Return Predictability Under the Alternative

Marco Rossi [1] Tim Simin [2] Daniel Smith[3]

[1]University of Notre Dame

[2]The Pennsylvania State University

[3]Queensland University of Technology

Second ITAM Finance Conference
June 8, 2013

Predictive Regressions

General Setup

$$
\begin{aligned}
y_{t+1} &= \alpha + \beta x_t + e_{t+1} \\
x_{t+1} &= (1-\rho)\mu_x + \rho x_t + u_{t+1}.
\end{aligned}
$$

NOTE:

1. dividend yield (DY) is "king" among predictors
2. predictors are typically persistent, i.e. high ρ
3. β is estimated with a lot of noise
4. $cov(e, u)$ is highly negative in the data (-0.712 for the DY with annual data)

Are Returns Predictable?

1. Optimists
   - evidence of long-horizon predictability (Fama and French, 1988)
   - “the dog doesn’t bark", i.e. dividend growth is not predictable, so returns must be (Cochrane, 2008)
   - more power at longer horizons (Campbell, 2001)
   - dividend yield predicts returns even in the presence of small sample bias (Lewellen, 2004)
   - macro fundamentals predict exchange rates (Mark (1995))
2. Skeptical
   - long-horizon predicability is a “myth": under the null, slope estimates are highly correlated (Boudouk and Richardson, 1994; Boudoukh, Richardson, and Whitelaw, 2006)
   - commonly used standard errors have poor size; some predictability only at short horizon (Ang and Bekaert, 2007)
   - lagged dependent variable problem (Goetzmann and Jorion, 1993)
   - Berkowitz and Giorgianni (2001) revisit Mark (1995) on exchange rates

Long-Horizon Predictability

"... researchers should be equally impressed by the short- and long-horizon evidence for the simple reason that the regressions are almost perfectly correlated. For an autocorrelation of 0.953 for annual dividend yields, we show analytically that the 1-year and 2-year predictive estimators are 98.8% correlated **under the null hypothesis** of no predictability."[a]

- Boudoukh, Richardson, and Whitelaw (2006, p.1578)

[a] The bold/increased font is mine.

Long-Horizon Predictability

"... researchers should be equally impressed by the short- and long-horizon evidence for the simple reason that the regressions are almost perfectly correlated. For an autocorrelation of 0.953 for annual dividend yields, we show analytically that the 1-year and 2-year predictive estimators are 98.8% correlated **under the null hypothesis** of no predictability."[a]

- Boudoukh, Richardson, and Whitelaw (2006, p.1578)

[a] The bold/increased font is mine.

Our paper provides

1. analytical standard errors derived under **any** hypothesis
2. extensive size and power analysis
3. revised empirical results in BRW

Preview of Findings

1. analytical standard errors derived under the alternative generate power gains
2. our theory suggests that power gains arise only if the alternative is given a chance
3. in practice, power gains are greatly reduced because of greater estimation risk and non-standard distribution or $\hat{\beta}$

Predictive Regressions with Overlapping Observations: Setup

Long-horizon returns ($y_t^N = \sum_{i=1}^N y_{t+i}$) as regressand

$$
\begin{aligned}
y_t^N &= \sum_{i=1}^{N} (\alpha + \beta_1 x_{t+i-1} + e_{t+i}) \\
&= \alpha_N + \beta_N x_t + \xi_t^N
\end{aligned}
$$

Predictive Regressions with Overlapping Observations: Setup

Long-horizon returns ($y_t^N = \sum_{i=1}^N y_{t+i}$) as regressand

$$\begin{aligned} y_t^N &= \sum_{i=1}^{N} (\alpha + \beta_1 x_{t+i-1} + e_{t+i}) \\ &= \alpha_N + \beta_N x_t + \xi_t^N \end{aligned}$$

Mapping with single-horizon regression

1. $\alpha_N = N\alpha + \mu_x(N\beta - \beta_N)$
2. $\beta_N = \beta\frac{1-\rho^N}{1-\rho}$
3. $\xi_t^N = \sum_{i=1}^{N-1} \beta_{N-i} u_{t+i} + \sum_{i=1}^{N} e_{t+i}$

System of Predictive Regressions

Following Boudoukh, Richardson, and Whitelaw (2006)

$$
\begin{aligned}
y_{t+1} &= \alpha_1 + \beta_1 x_t + \xi^1_{t+1} \\
y^2_{t+1} &= \alpha_2 + \beta_2 x_t + \xi^2_{t+1} \\
&\vdots \\
y^n_{t+1} &= \alpha_n + \beta_n x_t + \xi^n_{t+1} \\
&\vdots \\
y^N_{t+1} &= \alpha_N + \beta_N x_t + \xi^N_{t+1}
\end{aligned}
$$

1. estimate slopes with separate OLS regressions
2. use GMM tools to get covariance of all slopes
3. this allows to get analytical correlations of slopes and implement WALD tests

Spectral Density and Standard Error

Spectral density matrix (using GMM)

$$S = \sigma_e^2 \times f(\mu_x, \sigma_x^2, \rho) \qquad (BRW)$$

Spectral Density and Standard Error

Spectral density matrix (using GMM)

$$
\begin{aligned}
S \quad = \quad & \sigma_e^2 \times f(\mu_X, \sigma_X^2, \rho) && (BRW) \\
& \sigma_u^2 \times f(\mu_X, \sigma_X^2, \rho, \beta) && (2 \textit{ extra terms}) \\
& \sigma_{eu} \times f(\mu_X, \sigma_X^2, \rho, \beta)
\end{aligned}
$$

Spectral Density and Standard Error

Spectral density matrix (using GMM)

$$S = \begin{array}{ll} \sigma_e^2 \times f(\mu_x, \sigma_x^2, \rho) & (BRW) \\ \sigma_u^2 \times f(\mu_x, \sigma_x^2, \rho, \beta) & (2 \text{ extra terms}) \\ \sigma_{eu} \times f(\mu_x, \sigma_x^2, \rho, \beta) & \end{array}$$

Covariance matrix

$$V([\alpha_1, \beta_1, \ldots, \beta_N]) = (D' S^{-1} D)^{-1}$$

where D is the derivative matrix of the moment conditions.

Some Intuition

Example

N=2

$$SE(\beta_2) = \sqrt{2\frac{\sigma_e^2}{\sigma_x^2}(1+\rho) + \frac{\sigma_u^2}{\sigma_x^2}\beta^2 + \frac{\sigma_{eu}}{\sigma_x^2}\beta(1+2\rho)}$$

1. if you impose the null ($\beta = 0$), you get BRW
2. SEs under the alternative are not always better
3. if innovations are uncorrelated ($\sigma_{eu} = 0$) power deteriorates
4. if $\sigma_{eu} < 0$, we can get lower SEs, i.e. more power

Correlation of Slopes Across Horizons

1. if $\beta = 0$ (H_0), slopes are highly correlated
2. if $\beta > 0$ (H_1), correlation decays faster
3. persistence (ρ) affects speed of decay

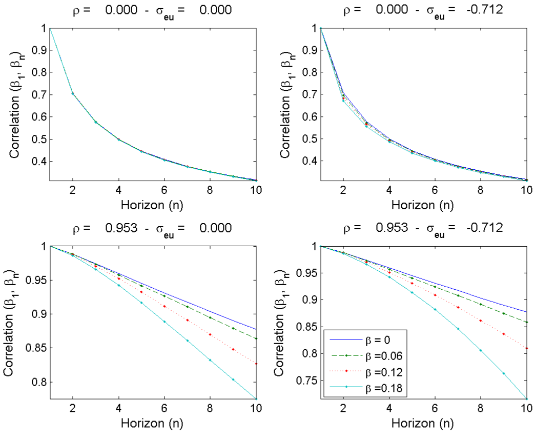

Figure: Correlations of β_1 and β_n

Relative Efficiency Across Horizons

Correlation is Key

1. if $\sigma_{eu} = 0$, imposing H_0 is better
2. if $\sigma_{eu} \neq 0$, **not** imposing H_0 might be better
3. in the data $\sigma_{eu} \approx -1$

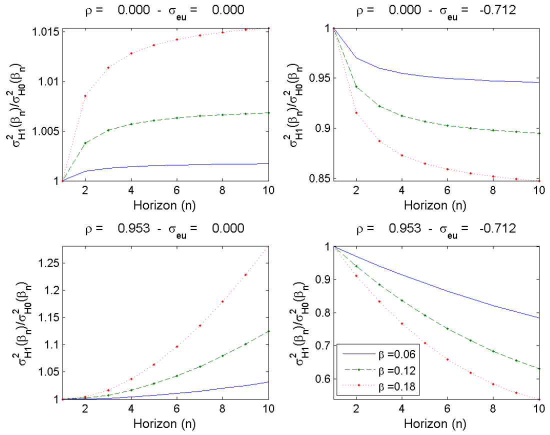

Figure: Ratio of Slope Variances (H_1 Vs H_0)

Simulation Under the Null

Data Generating Process

$$\begin{aligned} y_{t+1} &= e_{t+1} \\ x_{t+1} &= \rho x_t + u_{t+1} \end{aligned}$$

Yearly Case (BRW, 2006)

T=75

$\sigma_e^2 = 0.212$, $\sigma_u^2 = 0.154$,,

$\sigma_{eu} = -0.712$, $\rho = 0.953$

#simulations=50,000

Monthly Case (Barberis, 2000)

T=1000

$\sigma_e^2 = 0.0017$, $\sigma_u^2 = 0.000003$,

$\sigma_{eu} = -0.9351$, $\rho = 0.9774$

#simulations=50,000

Simulated Estimates and Small Sample Bias (yearly)

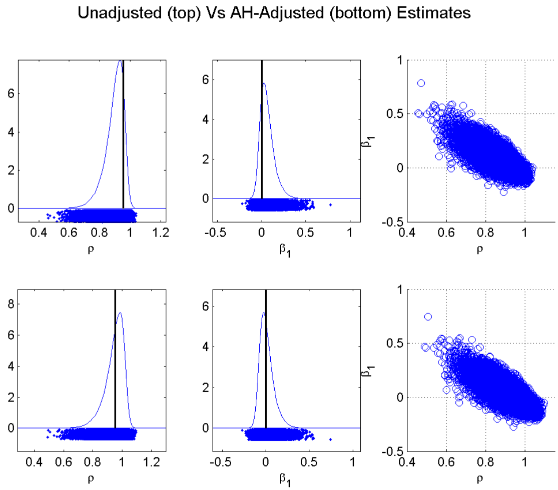

Figure: Distribution of Slope and Persistence Estimates

Simulated Estimates and Small Sample Bias (monthly)

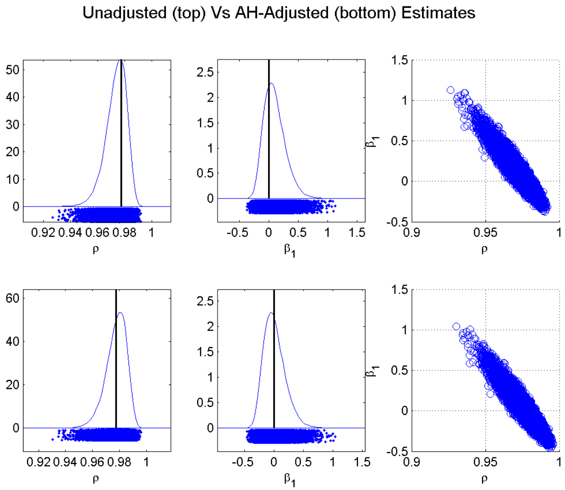

Figure: Distribution of Slope and Persistence Estimates

Rejection Probabilities Under the Null ($\sigma_{eu} = -0.712$)

n	β		$t - stat$						
	mean	SD	OLS	NW	HH	HOD	BRW	H1	H1-AH
				ρ =0.000					
1	0.013	0.159	5.2	7.0	6.7	4.9	4.9	5.3	5.7
			[-1.88, 2.06]	[-2.03, 2.22]	[-2.00, 2.19]	[-1.87, 2.03]	[-1.86, 2.03]	[-1.90, 2.07]	[-1.92, 2.10]
3	0.039	0.271	5.6	9.1	12.0	5.6	4.8	5.3	5.3
			[-1.77, 2.16]	[-1.85, 2.63]	[-1.86, 3.22]	[-1.81, 2.15]	[-1.79, 2.04]	[-1.69, 2.17]	[-1.69, 2.19]
5	0.064	0.342	5.7	11.2	16.8	6.3	4.1	4.7	4.7
			[-1.72, 2.21]	[-1.80, 2.86]	[-1.96, 3.99]	[-1.77, 2.22]	[-1.80, 1.94]	[-1.70, 2.07]	[-1.69, 2.08]
10	0.122	0.463	6.5	16.1	23.9	8.6	3.5	3.9	3.8
			[-1.60, 2.33]	[-1.78, 3.32]	[-2.22, 5.26]	[-1.77, 2.44]	[-1.78, 1.85]	[-1.70, 1.95]	[-1.68, 1.95]
				ρ =0.953					
1	0.058	0.080	9.6	12.1	11.6	8.9	9.3	10.0	10.3
			[-1.21, 2.60]	[-1.31, 2.82]	[-1.29, 2.78]	[-1.28, 2.45]	[-1.23, 2.53]	[-1.23, 2.62]	[-1.24, 2.65]
3	0.159	0.212	34.0	24.1	17.8	9.6	9.1	12.0	10.8
			[-2.03, 4.52]	[-1.68, 3.89]	[-1.46, 3.47]	[-1.29, 2.50]	[-1.20, 2.52]	[-1.17, 2.88]	[-1.12, 2.80]
5	0.245	0.321	45.8	29.0	23.9	10.1	8.6	13.0	10.7
			[-2.56, 5.80]	[-1.84, 4.54]	[-1.62, 4.33]	[-1.29, 2.54]	[-1.17, 2.45]	[-1.11, 2.95]	[-1.01, 2.79]
10	0.411	0.528	59.3	37.6	37.3	11.6	6.5	13.0	8.8
			[-3.41, 8.15]	[-2.17, 5.93]	[-2.13, 7.37]	[-1.32, 2.67]	[-1.11, 2.30]	[-1.04, 2.89]	[-0.85, 2.61]

1. slope estimates are biased upward
2. all tests have wrong size, but HOD and BRW do better
3. power calculations require size-adjustment
4. monthly simulations are better, but size-adjustment still necassary

Simulation Under the Alternative

Data Generating Process[a]

[a] Other parameters are equal to the null case.

$$
\begin{aligned}
y_{t+1} &= \beta_1 x_t + e_{t+1}, \quad \beta_1 > 0 \\
x_{t+1} &= \rho x_t + u_{t+1}
\end{aligned}
$$

1. we consider only "promising" standard errors
2. rejection cutoffs obtained from size analysis
3. that is, we raise the bar for our standard error

Distribution of t-tests (yearly case)[1]

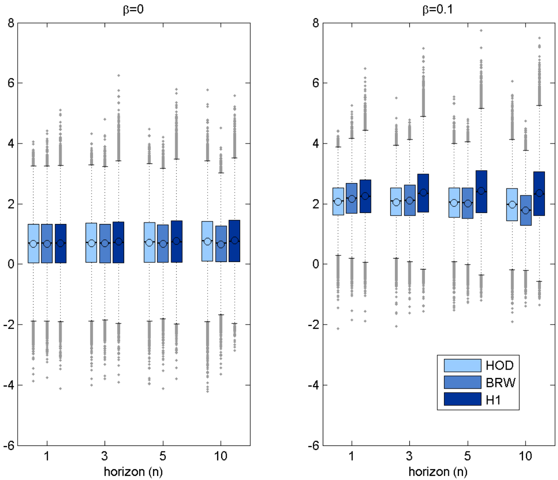

Figure: Distribution Under the Null and Alternative

Power Gains (yearly case)

1. gains in the order of 5-20%
2. for small betas, slight power deterioration
3. relative gains arise mostly at longer horizons, i.e. N=10

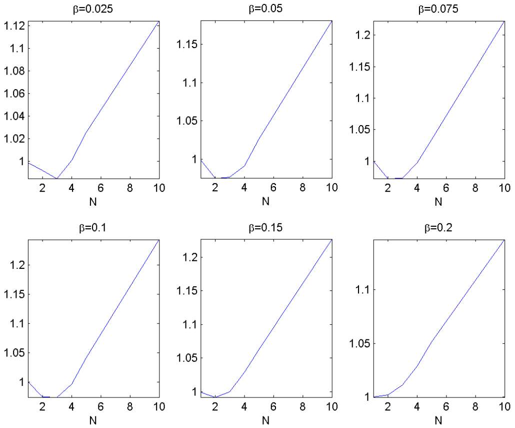

Figure: Power Gains Relative to BRW

Power Gains (monthly case)

1. gains in the order of 100%
2. no power deterioration
3. relative gains arise at very short horizons already

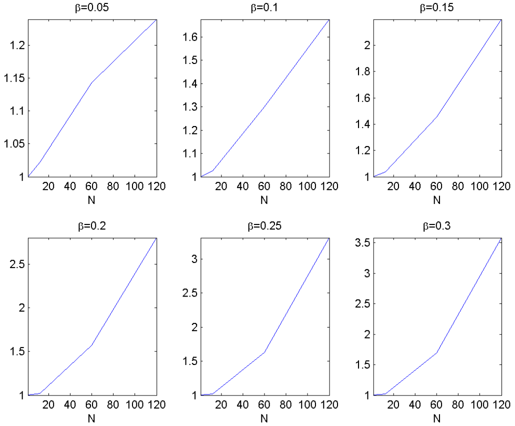

Figure: Power Gains Relative to BRW

Empirical Evidence

1. we replicate BRW (yearly frequency)
2. additionally, we consider Hodrick and proposed (H1) standard errors
3. unlike BRW, we also consider N=10 years
4. tables report both asymptotic and simulated p-values

Predictors

1. Log dividend yield, CRSP VW
2. Log payout yield, CRSP VW, cash flow
3. Log payout yield, CRSP VW, Treasury stock
4. Log net payout yield, CRSP VW, cash flow
5. Log earnings yield, S&P 500
6. Default yield spread
7. Term yield spread
8. Log book-to-market ratio
9. Equity share of new issuances
10. Risk-free rate

Empirical Application

Log dividend yield, CRSP VW (ρ= 0.951 - σ_{eu}= -0.712)

		HOD			BRW			H1			
n	β	SE	Asy. p	Sim. p	SE	Asy. p	Sim. p	SE	Asy. p	Sim. p	R^2
1	0.15	0.062	0.015	0.026	0.060	0.011	0.024	0.057	0.009	0.024	0.08
2	0.25	0.118	0.033	0.061	0.118	0.032	0.061	0.109	0.021	0.058	0.11
3	0.33	0.168	0.047	0.088	0.175	0.057	0.099	0.155	0.032	0.088	0.14
4	0.40	0.213	0.062	0.115	0.231	0.084	0.137	0.198	0.044	0.114	0.17
5	0.46	0.254	0.068	0.128	0.286	0.105	0.162	0.237	0.050	0.128	0.21
10	0.84	0.384	0.028	0.074	0.549	0.124	0.168	0.399	0.034	0.101	0.43

1. slope estimates and R^2 increase with horizon
2. relative to BRW, log dividend yield more significant with H1
3. at yearly frequency, log dividend yield still not significant

Other payout measures are significant

Predictors in Boudoukh, Michaely, Richardson, and Roberts (2007)

		HOD			BRW			H1			
n	β	SE	Asy. p	Sim. p	SE	Asy. p	Sim. p	SE	Asy. p	**Sim. p**	R^2
Log payout yield, CRSP VW, cash flow (ρ= 0.862 - σ_{eu}= -0.698)											
1	0.24	0.084	0.004	0.018	0.083	0.004	0.016	0.078	0.002	**0.017**	0.11
2	0.41	0.158	0.010	0.044	0.159	0.010	0.043	0.142	0.004	**0.041**	0.16
3	0.54	0.222	0.015	0.064	0.233	0.020	0.073	0.196	0.006	**0.063**	0.20
4	0.63	0.278	0.023	0.085	0.303	0.036	0.102	0.243	0.009	**0.080**	0.23
5	0.73	0.337	0.030	0.096	0.370	0.049	0.123	0.285	0.011	**0.089**	0.28
10	1.35	0.556	0.015	0.053	0.666	0.043	0.125	0.469	0.004	**0.058**	0.60
Log payout yield, CRSP VW, Treasury stock (ρ= 0.908 - σ_{eu}= -0.725)											
1	0.21	0.077	0.007	0.020	0.073	0.004	0.019	0.069	0.003	**0.019**	0.11
2	0.35	0.145	0.016	0.048	0.142	0.014	0.049	0.128	0.006	**0.048**	0.15
3	0.46	0.204	0.026	0.072	0.210	0.030	0.082	0.179	0.011	**0.074**	0.18
4	0.54	0.257	0.035	0.095	0.275	0.048	0.115	0.223	0.015	**0.096**	0.21
5	0.63	0.310	0.041	0.107	0.338	0.061	0.137	0.263	0.016	**0.107**	0.27
10	1.16	0.485	0.017	0.060	0.629	0.066	0.140	0.431	0.007	**0.078**	0.55
Log net payout yield, CRSP VW, cash flow (ρ= 0.658 - σ_{eu}= -0.303)											
1	0.76	0.274	0.005	0.013	0.176	0.000	0.011	0.154	0.000	**0.011**	0.25
2	1.36	0.518	0.009	0.032	0.321	0.000	0.032	0.269	0.000	**0.024**	0.39
3	1.54	0.616	0.013	0.050	0.450	0.001	0.056	0.380	0.000	**0.034**	0.37
4	1.55	0.670	0.021	0.066	0.565	0.006	0.081	0.491	0.002	**0.043**	0.31
5	1.55	0.609	0.011	0.076	0.669	0.021	0.100	0.602	0.010	**0.047**	0.28
10	2.33	0.859	0.007	0.041	1.080	0.031	0.103	1.105	0.035	**0.024**	0.40

Using Earnings Instead of Dividends

Log earnings yield, S&P 500 (ρ= 0.799 - σ_{eu}= -0.583)

		HOD			BRW			H1			
n	β	SE	Asy. p	Sim. p	SE	Asy. p	Sim. p	SE	Asy. p	**Sim. p**	R^2
1	0.13	0.055	0.016	0.015	0.061	0.031	0.013	0.059	0.027	**0.013**	0.06
2	0.22	0.098	0.024	0.039	0.115	0.055	0.036	0.108	0.040	**0.032**	0.08
3	0.29	0.141	0.040	0.057	0.166	0.082	0.062	0.149	0.053	**0.048**	0.10
4	0.36	0.182	0.050	0.076	0.214	0.095	0.091	0.187	0.056	**0.061**	0.13
5	0.38	0.214	0.075	0.087	0.258	0.141	0.111	0.221	0.085	**0.068**	0.14
10	0.70	0.337	0.038	0.048	0.447	0.118	0.112	0.364	0.054	**0.037**	0.29

1. always significant with H1
2. this measure is very noisy at the yearly frequency
3. Campbell and Shiller (1988) suggests using long historical averages of real earnings (10 or even 30 years)
4. this extension will be in the next draft

Bond Market Predictors

n	β	HOD SE	HOD Asy. p	HOD Sim. p	BRW SE	BRW Asy. p	BRW Sim. p	H1 SE	H1 Asy. p	H1 Sim. p	R^2
Default yield spread (ρ= 0.788 - σ_{eu}= -0.554)											
1	1.60	3.758	0.671	0.015	3.003	0.595	0.013	3.018	0.597	**0.013**	0.00
2	5.49	8.143	0.500	0.037	5.680	0.334	0.036	5.649	0.331	**0.031**	0.02
3	7.25	10.503	0.490	0.056	8.168	0.374	0.062	8.052	0.368	**0.047**	0.03
4	9.99	13.305	0.453	0.073	10.489	0.341	0.089	10.262	0.330	**0.060**	0.04
5	12.47	15.589	0.424	0.083	12.660	0.325	0.107	12.307	0.311	**0.064**	0.06
10	12.14	18.541	0.513	0.046	21.775	0.577	0.109	20.722	0.558	**0.035**	0.04
1	2.43	1.853	0.190	0.013	1.767	0.169	0.011	1.757	0.167	**0.011**	0.02
2	4.66	3.045	0.126	0.032	3.134	0.137	0.032	3.100	0.133	**0.022**	0.04
3	6.54	4.103	0.111	0.049	4.309	0.129	0.056	4.261	0.125	**0.032**	0.06
4	9.37	4.555	0.040	0.067	5.338	0.079	0.081	5.283	0.076	**0.040**	0.11
5	10.75	4.891	0.028	0.075	6.252	0.086	0.099	6.198	0.083	**0.044**	0.13
10	11.50	7.608	0.131	0.042	9.770	0.239	0.104	9.749	0.238	**0.025**	0.09
Risk-free rate (ρ= 0.898 - σ_{eu}= 0.055)											
1	-0.94	0.796	0.237	0.013	0.765	0.218	0.011	0.763	0.216	**0.011**	0.02
2	-1.54	1.387	0.266	0.033	1.491	0.301	0.031	1.484	0.299	**0.023**	0.03
3	-2.36	1.929	0.221	0.050	2.192	0.281	0.055	2.182	0.279	**0.034**	0.05
4	-3.35	2.430	0.168	0.068	2.870	0.244	0.083	2.859	0.242	**0.043**	0.08
5	-4.11	2.943	0.163	0.078	3.526	0.244	0.102	3.517	0.243	**0.046**	0.11
10	-6.16	5.057	0.223	0.041	6.512	0.344	0.108	6.584	0.349	**0.023**	0.15

“Sentiment” Measures

		HOD			BRW			H1			
n	β	SE	Asy. p	Sim. p	SE	Asy. p	Sim. p	SE	Asy. p	**Sim. p**	R^2
Log book-to-market ratio (ρ= 0.900 - σ_{eu}= -0.858)											
1	0.11	0.053	0.035	0.024	0.048	0.019	0.022	0.047	0.016	**0.022**	0.07
2	0.20	0.100	0.041	0.056	0.094	0.029	0.056	0.086	0.017	**0.058**	0.12
3	0.27	0.140	0.054	0.081	0.138	0.052	0.091	0.120	0.025	**0.089**	0.15
4	0.32	0.170	0.063	0.107	0.181	0.081	0.128	0.150	0.035	**0.117**	0.17
5	0.32	0.192	0.092	0.119	0.222	0.145	0.151	0.175	0.065	**0.132**	0.16
10	0.45	0.273	0.097	0.068	0.411	0.269	0.150	0.267	0.089	**0.106**	0.20
Equity share of new issuances (ρ= 0.475 - σ_{eu}= 0.152)											
1	-0.69	0.308	0.026	0.013	0.221	0.002	0.011	0.208	0.001	**0.011**	0.13
2	-1.13	0.604	0.060	0.032	0.380	0.003	0.032	0.355	0.001	**0.022**	0.17
3	-1.29	0.776	0.098	0.051	0.511	0.012	0.058	0.484	0.008	**0.033**	0.16
4	-1.32	0.775	0.088	0.067	0.624	0.034	0.082	0.599	0.027	**0.043**	0.14
5	-1.17	0.666	0.079	0.078	0.722	0.106	0.101	0.703	0.097	**0.045**	0.10
10	-0.80	0.736	0.278	0.039	1.097	0.467	0.103	1.106	0.470	**0.025**	0.03

Conclusion

Give the alternative a chance!

1. at year frequency, we have substantial improvement with respect to BRW
2. results here are lower bound, since power gains are higher at monthly frequency
3. STATISTICAL CATCH 22
   1. power gains if both conditions are met
      - tests are derived under the alternative
      - correlated shocks
   2. but tests derived under the alternative are more problematic if shocks are correlated

References I

Ang, A., and G. Bekaert, 2007, "Stock Return Predictability: Is it There?," *Review of Financial Studies*, 20(3), 651–707.

Barberis, N., 2000, "Investing for the long run when returns are predictable," *The Journal of Finance*, 55(1), 225–264.

Berkowitz, J., and L. Giorgianni, 2001, "Long-Horizon Exchange Rate Predictability?," *The Review of Economics and Statistics*, 83(1), 81–91.

Boudouk, J., and M. Richardson, 1994, "The Statistics of Long-Horizon Regressions Revisited," *Mathematical Finance*, 4(2), 103–119.

Boudoukh, J., R. Michaely, M. Richardson, and M. R. Roberts, 2007, "On the importance of measuring payout yield: Implications for empirical asset pricing," *The Journal of Finance*, 62(2), 877–915.

Boudoukh, J., M. Richardson, and R. F. Whitelaw, 2006, "The Myth of Long-Horizon Predictability," *Review of Financial Studies*, p. hhl042.

Campbell, J., 2001, "Why long horizons? A study of power against persistent alternatives," *Journal of Empirical Finance*, 8(5), 459–491.

References II

Campbell, J. Y., and R. J. Shiller, 1988, “Stock prices, earnings, and expected dividends,” *The Journal of Finance*, 43(3), 661–676.

Cochrane, J. H., 2008, “The Dog That Did Not Bark: A Defense of Return Predictability,” *Review of Financial Studies*, 21(4), 1533–1575.

Fama, E. F., and K. R. French, 1988, “Permanent and Temporary Components of Stock Prices,” *The Journal of Political Economy*, 96(2), 246–273.

Goetzmann, W. N., and P. Jorion, 1993, “Testing the Predictive Power of Dividend Yields,” *The Journal of Finance*, 48(2), pp. 663–679.

Lewellen, J., 2004, “Predicting returns with financial ratios,” *Journal of Financial Economics*, 74(2), 209–235.

Mark, N. C., 1995, “Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability,” *The American Economic Review*, 85(1), 201–218.