F500: Empirical Finance
Lecture 7: Present Value Relations

Oliver Linton
obl20@cam.ac.uk

February 25, 2020

Outline

1. Fundamentals versus Bubbles
2. Cash Flow model
3. Rational Bubbles and other Bubble Models
4. Excess Volatility tests
5. Campbell's Approximate Model of Log Returns
6. The Impact of Cash Flow and Discount Rate Innovations
7. Return Predictability

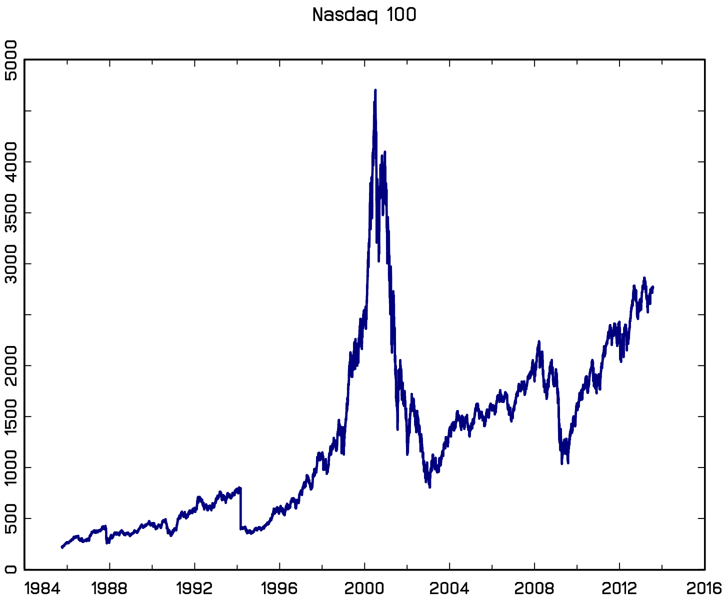
Reading: Linton (2019), Chapter 9

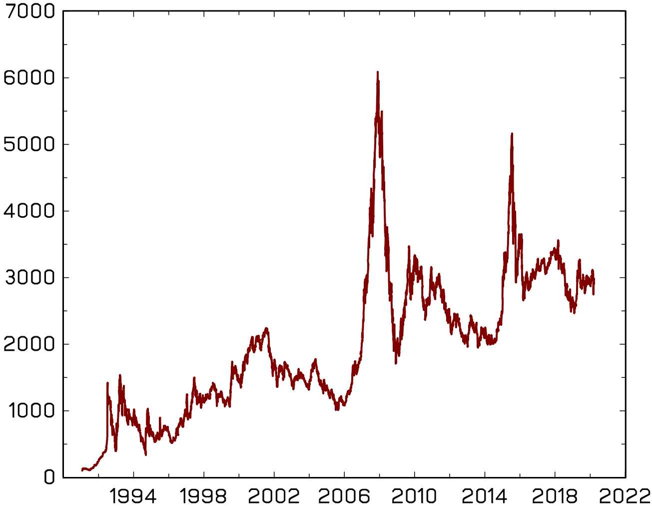
Fundamentals versus Bubbles

- Are stock prices (We now focus on prices rather than returns) driven by fundamental values or rational expectations of what those fundamental values will be in the future?
  - Efficient markets hypothesis, CAPM, etc
- Or are stock prices driven by "animal spirits", fads, bubbles, and irrational exuberance?
  - Behavioural finance, Kahneman and Tversky. Shiller (2000). Irrational Exuberance. Thaler (2016).
- Sometimes one and sometimes other.

Bubbles and Crashes

- Many examples of "bubbles" in financial history:
  - tulips in Amsterdam in 17th century
  - south sea bubble 18th century
  - 1920s? 1987? 2000?. There is less agreement on these.
  - Bitcoin?
- Typically think of bubbles as a pervasive market wide phenomenon with rapid increases ultimately followed by a crash.
- Is there a rational explanation for them, or are they only explicable due to animal spirits and irrationality?


Nasdaq 100
0
500
1000
1500
2000
3000
3500
4000
5000
1984
1988
1992
1996
2000
2004
2008
2012
2016



Shanghai stock market index

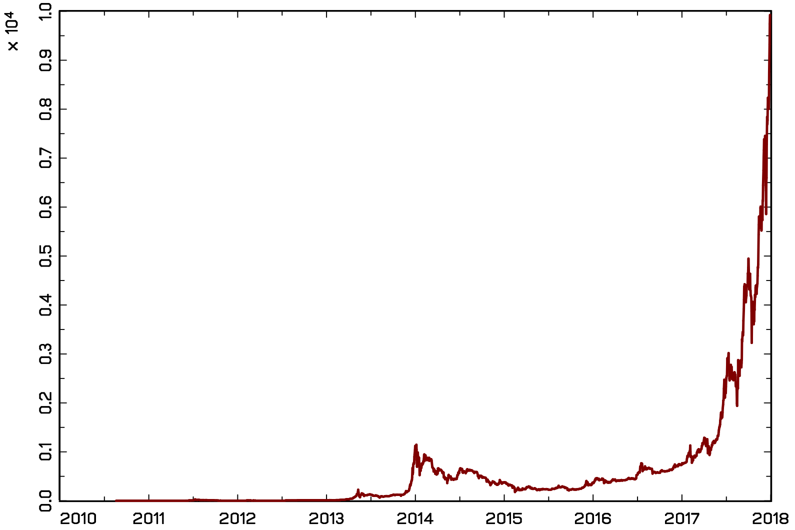


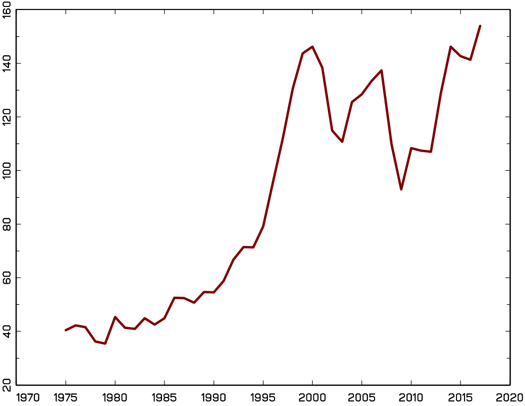
Figure: Bitcoin price level from 2010-2017

Shiller (2000, Irrational Exuberance) Table 6.1

Largest Recent One-Year Real Stock Price Index Increases

	Percentage	One year period	Subsequent ΔPrice
Philippines	683.4	Dec 1985-Dec 1986	28.4
Taiwan	400.1	Oct 1986-Oct 1987	65.7
Venezuela	384.6	Jan 1990-Jan 1991	33.1
Peru	360.9	Aug 1992-Aug 1993	15.8
Columbia	271.3	Jan 1991-Jan 1992	-19.9
Jamaica	224.5	April 1992-April 1993	-59.2
Chile	199.8	Jan 1979-Jan 1980	38.9
Italy	166.4	May 1985-May 1986	-15.7
Jamaica	163.4	Aug 1985-Aug 1986	8.7
Thailand	161.9	Oct 1986-Oct 1987	-2.6
India	155.5	April 1991-April 1992	-50.3
Italy	147.3	April 1980-April 1981	-32.1
Austria	145.4	Feb 1989-Feb 1990	-19.8
Finland	128.3	Sept 1992-Sept 1993	46.3
Denmark	122.9	April 1971-April 1972	-12.4

- Warren Buffet: Total market cap to GDP (%). Shows the US might be in a bubble



- But the measure is flawed:
  - GDP is a flow, market valuation is a stock. We see below how these may relate
- There are several further arguments in this particular case
  - Companies that make up the US market earn a substantial amount of profit overseas.
  - Corporate margins and thus profits as a percent of GDP fluctuate over time.
  - The proportion of public companies to private companies also fluctuates over time and impacts the total market cap calculation.

After the bubble comes the crash...

- Famous examples of stock market crashes include: the US markets on October 24th 1929 and October 19th 1987. May 6th, 2010 Flash Crash (no prior bubble)

BROOKLYN DAILY EAGLE

And Complete Long Island News

LATE NEWS

WALL ST. IN PANIC AS STOCKS CRASH

Attempt Made to Kill Italy's Crown Prince

STOCKS CRASH IN RUSH TO SELL; BILLIONS LOST

ASSASSIN CAUGHT IN BRUSSELS MOB; PRINCE UNHURT

Hollywood Fire Destroys Films Worth Millions

High Duty Group Gave $700,000 to Coolidge Drive

HOOVER'S TRAIN HALTED BY AUTO PLACED ON RAILS

- Each of these market events have been the subject of a lot of research and newspaper coverage as well as government reports. Typically these reports do not identify a single causal explanation but several factors that played a role. So even in these well known extreme cases it is hard to pin specific blame.

- Scheinkman (2014, Arrow Lecture). Three stylized facts about bubbles:
  - Asset price bubbles coincide with increases in trading volume
  - Asset price bubbles often coincide with financial or technological innovation (railways, tech bubble)
  - Asset price implosions seems to coincide with increases in assets's supply
- There are authors who dispute the bubble explanations in particular cases:
  - Garber (1990) proposes market fundamental explanations for the three famous historical bubbles.
  - Pastor and Veronesi (2006) argue that the turn of the 20th century tech bubble was at least partly explained by an increase in the uncertainty about average future profitability in the late 1990s.

What are fundamentals?

- Regulation, competition, innovation
- Macroeconomic news affecting all firms
- In some sectors/countries exchange rates are important
- Weather can be important for retail and agriculture,
- Earnings, Dividends of individual firms

Present Value Relations

- Dividend Discount model - Prices depend on cashflows and discount rates in a rational way.
- Letting R denote the required one-period return (discount rate) at time t, and suppose D_{t+1} arrives in the interval $(t, t+1]$

$$R = E_t R_{t+1} = E_t \left[\frac{P_{t+1} + D_{t+1}}{P_t} - 1 \right] = \overbrace{E_t \left[\frac{P_{t+1}}{P_t} - 1 \right]}^{\text{capital gain}} + \overbrace{E_t \left[\frac{D_{t+1}}{P_t} \right]}^{\text{dividend}}$$

Therefore,

$$P_t = E_t \left[\frac{P_{t+1} + D_{t+1}}{1+R} \right] = \frac{E_t P_{t+1}}{1+R} + \frac{E_t D_{t+1}}{1+R}.$$

Applying the same logic to future prices we have

$$P_{t+1} = \frac{E_{t+1} P_{t+2}}{1+R} + \frac{E_{t+1} D_{t+2}}{1+R}$$

so that (using law of iterated expectation) and continuing

$$\begin{aligned} P_t &= \frac{E_t \left[E_{t+1} P_{t+2} + E_{t+1} D_{t+2} \right]}{(1+R)^2} + \frac{E_t D_{t+1}}{1+R} \\ &= \frac{E_t D_{t+1}}{1+R} + \frac{E_t D_{t+2}}{(1+R)^2} + \frac{E_t \left[P_{t+2} \right]}{(1+R)^2} \\ &= \frac{E_t D_{t+1}}{1+R} + \frac{E_t D_{t+2}}{(1+R)^2} + \ldots + \frac{E_t D_{t+k}}{(1+R)^k} + \frac{E_t \left[P_{t+k} \right]}{(1+R)^k}. \end{aligned}$$

- Assuming the absence of price bubbles at infinity the discounted terminal price goes to zero with k, i.e.,

$$\lim_{k\to\infty} \frac{E_t\left[P_{t+k}\right]}{(1+R)^k} = 0$$

the sum converges to

$$P_t = E_t\left[\sum_{i=1}^{\infty}\left(\frac{1}{1+R}\right)^i D_{t+i}\right]$$

- Prices are set equal to the present discounted value of expected future cash flows.

The Gordon Growth Model special case

We rewrite the pricing equation as

$$P_t = \sum_{i=1}^{\infty} \left(\frac{1}{1+R} \right)^i E_t D_{t+i}$$

Special case with no risk and assuming that dividends grow at a constant rate $G < R$ ($D_{t+1} = D_t(1+G)$) gives the Gordon Growth Model:

$$P_t = D_t \sum_{i=1}^{\infty} \left(\frac{1+G}{1+R} \right)^i = \frac{(1+G)\, D_t}{R-G}$$

Hence with a constant dividend rate stock prices depend only upon innovations to expected dividends

$$\frac{\partial \log P_t}{\partial \log G} = \frac{G(1+R)}{(R-G)(1+G)} > 0 \quad ; \quad \frac{\partial \log P_t}{\partial \log R} = \frac{-R}{R-G} < 0$$

Can be very sensitive when $R \approx G$

Cointegrated Prices and Dividends

- Stock prices are not martingales since

$$E_t P_{t+1} = (1+R)P_t - E_t D_{t+1}$$

However, the implied price where dividends are reinvested is a martingale.

Theorem

Suppose that $\triangle D_t$ is stationary and the discount rate is constant. Then P_t and D_t are cointegrated of order one with cointegrating vector $(1, -\frac{1}{R})$ since

$$P_t - \frac{D_t}{R} = \left(\frac{1}{R}\right) E_t \left[\sum_{i=0}^{\infty} \left(\frac{1}{1+R}\right)^i \Delta D_{t+1+i}\right]$$

is a stationary process.

Rational Bubbles

- There are models of rational bubbles which start from the same pricing equation but do not assume the condition

$$\lim_{k\to\infty} \frac{E_t\left[P_{t+k}\right]}{(1+R)^k} = 0$$

- Specifically write

$$P_t = P_t^* + B_t$$

where P_t^* is the fundamental price and B_t is the bubble process that satisfies

$$B_t = \frac{E_t B_{t+1}}{1+R}.$$

This means that bubble component is growing in expectation.

- Some similarities with the fads model, but nonstationary

Example

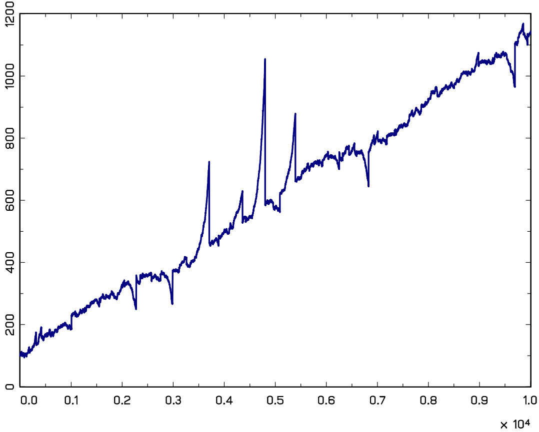
Blanchard and Watson (1982) model. A switching AR(1) process

$$B_{t+1} = \begin{cases} \frac{1+R}{\pi} B_t + \eta_{t+1} & \text{with probability } \pi \\ \eta_{t+1} \text{ or zero} & \text{with probability } 1-\pi \end{cases}$$

where η_{t+1} is iid with mean zero.
The bubble collapses with constant probability $1-\pi$ each period (stationary fad component η_{t+1}).
Expected dration of bubble is $1/(1-\pi)$. Satisfies $B_t = \frac{E_t B_{t+1}}{1+R}$

- Some logical issues with rational bubbles, see below

Example of bubble process: $\pi = 0.99$, $R = 0.001$ $\eta \sim N(0,1)$ superimposed on $p_t^* = 0.1 + p_{t-1}^* + N(0,1)$



Some Critiques of Rational Bubbles

- Bubbles cannot exist on finite-lived assets, it is all about infinity
- Negative bubbles cannot exist if there is a lower bound on the asset price (e.g. zero, for assets with limited liability). If $B_t < 0$ at any time t, then
$$E_t B_{t+j} = B_t(1+R)^j \to -\infty.$$
Therefore, $B_t \geq 0$ for all t. This means that $\eta_t \geq 0$ with probability one and this is not possible if it is mean zero unless it is always zero. If positive bubbles can exist, but negative bubbles are ruled out, then a bubble can never start. It must exist from the beginning of trading. Diba and Grossman 1988
- Positive bubbles cannot exist if there is an upper bound on the asset price (e.g. there exists a high-priced substitute in perfectly elastic supply).
- Brunnermeier (2012, NBER) discusses other bubble models including rational and irrational participants

Some evidence questioning the connection between prices and fundamentals in normal times

- Cutler, Poterba, and Summers (1989) and Fair (2002) investigated large price moves on the S&P500 (daily and intradaily in the latter case) and tried to match the movements up with news stories reported in the NYtimes/WSJ the following day: many large movements were associated with monetary policy, but there remained many significant movements that they could not find explanations for.
- Orange juice futures and weather, Roll (1984, AER).

Shiller Variance Bounds Tests

- Dividends are very stable; they fluctuate very little about an upward trend. Expected dividends should therefore also fluctuate little.
- Consequently stock prices should be stable. In fact, stock prices fluctuate wildly.
- Shiller (1981) shows how the variability of the dividend sets an upper bound to the variability of the stock price.

Theorem

Suppose that P_t^* *is an unbiased forecast of* P_t*, i.e.,* $E_t P_t^* = P_t$*. Then*

$$P_t^* = P_t + e_t$$

where $\mathrm{cov}(e_t, P_t) = 0$*, and so*

$$\mathrm{var}\left[P_t^*\right] = \mathrm{var}\left[P_t\right] + \mathrm{var}\left[e_t\right] \geq \mathrm{var}\left[P_t\right]$$

- For example

$$P_t^* = \sum_{i=1}^{\infty} \left(\frac{1}{1+R}\right)^i D_{t+i}$$

is an unbiased forecast of P_t.

- Shiller assumed that aggregate real dividends follow a finite variance stationary stochastic process around a deterministic growth rate g.
- Then the detrended variables

$$p_t = \frac{P_t}{(1+g)^t}, \; d_t = \frac{D_t}{(1+g)^{t+1}},$$

both are stationary processes.

- He obtains a relation between the real price and the real dividend stream

$$\begin{aligned} p_t &= \sum_{s=0}^{\infty} \left(\frac{1+g}{1+R} \right)^{s+1} d_{t+s} = \sum_{s=0}^{\infty} \left(\frac{1}{1+\overline{R}} \right)^{s+1} d_{t+s} \\ &= \sum_{s=0}^{M} \left(\frac{1}{1+\overline{R}} \right)^{s+1} d_{t+s} + \frac{E_t \left[p_{t+M} \right]}{(1+\overline{R})^M}, \end{aligned}$$

provided $g < R$, where $\overline{R} = (R-g)/(1+g)$ is the discount rate for the detrended series. He argues that $\overline{R} \simeq E(d_t)/E(p_t)$, see below.

- Shiller computes the perfect foresight price

$$p_t^* = \sum_{s=0}^{T-t} \frac{d_{t+s}}{(1+\overline{R})^s} + \frac{\widehat{p}_T}{(1+\overline{R})^T},$$

  where the terminal price $\widehat{p}_T$ is estimated as the average price over his hundred year sample.

- He estimates the trend rate g by regressing $\ln(P_t)$ on a constant and time trend. The discount rate is estimated as the average of detrended real dividend divided by the average detrended real price, around 0.045 in his sample.
- He uses the real S&P500 and Dow Jones index and dividend series from 1871-1979.
- He finds that the variance inequality is substantially violated - the variance of detrended prices is 5-13 times too high relative to the perfect foresight price variance.

This paper was very influential, but there was a lively debate for many years as to the methods and results. Subsequently, people argued these results are not robust because

- Estimates of var($p^*_{t,T}$) are downward biased even under stationary prices.
- Stock prices likely to be nonstationary random walks not deterministic trend with stationary components
- Dividends highly persistent/nonstationary random walks
- Alternative dividend models. Marsh and Merton (1987). Permanent income approach implies that changes in dividend payouts respond to lagged changes in prices.
- Violations could be due to time varying expected returns. Expected returns are not constant

Campbell's Approximate Model of Log Returns

Fact

Perhaps explanation of excess volatility is from the assumption that expected returns are constant. Relax the assumption

Calculating log return from the formula for one-period arithmetic return gives

$$\begin{aligned} r_{t+1} &\equiv \log\left(P_{t+1} + D_{t+1}\right) - \log\left(P_t\right) \\ &= \log\left(P_{t+1}(1 + D_{t+1}/P_{t+1})\right) - \log\left(P_t\right) \\ &= p_{t+1} - p_t + \log\left(1 + \exp\left(d_{t+1} - p_{t+1}\right)\right) \end{aligned}$$

He "linearized" this equation using the first-order Taylor approximation around average values

$$f\ (x_{t+1}) \approx f\ (\overline{x}) + f\ '(\overline{x})\,(x_{t+1} - \overline{x})$$

Letting $x = d_{t+1} - p_{t+1}$, $\overline{x} = \overline{d} - \overline{p}$ and $f(x) = \log(1 + \exp(x))$ (the `softplus` function) we get

$$\begin{aligned} r_{t+1} &\approx k + \rho\ p_{t+1} + (1-\rho)\ d_{t+1} - p_t \\ &= k + \overbrace{p_{t+1} - p_t}^{\text{capital gain}} + (1-\rho)\ \overbrace{(d_{t+1} - p_{t+1})}^{\text{dividend payout}} \end{aligned}$$

where $k = -\log(\rho) - (1-\rho)\log(1/\rho - 1)$ with

$$\rho = \frac{1}{1 + \exp(\overline{d} - \overline{p})} \approx \frac{1}{1 + \overline{D}/\overline{P}}$$

For US data, CLM argue that $\rho \simeq 0.96$ in annual data and so $k \simeq 3.1$

- Rewriting this equation

$$p_t = k + \rho \ p_{t+1} + (1-\rho) \ d_{t+1} - r_{t+1}$$

- Solving forward and imposing the no-bubbles condition gives

$$p_t = \frac{k}{1-\rho} + \underbrace{(1-\rho)\sum_{j=0}^{\infty}\rho^j E_t d_{t+1+j}}_{p_{dt}} - \underbrace{\sum_{j=0}^{\infty}\rho^j E_t r_{t+1+j}}_{p_{rt}}$$

Current prices reflect expected future dividend flow (cash flow) and expected future returns (risk premium). If $E_t r_{t+1+j}$ goes down then current prices go up

The Impact of Cash Flow and Discount Rate Innovations

- These equations can also be rearranged to show that

$$r_{t+1} - E_t\left[r_{t+1}\right] = \overbrace{\left(E_{t+1} - E_t\right)\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j}}^{\eta_{d,t+1}} - \overbrace{\left(E_{t+1} - E_t\right)\sum_{j=1}^{\infty}\rho^j r_{t+1+j}}^{\eta_{r,t+1}}$$

  where $E_t(\eta_{d,t+1} + \eta_{r,t+1}) = 0$
- An unexpectedly good stock return must occur because either:
  - the expectations of future dividends go up
  - or because expectations of future returns go down.
- The first term is a standard cash flow effect and the second is an expected return or risk premium effect: the price goes up if the risk premium or risk-free interest rate go down. This is important since one large anomaly in observed behaviour is the large size of realized return innovations relative to realized dividend innovations.

Example

Consider the simple case where expected returns are a constant plus AR(1) process

$$E_t\left[r_{t+1}\right] = r + x_t$$

$$x_{t+1} = \phi x_t + \xi_{t+1}, \quad -1 < \phi < 1$$

Inserting into the formula for realized return innovation as a combination of dividend and expected return innovations gives

$$p_{rt} \equiv \sum_{j=0}^{\infty} \rho^j E_t r_{t+1+j} = \frac{r}{1-\rho} + \frac{x_t}{1-\rho\phi}$$

- If the innovations are very persistent (ϕ close to 1) then stock price is sensitive to innovations to expected returns.
- If average dividend yield is low, then ρ is close to one, increasing the sensitivity of stock price to an innovation to expected returns.
- If $\rho\phi \approx 1$ this might help to explain the findings of excess return volatility.

- Restating in terms of realized returns

$$r_{t+1} = E_t\left[r_{t+1}\right] + \eta_{d,t+1} + \eta_{r,t+1} = r + x_t + \eta_{d,t+1} - \frac{\rho\xi_{t+1}}{1-\rho\phi}$$

- Making the (unrealistic) assumption that the innovations to dividends and the innovations to expected returns are uncorrelated gives

$$\mathrm{var}\left[r_{t+1}\right] = \sigma_d^2 + \sigma_x^2\left[\frac{1+\rho^2-2\rho\phi}{\left(1-\phi\rho\right)^2}\right] \approx \sigma_d^2 + \frac{2\sigma_x^2}{1-\phi}$$

using $\rho \approx 1$ so that $\frac{1+\rho^2-2\rho\phi}{(1-\phi\rho)^2} \approx \frac{2}{1-\phi}$.

- If $\sigma_x^2 = 0$, then $\mathrm{var}\left[r_{t+1}\right] = \sigma_d^2$. If $\sigma_x^2 > 0$, then $\mathrm{var}\left[r_{t+1}\right] > \sigma_d^2$.

Solution

So the big volatility in returns could be due to persistent time varying expected returns process.

Predictive regressions

- Campbell equations imply long horizon returns related to fundamentals

$$\overbrace{\sum_{j=0}^{\infty}\rho^j E_t r_{t+1+j}}^{\text{long horizon returns}} = d_t - p_t + \frac{k}{1-\rho} + \overbrace{\sum_{j=0}^{\infty}\rho^j E_t \Delta d_{t+1+j}}^{\text{small?}}$$

- Common starting point for modelling time varying expected returns is linear predictive regression

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

  Typically, x_t is dividend/price ratio, earnings/price ratio, or term structure variables lagged by one period.
- Dividends are measured at a monthly frequency over a trailing annual horizon, so it has been common practice to similarly aggregate returns.

- Suppose that

$$x_{t+1} = \phi x_t + \eta_{t+1}$$

- Then long horizon returns aggregate to give

$$\begin{aligned} r_{t+1} + r_{t+2} &= \beta x_{t+1} + \varepsilon_{t+2} + \beta x_t + \varepsilon_{t+1} \\ &= (1+\phi)\,\beta x_t + \varepsilon_{t+2} + \varepsilon_{t+1} + \beta \eta_{t+1} \end{aligned}$$

$$\begin{aligned} r_{t+1} + \cdots + r_{t+K} &= \left(1 + \phi + \ldots + \phi^K\right) \beta x_t + + \varepsilon_{t+1} + \ldots + \varepsilon_{t+K} \\ &\quad + \beta \eta_{t+1} + \cdots + \beta \phi^{K-1} \eta_{t+K-1} \end{aligned}$$

- We can write this model as (for $t = 1, \ldots, T - K$ and for vector case, including intercept)

$$r_{t+1:t+K} = r_{t+1} + \ldots + r_{t+K} = \beta(K)^{\intercal} x_t + u_{t:t+K},$$

where $u_{t:t+K}$ is an innovation that satisfies $E(u_{t:t+K} | x_t) = 0.$ Valid regression.

Econometric Issues

- Suppose that returns r_{t+1} obey **rw**1, then the aggregated returns $r_{t+1:t+K}$, for $K > 1$, will generally not be independent, in fact $\mathrm{cov}(r_{t+1:t+K}, r_{t+j+1:t+j+K}) \neq 0$ for $j = 1, \ldots, K$. Likewise the error terms satisfy

$$\mathrm{cov}(u_{t:t+K}, u_{t+j:t+j+K}) \neq 0,$$

  and in fact form an $MA(K)$ process.
- The OLS estimator of $\beta(K)$ is consistent, but the standard errors need to be adjusted to take account of the serial correlation induced by the overlapping observations in the standard errors.

Under the null hypothesis that $\beta = 0$, $r_{t+1:t+K} = \varepsilon_{t+1} + \ldots + \varepsilon_{t+K}$.

Definition

For $K = 1, 2, \ldots$, define the OLS estimator of $\beta(K)$

$$\widehat{\beta}(K) = \left(\sum_{t=1}^{T-K} x_t x_t^\intercal \right)^{-1} \sum_{t=1}^{T-K} x_t r_{t+1:t+K}$$

and let $\widehat{u}_t(K) = r_{t+1:t+K} - \widehat{\beta}^\intercal(K) x_t$ be the residuals. The Hansen and Hodrick (1980) standard errors are based on estimating the K autocovariances of $x_t, u_t(K)$. The Hodrick (1992) standard errors for $\widehat{\beta}(K)$ are derived from the matrix

$$\widehat{V}(K) = \left(\sum_{t=1}^{T-K} x_t x_t^\intercal \right)^{-1} \sum_{t=K}^{T} \widehat{u}_t^2(1) \left(\sum_{k=0}^{K-1} x_{t-k} \right) \left(\sum_{k=0}^{K-1} x_{t-k} \right)^\intercal \left(\sum_{t=1}^{T-K} x_t x_t^\intercal \right)^{-}$$

- Stambaugh (1999) considered the effect that **persistent predictors** (near unit roots) might have on this regression. He supposed that

$$x_{t+1} = \phi x_t + \eta_{t+1}, \quad \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \overbrace{\sim}^{iid} 0, \begin{pmatrix} \sigma_{\varepsilon\varepsilon} & \sigma_{\varepsilon\eta} \\ \sigma_{\varepsilon\eta} & \sigma_{\eta\eta} \end{pmatrix}.$$

- Stambaugh argued that $\widehat{\beta}_{OLS}$ is consistent but biased in finite samples and when ϕ is close to one, this bias can be very large.
- We have

$$E(x_t \varepsilon_{t+j}) = \begin{cases} 0 & \text{if } j \geq 1 \\ \neq 0 & \text{if } j \leq 0 \end{cases}$$

ie the regressors are predetermined but not strictly exogenous. OLS is biased in small samples

- The standard errors are themselves biased and noisy, especially when ϕ is large and/or K is large.
  - HH not guaranteed to be positive definite, and in practice are not when K is large.
  - H standard errors are always pd, but only valid under null.
  - Newey-West standard errors are guaranteed to be pd under null and alternative, but nasty.

- Empirically we often find that β is quite large and statistically significant; but we found that returns are almost unautocorrelated, e.g., $\rho(1) \sim 0.01$. How can that be?

Theorem

For all k

$$\mathrm{cov}(r_t, r_{t-k}) = 0.$$

if and only if

$$\frac{\sigma_{\varepsilon\eta}}{\sigma_{\eta\eta}} = -\frac{\beta\phi}{1-\phi^2}.$$

- Shows that the two findings are not necessarily incompatible. This corresponds to the case where shocks to dividend/price ratio are contemporaneously negatively correlated with innovations to returns.

Empirical Results

- Log stock returns on the log dividend price ratio

$$r_{t+1} + \ldots + r_{t+K} = \alpha(K) + \beta(K)\left(d_t - p_t\right) + u_{t:t+K}$$

- CLM report results for 1927-1994 and two subperiod 1927-1951 and 1951-1994. Table 7.1. Use real variables, ie adjusted for inflation
- Econometric issue due to **overlapping observations**: adjust standard errors to reflect error autocorrelation

"At a horizon of 1-month, the regression results are rather unimpressive: The R^2 *statistics never exceed 2%, and the t-statistics exceed 2 only in the post-World War II subsample. The striking fact about the table is how much stronger the results become when one increases the horizon. At a 2-year horizon the* R^2 *statistic is 14% for the full sample...at a 4-year horizon the* R^2 *statistic is 26% for the full sample."*

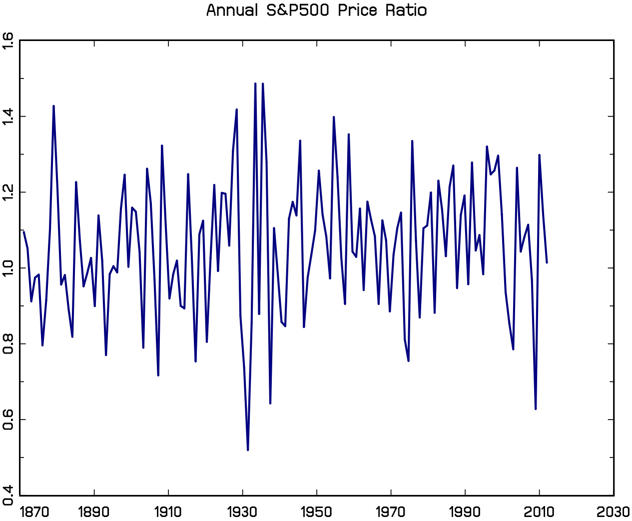
Subsequent work has cast doubt on the robustness of these findings. If $\phi \sim 1$, then the coefficients increase linearly in horizon. R^2 do too. Mechanically, spurious regression

Boudoukh, Richardson, and Whitelaw (2008). The myth of long horizon predictability.

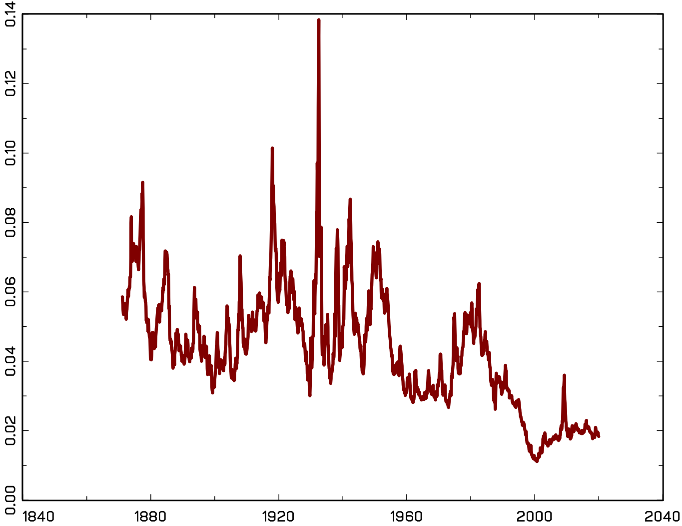
> *"The prevailing view in finance is that the evidence for long-horizon stock return predictability is significantly stronger than that for short horizons. We show that for persistent regressors, a characteristic of most of the predictive variables used in the literature, the estimators are almost perfectly correlated across horizons under the null hypothesis of no predictability. For the persistence levels of dividend yields, the analytical correlation is 99% between the 1- and 2-year horizon estimators and 94% between the 1- and 5-year horizons. Common sampling error across equations leads to ordinary least squares coefficient estimates and R^2s that are roughly proportional to the horizon under the null hypothesis. This is the precise pattern found in the data."*

Shiller's Data

http://www.irrationalexuberance.com/index.htm. Shows P_{t+1}/P_t at annual frequency for "S&P500" index. Average 1.057 std 17.8%



"S&P 500" dividend yield — (12 month dividend per share)/price (D_{t+1}/P_t). Currently 2.05%. Persistent series.



Revaluation of Shiller real dividend/price monthly data using Rolling Window

- Log stock returns on the log dividend price ratio

$$r_{t+1} + \ldots + r_{t+K} = \alpha(K) + \beta(K)\left(d_t - p_t\right) + u_{t:t+K}$$

- Recalculate using rolling forty year windows, $t = i - 240, \ldots, i + 240$, $i = 1890, \ldots, 1990$; $K \in \{1, 3, 12, 24, 36, 48, 60\}$
- Find $R^2(i, 1) \leq \ldots \leq R^2(i, 60)$, i.e., predictability increases with horizon
- However, we find a lot of variability over i, the 40 year periods centred on 1970 and 1940 both have high R^2 at long horizon, but the intermediate period and subsequent periods do not.

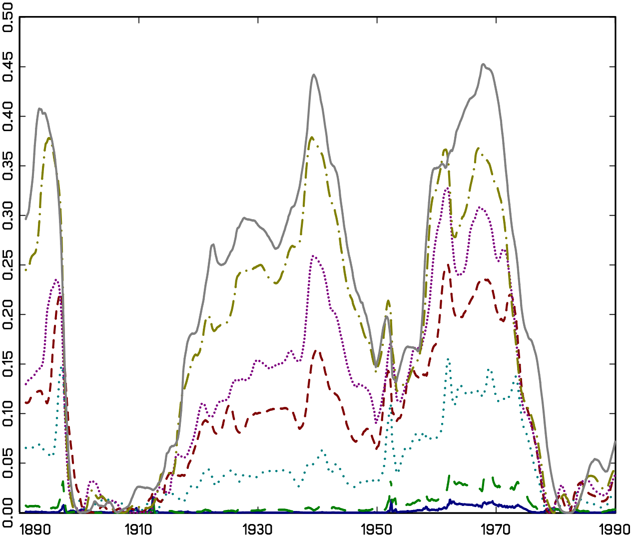


Figure: Rolling window (±20 years) R^2 of predictive regression

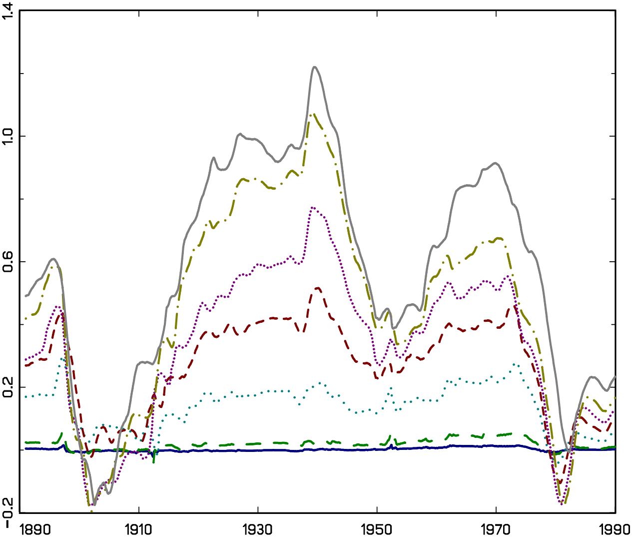


Figure: Rolling window (±20 years) slope coefficent of predictive regression

Conclusion

- There is an active area of research to establish long run predictability. Cochrane (1999, p.37) describes this as one of the three most important facts in Finance in his survey *New Facts in Finance*

  Now, we know that . . .
  [Fact] 2. Returns are predictable. In particular: Variables including the dividend/price (d/p) ratio and term premium can predict substantial amounts of stock return variation. This phenomenon occurs over business cycle and longer horizons. Daily, weekly, and monthly stock returns are still close to unpredictable. . .

- This claim is debatable. There is a lot of time variation in the predictability, and whether it is real or spuriously induced by aggregation has not been settled. In my view.

Dividend payouts vary considerably across firms. Price and Annual Dividend Yield for Dow stocks 2013

	P	$\%D/P$		P	$\%D/P$
Alcoa Inc.	9.32	1.39	JP Morgan	48.88	2.86
AmEx	61.69	1.30	Coke	37.42	2.94
Boeing	75.03	2.65	McD	93.90	3.35
Bank of America	12.03	0.67	MMM	103.23	2.50
Caterpillar	95.61	2.36	Merck	41.42	4.18
Cisco Systems	20.99	2.86	MSFT	28.01	3.57
Chevron	114.96	3.38	Pfizer	27.29	3.59
du Pont	46.94	3.66	P&Gamble	76.54	3.08
Walt Disney	55.61	1.71	AT&T	35.36	5.12
General Electric	23.29	3.39	Travelers	80.39	2.44
Home Depot	67.52	1.98	United Health	57.32	1.73
HP	16.79	3.14	United Tech	90.78	2.46
IBM	200.98	1.89	Verizon	44.40	4.71
Intel	21.12	4.47	Wall Mart	69.30	2.51
Johnson[2]	76.16	3.36	Exxon Mobil	88.36	2.76