Lecture 5

Predictability

Traditional Views of Market Efficiency (1960-1970)

- CAPM is a good measure of risk
- Returns are close to unpredictable
- (a) Stock, bond and foreign exchange changes are not predictable
  (b) Constant equity premium
- Market volatility does not change much through time

- Professional managers do not reliably outperform simple indices and passive portfolios once one corrects for risk

- Jensen (1978): "I believe there is no other proposition in economics which has more solid evidence supporting it than the Efficient Markets Hypothesis."

Modern Empirical Research (1980-to-present)

- Returns are predictable
  1. Valuation ratios (D/P, E/P, B/M ratios and)
  2. Interest rates (term spread, short-long T-bill rates, etc.)
  3. Decision of market participants (corporate financing, consumption).
  4. Cross-sectional equity pricing.
  5. Bond and foreign exchange returns are also predictable.
- Some funds seem to outperform simple indices, even after controlling for risk through market betas
- Development of equilibrium models with time-varying equity premium.

Predictive Regressions I

Fama and French (1989), JFE: economic questions

- Economic questions:

  1. Do the expected returns on bonds and stocks move together? Do the same variables forecast bond and stock returns?

  2. Is the variation in expected returns related to business cycles?

- Motivation:

  1. Mounting evidence that stock and bond returns are predictable

  2. Interpretations: market inefficiency versus rational variation in expected returns

- Framework: Regress future returns on variables X(t) known at time t.
  $$r(t, t+\tau) = \alpha(\tau) + \beta(\tau) X(t) + \varepsilon(t, t+\tau) \quad (1)$$
  where τ can be one month, one quarter, and one to four years.
- $r(t, t+\tau)$: value- and equal-weighted market portfolios of NYSE; value-weighted corporate bond portfolios.
- X(t) variables:
  - Dividend yields D(t)/P(t): summing monthly dividends for the year preceding time t divided by the value of the portfolio at time t (Discount rate intuition)
  - Term Premium TERM(t) –from Keim and Stambaugh (1986)
  - Default premium DEF(t) –from Keim and Stambaugh (1986)
- Fama and French (1989) sample non-overlapping data for quarterly and annual frequencies (240 quarterly and 60 annual observations) and use traditional OLS standard errors. For longer horizons, they use annual overlapping observations and modify the standard errors.

Table 2

Slopes, t-statistics, and R^2 from multiple regressions of excess returns on the term spread ($TERM$) and the value-weighted dividend yield (D/P) or the default spread (DEF); 1927–1987.[a]

$$r(t, t+T) = a + bD(t)/P(t) + cTERM(t) + e(t, t+T)$$

T	Aaa	Aa	A	Baa	LG	VW	EW	Aaa	Aa	A	Baa	LG	VW	EW
	Slopes for D/P							t-statistics for D/P slopes						
M	0.04	0.01	−0.05	−0.03	0.07	0.21	0.43	0.96	0.21	−0.55	−0.23	0.43	0.78	1.26
Q	0.20	0.14	0.03	0.14	0.48	1.09	2.05	1.07	0.69	0.09	0.31	0.79	1.09	1.58
1	0.30	−0.21	−0.64	−0.49	0.39	2.79	5.75	0.65	−0.45	−0.87	−0.68	0.30	1.31	2.31
2	1.09	0.25	0.99	0.91	3.92	8.89	15.66	1.38	0.31	0.95	0.62	1.62	3.13	4.98
3	1.83	1.17	2.91	2.87	7.83	12.20	20.21	2.24	1.66	6.49	2.62	3.21	3.77	4.91
4	2.72	2.16	4.14	4.47	10.54	15.37	23.29	3.91	2.76	5.20	6.22	5.94	5.08	4.74
	Slopes for $TERM$							t-statistics for $TERM$ slopes						
M	0.23	0.24	0.26	0.26	0.26	0.31	0.47	3.12	3.57	3.23	3.19	2.27	1.68	1.93
Q	0.57	0.52	0.59	0.55	0.60	0.65	0.96	1.75	1.63	1.83	1.74	1.37	1.03	1.10
1	3.36	3.30	3.92	3.36	3.61	1.56	2.69	5.44	5.25	5.81	4.88	3.78	0.82	1.01
2	4.23	4.22	4.46	4.15	4.41	−1.27	−0.38	3.98	3.83	2.83	2.80	1.85	−0.38	−0.07
3	5.01	4.62	4.86	4.63	4.66	−1.54	−0.17	3.52	3.27	2.56	2.47	1.70	−0.36	−0.03
4	5.09	4.61	5.50	5.22	6.06	0.32	4.07	2.39	2.15	2.36	2.66	2.95	0.09	0.73
	Regression R^2													
M	0.04	0.04	0.03	0.02	0.01	0.01	0.01							
Q	0.06	0.04	0.03	0.02	0.02	0.02	0.04							
1	0.39	0.33	0.30	0.20	0.11	0.02	0.07							
2	0.28	0.22	0.17	0.13	0.15	0.12	0.22							
3	0.26	0.19	0.21	0.18	0.24	0.18	0.27							
4	0.24	0.19	0.28	0.26	0.36	0.25	0.34							

T	Aaa	Aa	A	Baa	LG	VW	EW	Aaa	Aa	A	Baa	LG	VW	EW
	Portfolios													
	$r(t,t+T) = a + bDEF(t) + cTERM(t) + e(t,t+T)$													
	Slopes for *DEF*							*t*-statistics for *DEF* slopes						
M	0.07	0.07	0.07	0.05	0.27	0.04	0.41	0.78	0.74	0.50	0.30	0.99	0.09	0.67
Q	0.31	0.34	0.47	0.54	1.30	0.99	2.78	0.85	0.91	0.84	0.85	1.30	0.56	1.12
1	0.76	0.41	0.76	1.49	4.12	4.38	11.59	0.79	0.46	0.57	1.12	1.62	1.15	2.38
2	4.18	3.51	6.41	6.70	13.49	14.62	29.22	1.96	1.75	2.86	2.47	3.04	2.18	2.73
3	7.21	7.14	11.12	11.83	22.25	19.96	37.61	2.01	2.14	3.21	5.51	4.61	2.25	2.52
4	10.11	9.66	13.62	15.29	27.07	24.56	41.41	2.11	2.07	3.20	4.46	4.64	2.42	2.80
	Slopes for *TERM*							*t*-statistics for *TERM* slopes						
M	0.22	0.23	0.24	0.24	0.22	0.34	0.47	2.81	3.21	3.02	3.04	1.99	1.83	2.00
Q	0.55	0.48	0.49	0.47	0.42	0.70	0.83	1.51	1.36	1.48	1.41	0.97	1.01	0.92
1	3.25	3.15	3.58	2.89	2.73	1.20	1.35	4.41	4.49	4.89	3.96	3.08	0.60	0.52
2	3.48	3.41	3.13	2.73	2.10	−2.53	−3.47	2.89	3.07	2.04	1.97	1.19	−0.74	−0.81
3	3.63	3.06	2.75	2.32	1.02	−3.30	−4.29	1.80	1.74	1.26	1.24	0.50	−0.84	−0.80
4	3.09	2.56	2.94	2.29	1.68	−1.89	−0.45	0.98	0.88	1.03	0.99	0.76	−0.51	−0.12
	Regression R^2													
M	0.04	0.04	0.03	0.02	0.01	0.00	0.01							
Q	0.05	0.04	0.03	0.03	0.03	0.01	0.02							
1	0.39	0.33	0.29	0.20	0.15	0.00	0.07							
2	0.32	0.26	0.25	0.22	0.25	0.05	0.16							
3	0.34	0.29	0.33	0.31	0.36	0.09	0.20							
4	0.34	0.31	0.40	0.41	0.45	0.13	0.23							

[a] The regressions for T = one month (M), one quarter (Q), and one year use nonoverlapping returns. The regressions for two- to four-year returns use overlapping annual observations. The numbers of observations in the regressions are (M) 732, (Q) 244, (1 yr) 61, (2 yr) 60, (3 yr) 59, and (4 yr) 58. The standard errors in the t-statistics for the slopes are adjusted for heteroscedasticity and (for two- to four-year returns) the sample autocorrelation of overlapping residuals with the method of Hansen (1982) and White (1980). See note to table 1 for definition of portfolios.

Table 5

Slopes, t-statistics, and R^2 from multiple regressions of excess returns on the term spread ($TERM$) and the value-weighted dividend yield (D/P) or the default spread (DEF); 1941–1987.[a]

T	Aaa	Aa	A	Baa	LG	VW	EW	Aaa	Aa	A	Baa	LG	VW	EW
	Portfolios													
	$r(t,t+T) = a + bD(t)/P(t) + cTERM(t) + e(t,t+T)$													
	Slopes for *D/P*							*t*-statistics for *D/P* slopes						
M	0.13	0.11	0.11	0.13	0.30	0.40	0.53	2.75	2.58	2.54	2.81	3.82	2.88	2.99
Q	0.36	0.34	0.36	0.42	0.94	1.31	1.78	1.91	1.89	1.97	2.42	3.28	2.93	3.03
1	0.40	0.27	0.74	1.23	3.33	5.49	7.96	0.75	0.47	1.31	2.14	3.91	3.45	3.67
2	1.00	0.62	2.05	3.15	7.67	11.84	16.70	0.87	0.49	1.58	2.56	3.97	4.18	3.87
3	1.41	0.91	2.93	4.34	10.88	15.65	21.22	1.37	0.78	2.27	3.11	3.65	4.94	3.31
4	2.41	1.76	3.87	5.29	12.66	18.48	23.43	3.78	1.94	3.65	3.83	4.21	5.26	3.18
	Slopes for *TERM*							*t*-statistics for *TERM* slopes						
M	0.25	0.28	0.31	0.32	0.31	0.48	0.51	2.77	3.55	4.35	4.81	3.32	3.29	2.97
Q	0.62	0.60	0.73	0.75	0.77	1.13	1.17	1.51	1.52	2.07	2.43	1.89	2.17	1.82
1	3.64	3.56	3.87	3.57	3.27	1.64	1.33	4.74	5.07	5.68	5.57	4.10	0.94	0.66
2	4.29	4.18	4.25	4.16	3.71	−1.34	−2.90	3.25	3.48	3.64	4.06	3.04	−0.63	−0.89
3	4.41	3.81	3.83	3.62	2.71	−3.95	−6.35	2.13	2.09	1.95	2.31	2.01	−1.23	−1.27
4	3.73	3.07	3.27	3.51	3.27	−2.40	−2.67	1.11	0.97	1.02	1.40	1.45	−0.75	−0.64
	Regression R^2													
M	0.04	0.06	0.08	0.08	0.05	0.03	0.03							
Q	0.06	0.05	0.08	0.10	0.10	0.06	0.06							
1	0.39	0.37	0.44	0.41	0.35	0.16	0.18							
2	0.21	0.18	0.24	0.30	0.44	0.36	0.37							
3	0.13	0.08	0.15	0.24	0.46	0.53	0.48							
4	0.09	0.04	0.13	0.25	0.51	0.60	0.50							

T	Aaa	Aa	A	Baa	LG	VW	EW	Aaa	Aa	A	Baa	LG	VW	EW
	Portfolios													
	$r(t, t+T) = a + bDEF(t) + cTERM(t) + e(t, t+T)$													
	Slopes for *DEF*							t-statistics for *DEF* slopes						
M	0.23	0.27	0.30	0.36	0.84	0.52	0.91	1.87	2.35	2.93	3.08	3.89	1.43	1.77
Q	0.57	0.68	0.80	1.09	2.72	2.18	3.70	1.22	1.49	1.80	2.31	3.25	1.61	1.84
1	1.42	1.11	2.87	4.51	11.15	10.98	18.62	1.08	0.73	1.90	2.79	4.37	2.12	2.79
2	6.25	5.13	9.01	11.73	24.48	24.83	39.56	1.62	1.20	2.24	3.74	5.79	3.01	3.82
3	10.01	8.66	13.83	17.45	36.15	36.07	54.33	1.65	1.31	2.49	5.00	9.42	4.17	4.04
4	13.35	11.16	16.32	19.90	40.18	41.99	57.36	1.85	1.44	2.56	4.26	11.52	3.94	3.78
	Slopes for *TERM*							t-statistics for *TERM* slopes						
M	0.25	0.27	0.30	0.31	0.29	0.46	0.48	2.64	3.39	4.13	4.52	2.99	3.21	2.83
Q	0.61	0.58	0.71	0.71	0.68	1.09	1.08	1.44	1.44	1.96	2.25	1.64	2.03	1.62
1	3.60	3.53	3.79	3.46	3.03	1.75	1.31	4.59	4.97	5.61	5.55	4.41	0.99	0.61
2	4.05	3.95	4.01	3.94	3.44	−0.89	−2.57	3.26	3.61	3.86	4.89	4.82	−0.38	−0.70
3	3.94	3.33	3.38	3.18	2.19	−3.47	−6.12	1.74	1.69	1.51	1.74	1.28	−0.96	−1.16
4	3.22	2.59	2.85	3.12	2.92	−1.60	−1.97	0.87	0.75	0.81	1.12	1.05	−0.40	−0.49
	Regression R^2													
M	0.04	0.06	0.08	0.08	0.06	0.02	0.02							
Q	0.05	0.05	0.08	0.10	0.11	0.04	0.04							
1	0.40	0.38	0.46	0.45	0.45	0.09	0.13							
2	0.27	0.22	0.32	0.41	0.59	0.21	0.29							
3	0.23	0.16	0.29	0.42	0.70	0.39	0.44							
4	0.21	0.14	0.27	0.43	0.71	0.43	0.42							

[a] The regressions for T = one month (M), one quarter (Q), and one year use nonoverlapping returns. The regressions for two- to four-year returns use overlapping annual observations. The numbers of observations in the regressions are (M) 564, (Q) 188, (1 yr) 47, (2 yr) 46, (3 yr) 45, and (4 yr) 44. The standard errors in the t-statistics for the slopes are adjusted for heteroscedasticity and (for two- to four-year returns) the sample autocorrelation of overlapping residuals with the method of Hansen (1982) and White (1980). See note to table 1 for definition of portfolios.

- D/P has strongest effect (high t-stats and high R^2)
- Regression coefficients and R^2 rise with the forecast horizon.
- Rational time-variation of expected return:
  - time-varying risk aversion
  - time-varying amount of risk
- A parallel explanation based on investor sentiment
  - Evidence does not distinguish among potential explanations

Predictive Regressions II

Lamont (1998), Journal of Finance.

- Lamont also uses a predictive regression framework. He adds Earning Yield and Payout Ratio to the Dividend Yield.
- No interest rate data. Tests for unit roots in levels but not in ratios(?!)

- Findings: Dividends and earnings contain information about future returns .
- Interpretation: Dividends contain information about future dividends (dividends measure the permanent component of stock prices?) and earnings contain information on business conditions

- Cross-sectional predictability with the dividend payout ratio.

Table IV

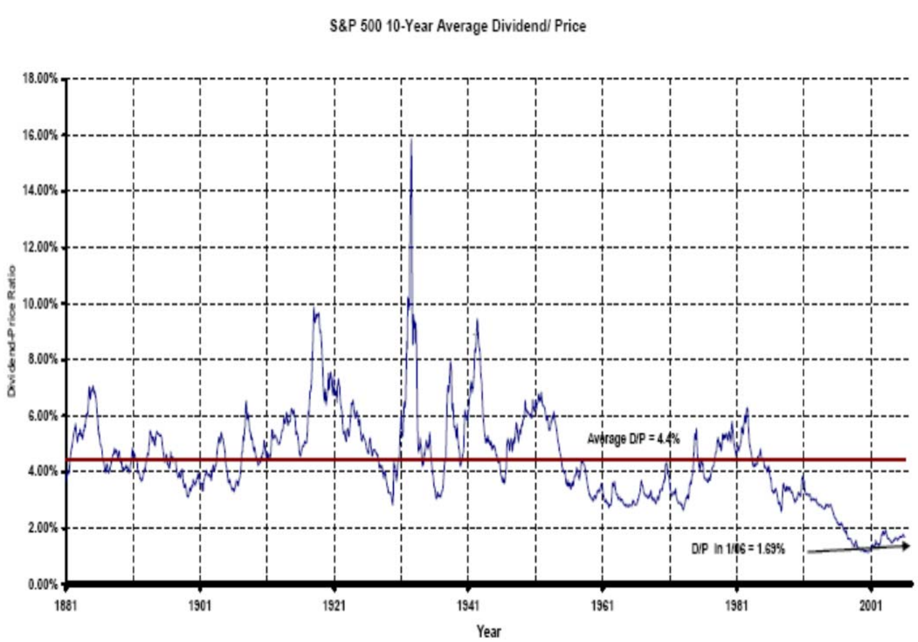
Forecasting Quarterly Excess Return Using Dividend Yield, Earnings Yield, and Payout, 1947–1994

Regressions of current stock returns on lagged dividend yields, earnings yields, and dividend payout ratios, 1947Q1–1994Q4. The dependent variable, $R_{m,t+1} - R_{f,t+1}$, is quarterly log excess returns on the S&P Composite Index. $d_t - p_t$ is the log dividend yield, $e_t - p_t$ is the log earnings yield, and $d_t - e_t$ is the log dividend payout ratio. RREL$_t$ is the relative bill rate. Forecast returns are the projected excess returns in 1995Q1, using 1994Q4 values of the regressors. Estimation error is the standard error of the point estimate, based on sampling error in the coefficients. Total forecast error includes both sampling error and residual variance. OLS standard errors are in parentheses below the coefficient estimates.

	Constant	$R_{m,t} - R_{f,t}$	RREL$_t$	$d_t - p_t$	$d_t - e_t$	$e_t - p_t$	R^2	Forecast Returns 1994Q4	
1	0.222			0.064			0.05	−0.004	Forecast
	(0.065)			(0.020)				(0.008)	Estimation error
								(0.074)	Total forecast error
2	0.062					0.012	0.00	0.015	Forecast
	(0.059)					(0.015)		(0.006)	Estimation error
								(0.076)	Total forecast error
3	0.207			0.194		−0.112	0.13	−0.037	Forecast
	(0.062)			(0.038)		(0.028)		(0.011)	Estimation error
								(0.072)	Total forecast error
4	0.207			0.083	0.112		0.13	−0.037	Forecast
	(0.062)			(0.020)	(0.028)			(0.011)	Estimation error
								(0.072)	Total forecast error
5	−0.042				0.083		0.04	−0.004	Forecast
	(0.020)				(0.028)			(0.009)	Estimation error
								(0.074)	Total forecast error
6	0.209	0.066	−0.892	0.078	0.086		0.15	−0.040	Forecast
	(0.062)	(0.069)	(0.465)	(0.020)	(0.030)			(0.011)	Estimation error
								(0.071)	Total forecast error

Methodological Concerns with Predictive Regression Framework (90's)

- Data snooping? Are D/P, TERM, Payout Ratios the only variables used in those regressions? Pre-testing Bias very likely.
- Regressors are only predetermined, but not exogenous. OLS slopes have a small bias: Stambaugh (1986). Traditional OLS S.E. are only appropriate asymptotically --if there is no serial correlation of the error term and if it is conditionally homoskedastic. Hodrick (1992).
- Valuation ratios are persistent and their innovations are correlated with returns, causing
  – biased predictive coefficients: Stambaugh (1999)
  – over-rejection by standard *t* test: Cavanagh-Elliott-Stock (1995)
- These problems are less relevant for interest rates and recently proposed predictor variables (persistent, but less correlated with returns).



• D/P is persistent, with a constant mean (mean reverting?) and no trend.

• Issue: How persistent is D/P?
- D/P is likely to be persistent: it reflects long-run expectations.
- But, is D/P stationary? unit root? explosive?

• Recall the Campbell-Shiller log-linear return formulas:

$$d_t - p_t = \frac{k}{1-\rho} + E_t \sum_{j=0}^{\infty} \rho^j [r_{t+1+j} - \Delta d_{t+1+j}].$$

• D/P is stationary if dividend growth and returns are stationary.

• J.Y. Campbell thinks that D/P might have a unit root, but
– It should not be explosive
– It should not have a trend (mean change = 0)

• Any return predictability that is not perfectly correlated with dividend predictability will show up in D/P.

Side literature: Dividend Growth Predictability

- Based on the Campbell-Shiller log-linear approximation, log(D/P) depends on future dividend growth and returns.
- Papers have also focused on dividend growth predictability. Cochrane (1992, 2001, 2006), Ang (2002), Lettau and Nieuwerburgh (2006), Ang and Bekaert (2007).
- <u>General finding</u>: aggregate stock returns are predictable by the D/P, not but dividend growth. Dividend growth rate is not predictable by D/P.
- Chen (2008) points out that results are sample dependent (some predictability before WWII) and also affected by the construction of the dividend growth data (with dividends reinvested or not).
- Potential Problem: Measurement error in d_t -too smooth. The observed data may not be the "true" series of interest. IVE: Earnings growth.

Stambaugh Bias
Stambaugh (1999): motivation

- One econometric problem in Fama and French (1989): the regressors are only predetermined but not exogenous.

- Start with predictive regression for returns, r(t+1):
  $r(t+1) = \alpha + \beta\, x(t) + u(t+1)$ (2)
  x(t): D(t)/P(t) –i.e., the dividend price ratio

- x(t) depends on the price at the beginning of t, the change of x at the end of t+1 reflects changes in price from t to t+1, as does r(t+1); $E[u(t+1)|x(t+1),x(t)] \neq 0$, more generally, $E[u(t)\,|x(s),x(w)] \neq 0$, $s<t<w$

- Stambaugh further assumes that
  $x(t) = \theta + \rho\, x(t-1) + v(t)$
  where (u(t), v(t)' follows a $N(0, \Sigma)$, independent across t.

- Results: b (OLS estimate) is biased upward, positively skewed, and has higher variance and kurtosis than the normal sampling distribution of the OLS estimator.

- Stambaugh bias:
  $E(b - \beta)= (\sigma_{uv}/ \sigma_v^2)\, E(p - \rho)$
  where p is the OLS estimator of ρ.

- It turns out p has a downward bias and $\sigma_{uv}$is negative
  => b shows an upward bias.
- Valuation ratios are sufficiently persistent. Conventional t-tests are misleading.

Table 1

Finite-sample properties of $\hat{\beta}$

The table reports finite-sample properties of the ordinary least squares (OLS) estimator $\hat{\beta}$ in the regression

$$y_t = \alpha + \beta x_{t-1} + u_t.$$

The sampling properties are computed under the assumption that x_t obeys the process

$$x_t = \theta + \rho x_{t-1} + v_t,$$

where $\rho^2 < 1$ and $[u_t\ v_t]'$ is distributed $N(0, \Sigma)$, identically and independently across t. The true bias and higher-order moments depend on ρ and Σ (with distinct elements σ_u^2, σ_v^2, and σ_{uv}). For each sample period, those parameters are set equal to the estimates obtained when y_t is the continuously compounded return in month t on the value-weighted NYSE portfolio, in excess of the one-month T-bill return, and x_t is the dividend–price ratio on the value-weighted NYSE portfolio at the end of month t. The moments in the standard setting are conditioned on $x_0, \ldots, x_{T-1}$ and ignore any dependence of u_t on those values. The p-values are associated with a test of $\beta = 0$ versus $\beta > 0$

	Sample period			
	1927–1996	1927–1951	1952–1996	1977–1996
A. True properties				
Bias	0.07	0.18	0.18	0.42
Standard deviation	0.16	0.33	0.27	0.45
Skewness	0.71	0.83	0.98	1.29
Kurtosis	3.84	4.14	4.62	5.83
p-value for $\beta = 0$	0.17	0.42	0.15	0.64
B. Properties in the standard regression setting				
Bias	0	0	0	0
Standard deviation	0.14	0.27	0.20	0.30
Skewness	0	0	0	0
Kurtosis	3	3	3	3
p-value for $\beta = 0$	0.06	0.22	0.02	0.26
C. Sample characteristics and parameter values				
$\hat{\beta}$	0.21	0.21	0.44	0.19
T	840	300	540	240
ρ	0.972	0.948	0.980	0.987
$\sigma_u^2 \times 10^4$	30.05	54.46	16.42	17.50
$\sigma_v^2 \times 10^4$	0.108	0.247	0.029	0.033
$\sigma_{uv} \times 10^4$	− 1.621	− 3.360	− 0.651	− 0.715

• The exact finite-sample moments and p-values in Table 1 depend on ρ and Σ (both unknown in practice). That is, in practice, we cannot know precisely the exact bias of b.

• The finite-sample properties in Table 1 are computed using the values of ρ and Σ obtained in the OLS estimation. (Many of the computations are relatively insensitive to small changes in the parameters.)

• Correcting the bias weakens the predictability evidence.

• **Result from Hodrick (1992) and Kim and Nelson (1993).** In a (1)-(2) framework Newey-West standard errors are not reliable in small samples.

Side Note about Long Horizon Results

- Recall that D/P and other ratios forecast excess returns on stocks. Regression coefficients and R^2 rise with the forecast horizon.
- This is a result of the fact that the forecasting variable is persistent

Consider the stylized model relating returns to a persistent valuation ratio like dividend/price

$$\begin{aligned} r_{t+1} &= \beta x_t + \varepsilon_{t+1},\ \beta > 0 \\ x_{t+1} &= \rho x_t + \eta_{t+1},\ 0 < \rho < 1 \end{aligned}$$

The relationship between $r_{t+2}(2) = r_{t+2} + r_{t+1}$ and x_t is

$$\begin{aligned} r_{t+2}(2) &= \beta x_{t+1} + \beta x_t + \varepsilon_{t+2} + \varepsilon_{t+1} \\ &= \beta(\rho x_t + \eta_{t+1}) + \beta x_t + \varepsilon_{t+2} + \varepsilon_{t+1} \\ &= \beta(1+\rho)x_t + w_{2t} \\ &= \beta_2 x_t + w_t,\ \beta_2 > \beta. \end{aligned}$$

In general,

$$\begin{aligned} r_{t+k}(k) &= \beta(1+\rho+\cdots+\rho^{k-1})x_t + w_{kt} \\ &= \beta_k x_t + w_{kt},\ \beta_k > \beta_{k-1} \end{aligned}$$

Recent Contributions

Baker and Wurgler, JF (2000)

- Baker and Wurgler (2000) also use a predictive regression framework. They add to the Fama and French (1989) variables B/M_{t-1}, S_{t-1} (equity share in new issues), and lagged R_{t-1}.
- S_{t-1} is a the most consistent and negative predictor of future returns.
- The S_{t-1} coefficient is over 20 times too large to be due to MM leverage effect: New issues represent only a small fraction of outstanding capital. Not enough influence on aggregate leverage to change expected returns.
- S_{t-1} could be related to future returns through investment, but in the aggregate investment is essentially unrelated to subsequent aggregate returns
- They discuss the Stambaugh bias, with a lukewarm approach.

Table V

Multivariate OLS Regressions for Predicting One-Year-Ahead Market Returns

OLS regressions of annual real equity market returns on multiple predictors:

$$R_{Et} = a + b_1 R_{Et-1} + b_2 BILL_{t-1} + b_3 TERM_{t-1} + b_4 D/P_{t-1} + b_5 B/M_{t-1} + b_6 S_{t-1} + u_t$$

where R_E denotes real percentage returns on the CRSP value-weighted (VW) or equal-weighted (EW) portfolio, *BILL* denotes the return on Treasury bills, *TERM* denotes the yield premium of long-term government bonds over treasuries, *D/P* denotes the dividend-to-price ratio, *B/M* denotes the book-to-market ratio, and *S* denotes the equity share in new issues. The dividend-to-price ratio, the book-to-market ratio, and the equity share are standardized to have zero mean and unit variance. *t*-statistics are shown in brackets using heteroskedasticity robust standard errors.

	1928–1997 Returns		1928–1962 Returns		1963–1997 Returns	
	VW CRSP	EW CRSP	VW CRSP	EW CRSP	VW CRSP	EW CRSP
Intercept	6.95	21.72	14.33	21.71	11.50	19.23
	[1.13]	[1.68]	[0.53]	[0.76]	[0.78]	[1.16]
R_E	0.05	0.08	0.27	0.20	−0.20	−0.09
	[0.39]	[0.82]	[1.12]	[1.09]	[−1.01]	[−0.68]
BILL	0.71	−0.85	4.96	9.60	0.66	4.20
	[0.89]	[−0.47]	[0.75]	[1.28]	[0.40]	[1.46]
TERM	−0.86	−3.66	−7.98	−10.86	0.15	6.09
	[−0.41]	[−0.96]	[−0.70]	[−0.84]	[0.08]	[1.45]
D/P	4.26	−1.58	−4.37	−9.17	14.51	63.21
	[1.13]	[−0.27]	[−0.51]	[−1.55]	[1.43]	[2.41]
B/M	1.51	13.50	19.59	34.10	−7.30	−14.30
	[0.38]	[2.38]	[1.99]	[6.34]	[−1.29]	[−1.47]
S	−7.88	−13.17	−8.84	−14.34	−8.27	−13.63
1	[−3.97]	[−3.77]	[−1.94]	[−2.21]	[−2.13]	[−2.48]
$\bar{R}^2$	0.12	0.28	0.27	0.51	0.12	0.29
N	70	70	35	35	35	35

Dealing with Stambaugh Bias

Lewellen (2004), JFE

- Lewellen (2004) conditions on estimated persistence and worst possible case for true persistence. Worst case: ρ=1.

$$E(b - \beta \mid \rho=1,p) = (\sigma_{uv}/ \sigma_v^2)\ (p - 1)$$

$$b_{adj} = b - (\sigma_{uv}/ \sigma_v^2)\ (p - 1)$$

- Estimated persistence is very close to one. The bias correction is small. Predictability survives:
  - D/P predicts market returns from 1946–2000 and sub-samples.
  - B/M and E/P predict returns during the shorter sample 1963–2000.

Dealing with Stambaugh Bias

Campbell and Yogo (2006), JFE

- Known result: Conventional t-test (when ρ is unknown) has good large-sample properties when x_t is I(0).
- But, even if the predictor variable is I(0), first-order asymptotics can be a poor approximation in finite samples when ρ is close to one.
- Two approaches with persistent regressors:
  - exact finite-sample theory (assume normality): Evans and Savin (1981, 1984) and Stambaugh (1999)
  - local-to-unity asymptotics (largest root is modeled as ρ=1+c/T, with c constant): Elliot and Stock (1994), Campbell and Yogo (2006).
- Propose a Q-statistics:

$$Q(\beta_0, \rho) = \frac{\sum_{t=1}^{T} x_{t-1}^{\mu}[r_t - \beta_0 x_{t-1} - \beta_{ue}(x_t - \rho x_{t-1})]}{\sigma_u (1-\delta^2)^{1/2} (\sum_{t=1}^{T} x_{t-1}^{\mu 2})^{1/2}}$$

where $\beta_{ue}=\sigma_{ue}/\sigma^2_u$. If we estimate ρ, then we can write:

$$Q(\beta_0, \rho) = \frac{(\widehat{\beta} - \beta_0) - \beta_{ue}(\widehat{\rho} - \rho)}{\sigma_u (1-\delta^2)^{1/2} (\sum_{t=1}^{T} x_{t-1}^{\mu 2})^{-1/2}}.$$

• Under the null of stationarity, the test is asymptotically normal. However, it is infeasible, since it depends on ρ –or in c= T(ρ-1)- and Σ. But, it can be easily made feasible (correcting for size, since we use α_1 to test ρ and α_2 to test β(ρ) , using Bonferroni bounds):

$$\beta(\rho) = \frac{\sum_{t=1}^{T} x_{t-1}^{\mu}[r_t - \beta_{ue}(x_t - \rho x_{t-1})]}{\sum_{t=1}^{T} x_{t-1}^{\mu 2}},$$

$$\underline{\beta}(\rho, \alpha_2) = \beta(\rho) - z_{\alpha_2/2}\sigma_u \left(\frac{1-\delta^2}{\sum_{t=1}^{T} x_{t-1}^{\mu 2}}\right)^{1/2},$$

$$\overline{\beta}(\rho, \alpha_2) = \beta(\rho) + z_{\alpha_2/2}\sigma_u \left(\frac{1-\delta^2}{\sum_{t=1}^{T} x_{t-1}^{\mu 2}}\right)^{1/2},$$

$$C_\beta(\alpha) = [\underline{\beta}(\overline{\rho}(\overline{\alpha}_1), \alpha_2), \overline{\beta}(\underline{\rho}(\underline{\alpha}_1), \alpha_2)].$$

Lewellen's (2004) test is a special case, when ρ is known.

• Note, that first we need to estimate ρ. For this, we can use the ADF test or the better behaved DF-GLS –see Elliot and Stock (1994).

• Campbell and Yogo (2006) compare the power of their Q-test to the Bonferroni t-test of Cavanagh et al. (1995) and Lewellen's (2004). The Bonferroni Q-test has good power over the other feasible tests.

• Findings:
  • Over the full sample, E/P has predictive power at various frequencies (annual to monthly), while D/P only at annual frequency. In the post-1952 sample, results are weaker. If we can rule out explosive root, D/P has predictive power.

  • t-test leads to valid inference for the interest rate data. Persistence is not a problem for interest rate variables because their innovations have low correlation with the innovations to stock returns.

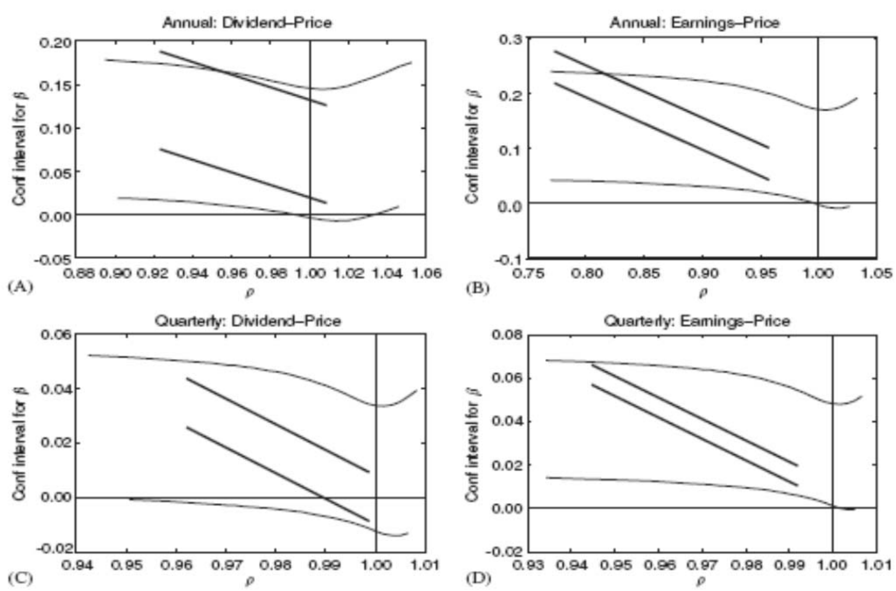


Fig. 4. Bonferroni confidence interval for the valuation ratios. This figure plots the 90% confidence interval for β over the confidence interval for ρ. The significance level for ρ is chosen to result in a 90% Bonferroni confidence interval for β. The thick (thin) line is the confidence interval for β computed by inverting the Q-test (t-test). Returns are for the annual and quarterly CRSP value-weighted index (1926–2002). The predictor variables are the log dividend–price ratio and the log earnings–price ratio.

Table 5
Tests of predictability

Series	Variable	t-stat	$\hat{\beta}$	90% CI: β		Low CI β
				t-test	Q-test	($\rho = 1$)
Panel A: S&P 1880–2002, CRSP 1926–2002						
S&P 500	d–p	1.967	0.093	[−0.040, 0.136]	[−0.033, 0.114]	−0.017
	e–p	2.762	0.131	[−0.003, 0.189]	**[0.042, 0.224]**	−0.023
Annual	d–p	2.534	0.125	[−0.007, 0.178]	**[0.014, 0.188]**	0.020
	e–p	2.770	0.169	[−0.009, 0.240]	**[0.042, 0.277]**	0.002
Quarterly	d–p	2.060	0.034	[−0.014, 0.052]	[−0.009, 0.044]	−0.010
	e–p	2.908	0.049	[−0.001, 0.068]	**[0.010, 0.066]**	0.002
Monthly	d–p	1.706	0.009	[−0.006, 0.014]	[−0.005, 0.010]	−0.005
	e–p	2.662	0.014	[−0.001, 0.019]	**[0.002, 0.018]**	0.001
Panel B: S&P 1880–1994, CRSP 1926–1994						
S&P 500	d–p	2.233	0.141	[−0.035, 0.217]	[−0.048, 0.183]	−0.081
	e–p	3.321	0.196	**[0.062, 0.272]**	**[0.093, 0.325]**	−0.030
Annual	d–p	2.993	0.212	**[0.025, 0.304]**	**[0.056, 0.332]**	0.011
	e–p	3.409	0.279	**[0.048, 0.380]**	**[0.126, 0.448]**	0.012
Quarterly	d–p	2.304	0.053	[−0.004, 0.083]	[−0.006, 0.076]	−0.027
	e–p	3.506	0.079	**[0.018, 0.107]**	**[0.027, 0.109]**	0.005
Monthly	d–p	1.790	0.013	[−0.004, 0.022]	[−0.007, 0.017]	−0.013
	e–p	3.185	0.022	**[0.002, 0.030]**	**[0.005, 0.028]**	0.000
Panel C: CRSP 1952–2002						
Annual	d–p	2.289	0.124	[−0.023, 0.178]	[−0.007, 0.183]	0.020
	e–p	1.733	0.114	[−0.078, 0.178]	[−0.031, 0.229]	−0.025
	r_1	−1.143	−0.095	[−0.229, 0.045]	[−0.231, 0.042]	—
	y–r_1	1.124	0.136	[−0.087, 0.324]	[−0.075, 0.359]	−0.156
Quarterly	d–p	2.236	0.036	[−0.011, 0.051]	[−0.010, 0.030]	0.005
	e–p	1.777	0.029	[−0.019, 0.044]	[−0.012, 0.042]	−0.003
	r_1	−1.766	−0.042	**[−0.084, −0.004]**	**[−0.084, −0.004]**	−0.086
	y–r_1	1.991	0.090	**[0.009, 0.162]**	**[0.006, 0.158]**	−0.002
Monthly	d–p	2.259	0.012	[−0.004, 0.017]	[−0.004, 0.010]	0.001
	e–p	1.754	0.009	[−0.006, 0.014]	[−0.004, 0.012]	−0.001
	r_1	−2.431	−0.017	**[−0.030, −0.006]**	**[−0.030, −0.006]**	−0.030
	y–r_1	2.963	0.047	**[0.020, 0.072]**	**[0.020, 0.072]**	0.016

This table reports statistics used to infer the predictability of returns. Returns are for the annual S&P 500 index and the annual, quarterly, and monthly CRSP value-weighted index. The predictor variables are the log dividend–price ratio (d–p), the log earnings-price ratio (e–p), the three-month T-bill rate (r_1), and the long-short yield spread (y–r_1). The third and fourth columns report the t-statistic and the point estimate $\hat{\beta}$ from an OLS regression of returns onto the predictor variable. The next two columns report the 90% Bonferroni confidence intervals for β using the t-test and Q-test, respectively. Confidence intervals that reject the null are in bold. The final column reports the lower bound of the confidence interval for β based on the Q-test at $\rho = 1$.

Understanding D/P: Back to the Gordon Growth Model – Campbell and Thompson (2007), RFS

- Assume, as in the Gordon growth model, that the dividend is known one period in advance.
- Assume that the log(D/P) follows a RW with normal innovations.
- Assume that the two-period ahead dividend growth rate is conditionally normal.

  - $D_{t+1}/P_t = \exp(x_t)$

  - $D_{t+1}/D_t = 1 + G_t = \exp(g_t)$

  - $x_t = x_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim N(0,\sigma^2_\varepsilon)$
- Use the definition of return and the formula for the conditional expectation of a lognormal R.V. --$E(Y) = \exp(\mu+\sigma^2/2)$

  $1 + R_{t+1} = (P_{t+1} + D_{t+1})/P_t = D_{t+1}/P_t + (D_{t+2}/D_{t+1})(D_{t+1}/P_t)\,(D_{t+2}/P_{t+1})^{-1}$

  $= \exp(x_t)[1+\exp(g_{t+1} - x_{t+1})]$

$$E_t[1 + R_{t+1}] = \exp(x_t)[1+E_t[\exp(g_{t+1} - x_{t+1})]]$$
$$= D_{t+1}/P_t + \exp(E_t[g_{t+1}]-0)\exp(\sigma^2{}_g/2 + \sigma^2{}_x/2 - \sigma_{gx})$$
$$= D_{t+1}/P_t + \exp(E_t[g_{t+1}])\exp(Var_t(p_{t+1}-p_t)/2)$$
$$\approx D_{t+1}/P_t + \exp(E_t[g_{t+1}]) + ½\, Var_t(r_{t+1})$$

• As in the Gordon model, the expected return is the level of D/P (not the log) plus expected dividend growth.

• The variance effect is subtle:
  – In the original Gordon model, returns and dividend growth have the same volatility.
  – In that case the expected return is level of D/P plus arithmetic average dividend growth.
  – In the data, stock returns are much more volatile.
  – In that case the expected return is level of D/P plus geometric average dividend growth plus one-half stock return volatility (not dividend volatility).

- Equivalently, the level of D/P plus geometric average dividend growth predicts the log stock return (instead of the simple stock return).

- Empirically, this approach has the advantage that we do not have to estimate the unconditional mean stock return from the noisy historical data

- Instead, we can use historical average growth, along with the current level of D/P.