice that from (7) the dividend yield depends on expected dividend growth, which itself is an AR(1) with coefficient b. So let's "force" dividend yield to also have an AR(1) with coefficient $\phi_1 = b$ and see if that works,

$$\begin{aligned}
dp_{t+1} - b dp_t &= -\frac{k}{1-\rho} - \frac{\mu_{t+1}}{1-\rho b} - b\left[-\frac{k}{1-\rho} - \frac{\mu_t}{1-\rho b}\right] && (9)\\
&= \frac{k}{1-\rho}[b-1] - \frac{1}{1-\rho b}[\mu_{t+1} - b\mu_t] && (10)\\
&= \frac{k}{1-\rho}[b-1] - \frac{1}{1-\rho b}[v_{t+1}] && (11)\\
\iff dp_{t+1} &= \underbrace{\frac{k}{1-\rho}[b-1]}_{\phi_0} + \underbrace{b}_{\phi_1} dp_t - \underbrace{\frac{1}{1-\rho b}[v_{t+1}]}_{\nu_{t+1}} && (12)
\end{aligned}$$

where we used (3) in (11). So we have

$$\begin{aligned}
\phi_0 &= \frac{k}{1-\rho}[b-1] && (13)\\
\phi_1 &= b && (14)\\
\nu_{t+1} &= -\frac{1}{1-\rho b}[v_{t+1}], && (15)
\end{aligned}$$

where ν_t is a composite error term.

Part d

Next, derive an expression for the coefficients in the following regression of real log dividend growth on the log dividend yield:

$$\Delta d_{t+1} = a_d + b_d dp_t + u^d_{t+1}. \tag{16}$$

Rewrite (7) for μ

$$\begin{aligned}
dp_t &= -\frac{k}{1-\rho} - \frac{\mu_t}{1-\rho b} && (17)\\
\iff \frac{\mu_t}{1-\rho b} &= -dp_t - \frac{k}{1-\rho} && (18)\\
\iff \mu_t &= -dp_t(1-\rho b) - \frac{k(1-\rho b)}{1-\rho} && (19)
\end{aligned}$$

then sub into dividend growth dynamics (2)

$$\begin{aligned}
\Delta d_{t+1} &= \mu_t + \varepsilon^d_{t+1} && (20)\\
&= -dp_t(1-\rho b) - \frac{k(1-\rho b)}{1-\rho} + \varepsilon^d_{t+1} && (21)
\end{aligned}$$

so we have

$$a_d = -\frac{k(1-\rho b)}{1-\rho} \tag{22}$$

$$b_d = -(1-\rho b). \tag{23}$$

Notice that $\rho b < 1$ so $b_d < 0$, which means that higher dividend yields are followed by lower dividend growth.

Part e

> Next, derive an expression for the coefficients in the following regression of returns on the log dividend yield:
>
> $$r_{t+1} = a_r + b_r dp_t + u^r_{t+1}. \tag{24}$$

Intuitively, from the dynamics (1) there is no predictability in returns, so our end answer is going to be that $a_r = b_r = 0$ and u^r_{t+1} is some composite error term that's just white noise.

More formally, start with the Campbell-Shiller approximation of log returns from slide 60, then substitute in (21) and (12), and collect terms

$$\begin{aligned} r_{t+1} &= \Delta d_{t+1} - \rho dp_{t+1} + k + dp_t & (25)\\ &= \left[-dp_t(1-\rho b) - \frac{k(1-\rho b)}{1-\rho} + \varepsilon^d_{t+1}\right] - \rho\left[\frac{k(b-1)}{1-\rho} + bdp_t - \frac{v_{t+1}}{1-\rho b}\right] + k + dp_t & (26)\\ &= \left[k - \frac{k(1-\rho b)}{1-\rho} - \rho\frac{k(b-1)}{1-\rho}\right] + \left[1-(1-\rho b)-\rho b\right] dp_t + \left[\varepsilon^d_{t+1} + \frac{\rho v_{t+1}}{1-\rho b}\right], & (27) \end{aligned}$$

so we have

$$a_r = k - \frac{k(1-\rho b)}{1-\rho} - \rho\frac{k(b-1)}{1-\rho} = 0 \tag{28}$$

$$b_r = 1-(1-\rho b) - \rho b = 0 \tag{29}$$

$$u^r_{t+1} = \varepsilon^d_{t+1} + \frac{\rho v_{t+1}}{1-\rho b} \tag{30}$$

so to conclude, we match our intuition that returns are not predictable (since it's just a combination of forecast errors):

$$r_{t+1} = \varepsilon^d_{t+1} + \frac{\rho v_{t+1}}{1-\rho b}. \tag{31}$$

Part f

> Derive a closed-form expression using these parameters for Campbell's decomposition of the variance of return innovations. Recall that
>
> $$\begin{aligned} r_t - \mathbb{E}_{t-1}[r_t] &= (\mathbb{E}_t - \mathbb{E}_{t-1})\left[\sum_{j=0}^{\infty}\rho^j \Delta d_{t+j} - \sum_{j=1}^{\infty}\rho^j r_{t+j}\right] & (32)\\ &= N_{\mathrm{DR}} + N_{\mathrm{CF}}, & (33) \end{aligned}$$

where the first part is CF news and the second part is DR news. The CF/DR decomposition of the variance is then given by:

$$\mathrm{Var}[r_t - \mathbb{E}_{t-1}[r_t]] = \mathrm{Var}[N_{\mathrm{CF}}] + \mathrm{Var}[N_{\mathrm{DR}}] + 2\mathrm{Cov}[N_{\mathrm{DR}}, N_{\mathrm{CF}}] \tag{34}$$

From (31), we see that the return realized at $t+1$ is only a function of ε^d_{t+1} (an expectational error on dividend growth) and v_{t+1} (an innovation to expected dividend growth). In other words, returns are not only not predictable, but only relate to the cash flow component (technically, the unforecastable component of cash flows). Intuitively, there is no discount rate news and all variation in the unexpected component of returns is due to news about cash flows.

More formally, take conditional expectations of (31)

$$\mathbb{E}_{t-1}[r_t] = \mathbb{E}_{t-1}\left[\varepsilon^d_t + \frac{\rho v_t}{1-\rho b}\right] = 0 \tag{35}$$

since ε^d_{t+1} and v_{t+1} are expectational errors. This means that $N_{\mathrm{CF}} = 0$ so

$$r_t - \mathbb{E}_{t-1}[r_t] = N_{\mathrm{DR}} \tag{36}$$

and

$$\mathrm{Var}\left[r_t - \mathbb{E}_{t-1}[r_t]\right] = \mathrm{Var}[N_{\mathrm{CF}}]. \tag{37}$$

Part g

Please report the ratio of discount rate news to total return innovation variance that is implied by your estimates:

$$\frac{\mathrm{Var}[N_{\mathrm{DR}}]}{\mathrm{Var}[r_t - \mathbb{E}_{t-1}[r_t]]} \tag{38}$$

It's zero since $\mathrm{Var}[N_{\mathrm{DR}}] = 0$ from part f.

Part h

Can you find parameter values that imply no dividend growth predictability in a forecasting regression of dividend growth on the log dividend yield?

From (16) we have the regression of dividend growth on dividend yield

$$\Delta d_{t+1} = a_d + b_d dp_t + u^d_{t+1} \tag{39}$$

$$b_d = -(1-\rho b). \tag{40}$$

In order to have no dividend growth predictability form dividend yield, we need $b_d = 0$ or $\rho b = 1$ or, importantly, since $\rho < 1$

$$b = \frac{1}{\rho} > 1. \tag{41}$$

This means that expected dividend growth is non-stationary and the cash flow process has an explosive variance. It also means that dividend yields are non-stationary, since from part c we have

$$dp_{t+1} = \phi_0 + \phi_1 dp_t + v_{t+1} \tag{42}$$

$$\phi_1 = b \tag{43}$$

and Campbell-Shiller relies crucially on a stationary price-dividend ratio.

2 Time-series predictability of returns and dividend growth

Part a

> Download monthly returns with and without dividends from CRSP in WRDS for the period 1945.01-2021.12.

I use the stock market indexes from CRSP. Alternatively, I could have used the monthly stock file and aggregated up prices, dividends and returns, but Bali, Engle, and Murray (2016) say the two methods are similar so I use the more convenient one.

Part b

> Compute monthly dividends.

CRSP gives total returns R_t and price returns Rx_t at time t on the market portfolio. I compute monthly dividends as

$$D_t = \frac{R_t - Rx_t}{P_{t-1}} \tag{44}$$

where P is the total market value. This follows the literature.

Part c

> Aggregate dividends within a year (from January to December) by investing them into cash or investing them into the aggregate stock market.

I follow Koijen and Van Nieuwerburgh (2011) and Chen (2009) in aggregating dividends. Specifically, consider three ways of reinvesting dividends: (1) not reinvesting them, (2) reinvesting them into cash (at the risk-free rate), and (3) reinvesting them into the aggregate stock market. The notation in this section follows Chen (2009).

Let D_t denote the dividend received in month t. With no reinvestment, the annual dividend is just $D_t^{12,\text{no}} = \sum_{i=0}^{11} D_{t-i}$.

With any reinvestment, the dividend received in (the end of) December cannot be reinvested to earn a return in the same year. But any dividend received earlier in the year can be. For example, the November dividend can be reinvested to earn the market return or in cash (the risk-free rate) for December. And the October dividend can be reinvested to earn the November and December returns. To that end, introduce the notation $D_{t,t-i}$ as the time t value of the dividend received at month $t-1$. For dividends reinvested in cash with (net) risk-free return R_t^f, we have

$$D_{t,t-i}^{\text{c}} = \begin{cases} D_{t-i}\prod_{j=1}^{i}(1+R_{t-i+j}^f) & \text{, if } i > 0 \\ D_t & \text{, if } i = 0. \end{cases} \tag{45}$$

Similarly, for dividends reinvested in the market with return R_t^m, we have

$$D_{t,t-i}^{\text{m}} = \begin{cases} D_{t-i}\prod_{j=1}^{i}(1+R_{t-i+j}^m) & \text{, if } i > 0 \\ D_t & \text{, if } i = 0. \end{cases} \tag{46}$$

Finally, the annual dividend is constructed as the sum $D_t^{12,\mathrm{c}} = \sum_{i=0}^{11} D_{t,t-i}^{\mathrm{c}}$ and $D_t^{12,\mathrm{m}} = \sum_{i=0}^{11} D_{t,t-i}^{\mathrm{m}}$.

Part d

Construct non-overlapping annual returns, annual dividend growth, and the log price- dividend ratio for cash-invested (i.e. in the risk-free) and for market-invested dividends. Compute returns and dividend growth in geometric terms. Report the mean and volatility of dividend growth from both methods. Explain the difference.

Table 1 shows the mean and standard deviation of dividend growth under the three reinvestment methods. The three methods have similar means but dividends reinvested in the market portfolio lead to a much more volatile dividend growth process. Intuitively (and written more precisely in Koijen and Van Nieuwerburgh (2011) and Chen (2009)) reinvesting dividends into the market portfolios imparts some of the properties of returns to the dividend growth process. The implication is the following: since dividend yields should predict returns with a positive sign and predict dividend growth with a negative sign, and since stock returns are more volatile than dividend growth, reinvesting dividends in the market portfolio leads to unpredictable dividend growth.

Table 1: Mean and Standard Deviation of Dividend Growth

Reinvestment Method for Dividend Growth	Mean	Standard Deviation
None	8.39	7.39
Cash	8.41	7.67
Market portfolio	8.87	13.45

Note: Units are in percent and annualized.

Part e

Continue with cash-invested dividends. Predict returns and dividend growth using the lagged log price-dividend ratio:

$$r_{t+1} = a_r + b_r pd_t + \epsilon_{t+1}^r \tag{47}$$
$$\Delta d_{t+1} = a_d + b_d pd_t + \epsilon_{t+1}^d. \tag{48}$$

Report the coefficients and the R-squared values. Repeated this exercise for the subsamples from 1945–1990 and 1990-2020. Try to explain the differences.

Table 2 reports the coefficients and R^2 for each regression and each sample. Price to dividend ratios have a negative sign in predicting returns and dividend growth is not predictable. Return predictability was higher in the early (post-war) sample as mentioned in Koijen and Van Nieuwerburgh (2011) and Chen (2009). Interestingly, return predictability has fallen in the late sample. Possible reasons why include a break in price-dividend ratio (Lettau and Van Nieuwerburgh, 2007), the global financial crisis and subsequent zero lower bound period.

Table 2: Regressions

	(1) Full Sample	(2) Early Sample (1945–1990)	(3) Late Sample (1990-2020)
		Return Regression	
a_r	0.42	0.91	0.71
	(0.15)	(0.27)	(0.45)
b_r	-0.09	-0.25	-0.16
	(0.04)	(0.09)	(0.12)
R^2	0.06	0.16	0.06
		Dividend Growth Regression	
a_d	0.16	0.13	0.01
	(0.07)	(0.11)	(0.22)
b_d	-0.02	-0.01	0.01
	(0.02)	(0.04)	(0.06)
R^2	0.02	0.00	0.00

Note:

Part f

Start from the Campbell and Shiller identity and estimate the equation

$$pd_{t+1} = a_{pd} + \phi pd_t + \epsilon^{pd}_{t+1} \tag{49}$$

We can obtain a variance decomposition of the log price-dividend ratio via:

$$\mathrm{Var}(pd_t) = \mathrm{Cov}\left(\sum_{s=1}^{\infty} \rho^{s-1} \mathbb{E}_t[\Delta d_{t+s}], pd_t\right) + \mathrm{Cov}\left(-\sum_{s=0}^{\infty} \rho^{s-1} \mathbb{E}_t[r_{t+s}], pd_t\right) \tag{50}$$

Divide both sides by $Var(pd_t)$ so that we can estimate how much of the variation in the log price-dividend ratio is due to discount rate news and cash flow news. Comment on the economic interpretation of the results.

Borrowing notation from Cochrane (2008), I divide through (50) by the left hand side to get the variance decomposition in terms of regression coefficients

$$1 = \beta\left(\sum_{s=1}^{\infty} \rho^{s-1} \Delta d_{t+s}, pd_t\right) - \beta\left(\sum_{s=0}^{\infty} \rho^{s-1} r_{t+s}, pd_t\right). \tag{51}$$

The second term is the regression coefficient of discount rates on the price-dividend ratio. Following the dynamics in (47)–(50), we have that

$$\beta\left(\sum_{s=0}^{\infty}\rho^{s-1}\Delta r_{t+s}, pd_t\right) = \sum_{s=1}^{\infty}\rho^{s-1}\beta(r_{t+s}, pd_t) \tag{52}$$

$$= \sum_{s=1}^{\infty}\rho^{s-1}\phi^{s-1}b_r \tag{53}$$

$$= \frac{b_r}{1-\rho\phi} \tag{54}$$

$$= b_r^{lr}. \tag{55}$$

Similarly, the first term is the regression of cash flows on the price-dividend ratio, and we get the similar expression for the first term which we'll call b_d^{lr}.

In the code, I estimate $\rho \approx 0.97$ and $\phi \approx 0.96$ which leads to $|b_r^{lr}| \approx 1.33$ and $|b_d^{lr}| \approx 0.34$. I report the absolute value of the coefficients just to make things comparable to Cochrane (2008) who uses dividend yields and $dp_t = -pd_t$. The takeaway is that most of the variation in price-dividend ratios comes from discount rates.

Part g

> The present-value identity implies restrictions between b_r, b_d, and ϕ. Derive the connection between the coefficients.

Start with the Campbell-Shiller return approximation

$$r_{t+1} = \rho pd_{t+1} + \Delta d_{t+1} - pd_t \tag{56}$$

then plug in (47), (48) and (50), assuming data are demeaned, and take time t expectations:

$$b_r pd_t + \epsilon^r_{t+1} = \rho\left[\phi pd_t + \epsilon^{pd}_{t+1}\right] + \left[b_d pd_t + \epsilon^d_{t+1}\right] - pd_t \tag{57}$$

$$b_r pd_t = \rho\phi pd_t + b_d pd_t - pd_t \tag{58}$$

$$b_r = \rho\phi + b_d - 1. \tag{59}$$

Note the signs are slightly different than Cochrane (2008) since he uses dividend yield and we used price-dividend ratio.

3 The Mankiw-Shapiro and Stambaugh bias.

Part a

Start from the predictive system that you estimated in the previous question for the sample from 1945-2020.

$$r_{t+1} = a_r + b_r pd_t + \epsilon^r_{t+1} \tag{60}$$
$$pd_{t+1} = a_{pd} + \phi pd_t + \epsilon^{pd}_{t+1} \tag{61}$$

We want to simulate from this model. Assume that the errors are normally distributed. Estimate and report the coefficients, including the covariance matrix of the shocks.

I demean the data so the constant terms can be ignored in the simulation. I estimate by OLS

$$b_r \approx -0.096 \tag{62}$$
$$\phi \approx 0.944 \tag{63}$$
$$\text{Cov}(e^r_{t+1}, e^{pd}_{t+1}) \approx 0.00694 \tag{64}$$
$$\text{Var}(e^r_{t+1}) \approx 0.0247 \tag{65}$$
$$\text{Var}(e^{pd}_{t+1}) \approx 0.0219. \tag{66}$$

Part b

Simulate 10,000 samples of 67 observations and estimate the model for each of the samples. Importantly, fix the seed for the subsequent exercises at the beginning of each of the 10,000 runs. Compute the average estimate of b_r and ϕ and compare them to the population values.

I follow the Monte Carlo simulation outlined in Mankiw and Shapiro (1986). Let all simulated variables be denoted with a tilde on top of their data counterparts. Let $N = 75$ denote the number of observations and $M = 10,000$ denote the number of runs of the simulation.

In each run of the simulation, I draw N observations of $\tilde{\epsilon}^r_{t+1}$ and $\tilde{\epsilon}^{pd}_{t+1}$ from a bivariate Normal distribution with mean 0, variance 1 and correlation equal to the sample correlation between ϵ^r_{t+1} and ϵ^r_{t+1}. Next, we generate simulated returns and price-dividend ratio. Set $\tilde{r}_{t+1} = \tilde{\epsilon}^r_{t+1}$. $\tilde{pd}_t$ is simulated from (61) using $\tilde{\epsilon}^{pd}_{t+1}$ and ϕ estimated from the data. The initial value $\tilde{pd}_0$ is drawn from a univariate Normal distribution with mean 0 and variance $1/(1-\phi^2)$. Then using the simulated data, we run the regressions (60) and (61). This completes one run of the simulation. Then repeat for a total of 10,000 runs.

From the simulation I obtain $\tilde{b}_r \approx 0.0382$ with a standard deviation of 0.0456, and $\tilde{\phi} \approx 0.923$ with a standard deviation of 0.0484. The persistence of the price-dividend ratio is quite close to the data, but the return predictability coefficient is a bit off.

Part c

Now repeat the exercise but setting the covariance

$$\text{Cov}(\epsilon_{t+1}^{r}, \epsilon_{t+1}^{pd}) = 0. \tag{67}$$

Comment on the difference.

From the simulation I obtain $\tilde{b}_r \approx 0$ with a standard deviation of 0.0437, and $\tilde{\phi} \approx 0.923$ with a standard deviation of 0.0483. The persistence of the price-dividend ratio is quite close to the data, but the return predictability coefficient is now closer.

From Stambaugh (1999), we know that when returns are regressed on lagged persistent variables, the disturbances are correlated with the regressor's innovation. In the case of price-dividend ratio regressions, this creates a downward bias. In part b, we allowed for covariance between the two error terms and found a much higher point estimate of the regression coefficient $\tilde{b}_r$. When we remove that covariance in part c, the bias decreases (we find a lower estimate of $\tilde{b}_r$.

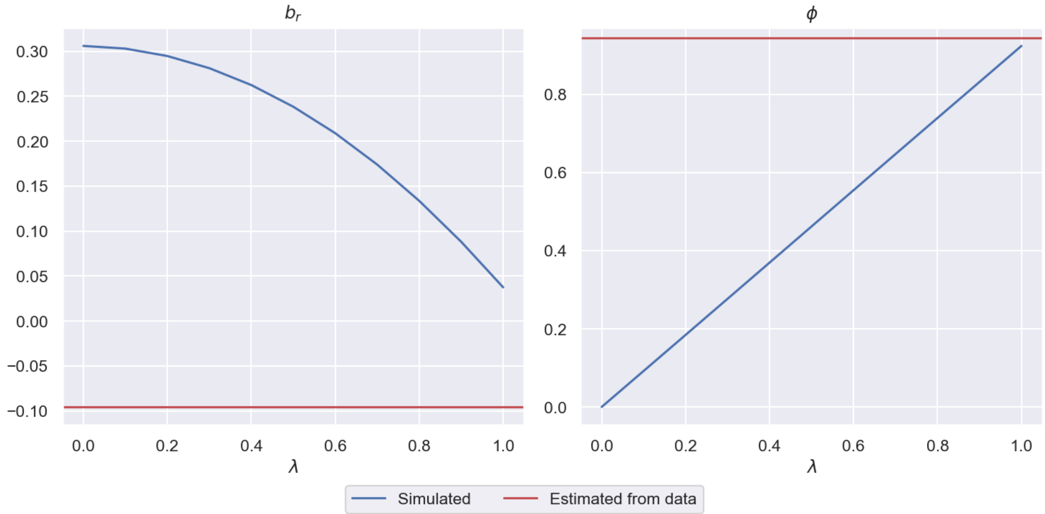
Part d

Set the covariance back to the sample estimate, but now plot the bias in b_r and ϕ as a function of ϕ where we vary $\phi = \lambda\hat{\phi}$ for $\lambda = 0, 0.1, \ldots, 0.9, 1$. $\hat{\phi}$ denotes the OLS estimate obtained in part 3a. Comment on the results.

To make the notation a bit clearer, now I vary the value of ϕ (estimated from the data) that is used as an input for the Monte Carlo simulation. I use $\lambda\phi$ for $\lambda = 0, 0.1, \ldots, 0.9, 1$. Figure 1 plots the values of b_r vs $\tilde{b}_r$ on the left, and ϕ vs $\tilde{\phi}$ on the right, both against λ. Simulated values are in blue, while red shows point estimates from the data.

When the predictor variable is less persistent, then $\tilde{\phi}$ diverges from ϕ mechanically. I find this plot a bit hard to read: as the true ϕ varies, then so would the true b_r. So as you vary the input ϕ, you would also want to vary b_r due the changing ϕ. However, we have not done that (the red lines are flat). From Stambaugh (1999), as the price-dividend ratio gets less persistent (ϕ decreases) then the bias would get closer to 0, especially when ϵ_{t+1}^{r} and ϵ_{t+1}^{pd} are correlated.

Figure 1: Plot of Simulated vs. Estimated Parameters



Note: Figure shows the simulation that estimates the expected time until vaccine deployment.

References

Turan Bali, Robert Engle, and Scott Murray. The market price of fiscal uncertainty. *Wiley*, 2016.

Long Chen. On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics*, 2009.

John H. Cochrane. The Dog That Did Not Bark: A Defense of Return Predictability. *The Review of Financial Studies*, 2008.

Ralph S.J. Koijen and Stijn Van Nieuwerburgh. Predictability of returns and cash flows. *Annual Review of Financial Economics*, 2011.

Martin Lettau and Stijn Van Nieuwerburgh. Reconciling the Return Predictability Evidence: The Review of Financial Studies: Reconciling the Return Predictability Evidence . *The Review of Financial Studies*, 2007.

N. Gregory Mankiw and Matthew D. Shapiro. Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models. *Economics Letters*, 1986.

Robert F. Stambaugh. Predictive regressions. *Journal of Financial Economics*, 1999.