

33 Course Review

1. The regressions.

- (a) *Forecasting regression* for predicting returns over time.

$$R_{t+1}^e = a + bx_t + \varepsilon_{t+1}; \quad t = 1, 2, \dots, T$$

- (b) *Time series regression* Explaining variation in returns over time; characterizing correlation between returns, finding betas for factor models,

$$R_t^{ei} = a_i + \beta_i r_{mrf_t} + \varepsilon_t^i; \quad t = 1, 2, \dots, T$$

- (c) *Cross-sectional regression* Explaining variation in *average* returns *across* *stocks* by variation in their betas (b,h,s,..) or characteristics

$$E(R^{ei}) = \alpha_i + \beta_i \lambda; \quad i = 1, 2, \dots, N$$

- (d) When a factor is itself a return, you can use the model prediction $\lambda = E(r_{mrf})$ and avoid actually running a cross sectional regression. Instead, you can find *the cross-sectional implications of the time-series regression*

$$E(R_t^{ei}) = a_i + \beta_i E(r_{mrf_t})$$

- (e) *Fama MacBeth Regressions* are a variant of cross sectional regression:

$$\begin{aligned} R_{t+1}^{ei} &= a_t + b_t z_{it} + \varepsilon_{t+1}^i \quad i = 1, 2, \dots, N \text{ at each } t, \text{ then} \\ \hat{b} &= \frac{1}{T} \sum_{t=1}^T \hat{b}_t, \quad \sigma(\hat{b}) = \frac{\sigma(\hat{b}_t)}{\sqrt{T}} \end{aligned}$$

z_{it} can be β_i , or characteristics such as size and book/market.

Week by week high points

1. Market return forecast

- (a) Issues:

- i. Does the expected return on the market $E_t(R_{t+1}^e)$ vary over time?
- ii. Why does P_t/D_t vary – future ΔD or future R ?

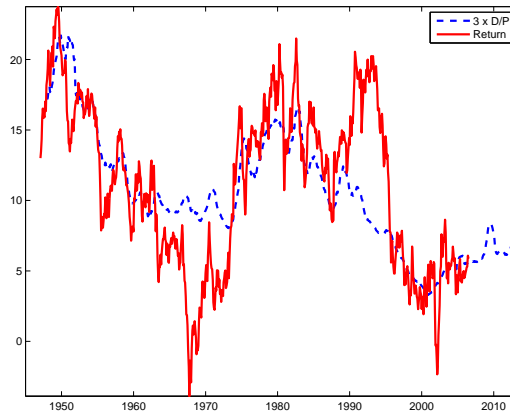
- (b) D/P can predict market returns. “Low” P gives high returns.

$$\begin{aligned} R_{t+1} &= a + b(D_t/P_t) + \varepsilon_{t+1}. \quad b \approx 4 \\ r_{t+1} &= a + b(d_t - p_t) + \varepsilon_{t+1} \quad b \approx 0.1 \end{aligned}$$

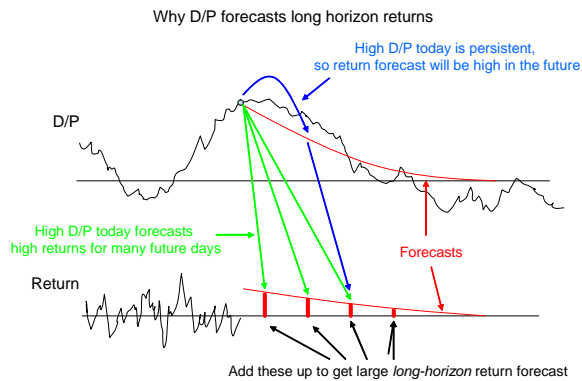
This means expected returns $E_t(R_{t+1}^e)$ vary over time

- (c) Economic significance:

- i. $\sigma[E_t(R_{t+1})] \approx 5\%$ is large relative to $E(R_{t+1}) \approx 7\%$. (R^2 measures $\sigma^2[E_t(R_{t+1})] / \sigma^2(R_{t+1})$, and $5^2/20^2 = 0.0625$.)
- ii. b, R^2 rise with horizon.



- iii. Stronger forecasts at long horizons result from a persistent forecasting variable (D/P) (we did this algebraically)



- iv. (Coming) return forecastability is “enough” to explain price volatility, that’s economic significance!
- (d) DP does *not* forecast Δd as it “should.”

$$\Delta d_{t+1} = a + 0 \times (d_t - p_t) + \varepsilon_{t+1}$$

- (e) Linearized present value formulas, useful tools.

$$\frac{P}{D} = \frac{1}{r - \Delta d}$$

$$p_t - d_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} (\Delta d_{t+j} - r_{t+j})$$

Much better than

$$\frac{P_t}{D_t} = E_t \sum_{j=1}^{\infty} \left(\frac{1}{R_{t+1}} \frac{1}{R_{t+2}} \cdots \frac{1}{R_{t+j}} \right) \frac{D_{t+j}}{D_t}$$

- i. Interpretation of the regressions: Prices today *reflect* expected dividend growth and expected returns for many periods in the future. If $E_t r_{t+j}$ rises, then $p_t - d_t$ will decline, and low $p_t - d_t$ will be followed by high returns on average, generating the regression.
- ii. Source: a useful return identity. Return must come from price rise or dividends!

$$\begin{aligned}
R_{t+1} &= \frac{D_{t+1} + P_{t+1}}{P_t} \dots (\text{algebra}) \dots \\
r_{t+1} &\approx \rho(p_{t+1} - d_{t+1}) - (p_t - d_t) + (\Delta d_{t+1}) \\
&= -\rho dp_{t+1} + dp_t + \Delta d_{t+1}
\end{aligned} \tag{51}$$

(f) Volatility and bubbles.

- i. If $d - p$ or $p - d$ vary, they must forecast long-run returns, long-run dividend growth, or their own long-run movements. The regression coefficients must add up. This means that we can account for price volatility with return forecasts, dividend growth forecasts or future prices.
- ii. Run both sides of

$$(d_t - p_t) = \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}$$

on $d_t - p_t$ and

$$1 = b_r^{lr} - b_d^{lr}$$

Long run return forecast and long run dividend growth forecast must add up.

- iii. 1 and 0. “should be” $b_d^{lr} = -1$, $b_r^{lr} = 0$. “Is” $b_r^{lr} = 1$ $b_d^{lr} = 0$
- iv. Regression coefficients are covariance over variance, so multiply through by $var(dp)$

$$var(dp_t) = var(p_t - d_t) = cov \left[dp_t, \sum_{j=1}^k \rho^{j-1} \Delta d_{t+j} \right] - cov \left[dp_t, \sum_{j=1}^k \rho^{j-1} r_{t+j} \right]$$

Measured variation in expected returns is just enough to account for all price-dividend volatility. Another measure of “economically large”

- v. If we only look out k steps

$$\begin{aligned}
(d_t - p_t) &= \sum_{j=1}^k \rho^{j-1} r_{t+j} - \sum_{j=1}^k \rho^{j-1} \Delta d_{t+j} + \rho^k (d_{t+k} - p_{t+k}) \\
1 &= b_r^{(k)} - b_d^{(k)} + \rho^k b_{dp}^{(k)}
\end{aligned}$$

High prices could mean prices that rise at ρ^k forever, a “bubble.” At 1 year, the last term is huge. At 15 years, it’s gone. This sense of “bubble” is not there.

- vi. Volatility Facts: Summary **Table II** from “Discount rates”

Method and horizon	Coefficient		
	$b_r^{(k)}$	$b_{\Delta d}^{(k)}$	$\rho^k b_{dp}^{(k)}$
Direct regression , $k = 15$	1.01	-0.11	-0.11
Implied by VAR, $k = 15$	1.05	0.27	0.22
VAR, $k = \infty$	1.35	0.35	0.00

$$\sum_{j=1}^k \rho^{j-1} r_{t+j} = a + b_r^{(k)} dp_t + \varepsilon_{t+k}^r$$

(g) VAR and impulse-response.

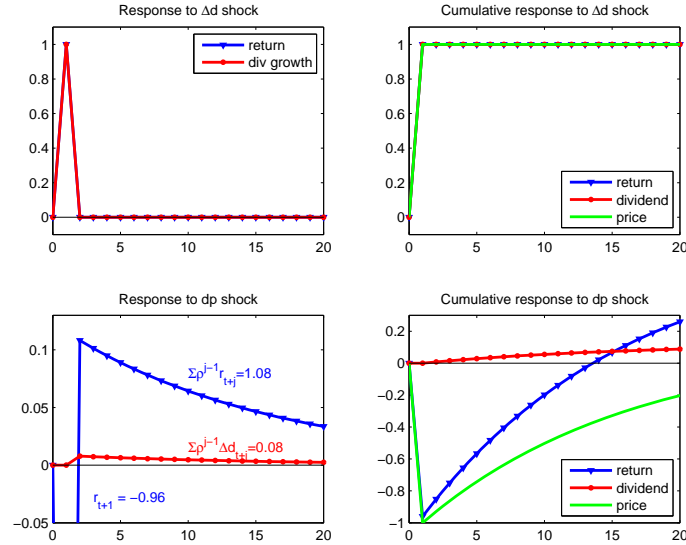
i. VAR

$$\begin{aligned} r_{t+1} &= b_r dp_t + \varepsilon_{t+1}^r \\ \Delta d_{t+1} &= b_d dp_t + \varepsilon_{t+1}^d \\ dp_{t+1} &= \phi dp_t + \varepsilon_{t+1}^{dp} \end{aligned}$$

ii. We can find implied long run forecasts and other statistics by iterating forward, for example,

$$r_{t+2} = b_r \phi dp_t + (b_r \varepsilon_{t+1}^{dp} + \varepsilon_{t+2}^r)$$

iii. “Impulse response.” Price movements *with no dividend change* melt away and correspond to higher expected returns. Here stocks are like bonds. Price movements *with a dividend change* are permanent and represent “cashflow risk”



Response to ε^d and ε^{dp} shocks from

$$\begin{aligned} \Delta d_{t+1} &= a_d + b_d dp_t + \varepsilon_{t+1}^d \\ r_{t+1} &= a_r + b_r dp_t + \varepsilon_{t+1}^r \\ dp_{t+1} &= a_{dp} + \phi dp_t + \varepsilon_{t+1}^{dp} \\ \text{with } \varepsilon_{t+1}^r &= \varepsilon_{t+1}^d - \rho \varepsilon_{t+1}^{dp} \end{aligned}$$

You can do this by hand: From

$$\begin{aligned} r_{t+1} &= b_r \times dp_t + \varepsilon_{t+1}^r = 0.1 \times dp_t + \varepsilon_{t+1}^r \\ \Delta d_{t+1} &= b_d \times dp_t + \varepsilon_{t+1}^d = 0 \times dp_t + \varepsilon_{t+1}^d \\ dp_{t+1} &= \phi \times dp_t + \varepsilon_{t+1}^{dp} = 0.94 \times dp_t + \varepsilon_{t+1}^{dp} \end{aligned}$$

and the identity (51)

$$\varepsilon_{t+1}^r = -\rho \varepsilon_{t+1}^{dp} + \varepsilon_{t+1}^d$$

plot the responses to $\varepsilon^{dp} = 1, \varepsilon^d = 0$ hence $\varepsilon^r = -0.96 = \rho$, and the response to $\varepsilon^d = 1, \varepsilon^{dp} = 0$ and hence $\varepsilon^r = 1$.

- (h) Good *expected* return news lowers *actual* returns (return decomposition).
- (i) Preview: The “random walk” is overturned in many markets. D/P forecasts stock returns, yield spreads forecast bond returns, interest rate spreads forecast fx returns.
- (j) Survey: Many other variables help to predict both stock returns and dividend growth. For example, cay (consumption/wealth)

$$R_{t+1} = a + b \times dp_t + c \times cay_t + \varepsilon_{t+1}$$

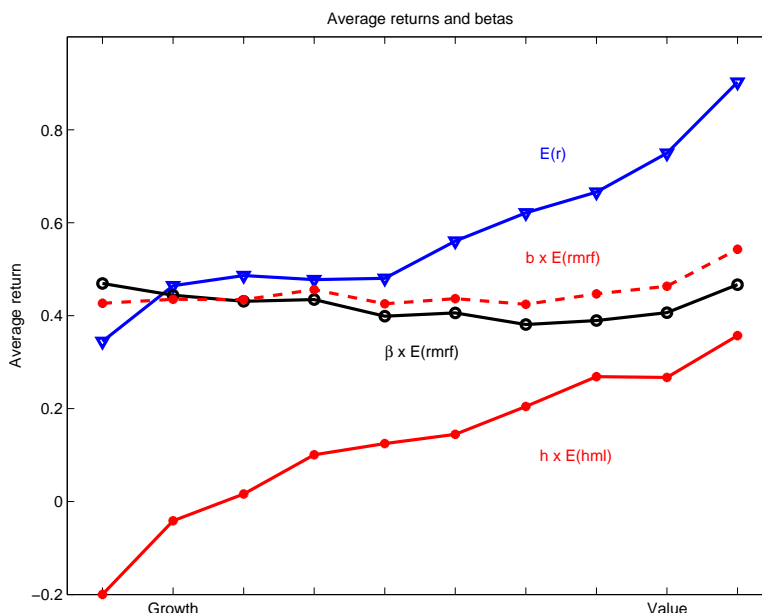
a high value of another variable raises *both* expected dividends and expected returns, or higher short run returns but lower long run returns, to leave dp unchanged.

$$(d_t - p_t) = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}$$

- (k) Interpretation: the premium for holding risk varies over time, higher in economic bad times (bottoms of recessions). We all want to sell; we can't; prices go down.

2. Fama-French and the cross section of average returns

- (a) Point: Explain the variation *across assets* in *average* returns.
- (b) Value and small stocks have higher average returns. This is not explained by CAPM betas. But it is pretty well explained by 3F betas - higher return portfolios have higher betas.



Average returns and betas for Fama - French 10 B/M sorted portfolios. Monthly data 1963-2010.

- (c) Portfolios sorted by E/P, Sales, 5 year return do well also in FF3F. The pattern in ER is matched by betas.

- (d) Portfolios sorted by momentum fail the FF3F. Past winners win, but have low, not high, hml loadings. A momentum factor?
- (e) Method: *Understand table 1. “Description” vs. “explanation”* Time series regression with 3 factors

$$R_{t+1}^{ei} = \alpha_i + b_i r_{mr} f_{t+1} + h_i hml_{t+1} + s_i smb_{t+1} \quad (52)$$

Then look at

$$E(R_{t+1}^{ei}) = \alpha_i + b_i E(r_{mr} f_{t+1}) + h_i E(hml_{t+1}) + s_i E(smb_{t+1}) \quad (53)$$

Look to see if assets with high $E(R^{ei})$ have high b_i, h_i or s_i ; look at α_i to see if they are small compared to $E(R^{ei})$. An “ocular cross-sectional regression.”

- (f) Also in Table 1: R^2 is very high. (52) is a great *3 factor model of return covariances*, as (53) is a great *three factor model of mean returns*. The difference? For covariances, we want a big R^2 even if there is big alpha. For means, we want small alpha even if there is small R^2 .
- (g) Another big point: forming *portfolios* to zoom the effects of regressions and turn them into simple average return questions.
 - i. Example: momentum takes a small autocorrelation, multiplies it by huge lagged returns, and produces a large average return spread.
 - ii. Equations: $R_{t+1} = a + 0.1R_t + \varepsilon_{t+1}$ $R^2 = 0.01$ means that the top 10% with $R_t = 100\%$ will have $E(R_{t+1}) = 10\%$.
- (h) FF works well on size, B/M, E/P, sales growth, and long-run reversal.
- (i) FF fails to explain momentum, because high momentum portfolios act like growth, but earn high returns. The betas go the wrong way.
- (j) *Fama and French, Dissecting Anomalies* takes on even more sorts (i.e. beyond B/M, size, E/P, sales growth, 5 year reversal, 1 year momentum.) It also asks what is pervasive in the market and what is only a feature of micro-cap stocks.
 - i. Main table: characteristic-adjusted 1-5 portfolio returns. Momentum, Net Stock issue are still strong. Asset growth, profitability only seem to work in tiny stocks.
 - ii. Table 4: cross sectional regressions of average returns on characteristics.

$$E(R^{ei}) = a + b \log(size_i) + c \log(bm_i) + \dots$$

These are a way of describing patterns in average returns, not an “explanation!” They let you see which anomalies survive in the presence of others — is net stock issues just the same as book/market, both revealing “high” prices? You can’t just stare at a table of portfolio means when there are more than two (size, bm) anomaly variables.

- iii. Unanswered: do we need new factors?

3. More anomalies, More factors

- (a) Novy-Marx, The Gross Profitability Premium

i. Idea:

$$pd_t \approx E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$$

$$\frac{M_t}{B_t} = E_t \sum_{j=1}^{\infty} \frac{Y_{t+j} - dB_{t+j}}{(1+r)^j B_t}$$

so, *variables that help to predict future earnings can help to forecast returns, given the M/B or D/P ratio.*

ii. Or, controlling for price, the good stock/good company fallacy is right!

iii. Novy-Marx: Gross profits today are a better forecast of net profits and dividends in the future, so better return forecasters.

iv. Evidence: Table 1 FMB regressions, profits forecast returns; Table 6: double sorts show higher returns to higher profitability; Table 9: PMU factor is not priced by the others; Table 10 PMU factor subsumes other expected return sorts.

(b) Fama, Eugene F., and Kenneth R. French 2013, "A five-factor asset pricing model"

i. A similar earnings factor, drives out HML!

(c) Frazzini, and Pedersen 2013 Betting against beta

i. Table III: Beta portfolios have about the same $E(R)$ but different betas: a puzzle (puzzles are always joint!) BAB portfolio generates positive alpha, zero beta "sml is too flat."

4. Asset pricing theory

(a) APT Assets with high expected returns must move together, or there would be huge Sharpe ratios from forming portfolios.

i. Lesson: regressions are also directions to form risk-minimizing portfolios, portable alpha!

ii. Math:

$$R_{t+1}^{ep} = R_{t+1}^{ei} = \alpha_i + \beta_{i1} f_{t+1}^1 + \beta_{i2} f_{t+1}^2 + \varepsilon_{t+1}^i$$

$$R_{t+1}^{ep} = R_{t+1}^{ei} - (\beta_{i1} f_{t+1}^1 + \beta_{i2} f_{t+1}^2) = \alpha_i + \varepsilon_{t+1}^i$$

$$\text{Sharpe} = \alpha_i / \sigma(\varepsilon^i)$$

$$\text{Portfolios} : w_\varepsilon = \frac{1}{\gamma} \frac{\alpha_i}{\sigma^2(\varepsilon^i)}$$

iii. You: buy large Sharpe ratios. Market equilibrium: when large Sharpes are gone, "small" ε^i imply "small" α_i .

iv. APT does not explain why factors get their premiums, why $E(f)$, why $E(hml)$? "Relative pricing."

v. APT logic does not extend past high R^2 securities. CAPM/multifactor logic can apply anywhere.

(b) Consumption-based / all asset pricing theory. Bottom line comes down to

$$p_t = E_t(m_{t+1}x_{t+1})$$

The basic m_{t+1}

$$m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)} = \beta \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \approx 1 - \delta - \gamma \Delta c_{t+1}.$$

Source: an investor thinks about how much of an asset to buy, maximizes utility.

(c) Classic issues:

i. Interest rates

$$R_t^f = 1/E(m) \approx 1 + \delta + \gamma E_t(\Delta c_{t+1})$$

Rates are higher if people are impatient, and in good ($E(\Delta c)$) times.

ii. Valuing risk, only covariance matters

$$p_t = E(m_{t+1}x_{t+1}) \approx \frac{E_t(x_{t+1})}{R^f} - \gamma \text{cov}(x_{t+1}, \Delta c_{t+1})$$

iii. The expected return premium depends on covariance of returns with consumption growth. Assets must pay a higher average return to investors (low price) if they tend to do badly in *bad* times. From $0 = E(mR^{ei})$,

$$E(R_{t+1}^{ei}) \approx \gamma \text{cov}(R_{t+1}^{ei}, \Delta c_{t+1}) = \beta_{i,\Delta c} \times \lambda_{\Delta c}$$

iv. “only systematic risk matters” “idiosyncratic variance isn’t priced.”

$$R^{ei} = \beta_{i,m}m + \varepsilon^i$$

$$E(R^{ei}) = \beta_{i,m}\lambda$$

$$\text{var}(R^{ei}) = \beta_{i,m}^2 \sigma_m^2 + \sigma_{\varepsilon^i}^2$$

($m = \Delta c$ too)

v. Mean-variance frontier

A. From $0 = E(mR^e)$

$$\frac{E(R^{ei})}{\sigma(R^{ei})} = \frac{\sigma(m)}{E(m)} \text{corr}(m, R^{ei}) \leq \frac{\sigma(m)}{E(m)} \approx \gamma \sigma(\Delta c)$$

B. Roll theorem: Frontier returns are perfectly correlated with m , hence carry all pricing information

$$E(R^{ei}) = \beta_{R^e, R^{emv}} \lambda_{mv} \leftrightarrow R^{emv} \text{ is on the mvf}$$

(d) **CAPM, ICAPM, Multifactor models, APT**

i. The models are all excuses for

$$m_{t+1} = \beta \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \approx b' f_{t+1} \leftrightarrow E(R^e) = \beta_{R^e, f} \lambda$$

- ii. Idea: We use other *proxies* for good and bad times and ignore consumption data and its measurement and definition problems. But a “model” also says that no *other* factors matter.
- iii. Example: log utility CAPM $u = \log(c)$, means $\frac{p_t}{c_t} = E_t \int e^{-\rho s} \left(\frac{c_{t+s}}{c_t} \right)^{-1} \frac{c_{t+s}}{c_t} ds = \frac{1}{\rho}$ constant, and thus the return on the consumption claim = market = consumption growth = reference for CAPM. Point: consumption = wealth, so you can measure consumption growth with wealth growth.
- iv. Multifactor “Theory”: market, news about future returns (news that $E_t R_{t+1}$ is low is bad news), non-market income, macro variables can all be factors.
- v. Most measures of “good and bad times” in practice are portfolio returns. Market return \rightarrow CAPM. If consumption growth is driven by rmrf, hml, smb, we get the FF3F model. These work better in practice, at the cost of a lot of theoretical purity. Checking whether factors (hml, smb) do represent proxies for identifiable macro risks is still going on with no big success to report.
- vi. Portfolio logic for multifactor models. A, B have the same mean, variance, beta. In a recession (“state variable”), A goes up while B goes down.
 - A. \Rightarrow People want more A/less B \Rightarrow Price of A goes up/ B down \Rightarrow Expected returns of A go down/ B up.
 - B. \Rightarrow Expected returns depend on recession sensitivity as well as market sensitivity. In equations,

$$E(R^{ei}) = \beta_{i,m} \lambda_m + \beta_{i,recess} \lambda_r$$

- vii. Tools:
 - A. $m = a - b'f \leftrightarrow E(R^e) = \beta \lambda$
 - B. R^e on MVF $\leftrightarrow E(R^e) = \beta \lambda$
 - C. Mimicking portfolio theorem: the regression of m on all asset returns R^e prices as well as m itself. “Proxies for state variables.”
 - D. Quadratic, log utility derivations of the CAPM. $c = kW$ to substitute consumption to market return.

5. Empirical methods – not covered 2014

- (a) Objective: See if

$$\begin{aligned} E(R^{ei}) &= (\alpha_i) + \beta'_i \lambda \\ y &= \varepsilon + x'b \end{aligned}$$

is true, i.e. if $\alpha_i = 0$.

- (b) All methods tell you how to *estimate* free parameters, $\alpha, \beta, \lambda, E(R^e)$; how to find *standard errors*, and how to *test* the model, primarily “are all the alphas = 0?”
- (c) Methods:
 - i. Time series regression (α, β from OLS TS regression, $\lambda = E(f)$)
 - ii. Cross section regression (β from OLS TS regression, α, λ from cross sectional)
 - iii. Fama-MacBeth.

- (d) A statistical test whether all the α are in fact equal to zero, and the α you see are just luck. The weighted sum of squared α s. $(\alpha' \Sigma^{-1} \alpha)$. “GRS” test.
- (e) When factors are also excess returns, you can just run *time-series regressions* to evaluate the model. The point of the model: the time series *intercepts* α are the cross sectional *errors* and should be zero – the *cross-sectional implication* of the time series regression. It’s not really about time-series R^2 , etc.
- (f) When factors are not excess returns, you can run a *cross-sectional regression* of average returns on betas *across assets*. Yucky formulas for standard errors and GRS-like test.
- (g) Should you drop smb? (for example)

$$\begin{aligned} E(R^{ei}) &= \alpha_i + b_i E(rmr f) + h_i E(hml) + s_i E(smb) \\ E(R^{ei}) &= \alpha_i + b_i E(rmr f) + h_i E(hml). \end{aligned}$$

- i. Not $E(smb) = 0$. Instead, test for $\alpha_s = 0$ in

$$smb_t = \alpha_s + b_s rmr f_t + h_s hml_t + \varepsilon_t^s$$

If $\alpha_s = 0$ then the α_i above are the same. Then dropping smb won’t change alphas.

- ii. *It depends on your purpose.* smb may be an important “unpriced factor.” Compare

$$R_t^{ei} = \alpha_i + b_i rmr f_t + h_i hml_t + s_i smb_t + \varepsilon_t^i \quad (54)$$

$$R_t^{ei} = \alpha_i + b_i rmr f_t + h_i hml_t + \varepsilon_t^i \quad (55)$$

Even if you don’t need smb for *average returns* (α), smb and other non-priced factors (industry) can be important for understanding *variation* in returns (risk management, performance evaluation, better t stats). (54) is very important to an investor who owns a small firm, and wants a portfolio tilted away from small stocks that will all tank when his company tanks, even if the average investor doesn’t care.

6. Mutual Funds.

- (a) Survivor bias, why we want to look at the average fund (in some group) not the “good” (lucky?) funds
- (b) Performance attribution – it’s ok to include factors whose means may be zero, and which don’t have deep economic foundation (umd, put premium, etc.)

$$R^{ei} = \alpha_i + b_i rmr f_t + h_i hml_t + s_i smb_t + u_i umd_t + p_i putwrite_t + \dots + \varepsilon_t^i$$

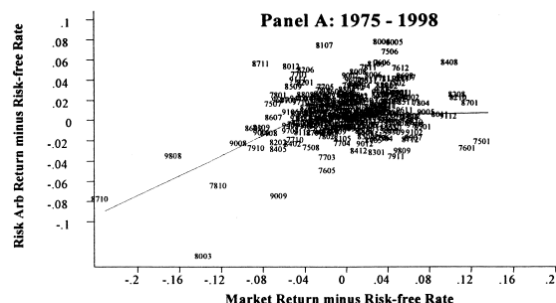
The question is only: can I do what the fund does with easy passive investments? Once again, the model you use depends on your purpose.

- (c) Carhart facts:
 - i. Good one-year fund returns persist.
 - ii. They are not explained by the CAPM. They are well explained by 4 factor model. -0.1% month alpha throughout. A small puzzle that losers keep losing so much.
 - iii. Portfolios sorted by past 5 year returns do not do well, another indication it’s momentum not skill.

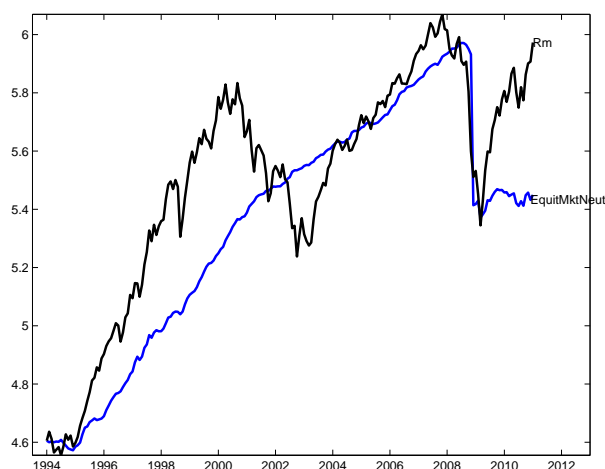
- iv. Funds that did well last year have stocks that will keep going up due to stock momentum, not momentum funds.
 - v. Cross sectional regressions show that turnover and fees cost investors. This is puzzling, even the Grumpy Economist thinks the effect should be zero.
- (d) Fama and French: By cleverly using a σ/\sqrt{T} type calculation they infer the distribution of sample alphas under the hypothesis that true alpha is zero. How often do we see alphas that we see in sample by chance? We can also infer the distribution of true alpha. They calculate, if true alpha is zero, how many funds with large α (really α t statistics) should we see? Roughly, only 5% should have α t stat > 2 . In fact, we see slightly more with t stat > 2 , and many more than 5% with a t-stat < -2 . Roughly it looks like the distribution of “true alpha” has a negative mean and a standard deviation of about 1%
- (e) Berk: An economic model with surprising results. If you have skill, investors give you money until returns to investors are back to normal. Thus, flows rationally follow performance, returns don’t persist, but there is skill. Example: Manager has 6% alpha only for the first \$1m. Attracts funds until he has \$6m AUM, then gets 60k fees.
- (f) Berk II: we should measure alpha x assets under management, so small alpha times large assets = skill. We should not benchmark to portfolios without transactions costs, and that nobody was treating passively at the time, value in 1968. With that, lots of sample “skill.” Fama French distribution? Not done.

7. Hedge funds

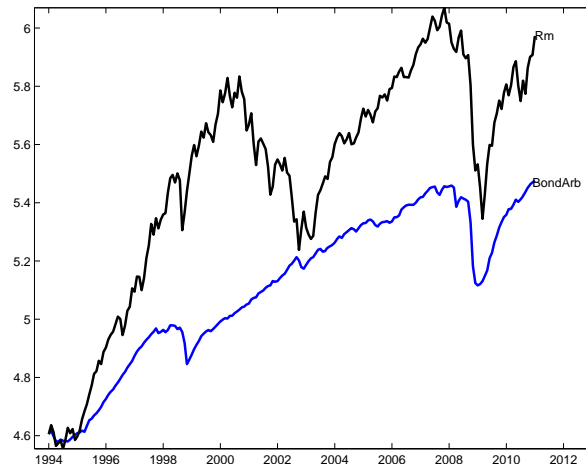
- (a) Complex strategies can work like writing put options.
- i. This makes statistical analysis hard. Sampling from +1 +1 +1 +1 +1 -1000 +1 +1...you don’t often see -1000
 - ii. Writing put options may be profitable even including the losses. “catastrophe insurance,” “providing liquidity” “short volatility” etc. The price (implied vol) of out of the money puts is high.
- (b) Mitchell/Pulvino
- i. What is merger “arbitrage?”
 - ii. Big picture: Merger arb returns looks like writing index puts, especially on cash offers. It makes a small profit if it works, loses a huge amount if merger is withdrawn. Mergers tend to be withdrawn in huge market losses.



- iii. See figure 4 of two-part beta regressions. Tables: about 0 beta in up markets and 0.5 beta in down markets overall. For cash mergers it's much stronger, (Table VI) 0.77 in down markets.
 - iv. The paper also shows how important transactions costs are. Profits without transactions costs, liquidity limits look great. This is an important cautionary tale to your back-testing efforts.
 - v. If you want to interpret the intercept as alpha, it has to be option *returns*, so you may have to make up prices via Black-Scholes or use actual option returns.
- (c) Asness et al “Do hedge funds hedge”
- i. 3 month lag betas matter too, and raise the overall beta to substantial levels.
 - ii. Interpretation: stale pricing of illiquid securities (and some optimistic pricing)
 - iii. Up/down betas are very different. There is much more down beta than up beta. Interpretation: optimistic stale pricing. Our interpretation: option-like character of returns
 - iv. Problem set: both issues extend to other factors, especially default. Up/down beta seems to be dying out for many styles?
 - v. Here's a big (and still strong) up/down beta, looking just like writing a put option.



- vi. Here's one that had an up/down beta in 1998, but recovered in the crash, so it has a big up beta too – bigger betas for big moves in both directions



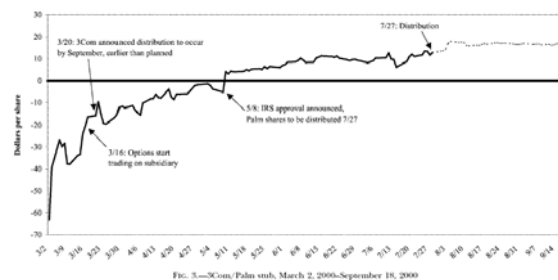
(d) My comments:

- i. Problems with the 2+20 option-like compensation scheme. Incentive to volatility, portfolio of options \neq option on portfolio.
- ii. Betas are in principle what we want, but short horizons, style drift, and exploding numbers of factors make it hard. Analyze portfolios instead, or require beta reporting?
- iii. Forming portfolios of hedge funds: an open question and minefield. Is the fund short what you are long? Is fund A short and fund B long?

8. Short sales and “overpricing”

(a) Lamont/Thaler

- i. 3Com/Palm story. 3com keeps 95% of Palm, and will spin off to shareholders. You can buy Palm cheaper by buying 3Com than by buying Palm. A “negative stub.”
- ii. It lasted a long time.



- iii. Mispricing is associated with huge volume. Palm has huge turnover but also larger bid/ask spreads. It looks like a big demand for information trading.
- iv. 3com *fell* as Palm exploded on first day.
- v. You can’t sell short to eliminate “mispricing.”
- vi. Put/call parity is violated, so you can’t synthesize a short in the options market.

vii. As short, options warm up “supply” increases a lot, disparity declines.

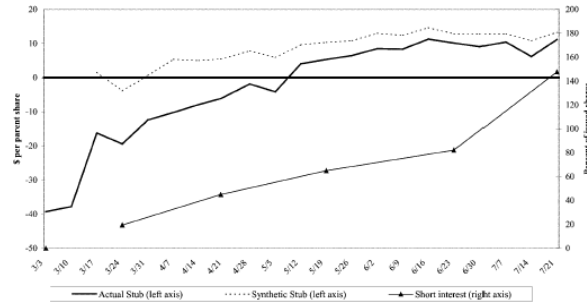


FIG. 5.—3Com/Palm: actual stub, synthetic stub, and short interest, March 5, 2000–July 21, 2000

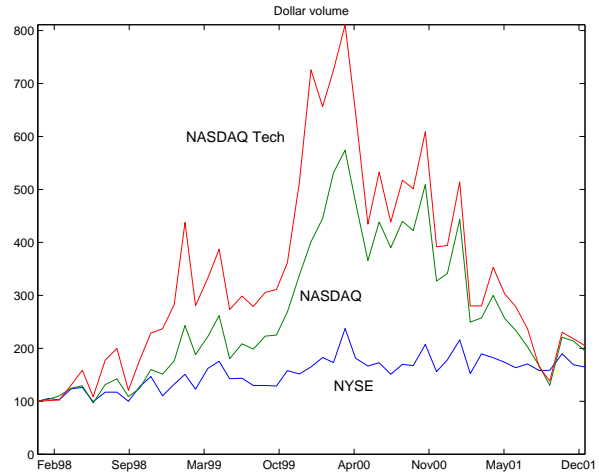
viii. Yes, you can’t arbitrage, but why are prices wrong in the first place? Who is buying Palm rather than 3 com and waiting? L&T: retail investors, “morons.”

(b) Cochrane, Stocks and Money

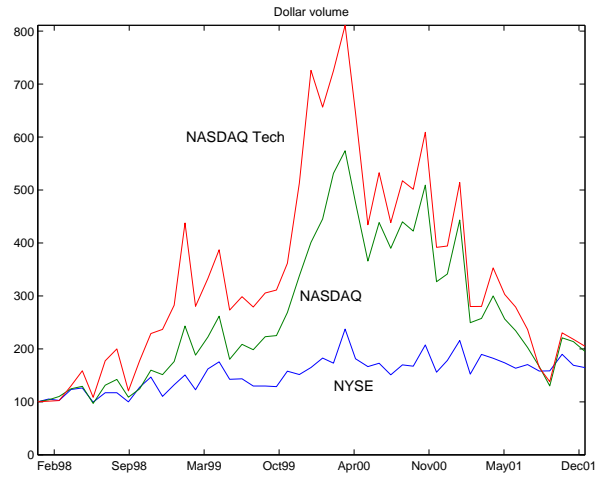
- i. 3com/Palm looks like bond/money in a hyperinflation.
- ii. Money, bond are both claims to \$1 in 6 months. Money is “overpriced.” This one, we understand. Nobody holds money for 6 months, it turns over.
- iii. Mispricing is large when
 - A. Higher turnover
 - B. Lower money supply
 - C. Short constraints (print money!)
 - D. Need a special demand for *this* security. Money: to make payment. Palm? Information trading.
- iv. Evidence: Fits all L&T facts. In particular
 - A. Turnover.
 - B. Fall of parent.
 - C. Very small initial float in 5% carve-out. Price decline as shorts add supply
- v. What is the special demand for Palm shares, that can’t be satisfied with 3Com and a little patience? Plots that 3com, palm are delinked in intraday data. You must daytrade palm stock.

	b	R^2	$\sigma(y)$	$\sigma(\varepsilon)$
One day returns:				
$\text{Palm}_t = a + b3\text{Com}_t + \varepsilon_t$	0.96	0.60	7.2	4.5
$(\text{Palm}_t - \beta\text{Nasdaq}_t) = a + b(3\text{Com}_t - \beta\text{Nasdaq}_t) + \varepsilon_t$	0.93	0.53	6.9	4.6
Five day returns:				
$\text{Palm}_t = a + b3\text{Com}_t + \varepsilon_t$	1.03	0.69	15.0	8.3
$(\text{Palm}_t - \beta\text{Nasdaq}_t) = a + b(3\text{Com}_t - \beta\text{Nasdaq}_t) + \varepsilon_t$	0.95	0.54	13.4	10.0

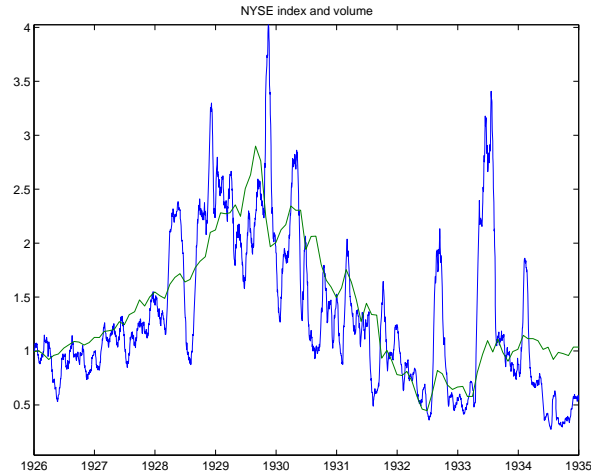
vi. More general liquidity premia in stocks?



Dollar volume on NYSE, NASDAQ and NASDAQ with SIC code 737. Series are normalized to 100 on Jan 1 1998.



Dollar volume on NYSE, NASDAQ and NASDAQ with SIC code 737. Series are normalized to 100 on Jan 1 1998.



Share volume (20 day moving average) and NYSE index in the great crash. Volume is the more volatile series. NYSE index from CRSP; volume from NYSE. Both series normalized to 1 in 1926.

vii. Who are all these information traders, and how can they all be smarter than average?

9. Liquidity and trading

- (a) Brandt and Kavajecz Signed order flow correlates with yield changes. But is it “price discovery” or “price impact” = “inventory?”
 - i. 2-5 year order flow predicts other yield changes
 - ii. On the run orderflow predicts off the run yield changes
 - iii. Yesterday’s orderflow does not signal a bounce-back..
- (b) Hasbrouk and Saar, HFT. Weird clock time clustering of messages. “Liquidity?”
- (c) Kyle et al flash crash. HFT’s do 4 seconds of momentum, 10 seconds of reversal. Extra graphs – weird stuff in high frequency markets leading to the suspicion of less liquidity?
- (d) Budish, Cramton and Shim. Chicago vs. NY arbitrage. Correlations that are 1 at hours go to zero for fine enough intervals. Correlations have moved to shorter times. “Arbitrages” are just as big but don’t last as long. Proposal for once per second batch auctions.

10. Financial Crisis

- (a) Background: Many stories/fancy words.
 - i. Bank runs: Illiquid assets, liabilities that give an incentive to run, if I run, the bank is more illiquid
 - ii. Debt overhang story. Bankruptcy as recapitalization
 - iii. Short term debt; the example of a 10 year project financed by rolling over short term debt.
 - iv. “Standard policy tools:” Deposit insurance, regulation, lender of last resort, resolution, and how they slowly blew up.

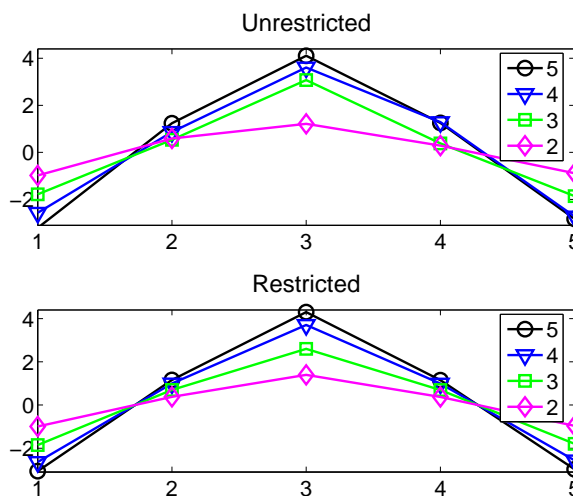
- (b) Duffie. How a Bear/Lehman fails. Repo, brokerage accounts and derivatives margins have “run” features: You have an incentive to pull out before bankruptcy, and if you do, that makes the bank less liquid because they rehypothecate your securities, i.e. use them as collateral for their own borrowing and trading.
- (c) Gordon and Metrick
- The run on repo markets was a (the) central part of the credit freeze.
 - Repo, how it works. Suddenly haircuts were much higher because people didn’t want to be stuck with bad collateral. This is like a rise in reserves requirements, drives a huge demand for “liquidity.”
 - “Information insensitive” securities in normal times become “information sensitive” and hence illiquid. “Repo is like e-coli”
 - The “systemically insolvent” problem. Failure at X makes us worry about Y, pull out our money, but we can’t do that in aggregate

11. Term structure: expectations and bond risk premia.

- (a) Definitions. Price, yield, return, forward rate, logs vs. levels.
- (b) Expectations. The *expected* return from two different ways of getting money across time should be the same
- Long maturity yield = average of expected future short rates (plus risk premium)*
 - Forward rate = expected future spot rate (plus risk premium)*
 - Expected holding period returns should be equal across maturities (plus risk premium)*
 - (Problem set: $f_t^{(n)} = E_t \left(f_{t+1}^{(n-1)} \right)$)
- (c) Strict: no risk premium. Usual: a constant risk premium
- (d) Similarly, *expected FX depreciation = interest differential across countries (plus risk premium)*
- (e) Empirical: you can see in a plot that an upward sloping yield curve is followed by a rise in yields – but a bit late.
- (f) Fama-Bliss:
- A forward rate 1% above a spot rate corresponds to 1% higher return on long term bonds!
 - Equivalently, a forward 1% above spot corresponds to 0% increase in spot next year. (Understand FB table!)

n	$rx_{t+1}^{(n)} = a + b \left(f_t^{(n)} - y_t^{(1)} \right) + \varepsilon_{t+1}$				$y_{t+n-1}^{(1)} - y_t^{(1)} = a + b \left(f_t^{(n)} - y_t^{(1)} \right) + \varepsilon_{t+1}$			
	b	$\sigma(b)$	R^2		a	b	$\sigma(b)$	R^2
2	0.83	0.27	0.11		0.17	0.27	0.01	
3	1.14	0.35	0.13		0.53	0.33	0.05	
4	1.38	0.43	0.15		0.84	0.26	0.14	
5	1.05	0.49	0.07		0.92	0.17	0.17	
	forecasting <i>one</i> year returns on <i>n</i> -year bonds				forecasting <i>one</i> year rates <i>n</i> years from now			

- iii. The expectations hypothesis works better at long (5 year) horizons. The forward rate does forecast the 1 year rate 4 years out. The required adjustment happens, just “sluggishly.”
- (g) Exchange rates: Foreign 1% over US seems to imply even more than 1% return for a year – though with really tiny R^2 (0.03 monthly). Equivalently, exchange rates seem if anything to go the wrong way.
- (h) Both: high expected returns when interest rates are low, in the depths of a recession.
- (i) Cochrane-Piazzesi update: forecast 1 year bond returns with *all* forward rates, not just matched forward rate. Result:



There is a single, tent-shaped, combination of forward rates that forecasts the returns of bonds of any maturity. Then it forecasts long term bonds more than short term bonds.

$$rx_{t+1}^{(n)} = b_n (\gamma' f_t) + \varepsilon_{t+1}^{(n)}, \quad n = 2, 3, 4, 5.$$

It drives out FB, forecasts stock returns, gets a 0.35-0.42 R^2 . Signal: a slope plus the 4-5 spread. Why?

- (j) (Omitted 2014) Lustig Roussanov and Verdelhan / Jurek FX update.
 - i. The regressions are strong through 2007, though still with very small R^2 . All “carry trade” did poorly in 2008, and this may be the “put option” data point that we’ve been waiting for all these years.
 - ii. Jurek looks at portfolio means and sharpe ratios (invest in high ($i^* - i$) countries) and they are high even though R^2 is low. Like momentum, forming portfolios shows you how even small R^2 can give large returns. Again, though, numbers are only through 2007.
 - iii. LRV apply FF procedures to FX markets. They form FF-like 1-6 portfolios based on country interest differentials vs. US. The high interest differential portfolio has a high average return, and vice versa.
 - iv. Then they factor-analyze these portfolio returns, finding a “level” (US dollar appreciation) factor, and a “carry trade” (high interest differential countries go down, low differential countries rise) factor. The “carry trade” factor exposures line up with mean returns. I.e. they did just like B/M portfolios and an HML factor for FX returns across countries.

12. Interest rates IIa: factor models.

- (a) Central trick: Eigenvalue decomposition produces a factor model from a covariance matrix.

$$\begin{aligned}\Sigma &= Q\Lambda Q' \\ &= \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} - & q'_1 & - \\ - & q'_2 & - \\ - & q'_3 & - \end{bmatrix} \\ &\quad Q'Q = QQ' = I\end{aligned}$$

- (b) The columns of Q tell you how to construct the factor x from data on y .

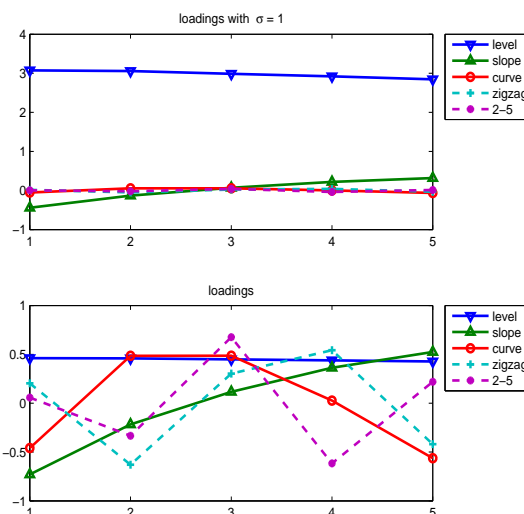
$$\begin{aligned}x_t^{(1)} &= \begin{bmatrix} - & q_1 & - \end{bmatrix} \begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \end{bmatrix} \\ x_t^{(2)} &= \begin{bmatrix} - & q_2 & - \end{bmatrix} \begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \end{bmatrix} \\ &\dots\end{aligned}$$

- (c) The columns of Q tell you how a movement in each factor moves the y in turn. These loadings give a multifactor model.

$$y_t = Qx_t$$

$$\begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \end{bmatrix} = \begin{bmatrix} | \\ q_1 \\ | \end{bmatrix} x_t^{(1)} + \begin{bmatrix} | \\ q_2 \\ | \end{bmatrix} x_t^{(2)} + \begin{bmatrix} | \\ q_3 \\ | \end{bmatrix} x_t^{(3)}$$

- (d) The x factors so constructed are uncorrelated and have $\sigma^2(x^{(i)}) = \lambda_i$.
- (e) A natural idea: drop factors with very small variance, and get an approximate model with a smaller number of factors. For example, we can describe movements in 30 bond yields by 3 factors, “level,” “slope” and “curve.”
- (f) The result for yields



13. The expectations-hypothesis factor model illustrates the logic of term structure models

$$y_{t+1}^{(1)} - \delta = \rho(y_t^{(1)} - \delta) + \varepsilon_{t+1}$$

$$f_t^{(2)} = E_t(y_{t+1}^{(1)}) = \delta + \rho(y_t^{(1)} - \delta)$$

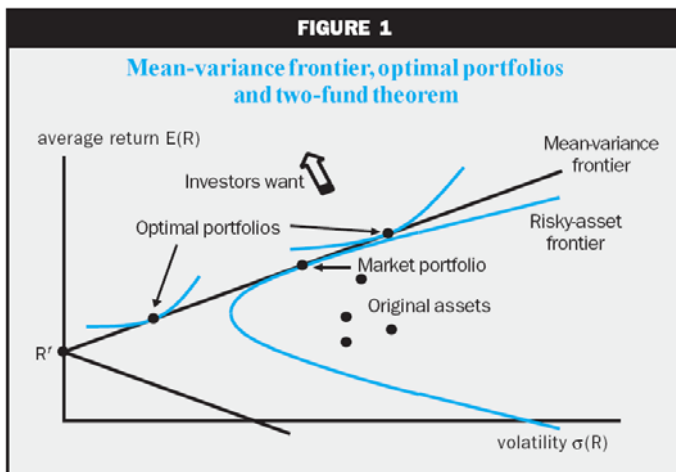
$$f_t^{(3)} = E_t(y_{t+2}^{(1)}) = \delta + \rho^2(y_t^{(1)} - \delta)$$

$$f_t^{(N)} = E_t(y_{t+N-1}^{(1)}) = \delta + \rho^{N-1}(y_t^{(1)} - \delta)$$

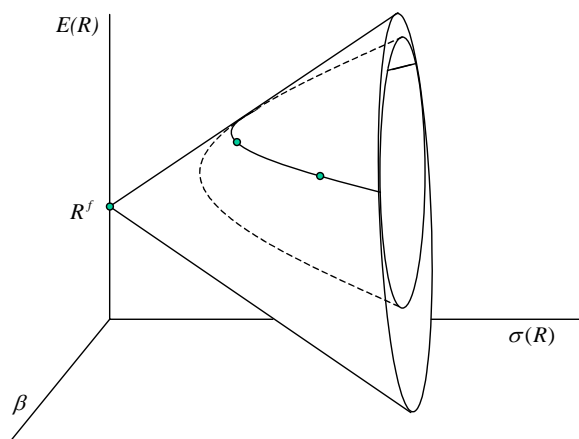
A single factor model, which you can extend to any maturity

14. Portfolio Theory.

(a) Graphical review of mean-variance, two fund theorem.



(b) How does a new factor, e.g. value, affect things?



Three (N) fund theorem. The market is no longer mean-variance efficient. A bit of a reason for management in “style coaching.”

(c) Math: all portfolio problems are simple (solveable!) versions of this:

$$\max_{\{c_t, w_t\}} E \sum_{j=0}^{\infty} \beta^j u(c_t) \text{ or } E \int_0^{\infty} e^{-\rho t} c_t dt;$$

$$\begin{aligned} W_{t+1} &= R_{t+1}^p (W_t + y_t - c_t); \\ R_{t+1}^p &= R_t^f + w_t' R_{t+1}^e \\ &\quad E_t(R_{t+1}^e), \text{ cov}_t(R_{t+1}^e) \text{ given} \end{aligned}$$

The new research just adds y_t and E_t, cov_t .

(d) Portfolio weights

$$w = \frac{1}{\gamma} \Sigma^{-1} E(R^e) + \frac{\eta}{\gamma} \beta_{R,z'}$$

Mean-variance plus hedging demand. Jobs / changing $E_t R_{t+1}^e$, etc. introduce the second term.

(e) The portfolio relative to the market (if everyone is like this)

$$\begin{aligned} R^i &= R^f + \frac{\gamma^m}{\gamma^i} R^{em} + \frac{1}{\gamma^i} (\eta^{i'} - \eta^{m'}) R^{ez} \\ R^{ez} &= \beta_{z,R'} R^e = \text{best hedge for state variable risk (hml)} \end{aligned}$$

(f) Relative to a factor model (say, capm), for the $\eta = 0$ case allows us to separate “policy” and “alpha chasing”

$$\begin{aligned} R_t^p &= R^f + w' R_t^{em} + w_{\varepsilon}' (R^e \alpha + \varepsilon_t) \\ w &= \frac{1}{\gamma} \Sigma_f^{-1} E(F) \\ w_{\varepsilon} &= \frac{1}{\gamma} \Sigma_{\varepsilon}^{-1} \alpha \end{aligned}$$

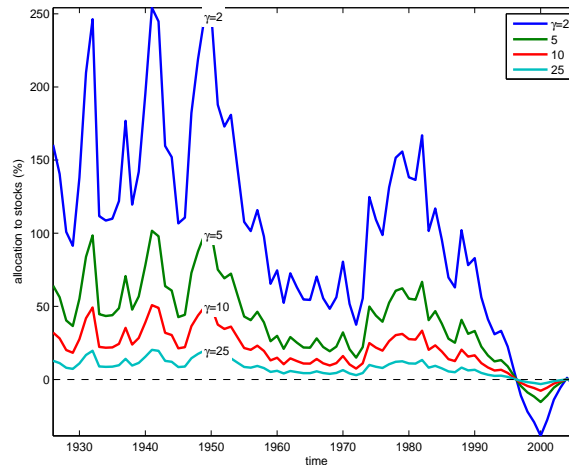
(g) Applications

- i. $\eta^i = \eta^m = 0$ the mean-variance case.
- ii. $\eta^i = 0$: MV investors can do better than market portfolio!
- iii. $\eta^i \neq 0$: Hedge labor or business income, even if $\eta^m = 0$.
- iv. Everyone: k funds, not 2 funds, and a hard problem.
- v. Hedging demand example: long term bonds – it totally changes the picture!
- vi. Market timing demand from the first term. Example

$$E_t(R_{t+1}^{em}) = a + b \left(\frac{D_t}{P_t} \right)$$

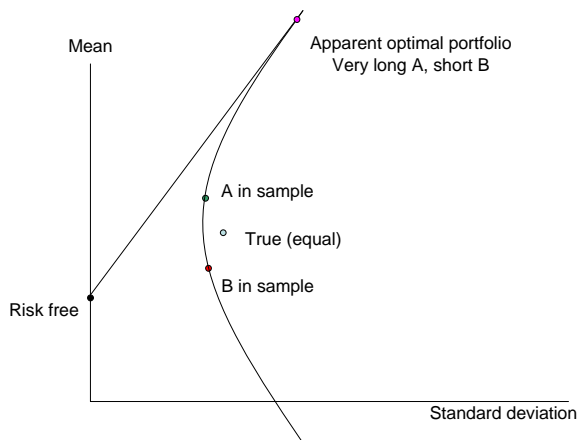
gives

$$w_t = \frac{1}{\gamma} \frac{a + b \left(\frac{D_t}{P_t} \right)}{\sigma_{\varepsilon}^2}$$



Wild, but do we believe it?

- (h) Wacky weights problems in all portfolio optimizers, since they are very sensitive to input assumptions.



- (i) You can mitigate this by shading inputs back towards inputs that give “hold the market” (Barberis/Pastor-Stambaugh.) See that uncertainty about parameters is a source of risk to you, the investor, and lower portfolio weights accordingly

$$w = \frac{1}{\gamma} \frac{E(R^e)}{\sigma^2(R^e) + \sigma^2(E(R^e))}$$

- (j) Closing thoughts:

- i. *The average investor holds the market.* Advice can only hold for those who are different from average.
- ii. Hedging is really poor right now. Styles (industry, etc.) advice, advice to hedge your other income does not depend on a premium (alpha) and is widely ignored.

15. Relax! Good luck on final.