Big Data in Finance

Asset Management - I

Tarun Ramadorai

Today's Agenda

- Predictability.
  - Definitions and preliminaries.
  - Non-return forecasting variables.
- The out-of-sample predictability debate.
  - Goyal and Welch (2008) and Campbell and Thompson (2008).
  - Economic significance of stock return predictability.
- New approaches and machine learning.

Market Efficiency

- Fama (1970) defines a market as efficient if "prices fully reflect all available information". In practice this means:

$$R_{i,t+1} = \Theta_{it} + U_{i,t+1}, \tag{1}$$

  where Θ_{it} is the rationally expected return on asset i, and $U_{i,t+1}$ has zero expectation with respect to the information set at t.

- This will only have content if we can restrict Θ_{it} using an economic model.

- Thus market efficiency is not testable except in combination with a model of expected returns.
  - This is called the joint hypothesis problem.

Cross-Sectional vs. Time-Series Efficiency

- Time-series efficiency. (What we are about to look at now)
  - Fix i, model returns over t.
  - Economic model for Θ_{it} in 1) is a time series model. The simplest such model is $\Theta_{it} = \Theta$, a constant.
  - Equilibrium models with time-varying expected returns also possible.
  - We will concentrate in this lecture on explaining (and more importantly, forecasting) aggregate stock index behaviour.

Cross-Sectional vs. Time-Series Efficiency

- Time-series efficiency. (What we are about to look at now)
  - Fix i, model returns over t.
  - Economic model for Θ_{it} in 1) is a time series model. The simplest such model is $\Theta_{it} = \Theta$, a constant.
  - Equilibrium models with time-varying expected returns also possible.
  - We will concentrate in this lecture on explaining (and more importantly, forecasting) aggregate stock index behaviour.

- Cross-sectional efficiency. (What we will look at next)
  - Take average returns over t and consider various partitions of i.
  - Economic model for Θ_{it} in (1) is a cross-sectional asset pricing model like the CAPM.
  - Tests of the CAPM can be thought of as joint tests of the CAPM and market efficiency.

Cross-Sectional vs. Time-Series Efficiency

- Time-series efficiency. (What we are about to look at now)
  - Fix i, model returns over t.
  - Economic model for Θ_{it} in 1) is a time series model. The simplest such model is $\Theta_{it} = \Theta$, a constant.
  - Equilibrium models with time-varying expected returns also possible.
  - We will concentrate in this lecture on explaining (and more importantly, forecasting) aggregate stock index behaviour.

- Cross-sectional efficiency. (What we will look at next)
  - Take average returns over t and consider various partitions of i.
  - Economic model for Θ_{it} in (1) is a cross-sectional asset pricing model like the CAPM.
  - Tests of the CAPM can be thought of as joint tests of the CAPM and market efficiency.

Defining Time-Series Efficiency

- Even with a model of expected returns, we need to specify what is in the information set used to form expectations of $U_{i,t+1}$ at time t.

- Fama defines 3 forms of the efficient market hypothesis, corresponding to what is included in the information set used to forecast $U_{i,t+1}$:

  - Weak form. Past returns.

  - Semi-strong form. Past publicly available information, e.g. stock splits, trading volume, dividends, earnings (also returns).

  - Strong form. Any past information, even if it is only available to insiders.

Examples

Michael Jensen (1978): “There is no other proposition in economics which has more solid evidence supporting it than the Efficient Markets Hypothesis”.

Examples

Robert Shiller (1984): “Returns on speculative assets are nearly unforecastable; this fact is the basis of the most important argument in the oral tradition against a role for mass psychology in speculative markets. One form of this argument claims that because real returns are nearly unforecastable, the real price of stocks is close to the intrinsic value, that is, the present value with constant discount rate of optimally forecasted future real dividends. This argument... is one of the most remarkable errors in the history of economic thought”.

- The debate continues to this day. Debates currently center around:
  - Whether predictability exists at all, in and out of sample.
  - If predictability exists, whether the proximate cause is frictions or irrationality.

Short and Long-Run Predictability

- Short-term return predictability is easy to detect if it is present, and hard to explain using a risk-based asset pricing model.
  - Can have modest effects on prices, and disappear quickly if arbitrageurs discover it.
  - Can also disappear if transactions costs decline making arbitrage cheaper (e.g., decimalization).
- Long-term return predictability can have large effects on prices; harder to detect without a very long time series.
  - Can be explained by a more sophisticated model of risk and return.

Alternative Time Series Hypotheses

- To devise meaningful time-series tests and interpret the results, it is useful to consider alternative hypotheses:

- Market prices contaminated by short-term noise, generating short-run reversals.
  - Noise could be caused by illiquidity (bid-ask bounce or other issues).

- Market reacts sluggishly to information, generating short-run predictability based on past returns or information releases.

- Market prices deviate substantially from efficient levels; long-lasting deviations are hard to arbitrage.
  - Generates long-run reversal and predictability based on price levels, and is difficult to detect in the short run.

Tests of Autocorrelation in Stock returns

- Do past returns predict future returns (weak-form market efficiency)?
  - One way to check is to inspect the autocorrelations of stock returns.

- There has been a significant amount of literature devoted to this topic, which includes (but is not limited to):
  - Box-Pierce statistics.
  - Variance ratio statistics.
  - Inference on long horizon and short-horizon return autocorrelations and asymptotic inference.
  - Cross-correlations and cross-autocorrelations (lead-lag effects).

- We won't spend too much time on this, but present analogous results on time-series momentum which employ regression analysis.
  - We can then simply apply machine learning techniques to these regression problems.

Time-Series Momentum

- Moskowitz, Ooi, and Pedersen (2012) investigate time-series regressions for a large set of assets, of the form:

$$r_t^s/\sigma_{t-1}^s = \alpha + \beta_h r_{t-h}^s/\sigma_{t-h-1}^s + \varepsilon_t^s,$$

  - where σ_{t-1}^s is the square-root of ex-ante volatility, which is estimated as:

$$\sigma_t^2 = 261 * \sum_{i=0}^{\infty}(1-\delta)\delta^i(r_{t-1-i} - \bar{r}_t)^2$$

  - the weights $(1-\delta)\delta^i$ add up to 1, and $\bar{r}_t$ is the exponentially weighted moving average return.

  - δ chosen such that $\sum_{i=0}^{\infty}(1-\delta)\delta^i i = \frac{\delta}{1-\delta} = 60$ days.

Time-Series Momentum

- These authors also estimate:

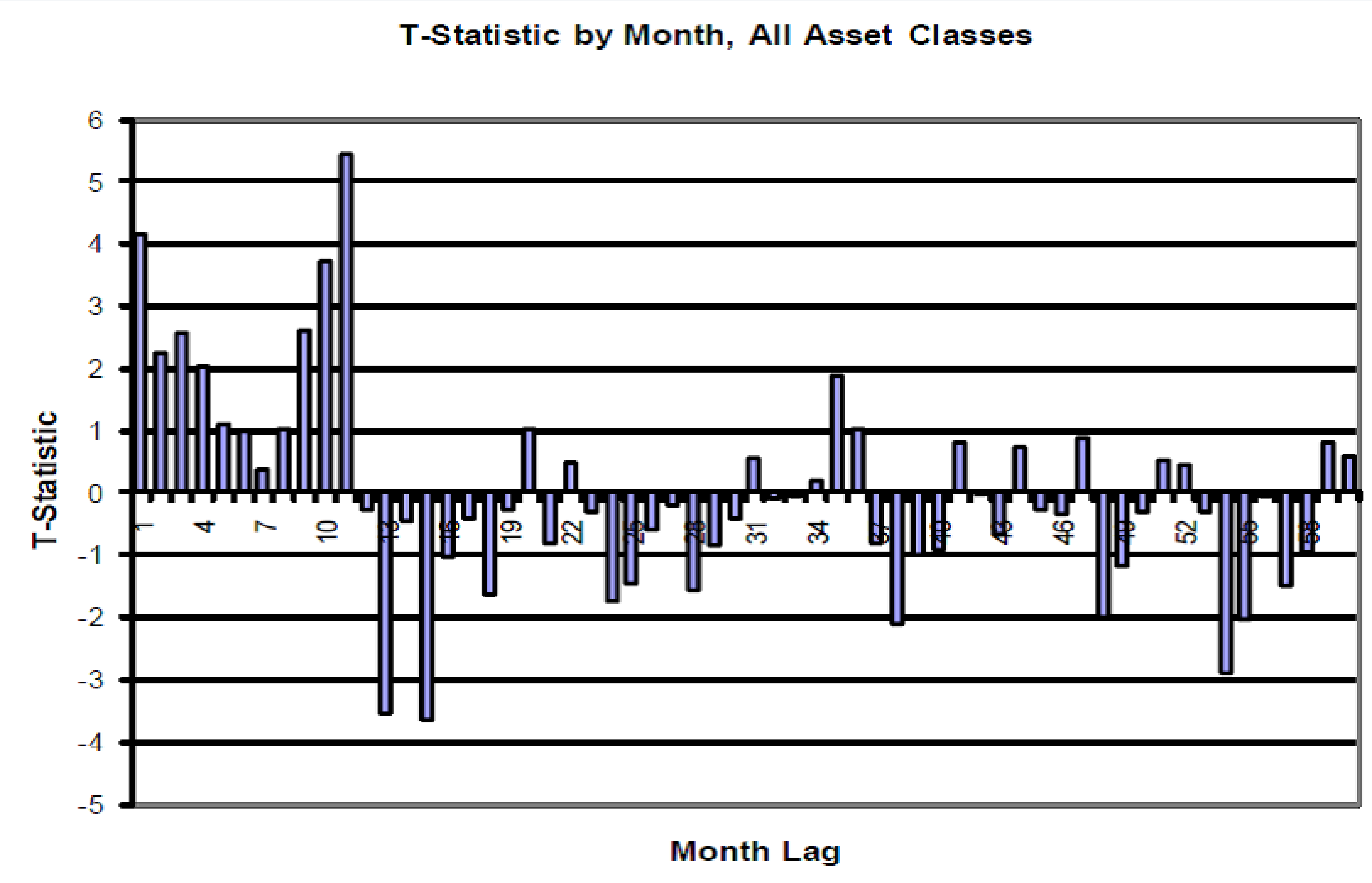
$$r_t^s / \sigma_{t-1}^s = \alpha + \beta_h sign(r_{t-h}^s) + \varepsilon_t^s.$$

- Note two things:

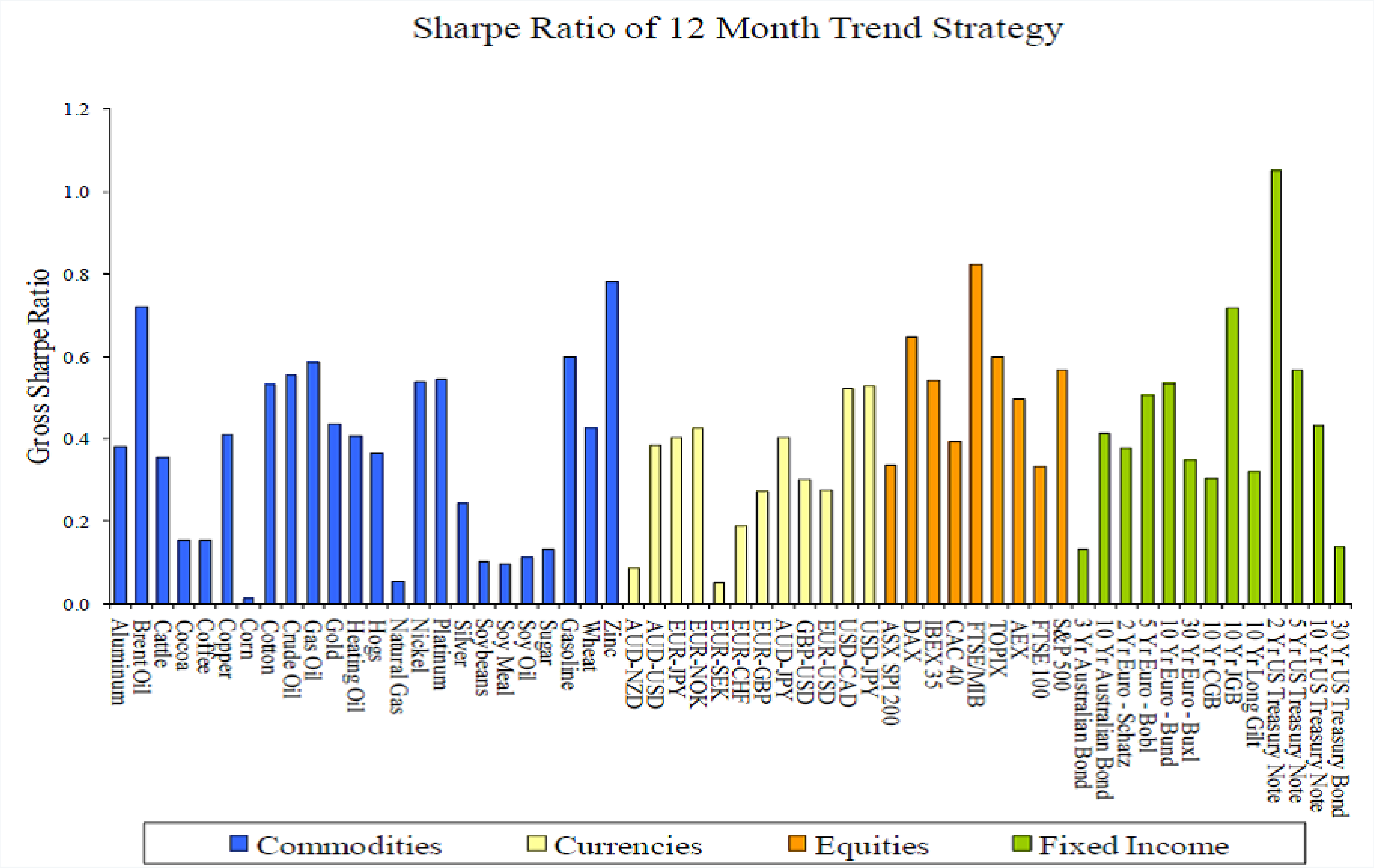
- We can easily apply techniques such as regression trees to this analysis to check for nonlinearities in the forecasting relationship.

- We can easily use test-training cycle and k-fold cross-validation to pick optimal choices of tuning parameters.

Time-Series Momentum
It's Not Just in Stocks.



Time-Series Momentum
It's Not Just in Stocks.



Semi-Strong Form Efficiency

- Authors have also tried many different forecasting variables for returns.
- As mentioned, debate in the literature about whether predictability actually exists, i.e., is predictability statistically and economically meaningful?
- We adopt a purely statistical approach, for now, but note that the economic magnitude of predictability can also be evaluated in the context of a simple present-value model.
- A fast summary of the evidence for non-return predictability follows.

A Fast Summary of Forecasting Variables

- Nominal Interest Rates and Inflation:
  - the short-term interest rate.
  - the long-term bond yield.
  - the term spread between long- and short-term Treasury yields.
  - the default spread between corporate and Treasury bond yields.
  - the lagged rate of inflation.

A Fast Summary of Forecasting Variables

- Valuation Ratios.
- Each of these ratios has some accounting measure of corporate value in the numerator, and market value in the denominator:
  - the price dividend ratio.
  - the earnings price ratio.
  - the smoothed price earnings ratio: proposed by Campbell and Shiller (1988a, 1998b) is the ratio of current price to a 10-year moving average of earnings.
  - the book to market ratio.

D/P Ratio and Stock Prices
Source: Cochrane textbook

Table 20.1. OLS regressions of percent excess returns (value weighted NYSE – treasury bill rate) and real dividend growth on the percent VW dividend/price ratio

Horizon k	$R_{t\to t+k} = a + b(D_t/P_t)$			$D_{t+k}/D_t = a + b(D_t/P_t)$		
(years)	b	$\sigma(b)$	R^2	b	$\sigma(b)$	R^2
1	5.3	(2.0)	0.15	2.0	(1.1)	0.06
2	10	(3.1)	0.23	2.5	(2.1)	0.06
3	15	(4.0)	0.37	2.4	(2.1)	0.06
5	33	(5.8)	0.60	4.7	(2.4)	0.12

$R_{t\to t+k}$ indicates the k-year return. Standard errors in parentheses use GMM to correct for heteroskedasticity and serial correlation. Sample 1947–1996.

A Fast Summary of Forecasting Variables

- Other Forecasting Variables:
  - the equity share of new issues proposed by Baker and Wurgler (2000).
  - net equity issuance $NTIS$, which represents the net issuing activity (IPOs, SEOs, stock repurchases, less dividends) of firms as a percentage of their market capitalization.
  - the consumption-wealth ratio of Lettau and Ludvigson (2001).
  - the aggregate portfolio illiquidity of hedge funds, proposed by Kruttli, Patton, and Ramadorai (2014).

Baker and Wurgler (2000)

Table III

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

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

$$R_{Et} = a + bX_{t-1} + u_t$$

where R_E denotes real percentage returns on the CRSP value-weighted (VW) or equal-weighted (EW) portfolio and X variously denotes the dividend-to-price ratio (D/P), the book-to-market ratio (B/M), or the equity share in total new equity and debt issues ($S = e/(e + d)$). The dividend-to-price ratio, the book-to-market ratio, and the equity share in new issues are standardized to have zero mean and unit variance. t-statistics are shown in brackets using heteroskedasticity robust standard errors.

	D/P			B/M			S		
	b	$t(b)$	$\bar{R}^2$	b	$t(b)$	$\bar{R}^2$	b	$t(b)$	$\bar{R}^2$
Panel A: 1928–1997 Returns									
VW CRSP	5.01	[2.12]	0.04	4.61	[1.79]	0.04	−7.42	[−3.86]	0.12
EW CRSP	5.75	[2.04]	0.02	13.06	[2.83]	0.16	−13.12	[−3.64]	0.16
EW − VW CRSP	2.87	[1.81]	0.02	8.45	[3.37]	0.26	−5.70	[−2.69]	0.11
Panel B: 1928–1962 Returns									
VW CRSP	8.36	[2.40]	0.08	12.24	[3.00]	0.19	−7.39	[−3.14]	0.11
EW CRSP	6.24	[1.08]	−0.01	26.64	[3.84]	0.41	−14.24	[−2.97]	0.19
EW − VW CRSP	0.88	[0.32]	−0.03	14.40	[3.69]	0.56	−6.84	[−2.33]	0.20
Panel C: 1963–1997 Returns									
VW CRSP	4.02	[0.95]	0.00	−0.73	[−0.29]	−0.03	−7.82	[−1.99]	0.11
EW CRSP	24.32	[2.95]	0.12	3.56	[1.13]	−0.01	−11.20	[−1.98]	0.09
EW − VW CRSP	21.48	[4.78]	0.30	4.29	[2.90]	0.07	−3.38	[−1.13]	0.00

Kruttli, Patton, and Ramadorai (2013)

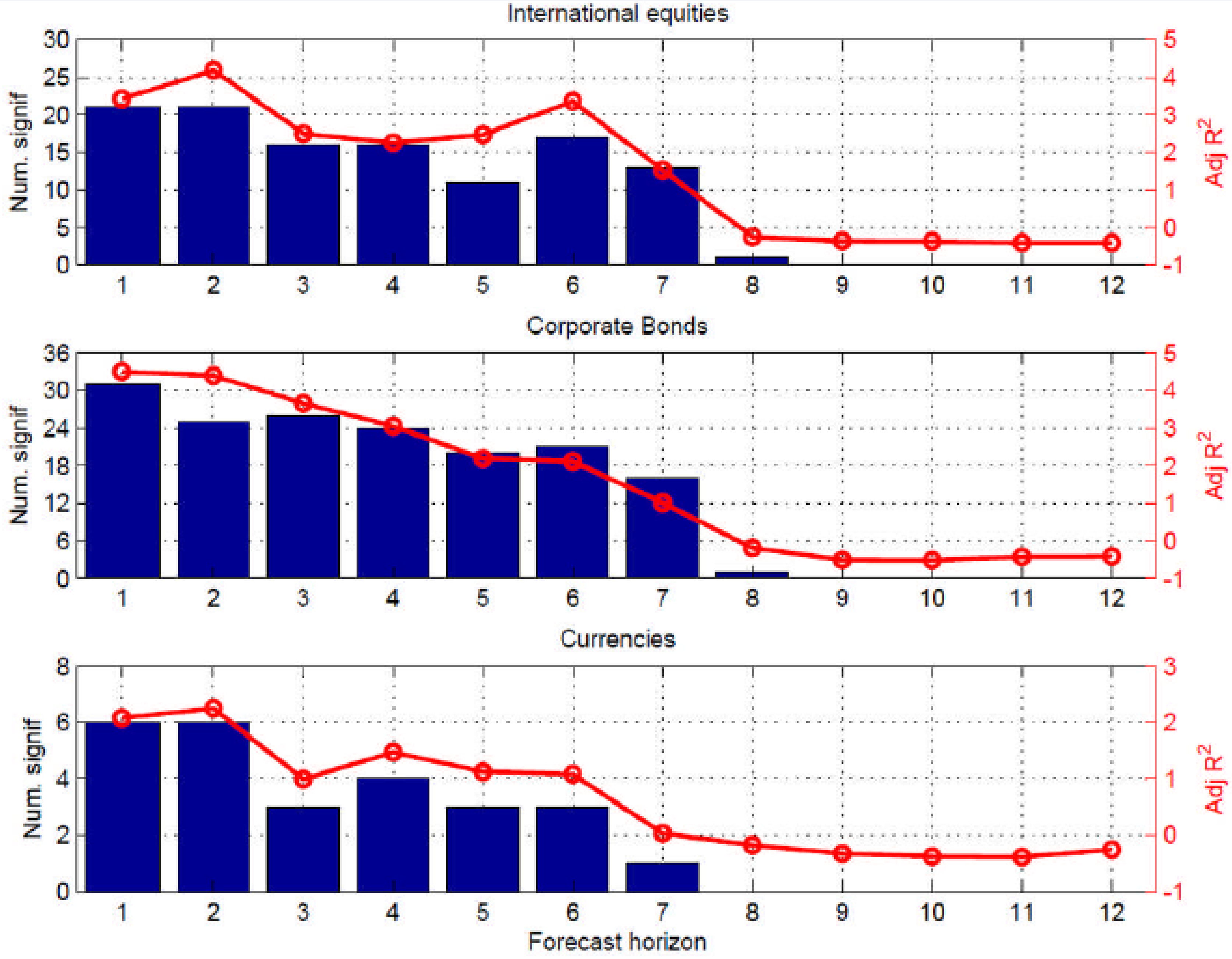


Figure 2 A: This figure shows the average adjusted R^2 and number of significant coefficients of the hedge fund illiquidity measure for different forecast horizons (in months). The results are for single predictor in-sample regressions, i.e. the only predictor is the hedge fund illiquidity measure.

Why Do We Care?

- One perspective: return predictability is evidence of market inefficiency; prices take long but temporary swings away from fundamental value (e.g., Baker and Wurgler (2000)).
- Another: predictability is evidence of time-varying equilibrium expected returns (note this is an excess return prediction).
  - Fama and French (1989) suggest that expected returns vary over business cycles, (i.e., takes a higher risk premium to induce stock holdings at the bottom of a recession).
  - Mechanically, when expected returns go up, prices go down.
  - We see low prices, followed by higher returns required by the market.
  - Changes in expected returns could be driven by changing investment opportunity sets or changing prices for risk.

Why Do We Care?

- A huge normative literature has sprung up around this, that purports to tell people how to invest.
  - If returns are predictable, especially out of sample, it helps us think about portfolio choice in the real world.
  - This is where we are headed next.
- An important debate in the literature (and in practice) about whether we can indeed find any variables that help us predict aggregate market returns.
- Our goal is to see if some of the machine learning techniques that we have discussed can help.

The Predictability Debate

Examples

Goyal and Welch (2008): “Our paper has systematically investigated the empirical real-world out-of-sample performance of plain linear regressions to predict the equity premium. We find that none of the popular variables has worked – and not only post-1990... Our profession has yet to find a variable that has had meaningful robust empirical equity premium forecasting power, at least from the perspective of a real-world investor.”

Examples

Campbell and Thompson (2008): “In this note we show that forecasting variables with significant forecasting power insample generally have a better out-of-sample performance than a forecast based on the historical average return, once sensible restrictions are imposed on the signs of coefficients and return forecasts. The out-of-sample predictive power is small, but we find that it is economically meaningful.”

Goyal and Welch (2008): The Method

- The Goyal and Welch procedure is to run a regression:
$$r_{t+1} = a + bx_t + \varepsilon_{t+1}$$
from time $t = 1$ to time $t = \tau$.

- Then use
$$\mathrm{E}_\tau[r_{\tau+1}] = a + bx_\tau$$
as a forecast of the next period return.

- This defines an out-of-sample forecast error of the forecasting model as:
$$f_{\tau+1} = r_{\tau+1} - \mathrm{E}_\tau[r_{\tau+1}]$$

Goyal and Welch (2008): The Method

- We can define the 'historical mean return' forecast error as:

$$m_{\tau+1} = r_{\tau+1} - \frac{1}{\tau}\sum_{t=1}^{\tau} r_t$$

  Note that this is just forecasting the next period return using the sample average.

- Define a Mean Squared Error (MSE) of each model as:

$$\frac{SSE(R)}{T} = \frac{1}{T}\sum_{\tau=start}^{end} (f_{\tau+1})^2 \text{ and } \frac{SSE(M)}{T} = \frac{1}{T}\sum_{\tau=start}^{end} (m_{\tau+1})^2$$

The Method

- Next, compute the difference between the (square roots of) these mean-squared errors:

$$\Delta RMSE = \sqrt{SSE(M)/T} - \sqrt{SSE(R)/T}$$

- Clearly, if the forecasting model is better, it will have smaller forecast errors, i.e., $\sqrt{SSE(M)/T} > \sqrt{SSE(R)/T}$, and thus $\Delta RMSE > 0$.

- In the results shown below, $\Delta RMSE\ D + 20$ begins the annual forecasting exercise twenty years after data first become available, and $\Delta RMSE$ 1965 begins the annual forecasting exercise in 1965.

Table 1: Forecasts at Annual Frequency

This table presents statistics on forecast errors in-sample (IS) and out-of-sample (OOS) for excess stock return forecasts at the annual frequency (both in the forecasting equation and forecast). Variables are explained in Section 2. Stock return is price changes, including dividends, of S&P500. Panel A presents the results for insignificant predictors while Panel B presents the results for significant predictors. The data period for each variable is indicated next to it. The column heading 'D+20' in Panel A begins forecast 20 years after the sample date while the column heading '1965-' in Panel A begins forecast in 1965. All numbers, except $\overline{R}^2$ and power, are in percent per year. A star next to $\overline{R}^2$ denotes significance of the in-sample regression (as measured by empirical F-statistic). ΔRMSE is the RMSE (root mean square error) difference between the unconditional forecast and the conditional forecast for the same sample/forecast period (positive numbers signify superior out-of-sample conditional forecast). The column 'IS for OOS' in Panel B gives the ΔRMSE of IS errors for OOS period. MSE-F is F-statistic by McCracken (2004), which tests for equal MSE of the unconditional forecast and the conditional forecast. One-sided critical values of MSE statistics are obtained empirically from bootstrapped distributions, except for **caya** and **all** models where they are obtained from McCracken (2004) (critical values for **ms** model are not calculated). 'Pwr' is the power of ΔRMSE and is calculated as the fraction of draws where the simulated ΔRMSE is greater than the empirically calculated 95% critical value and is reported in percent. The two numbers under the power column are power for all simulations and simulations that are found to be in-sample significant at the 95% level. Significance levels at 90%, 95%, and 99% are denoted by one, two, and three stars, respectively.

Many Predictors Insignificant Insample

Panel A: Insignificant in-sample predictors

	Variable	Full Data			
		Data	IS	OOS	
			$\overline{R}^2$	ΔRMSE D+20	ΔRMSE 1965–
d/p	Dividend Price Ratio	1872–2004	0.47	–0.1092	–0.0908
d/y	Dividend Yield	1872–2004	0.89	–0.0971	–0.3162
e/p	Earning Price Ratio	1872–2004	1.00	–0.0886	0.0899
d/e	Dividend Payout Ratio	1872–2004	–0.75	–0.3140	–0.1846
svar	Stock Variance	1885–2004	–0.76	–2.3405	0.0104
ntis	Net Equity Expansion	1927–2004	–0.03	–0.0352	–0.2303
tbl	T-Bill Rate	1920–2004	0.57	–0.1083	–0.1318
lty	Long Term Yield	1919–2004	–0.53	–0.4638	–0.7499
ltr	Long Term Return	1926–2004	1.00	–0.7696	–1.2016
tms	Term Spread	1920–2004	0.30	–0.0488	–0.0008
dfy	Default Yield Spread	1919–2004	–1.20	–0.1376	–0.1249
dfr	Default Return Spread	1926–2004	0.38	–0.0330	–0.0194
infl	Inflation	1919–2004	–0.98	–0.1939	–0.0714

Significant IS Predictors Insignificant OOS

			IS		Forecasts begin 1965			
					IS for OOS	OOS		
	Variable	Data	$\overline{R}^2$	ΔRMSE	ΔRMSE	ΔRMSE	MSE-F	Pwr
d/y	Dividend Yield	1927–2004	2.65*	0.3799*	0.2421	−0.3121	−1.51	30 (**72**,6)
e/p	Earning Price Ratio	1927–2004	3.01*	0.4147*	0.1635	−0.0861	−0.42	38 (**64**,11)
b/m	Book to Market	1927–2004	3.97*	0.5079*	−0.4097	−1.3070	−5.79	42 (**61**,13)
eqis	Pct Equity Issuing	1927–2004	9.62***	1.0641***	0.4591*	0.2313	1.18*	68 (**78**,15)
cayp	Cnsmptn, Wlth, Incme	1948–2001	24.89***	2.2188***	2.3357***	2.2418	11.83***	88 (**90**,12)
all	Kitchen Sink	1927–2004	15.61**	3.0042	1.5728	−6.8113	−20.26	— (—,—)
caya	Cnsmptn, Wlth, Incme	1948–2001	—	—	—	−0.4958	−1.99	— (—,—)
ms	Model Selection	1927–2004	—	—	—	−1.1309	−5.06	— (—,—)

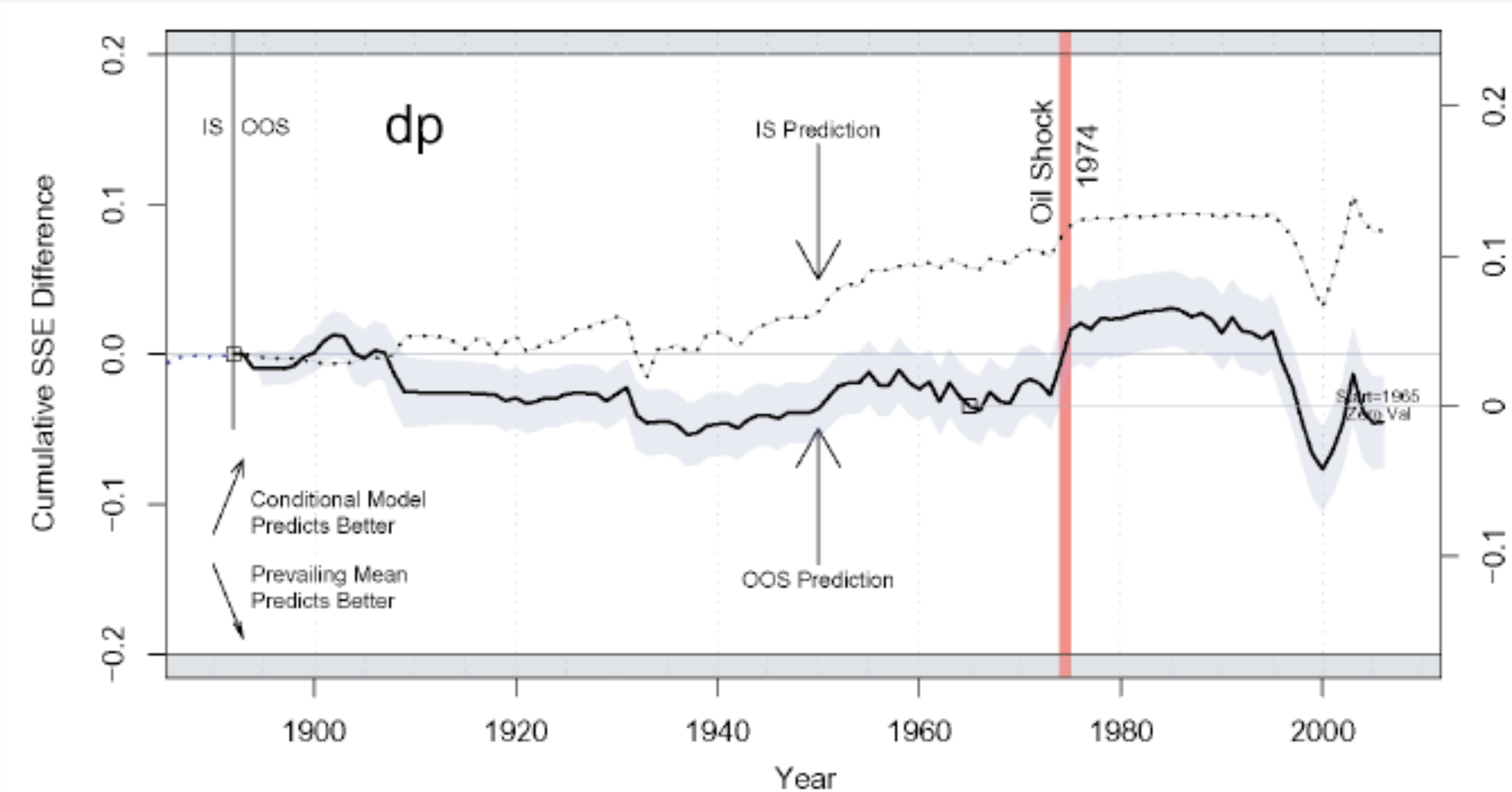
Desirable Predictor Characteristics?

A well-specified signal would inspire confidence in a potential investor if it had

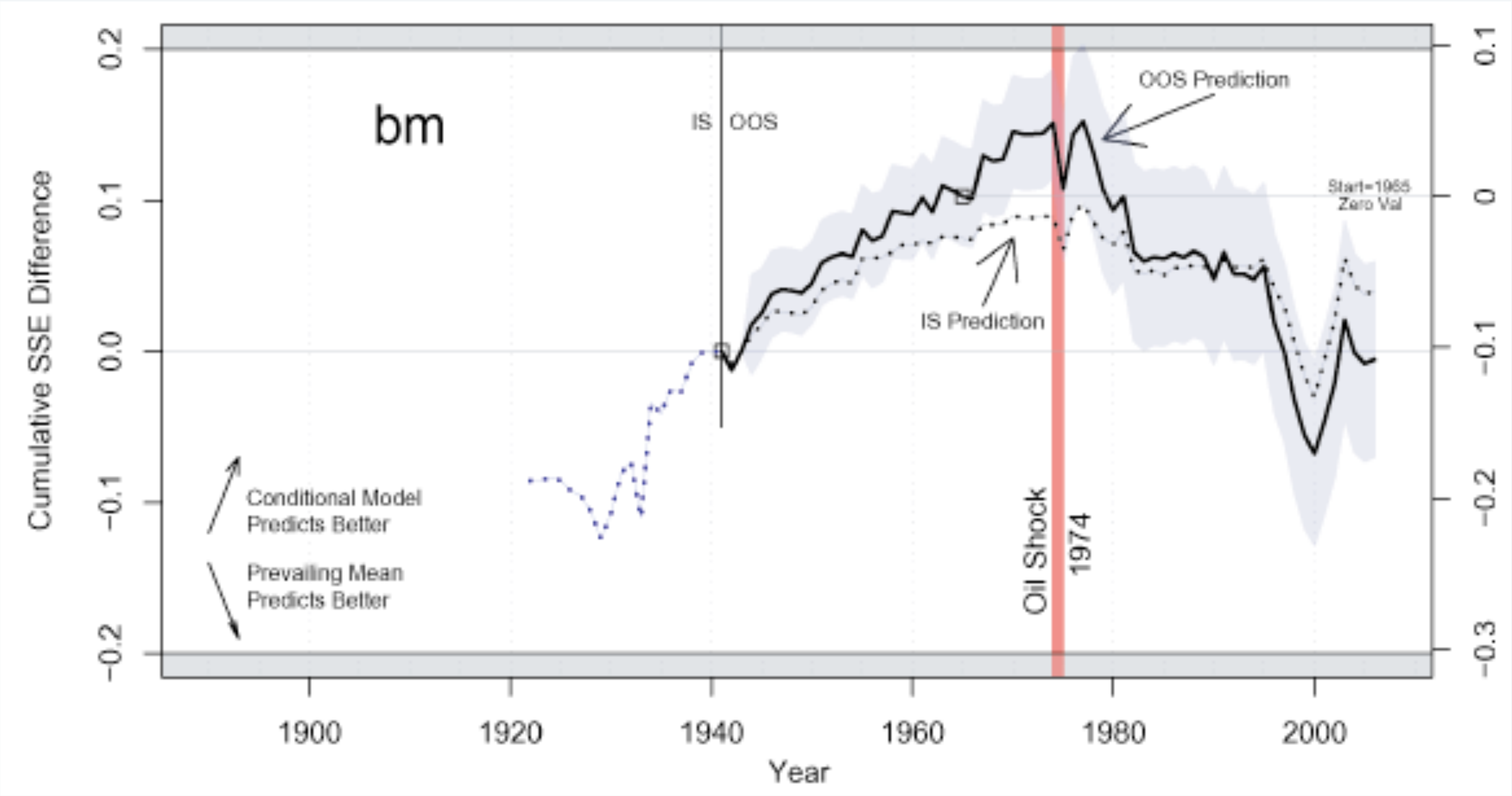
1. both significant IS and OOS performance;
2. an irregular but upward drift;
3. an upward drift not just in one short or unusual sample period—say just the two years around the 1974 Oil Shock; and
4. an upward drift that remains positive over the most recent several decades—otherwise, even a reader taking the long view would have to be concerned with the alternative explanation that the underlying model has changed.

Bad News for D/P

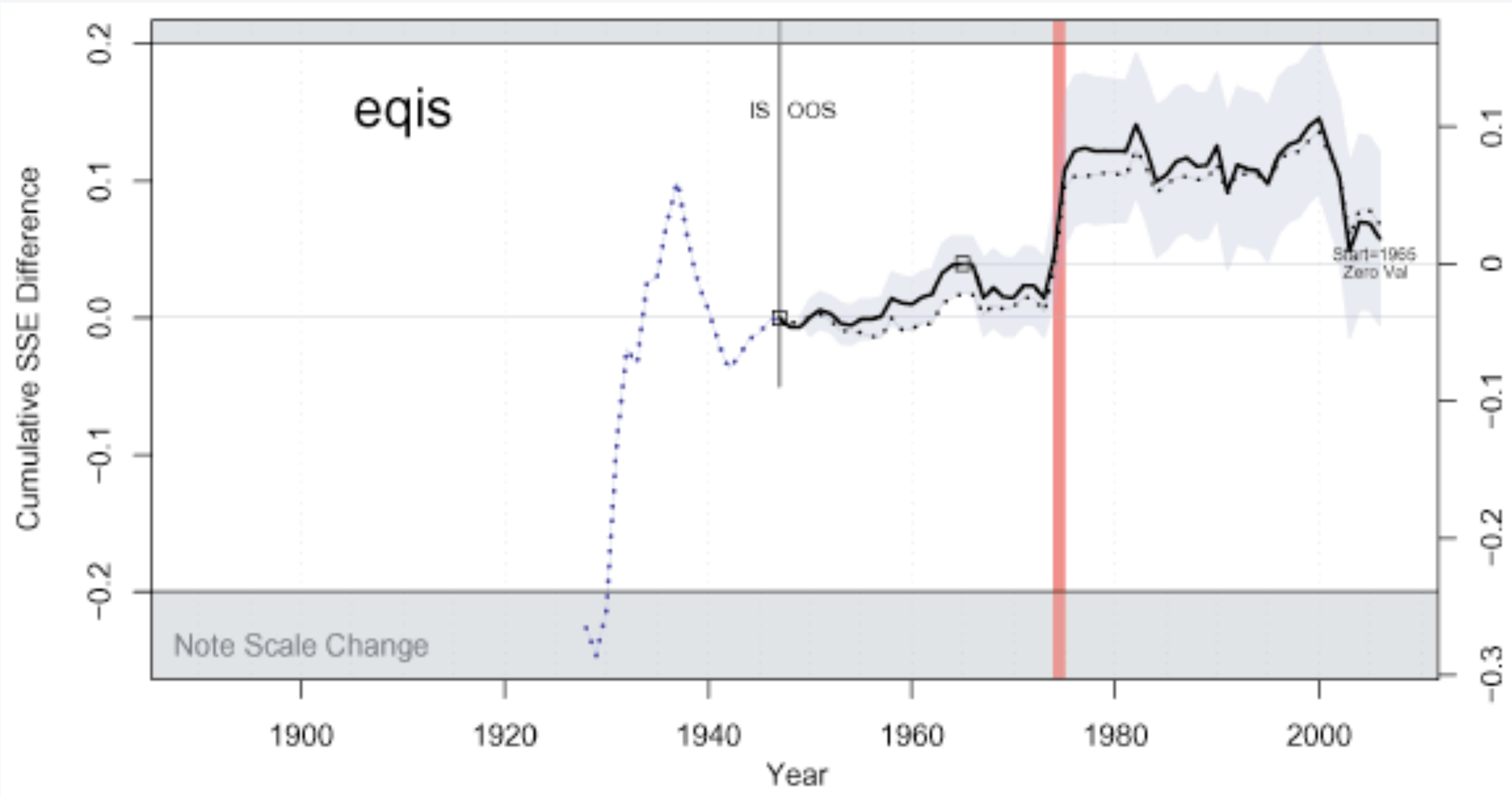
Blue Band shows statistical significance at the 95% level.



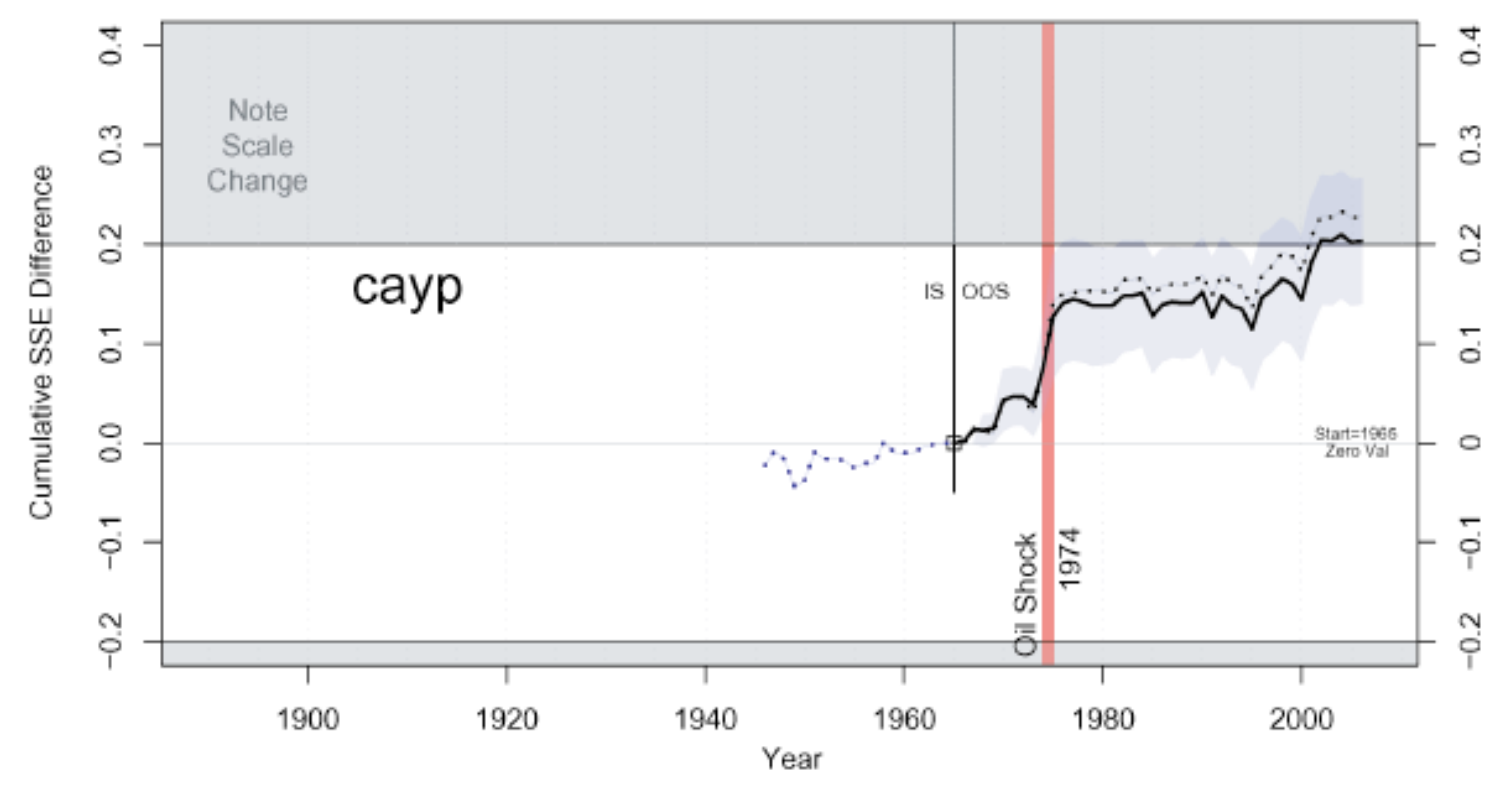
Bad News for B/M



Better OOS Prediction: The Equity Share



Better OOS Prediction: CAY



Campbell and Thompson (2008)

Examples

"In this note we show that forecasting variables with significant forecasting power insample generally have a better out-of-sample performance than a forecast based on the historical average return, once sensible restrictions are imposed on the signs of coefficients and return forecasts. The out-of-sample predictive power is small, but we find that it is economically meaningful. We also show that a variable is quite likely to have poor out-of-sample performance for an extended period of time even when the variable genuinely predicts returns with a stable coefficient."

Out-Of-Sample R-Squared Statistic

- Campbell and Thompson use an out-of-sample R^2 statistic (which can be compared to the usual in-sample R^2 statistic) to evaluate the performance of forecasting variables.

$$R^2_{oos} = 1 - \frac{\Sigma_{t=1}^{T}(r_t - \hat{r}_t)^2}{\Sigma_{t=1}^{T}(r_t - \bar{r}_t)^2}$$

  $\hat{r}_t$ is the fitted value from a predictive regression estimated through period $t-1$, and $\bar{r}_t$ is the historical average return estimated through period $t-1$.

- $\Delta RMSE$ and R^2_{oos} are similar. They have the same sign. If R^2_{oos} falls below zero, therefore, we need to worry.

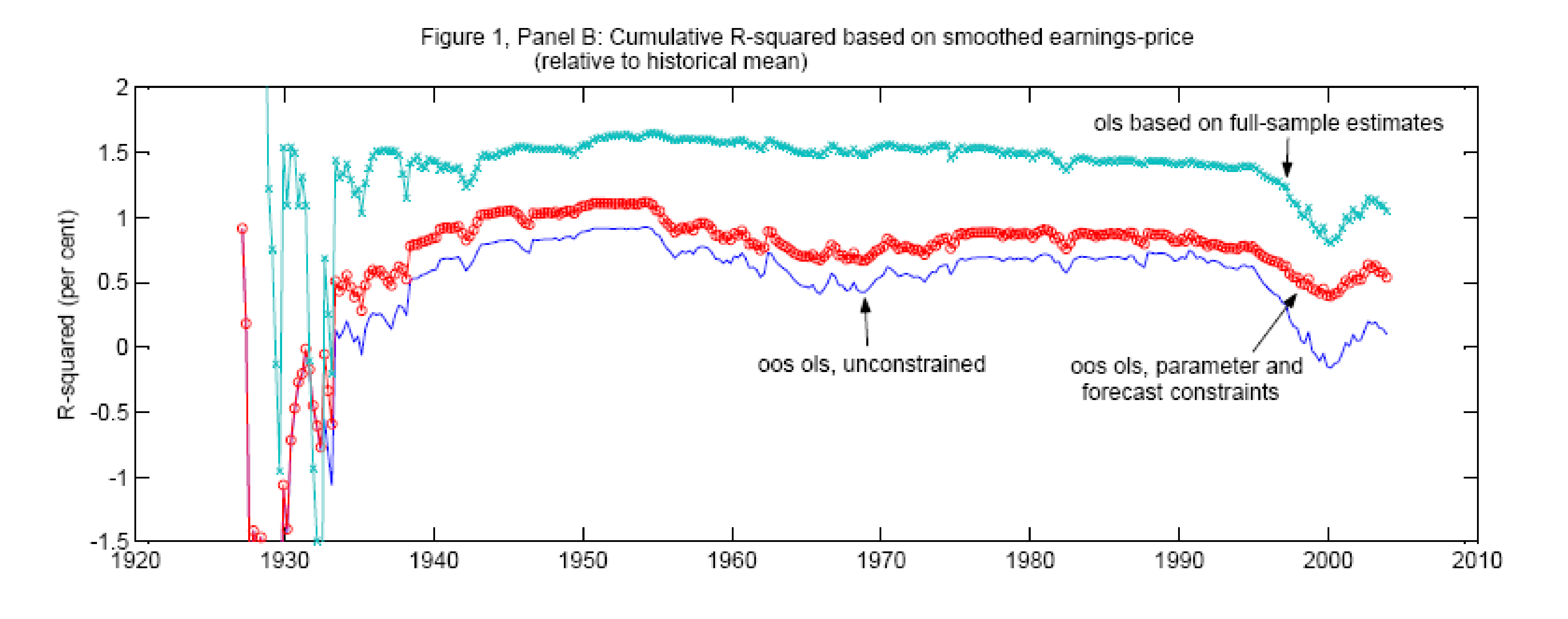
Campbell and Thompson

- Campbell and Thompson's basic approach is to impose 'sensible restrictions' on parameters and forecasts, based on what they think is *economically* meaningful in the forecasting regressions.

- For example, if a variable forecasts with the wrong theoretically predicted sign, they set the forecast to zero in that period.

- Or if a variable forecasts a negative equity premium, they set the forecast to zero in that period.

Campbell and Thompson (2008)



Economic Significance of Predictability

- How big an R^2_{oos} do we need for predictability to mean something important to an investor?

- Assume:

$$r_{t+1} = \mu + x_t + \varepsilon_{t+1}$$

  r_{t+1} represents the excess return, x_t the (mean-zero) predictor variable and μ the unconditional mean of the excess return.

Economic Significance of Predictability

- Consider an investor with a single-period horizon, and M-V preferences with a risk-aversion of γ. Assume she does not observe x_t.

- From the standard portfolio choice problem, her portfolio weight in the risky asset can be represented as:

$$\omega_t = \omega = \frac{\mu}{\gamma \sigma_r^2} = \frac{\mu}{\gamma(\sigma_x^2 + \sigma_\varepsilon^2)}$$

Economic Significance of Predictability

- The average excess return on the investor's portfolio is:

$$\mathrm{E}[\omega r_{t+1}] = \frac{\mu}{\gamma\sigma_r^2}\mathrm{E}[r_{t+1}] = \frac{\mu^2}{\gamma\sigma_r^2} = \frac{1}{\gamma}S^2 \tag{2}$$

  Where S^2 is the squared Sharpe ratio on the risky asset.

- However, if the investor observes the predictor variable x_t, his (time-varying) portfolio weight in the risky asset becomes:

$$\omega_t = \frac{\mu + x_t}{\gamma\sigma_\varepsilon^2}$$

- Why? Numerator changes because the investor scales up or down his investment based on the realization of x_t in the prior period.

- Denominator changes because the investor observes x_t in each period, therefore it no longer contributes to the variance of the portfolio weight in each period.

Economic Significance of Predictability

- Now, the average excess return on the investor's portfolio is:

$$
\begin{aligned}
\mathrm{E}[\omega_t r_{t+1}] &= \frac{1}{\gamma\sigma_\varepsilon^2}\mathrm{E}[(\mu + x_t)(\mu + x_t + \varepsilon_{t+1})] \qquad (3)\\
&= \frac{\mu^2 + \sigma_x^2}{\gamma\sigma_\varepsilon^2}\\
&= \frac{1}{\gamma}\frac{S^2 + R_{oos}^2}{1 - R_{oos}^2} \qquad (4)
\end{aligned}
$$

- Where $R_{oos}^2 = \frac{\sigma_x^2}{\sigma_x^2+\sigma_\varepsilon^2}$ is the R^2 statistic from the predictive regression determining the future excess return.

Economic Significance of Predictability

- Ratio of the two expected returns (with and without predictability) gives the proportional increase in the investor's wealth from factoring in predictability.

- Ratio is very close to

$$\frac{R^2_{oos}}{S^2}$$

for small R^2_{oos} and S^2, and when the time interval is short.

Quantifying the Economic Significance of Predictability

- In monthly data since 1871, the monthly Sharpe ratio for stocks is around 0.108. This means $S^2 \approx 0.012$.

- Say average R^2_{oos} for the earnings-price ratio predictive regression is ≈ 0.0025 (0.25%).

- In our hypothetical example of a one-period M-V investor, means that average excess return will proportionally improve by approximately $0.0025/0.012 \approx 21\%$!

- Absolute increase in returns is around 25 basis points per month ($\approx 3\%$ per year) for an investor with unit risk aversion, and $\approx 1\%$ per annum for an investor with risk aversion $\gamma = 3$.
  - Is this factor large or small?

Johannes, Korteweg, Polson (2014)

- In a recent addition to the debate, Johannes et al. (2014) point out two flaws with the standard predictive regression models.

- Propose an alternative:

$$
\begin{aligned}
r_{t+1} &= \alpha + \beta_t x_t + \sqrt{V_t^r}\varepsilon_{t+1}^r \\
\ln V_{t+1}^r &= \alpha_r + \beta_r \ln V_t^r + \sigma_r \eta_{t+1}^r \\
\beta_{t+1} &= \beta_0 + \beta_\beta \beta_t + \sigma_\beta \varepsilon_{t+1}^\beta
\end{aligned}
$$

- Note the two changes: i) Stochastic Volatility (SV). ii) Dynamic coefficient (DC).

Johannes, Korteweg, Polson (2014)

Table I

Portfolio Returns: Dividend Yield Data

This table shows out-of-sample portfolio returns using the dividend yield as the predictor for one-month, one-year, and two-year investment horizons. Investors have power utility with risk aversion parameter γ, and allocate their wealth between the market portfolio of stocks and a risk-free one-period bond. The certainty equivalent returns in Panel A represent the annualized risk-free return that gives the investor the same utility as the risky portfolio strategy. Panel B shows monthly Sharpe ratios. CM stands for a model with constant mean (i.e., no predictability), and CV and SV stand for constant and stochastic volatility, respectively. Hence, CV-CM represents a model with constant mean and constant volatility of returns. DC denotes drifting coefficients and represents models where the coefficient on net payout is allowed to vary over time. CV-OLS uses the OLS point estimates of equation (1), with data up to time t. CV-rolling OLS uses a 10-year rolling regression model to form portfolios. * and ** indicate that the result is significant at the 10% and 5% level, respectively, based on 500 simulated data sets with constant mean and volatility. † and †† indicate that the result is significant at the 10% and 5% level, respectively, based on 500 simulated data sets with constant mean and stochastic volatility.

	$\gamma = 4$			$\gamma = 6$		
	1m	1y	2y	1m	1y	2y
Panel A: Certainty Equivalent Returns (in % per annum)						
Constant volatility models						
CV-CM	4.77	5.05	5.05	4.38	4.61	4.61
CV-OLS	−8.60	1.14	−0.17	−25.22	1.70	−1.16
CV-rolling OLS	−19.68	−11.48	−12.59	−46.01	−24.71	−30.25
CV	−7.04	4.42	4.47	−2.47	4.09	4.03
CV-DC	−7.64	4.04	3.95	−3.29	3.72	3.51
Stochastic volatility models						
SV-CM	5.52	5.92	5.94	4.89	5.14	5.14
SV	6.43**,††	6.51**,††	6.53**,††	5.52**,††	5.64**,††	5.64**,††
SV-DC	6.56**,††	5.85	5.68	5.64**,††	5.17	5.02

Applying Machine Learning

- The Johannes et al. results suggest that amending the basic linear model to include time-varying coefficients helps in out of sample prediction.

- More generally, the stock return prediction problem falls into the class of standard prediction problems, and can easily be handled by using ML in a regression context.
  - Evaluation techniques are straightforward here. We can simply compare k-fold cross-validation results from ML on the test sample, with the historical mean return model or other models.

- Are there improvements generated by ML-based approaches? Key is whether there are nonlinearities in the forecasting relationship.

- We use the Goyal-Welch data (and measurement approach) to check the performance of standard ML models.

Imperial College
Business School

Imperial means
Intelligent Business