Empirical Asset Pricing

Francesco Franzoni

Swiss Finance Institute - University of Lugano

ICAPM, Recent Trends, and Skeptical Appraisal

Lecture Outline

1. Bad beta, good beta
2. Growth or Glamour
3. Recent trends
4. A skeptical look at AP tests

Relevant readings:

- Campell and Vuolteenaho, 2004, "Bad beta, Good beta", American Economic Review
- Quickly look at: Campbell, Polk, and Vuolteenaho, 2009, "Growth or Glamour? Fundamentals and systematic risk in stock returns", Review of Financial Studies
- Take a look at: Zhang (2005, JF), Bansal, Dittmar, and Lundblad (2005, JF)
- Also look at: Lewellen, Nagel, and Shanken, 2009, "A skeptical appraisal of asset pricing tests", Journal of Financial Economics

1. Bad beta, Good beta

Motivation

- CAPM failure: low unconditional beta stocks earn high returns
- These stocks are typically value and small portfolios
- In a CAPM world, investors should overweight these stocks and the premium should disappear
- But the premium has been there for a while
- We focus on this paper because it points out that cash flow betas could solve the value premium
- Also the methodology for return decomposition has been very influential and has raised recent criticism
- There are a few aspects that are subject to the Lewellen and Nagel critique

Possible Explanation

- The inspiration comes from Merton's (1973) ICAPM
- Possibly, these stocks are not good hedges for changes in investment opportunity set
- Hence, they should pay a high risk premium
- The opposite is true for growth stocks
- In other words, value stocks are more risky than growth stocks along some dimension and this fact prevents investors from overweighting value stocks in their portfolios

Small detour: ICAPM intuition

- In a two-period setting, the maximization problem is

$$\begin{aligned} \max_{c_t, c_{t+1}} \; U(c_t, c_{t+1}) &= u(c_t) + \beta E_t [u(c_{t+1})] \\ \text{s.t.} \quad c_t &= w_t - \iota' a \\ c_{t+1} &= a' (\iota + R_{t+1}) \end{aligned}$$

  where a is the vector of wealth amounts invested into each asset, R_{t+1} is the return on the vector of risky assets, and ι is a vector of ones

- The first order condition for asset i is (assume $\beta = 1$ for simplicity):

$$E_t \left[\underbrace{\frac{u'(c_{t+1})}{u'(c_t)}}_{m_{t+1}} \left(1 + R^i_{t+1}\right) \right] = 1$$

  where m_{t+1} is the stochastic discount factor (SDF)

- From the f.o.c. you get a policy rule:

$$c_t = f(w_t, z_t)$$

  where z_t are state variables that belong to the investor's information set at time t

- Now, extend the model to a multi-period (i.e. ICAPM) setting. In this case, you have a policy rule for period $t+1$ as well:

$$c_{t+1} = f(w_{t+1}, z_{t+1})$$

- Therefore the SDF can be linearized as follows

$$\begin{aligned} m_{t+1} &= b_{0,t} + b_{1,t} R^w_{t+1} + b_{2,t} z_{t+1} \\ b_{1,t} &< 0 \quad , \quad b_{2,t} < 0 \end{aligned}$$

  where R^w_{t+1} is the return on the total wealth portfolio

$$R^w_{t+1} = a' R_{t+1} / \iota' a$$

- $b_{2,t} < 0$ comes from the assumption: $z_{t+1} \uparrow \Longrightarrow c_{t+1} \uparrow$

- So, we can re-write the f.o.c. for asset i as

$$E_t \left[\left(b_{0,t} + b_{1,t} R^w_{t+1} + b_{2,t} z_{t+1} \right) R^i_{t+1} \right] = 1$$

  or re-arranging terms and using the f.o.c for the risk-free rate:

$$\begin{aligned} & E_t \left[R^i_{t+1} \right] \qquad (1) \\ = & \left(1 + R^f_{t+1} \right) [1 - b_{1,t} Cov_t \left(R^i_{t+1}, R^w_{t+1} \right) \\ & - b_{2,t} Cov_t \left(R^i_{t+1}, z_{t+1} \right)] \end{aligned}$$

- Hence, from equation (1) we conclude that investors require higher expected return from asset i if $Cov_t\left(R^i_{t+1}, z_{t+1}\right) > 0$.

- That is, an asset whose payoff covaries positively with good news about future investment opportunities (i.e. $R^i_{t+1} \uparrow$ when $z_{t+1} \uparrow$) is the opposite of insurance and needs to pay a premium

- This component of the risk premium derives from what is called 'hedging demand'

Back to the paper: Return decomposition

- Following Campbell and Shiller (1988), one can decompose asset returns into two components (see CLM, chapter 7):
  1. Cash flow news: N_{CF}
  2. Discount rate news: N_{DR}

$$
\begin{aligned}
r_{t+1} - E_t(r_{t+1}) &= (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} \\
&\quad - (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r_{t+1+j} \\
&= N_{CF} - N_{DR} \qquad (2)
\end{aligned}
$$

- This equation derives from a log-linearization of the definition of returns, which is then iterated forward ad infinitum

- It is an identity and it holds under any circumstances (up to approximations errors)

- ρ is a linearization constant and is related to the average $\log(D/P)$, which calibrates $\rho = 0.95^{1/12}$ in monthly data

- Equation (2) says that returns and prices increase in $t+1$ relative to the expectation in period t if:

  1. There is an upward revision in the expected growth rate of dividends (Δd): cash flow news

  2. There is a downward revision in the rate at which future dividends are discounted, that is in the expected return: discount rate news

A two beta model

- One can define a beta relative to each of the two components of the market return
- The bad beta, or cash flow beta is

$$\beta_{i,CF} = Cov(r_{it}, N_{CFt})/Var(r^e_{Mt} - E_t r^e_{Mt}) \quad (3)$$

  where r^e_{Mt} is the excess return on the market
- The good beta, or discount rate beta is

$$\beta_{i,DR} = Cov(r_{it}, -N_{DRt})/Var(r^e_{Mt} - E_t r^e_{Mt}) \quad (4)$$

- The sum of the two betas gives the CAPM beta:

$$\beta^i_{total} = \beta_{i,CF} + \beta_{i,DR}$$

- Notice, these are unconditional betas
- The bad beta is called like that because (if $\beta_{CF} > 0$) a negative change in the growth rate of dividends is associated with a drop in the stock's price. This is a permanent loss in investors' wealth that is not going to revert in the future. They command a high premium

- Reason for the label "good beta": on the one hand, if $\beta_{DR} > 0$ the price of asset i decreases when prices in the market decrease ($N_{DR} \uparrow$), which is bad. But, on the other hand, future investment opportunities improve ($r_{M,t+j} \uparrow$, $j > 1$): investors will earn higher returns on future investments (good news). So, shocks related to discount rates news are only temporary and, as such, they are not so painful to investors

- In other words, N_{DR} is a state variable that predicts improvements in future investment opportunities. Negative covariance with N_{DR}, that is $\beta_{DR} > 0$, is a characteristic that attracts hedging demand and commands lower expected returns (ICAPM)

- Going back to the ICAPM slides, consider N_{DR} as z_{t+1}: a positive shock to N_{DR} means higher future returns on total wealth. So, negative covariance with z_{t+1} (or positive covariance with $-N_{DR}$), that is $\beta_{DR} > 0$, denotes an asset that acts as insurance against future changes in investment opportunities. This asset commands a lower risk premium

Computing the two return components

- The authors use a VAR to predict $r^e_{M,t+1}$

$$z_{t+1} = a + \Gamma z_t + u_{t+1}$$

where z_t is an m–vector of variables that are useful in predicting the market return. Moreover, the first element of z_t is $r^e_{M,t}$

- Once estimated, one can use the VAR coefficients to form expectations (=forecasts) of $r^e_{M,t+j}$:

$$E_t\left(r^e_{M,t+j}\right) = 1\hat{\Gamma}^j z_t$$

where 1 is an m–vector with one as first element and zero elsewhere

- Then, one can compute discount rate news

$$N_{DR,t+1} = (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r^e_{M,t+1+j}$$

using the VAR coefficients and z_t

- Finally, cash flow news are

$$N_{CF,t+1} = r_{M,t+1} - \underbrace{E_t\left(r_{M,t+1}\right)}_{1\hat{\Gamma} z_t} + N_{DR,t+1}$$

- Discount rate news are computed using innovations in the forecasting VAR for the market return

- Cash flow news are computed residually: they are the component of unexpected returns that is not due to discount rate news

The state variables z_t

- These variables include $r^e_{M,t}$ and variables that predict the excess market return

- Besides $r^e_{M,t}$, z_t include:
  - The term spread: the difference in yields between long-term (10-year) bonds and short term bonds. The yield curve tracks the business cycle
  - The smoothed price-earnings ratio for the market. Intuition from Gordon's model
  - The small-stock value spread (VS): difference in log B/M ratios for small-value and small-growth stocks

- The first two variables are known predictors of the market return (see Fama and French 1989)

- The motivation for VS is based on ICAPM. Growth stocks need to have intertemporal hedging value in order to earn lower returns than value stocks.

Hence, prices of growth firms are high when market return is expected to be low. This implies that a high VS (which obtains with high prices of growth firms and low prices of value firms) predicts a low market return

- The estimated VAR produces the expected signs for the coefficients on z_t

Estimates of the two betas

- The two beta components are computed replacing the estimated news components in equations (3) and (4)
- Their test assets are:
  - 25 B/M and size sorted portfolios
  - 10 portfolios sorted according to stock sensitivities to innovations in VS and $r_{M,t}$
  - 10 portfolios sorted according to stock sensitivities to innovations in term-spread and $r_{M,t}$
- The choice of the additional portfolios is dictated by Daniel and Titman's (1997) argument that the B/M sort produces a sort on factors in returns (possibly industry factors). Hence, one needs to test models on different samples (see also Lewellen, Nagel, and Shanken, 2009)
- It turns out that the additional portfolios do not produce large pricing errors. So, considering them in the tests favors the null of correct functioning of the two beta model (is it an ad-hoc choice of the test assets?)

The early sample: 1929-63

- Estimated betas are in Table 4:

$\hat{\beta}_{CF}$	Growth		2		3		4		Value		Diff.	
Small	0.53	[0.11]	0.46	[0.09]	0.40	[0.08]	0.42	[0.07]	0.49	[0.08]	−0.04	[0.07]
2	0.30	[0.06]	0.34	[0.06]	0.36	[0.06]	0.38	[0.06]	0.45	[0.08]	0.16	[0.04]
3	0.30	[0.06]	0.28	[0.05]	0.31	[0.06]	0.35	[0.06]	0.47	[0.08]	0.18	[0.04]
4	0.20	[0.05]	0.26	[0.05]	0.31	[0.05]	0.35	[0.07]	0.50	[0.09]	0.30	[0.05]
Large	0.20	[0.05]	0.19	[0.05]	0.28	[0.06]	0.33	[0.07]	0.40	[0.09]	0.19	[0.06]
Diff.	−0.33	[0.09]	−0.26	[0.06]	−0.12	[0.05]	−0.09	[0.04]	−0.10	[0.04]		

$\hat{\beta}_{DR}$	Growth		2		3		4		Value		Diff.	
Small	1.32	[0.18]	1.46	[0.19]	1.32	[0.15]	1.27	[0.14]	1.27	[0.15]	−0.06	[0.15]
2	1.04	[0.11]	1.15	[0.11]	1.09	[0.11]	1.25	[0.11]	1.25	[0.13]	0.21	[0.08]
3	1.13	[0.10]	1.01	[0.08]	1.08	[0.09]	1.05	[0.10]	1.27	[0.12]	0.14	[0.06]
4	0.87	[0.07]	0.97	[0.08]	0.97	[0.09]	1.06	[0.10]	1.36	[0.13]	0.49	[0.10]
Large	0.88	[0.07]	0.82	[0.07]	0.87	[0.08]	1.06	[0.09]	1.18	[0.12]	0.31	[0.10]
Diff.	−0.45	[0.15]	−0.64	[0.15]	−0.43	[0.10]	−0.21	[0.08]	−0.08	[0.10]		

- Value stocks have higher β_{CF} and β_{DF} (exception small-growth)

- This implies that in the early sample value stocks have higher total beta

- Because value stocks earn higher average returns than growth stocks, CAPM is not rejected

The modern sample: 1963-2001

- Estimated betas are in Table 5:

$\hat{\beta}_{CF}$	Growth		2		3		4		Value		Diff.	
Small	0.06	[0.07]	0.07	[0.06]	0.09	[0.05]	0.09	[0.04]	0.13	[0.04]	0.07	[0.04]
2	0.04	[0.06]	0.08	[0.05]	0.10	[0.04]	0.11	[0.04]	0.12	[0.04]	0.09	[0.03]
3	0.03	[0.05]	0.09	[0.04]	0.11	[0.04]	0.12	[0.03]	0.13	[0.04]	0.09	[0.04]
4	0.03	[0.05]	0.10	[0.04]	0.11	[0.03]	0.11	[0.03]	0.13	[0.04]	0.10	[0.04]
Large	0.03	[0.04]	0.08	[0.03]	0.09	[0.03]	0.11	[0.03]	0.11	[0.03]	0.09	[0.03]
Diff.	−0.03	[0.05]	0.02	[0.05]	−0.01	[0.04]	0.02	[0.04]	−0.01	[0.04]		

$\hat{\beta}_{DR}$	Growth		2		3		4		Value		Diff.	
Small	1.66	[0.13]	1.37	[0.11]	1.18	[0.10]	1.12	[0.09]	1.12	[0.10]	−0.54	[0.08]
2	1.54	[0.11]	1.22	[0.09]	1.07	[0.08]	0.96	[0.08]	1.03	[0.09]	−0.52	[0.08]
3	1.41	[0.10]	1.11	[0.08]	0.95	[0.08]	0.82	[0.07]	0.94	[0.09]	−0.47	[0.09]
4	1.27	[0.09]	1.05	[0.08]	0.89	[0.07]	0.79	[0.07]	0.87	[0.08]	−0.41	[0.09]
Large	1.00	[0.07]	0.87	[0.07]	0.74	[0.06]	0.63	[0.07]	0.68	[0.07]	−0.33	[0.08]
Diff.	−0.66	[0.12]	−0.50	[0.11]	−0.44	[0.10]	−0.49	[0.09]	−0.44	[0.08]		

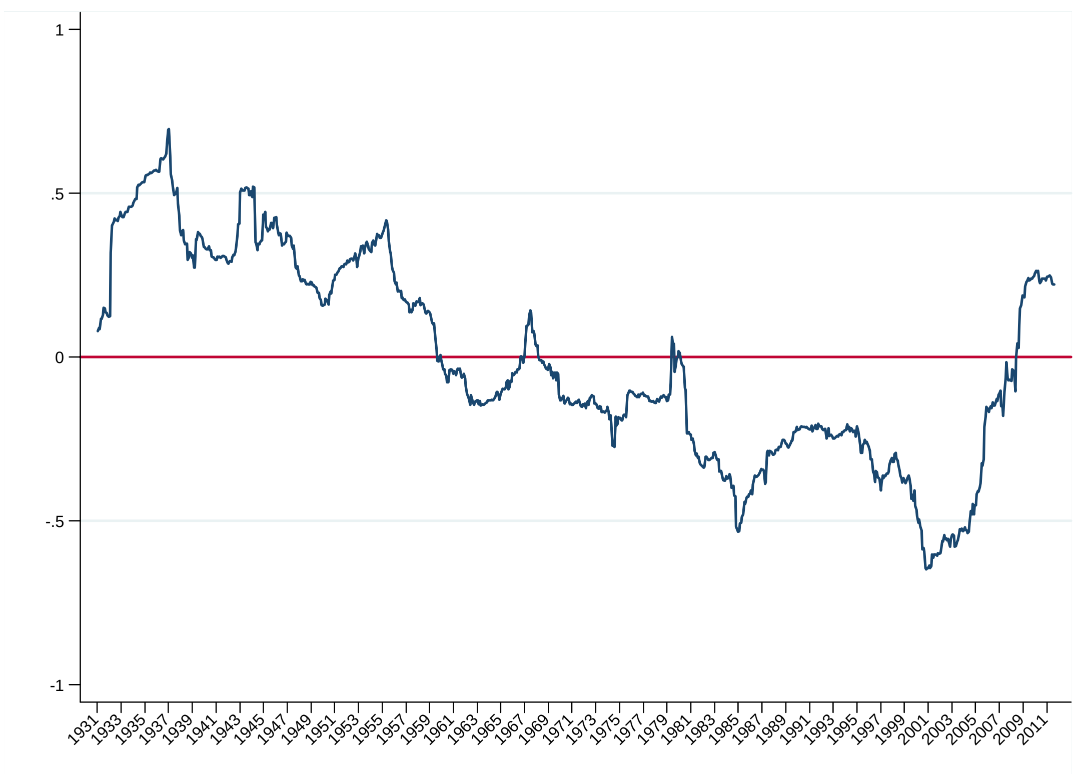
- Value stocks still have somewhat higher β_{CF}, but lower β_{DR}

- Given that β_{DR} is a higher fraction of total beta, in modern sample: $\beta_{total}^{Gr} > \beta_{total}^{Value}$

- Value stocks earn higher returns. Hence, you see how the ranking in β_{CF} is what matters for explaining the cross-section of returns

The decrease in beta of value stocks

- The comparison of Tables 4 and 5 highlights a major structural development
- In terms of total beta, value stocks have become less risky than growth stocks
- This finding confirms the evidence pointed out by Franzoni (2002)
- The following graph plots five-year rolling-window estimates of the beta of the HML portfolio (Value – Growth)



- Campbell and Vuolteenaho argue that:
  - High leverage of distressed value firms explains their high β_{CF} during the Great Depression and War
  - Growth opportunities and equity dependence explain sensitivity to discount rate shocks (β_{DR}) for growth companies
  - In early sample, strict listing requirements implied that growth companies were not so small and equity dependent (lower β_{DR})
  - The IPO waves in 1960's and 1990's, the Nasdaq in the 1970's brought many smaller and equity dependent firms into the market. These were growth firms with high β_{DR}

An asset pricing model

- The authors base their tests on an AP model derived from the ICAPM (see also Campbell 1992)

- In a context with log-normal returns and CRRA utility, the premium on each asset is

$$
\begin{aligned}
& E_t(r_{i,t+1}) - r_{f,t} + \frac{\sigma^2_{i,t}}{2} \\
= \; & \gamma \sigma^2_{P,t} \beta_{i,CF_P} + \sigma^2_{P,t} \beta_{i,DR_P}
\end{aligned}
$$

  where P is the investor's optimal portfolio and γ is the risk aversion coefficient

- They begin by assuming that P is the market index

- Observe that if $\gamma = 1$ CAPM obtains

- Otherwise, if $\gamma > 1$, the risk premium on the bad beta is higher

- The coefficient on β_{DR} is constrained by the theory to be equal to σ^2_M. That is, 0.05 in early sample and 0.025 modern sample

- They also allow for the possibility that P invests a fraction w in market index and a fraction $1 - w$ in risk free rate. In this case, there are two free parameters γ and w. Still, the ratio of the risk premia of the two betas is equal to γ

Testing the AP model

- They test the unconditional version of the model
- They do cross-sectional tests by running the regression:

  $$\bar{R}_i^e = g_0 + g_1 \hat{\beta}_{i,CF} + g_2 \hat{\beta}_{i,DR} + e_i \qquad i = 1...N$$

  where the test assets are the 45 portfolios
- Three specifications:
  1. CAPM: restrict $g_1 = g_2$
  2. ICAPM: restrict $g_2 = \sigma_M^2$
  3. Two-factor model: unrestricted
- Also, they test with g_0 restricted to be zero and unrestricted (Black version)
- They evaluate performance by looking at R^2
- They also look at a quadratic form in the pricing errors

  $$\hat{e}' \hat{\Omega}^{-1} \hat{e}$$

  where $\hat{e}$ is the vector of N pricing errors and $\hat{\Omega}$ is a diagonal matrix of estimated return volatilities

- The metric gives weighted sum of squared pricing errors: more volatile portfolios receive less weight

- What are the volatilities of the twenty additional portfolios?

- Standard errors are bootstrapped to take into account all issues: correlation, heteroskedasticity, and estimated betas

Results: early sample

- Table 6:

Parameter	Factor model		Two-beta ICAPM		CAPM	
R_{zb} less R_{rf} (g_0)	0.0042	0	0.0023	0	0.0023	0
% per annum	4.98%	0%	2.76%	0%	2.74%	0%
Std. err. A	[0.0032]	N/A	[0.0024]	N/A	[0.0028]	N/A
Std. err. B	(0.0029)	N/A	(0.0030)	N/A	(0.0028)	N/A
$\hat{\beta}_{CF}$ premium (g_1)	0.0173	0.0069	0.0083	0.0148	0.0051	0.0067
% per annum	20.76%	8.22%	9.91%	17.80%	6.11%	8.00%
Std. err. A	[0.0231]	[0.221]	[0.0167]	[0.0175]	[0.0046]	[0.0034]
Std. err. B	(0.0266)	(0.0248)	(0.0221)	(0.0442)	(0.0046)	(0.0034)
$\hat{\beta}_{DR}$ premium (g_2)	−0.0003	0.0066	0.0041	0.0041	0.0051	0.0067
% per annum	−0.41%	7.93%	4.95%	4.95%	6.11%	8.00%
Std. err. A	[0.0092]	[0.0067]	[0.0006]	[0.0006]	[0.0046]	[0.0034]
Std. err. B	(0.0088)	(0.0071)	(0.0006)	(0.0006)	(0.0046)	(0.0034)
$\hat{R}^2$	48.08%	40.26%	45.85%	37.98%	44.52%	40.26%
Pricing error	0.0117	0.0126	0.0119	0.0133	0.0127	0.0126
5% critic. val. A	[0.019]	[0.024]	[0.024]	[0.033]	[0.021]	[0.027]
5% critic. val. B	(0.019)	(0.024)	(0.031)	(0.099)	(0.021)	(0.027)

- The three models perform equally well, as expected

- Focus on pricing errors: the statistic is always below critical value. The models are not rejected

Results: modern sample

- Table 7:

Parameter	Factor model		Two-beta ICAPM		CAPM	
R_{zb} less R_{rf} (g_0)	0.0009	0	−0.0009	0	0.0069	0
% per annum	1.05%	0%	−1.04%	0%	8.24%	0%
Std. err. A	[0.0029]	N/A	[0.0031]	N/A	[0.0026]	N/A
Std. err. B	(0.0033)	N/A	(0.0031)	N/A	(0.0026)	N/A
$\hat{\beta}_{CF}$ premium (g_1)	0.0529	0.0572	0.0575	0.0483	−0.0007	0.0051
% per annum	63.47%	68.59%	69.04%	57.92%	−0.83%	6.10%
Std. err. A	[0.0178]	[0.0163]	[0.0182]	[0.0272]	[0.0034]	[0.0023]
Std. err. B	(0.0325)	(0.0444)	(0.0262)	(0.0423)	(0.0034)	(0.0023)
$\hat{\beta}_{DR}$ premium (g_2)	0.0007	0.0012	0.0020	0.0020	−0.0007	0.0051
% per annum	0.88%	1.44%	2.43%	2.43%	−0.83%	6.10%
Std. err. A	[0.0033]	[0.0031]	[0.0002]	[0.0002]	[0.0034]	[0.0023]
Std. err. B	(0.0085)	(0.0099)	(0.0002)	(0.0002)	(0.0034)	(0.0023)
$\hat{R}^2$	52.10%	51.59%	49.26%	47.41%	3.10%	−61.57%
Pricing error	0.0271	0.0269	0.0272	0.0275	0.0592	0.0875
5% critic. val. A	[0.028]	[0.042]	[0.051]	[0.314]	[0.032]	[0.046]
5% critic. val. B	(0.030)	(0.071)	(0.051)	(0.488)	(0.032)	(0.046)

- CAPM has worst performance: low R^2 and pricing errors different from zero

- ICAPM is second best performer: $R^2 \approx 48\%$ and pricing errors not statistically different from zero

- Two-factor model (unrestricted) is best performer: $R^2 \approx 52\%$ and pricing errors not statistically different from zero

How about the restrictions?

- Notice that the theory predicts that g_1/g_2 is equal to γ, the risk aversion coefficient
- In the modern sample, $\hat{\gamma}$ is extremely high: $[23.8; 72.1]$. These values seem implausible
- If you compare Tables 6 and 7 you observe that $\hat{\gamma}$ has increased substantially from the first to the second subsample
- This increase in risk aversion goes against the common sense and the empirical evidence that points towards a decrease in the equity premium (Fama and French, 2002, Polk, Thompson, and Vuolteenaho, 2006)
- Hence, it seems that the the cross-sectional coefficients are treated as free parameters that are used to improve the fit of the model
- This practice runs into Lewellen and Nagel's critique

New definitions of beta

- The paper points out that for an AP model to work, value firms need to be riskier than growth stocks along some dimension

- As a consequence, different papers try to provide a definition of beta that yields large estimates of risk for values stock and low estimates for growth stocks

- This definition of beta could be a cash flow beta

- Cohen, Polk, and Vuolteenaho (2009, "The price is (almost) right", Journal of Finance) successfully define the cash flow beta by looking at ROE

- In a somewhat different spirit, Bansal, Dittmar, and Lundblad (2005) suggest that value stocks have higher consumptions betas. That is, their cash flows covary more with consumption growth

- In any case, the cross-sectional tests have to take into account the theoretical restrictions

Further Caveats

- Chen and Zhao (2009, RFS) criticize the return decomposition approach
- CF news are computed residually relative to DR news. Hence, CF news capture all the noise in the VAR
- If a variable is omitted from z_t but it belongs to investors' information set, it ends up in CF news. So, CF news are really sensitive to the VAR specification
- For a portfolio of treasury bonds, where CF news are virtually zero, the authors show that Campbell and Vuolt. decomposition produces a large CF news component of returns
- Chen and Zhao also show that value companies do not have larger CF betas with alternative VAR specifications
- Indeed, if one omits VS from z_t, value stocks don't have higher β_{CF}

- As a solution, they propose modelling directly the CF and DR components of returns

- Campbell's reply to Chen and Zhao experiment with bonds is that they mistakenly compute CF news at portfolio level, which does not account for portfolio rebalancing

- Still, Chen and Zhao's criticism is very powerful

2. Growth or Glamour?

Motivation

- Campbell and Vuolteenaho (2004) find that the covariance with market CF news matters for pricing
- At the same time, the covariance with market DR news is the main source of stock return volatility
- The question Campbell, Polk, and Vuolteenaho (CPV) ask is what determines a firm's exposure to either source of news
- If the sensitivity arises from firm cash flows, the explanation of volatility has to be founded on firm fundamentals. One view is that value and growth firms are exposed to different cash flow risks
  - Value firms are in financial distress and more likely to go bankrupt
  - Growth firms have investment opportunities that become profitable if they can find sufficiently cheap financing

- According to this view, value and growth firms returns move together because their fundamentals move together. Firm discount rates could even be constant

- An alternative possibility is that volatility arises because of the discount rates applied to the firm cash flows:

  – One story suggests that growth firms are long duration assets whose values are more sensitive to changes in the market discount rate

  – Another view is that investors price assets according to sentiment (Barberis and Shleifer, 2003, Barberis, Shleifer and Wurgler, 2005). Sentiment has nothing to do with fundamentals and shows up has discount rate news

- The paper tries to disentangle the two explanations

Four-way decomposition

- First, as in C&V, they decompose market returns into CF and DR news $N_{M,CF}$ and $N_{M,DR}$

- Next, using the same methodology, they decompose firm level returns into $N_{i,CF}$ and $N_{i,DR}$

- They aggregate the firm news into five portfolios based on B/M

- Finally, the compute four components of the total beta:

$$
\begin{aligned}
\beta_{CFi,CFM} &= Cov(N_{i,CFt}, N_{M,CFt})/Var(r^e_{Mt}) \\
\beta_{DRi,CFM} &= Cov(-N_{i,DRt}, N_{M,CFt})/Var(r^e_{Mt}) \\
\beta_{CFi,DRM} &= Cov(N_{i,CFt}, -N_{M,DRt})/Var(r^e_{Mt}) \\
\beta_{DRi,DRM} &= Cov(-N_{i,DRt}, -N_{M,DRt})/Var(r^e_{Mt})
\end{aligned}
$$

- The four components add up to the total market beta

$$\beta_i = \beta_{CFi,CFM} + \beta_{DRi,CFM} + \beta_{CFi,DRM} + \beta_{DRi,DRM}$$

- The first two components give the C&V's bad beta

$$\beta_{i,CF} = \beta_{CFi,CFM} + \beta_{DRi,CFM}$$

- The other two coponents add to the good beta

$$\beta_{i,DR} = \beta_{CFi,DRM} + \beta_{DRi,DRM}$$

Firm level VAR

- The z_t state variables in the firm-level VAR are different from the variables in the aggregate VAR:
  - Firm level log return r_i
  - Firm log B/M
  - Firm long term profitability ROE over five years
- They assume that the VAR parameters are the same for all firms and estimate the VAR on pooled data at annual frequency
- Variance decomposition: compute the share of the return variance due to N_{iCF} and N_{iDR}
- Unlike the aggregate VAR, most (80%) of total return variance at firm level comes from N_{iCF}
- The result is likely different for the aggregate VAR (in which 2/3 of variance is due to aggregate DR news) because idiosyncratic CF news tend to wash out at market level

Main Results

- Using the four components, they compute the four betas. Results in Table 4:

	Growth	2	3	4	Value	G-V
Good beta components						
$\beta_{DRi,DRM}$: Growth and value $-N_{DR}$ on the market's $-N_{DR}$						
1929-2000	**.74**	**.77**	**.77**	**.80**	**.80**	**-.06**
	(60)	(78)	(85)	(74)	(98)	(-5.7)
1929-1962	**.66**	**.69**	**.69**	**.72**	**.73**	**-.07**
	(45)	(69)	(78)	(53)	(88)	(-5.3)
1963-2000	**.94**	**.95**	**.95**	**.97**	**.97**	**-.03**
	(68)	(103)	(114)	(93)	(91)	(-1.8)
$\beta_{CFi,DRM}$: Growth and value N_{CF} on the market's $-N_{DR}$						
1929-2000	**.03**	**-.00**	**-.04**	**-.09**	**.01**	**.02**
	(.86)	(-.06)	(-.99)	(-1.8)	(.18)	(.34)
1929-1962	**.00**	**.03**	**.01**	**-.03**	**.16**	**-.16**
	(-.02)	(.63)	(.29)	(-.42)	(2.3)	(-2.7)
1963-2000	**.13**	**-.05**	**-.15**	**-.24**	**-.35**	**.48**
	(1.92)	(-.89)	(-2.4)	(-2.6)	(-3.6)	(4.0)
Bad beta components						
$\beta_{DRi,CFM}$: Growth and value $-N_{DR}$ on the market's N_{CF}						
1929-2000	**.04**	**.04**	**.04**	**.05**	**.05**	**-.01**
	(1.2)	(1.2)	(1.3)	(1.4)	(1.5)	(-3.1)
1929-1962	**.08**	**.08**	**.08**	**.08**	**.09**	**-.02**
	(1.8)	(1.7)	(1.9)	(1.8)	(2.0)	(-2.7)
1963-2000	**-.04**	**-.04**	**-.04**	**-.04**	**-.04**	**-.00**
	(-.80)	(-.78)	(-.72)	(-.71)	(-.69)	(-.87)
$\beta_{CFi,CFM}$: Growth and value N_{CF} on the market's N_{CF}						
1929-2000	**.06**	**.08**	**.10**	**.12**	**.14**	**-.09**
	(4.8)	(11)	(11)	(10)	(11)	(-4.1)
1929-1962	**.06**	**.07**	**.08**	**.09**	**.13**	**-.07**
	(4.9)	(7.0)	(6.8)	(6.3)	(8.7)	(-3.3)
1963-2000	**.03**	**.09**	**.12**	**.17**	**.16**	**-.13**
	(1.2)	(8.8)	(10)	(9.3)	(5.8)	(-3.0)

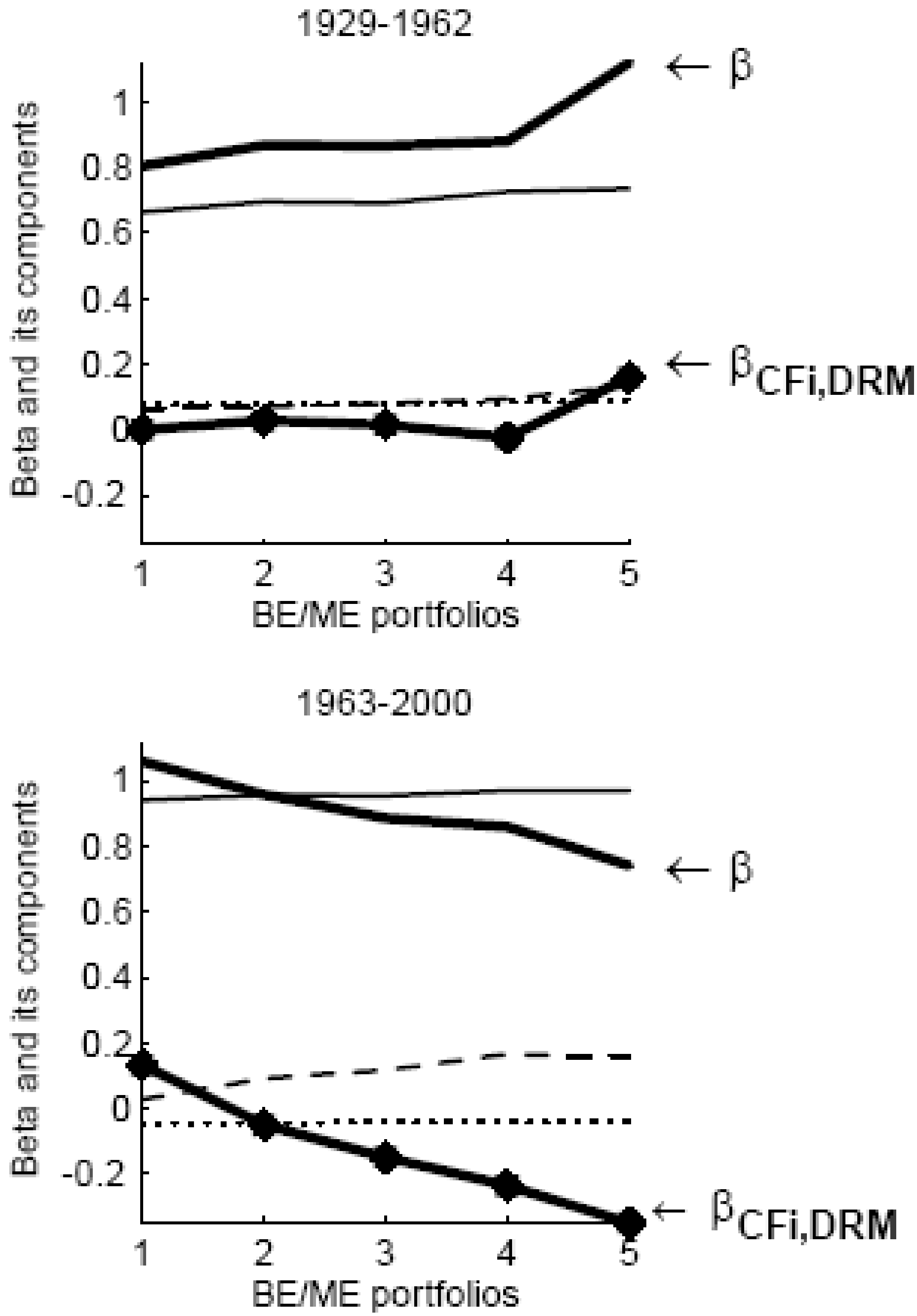
- Remember from C&V: value firms have higher bad beta and growth firms have higher good beta (in modern sample only)

- The results suggest that this heterogeneity is mainly due to firm fundamentals

- That is: the spread in $\beta_{i,CF}$ is due to a spread in $\beta_{CFi,CFM}$. The spread in $\beta_{i,DR}$ is mainly due to a spread in $\beta_{CFi,DRM}$

- Across the two subsamples: the increase in the good beta of growth stocks is due to the increase in the cash flow component of the good beta

- The results are summarized in Figure 2


1929-1962
Beta and its components
1
0.8
0.6
0.4
0.2
0
-0.2
1
2
3
4
5
BE/ME portfolios
← β
← β CFi,DRM
1963-2000
Beta and its components
1
0.8
0.6
0.4
0.2
0
-0.2
1
2
3
4
5
BE/ME portfolios
← β
← β CFi,DRM

Direct proxies of CF news

- Given Chen and Zhao (2007) critique, the authors decide to provide direct measures of CF news

- For each portfolio, they construct the discounted sum of future ROE

$$N_{i,CF,t} = \sum_{k=1}^{K} \rho^{k-1} roe_{i,t,t+k}$$

  where $roe_{i,t,t+k}$ is profitability for portfolio i, sorted in period t, and measured in period $t+k$.

- Portfolio profitability is computed by value-weighting earnings and book-value of the stocks in the portfolio and taking the ratio

- Portfolios are re-sorted annually, but stocks are followed for K years (K from 3 to 5) to compute the ROE

- Results in Table 5 confirm the previous evidence: the spread in good and bad betas is due to the firm CF news components

Other results and perspective

- Given that the fundamental view seems to prevail, the question is which characteristic of cash flows explains the betas

- This knowledge would help to form simple estimates of the cost of capital

- They regress the beta components on firm level explanatory variables: ROA, volatility of ROA, beta of ROA, Leverage, Capital Expenditures

- They find evidence suggesting that these variables drive the cash flow components of the good and bad beta

- Hence, once again, fundamental variables rather than sentiment seem to matter for firms riskiness

- Still, one would like to know which economic story is behind the surge of the good beta as a component of the total beta, and the related fact that the good beta is larger for growth stocks (in the modern sample)

- Every story should be founded on the effect of shocks to discount rates to firm cash flows

- Two main forces could be at work:

  1. Growth firms are financially constrained (equity dependent). Hence, aggregate DR news is relevant for their cost of capital and their possibility of undertaking new projects (which ultimately translates into firm cash flows)

  2. Growth firms have more growth options. Aggregate DR news determines whether the options are going to be exercised (in turn, exercising the option affects the profile of future cash flows)

3. Recent Trends

Other stories for the Value Premium

- Building on previous evidence, the literature has recently proposed other rational based stories for the value premium

- One direction is *production based* models. This literature tries to found the different riskiness of B/M sorted portfolios on the different underlying production technologies (e.g. Zhang, 2005, "The value premium", JF)

- Also related: Berk, Green, and Naik, 1999, JF; Gomes, Kogan, and Zhang, 2003, JPE; Lettau and Wachter, 2007, JF; Obreja, 2007, "Financial Leverage and the cross-section of stock returns", WP

- Another direction is related to the literature on *long-run risk* and predictability (see classes on predictability). This literature revives the Consumption CAPM by arguing that investors care about the growth rate of long-run consumption (e.g. Bansal, Dittmar, and Lundblad, 2005, JF)

- Also related: Parker and Julliard, 2005, JPE; Hansen, Heaton, and Li, 2008, JPE

Zhang's "The Value Premium"

- This is a neoclassical general equilibrium model that endogenously generates the value premium

- It is a theoretical paper with a calibration exercise. It generates testable empirical predictions

- There are two driving factors:
  1. Costly reversibility (that is, asymmetric adjustment costs for the stock of capital)
  2. A time-varying price of risk (which is higher in recessions)

- The value of a firm derives from:
  - Assets in place (the existing capital stock)
  - Growth options (the NPV of potential investment projects)

- Assets in place represent the main part of the price of value firms

- Growth options are the predominant component of growth firms' valuations

- Growth options are risky because they are more valuable in good times. So, one would tend to conclude that growth stocks are riskier and should earn higher returns

- Instead, in the model assets in place are more risky

- So, value stocks earn higher risk premium

Model's intuition

- Exogenous productivity shocks determine booms and recessions
- In recessions, firms would like to reduce the stock of unproductive capital
- This is especially the case for value firms that have more assets in place
- So, value firms have to cut dividends more than growth firms to face adjustment costs in bad times
- Hence, value firms returns covary strongly with aggregate shocks in recessions $\Longrightarrow$ higher value beta than growth beta in recessions (as in Lettau and Ludvigson, 2001)
- In booms, value firms do not need to expand capital, because it is already there. Instead, growth firms would like to expand. They are now the ones that bear higher adjustment costs

- However, adjustment costs are smaller in good times (assumption of asymmetry). So, growth firms dividends do not strongly covary with productivity shocks:
  - In booms, following a shock, a firm can modify its capital stock without incurring high adjustment costs: dividends are smooth
- As a result, the spread in betas between value and growth firms is not large in good times
- Consistent with empirical evidence, the model produces:
  - Large spread in betas conditional on recessions
  - Small or even negative unconditional spread in betas

The role of the price of risk

- Because the unconditional spread in betas is almost zero, a constant price for risk would generate an almost zero spread in returns
- Instead, he assumes that the price of risk is higher in recessions (countercyclical). For example, investors are more risk averse when their consumption is far below the habit level (see Campbell and Cochrane, 1999)
- It follows that the large spread in betas in bad times is interacted with a large spread in risk
- This produces an unconditional large value premium
- The effect is reinforced by the fact that a large cost of capital in bad times induces value firms to cut investments even more in recessions

The testable predictions

- The model produces yet untested predictions

  1. Value firms divest more than growth firms in bad times, and vice versa in good times
  2. The expected value premium and value spread are both countercyclical
  3. The degree of asymmetry in adjustment costs correlates positively with the value premium across industries

Risks for the long-run

- This strand of literature started with a focus on the equity premium and predictability (Bansal and Yaron, 2004, JF)

- Still, it can generate cross-sectional predictions

- The idea is that there are small persistent shocks to dividend growth (long run risks). These are like CF news

- At high frequencies, these shocks are not identifiable and consumption growth appears to be i.i.d., that is, in finite sample dividends appear to be unpredictable and price volatility is erroneously imputed to DR news

- But these shocks have large implications for the present value of dividends and generate large price volatility

- The higher the correlation between stock level CF and consumption growth news (the cash flow beta) the higher the risk premium a stock should pay

- Unlike the first papers we saw in these class notes, the CF beta is defined as covariance of dividend growth with consumption growth

- Bansal, Dittmar, and Lundblad (2005, JF) run this regression on quarterly data:

$$g_{i,t} = \gamma_i \left(\frac{1}{K} \sum_{k=1}^{K} g_{c,t-k} \right) + u_{i,t}$$

  where $g_{i,t}$ is quarterly portfolio dividend growth, $g_{c,t}$ is quarterly consumption growth, and γ_i captures the CF beta

- To capture the fact that consumption reacts to news on future dividend growth, in their preferred specification K is 8 quarters

- They test the model on 10 B/M, 10 Size, and 10 Momentum portfolios

- High B/M, Small, and Winner portfolios have high CF betas

- The model explains the cross-section of returns with $R^2 = 62\%$

- Problems with this approach:
  1. They define dividend growth at portfolio level, not at firm level
     - Reinvestment of capital gains in the portfolio can have perverse effects on dividend growth
     - For example, average growth rate of dividends is 4% per year for value and 0.76% for growth portfolio: counterfactual. We know that growth firms grow more than value firms
  2. As other papers, they compare models based on R^2. This can be highly misleading (see Lewellen, Nagel, and Shanken, 2009, JFE)

4. A skeptical look at AP tests

Motivation

- Lewellen, Nagel, and Shanken (2009) observe that in recent times many different AP models seem to successfully explain the size and value premia
- The models are often unrelated to one another: how can they all be right?
- They suggest that the problem lies in the choice of testing the models on the same set of assets, the 25 size and B/M sorted portfolios, and in the fact that the tests relie too much on cross-sectional R^2
- The paper is extremely interesting because it brings statistical discipline back into AP tests

The fallacy of the R^2

- The core of their arguement is the following:
  - The 25 portfolios have a well-defined factor structure. That is, the 3 FF factors explain over 90% of their variance
  - As long as the new factors that a model proposes are even slightly correlated with the 3 FF factors, and they are not correlated with the residuals, then the new factors will also produce high R^2 in cross-sectional regressions of average returns onto factors loadings
- Let me try to reformulate their statement:
  - The 25 portfolios are (almost completely) spanned by a set of vectors (the FF factors)
  - If you take another set of vectors that are not linearly independent of the 3 FF factors, then the new set of vectors will also span the 25 portfolios

A simple proof

- Assume the true model is:

$$R = BF + e$$
$$E(e) = 0$$

where R is a set of N excess returns and F is a set of K factors that explain the expected returns. That is:

$$\mu = B\mu_F$$

with $B = Cov(R, F)\,Var^{-1}(F)$

- So, the true model has a cross-sectional $R^2 = 1$

- Assume the proposed model P contains J factors. You want to test whether the betas on the J factors explain the cross-section of average returns

- The betas for the new model are

$$C = Cov(R, P)\,Var^{-1}(P)$$

- We say that the model explains the cross-sections of expected returns if

$$\mu = C\gamma$$

for some vector γ that should capture the price of risk on the J factors

Proposition 1 *Suppose $J = K$. Suppose P is correlated with R only through F, that is, $Cov(P, e) = 0$. Assume that $Cov(F, P)$ is non singular. Then expected returns are exactly linear in stocks' loadings on P - even if P has very small correlation with F and explains very little of the time-series variation in returns*

Proof. The assumption of $Cov(P, e) = 0$, implies that $Cov(R, P) = BCov(F, P)$. So, the loadings on P are $C = BCov(F, P)Var^{-1}(P)$. Let $Q = Cov(F, P)Var^{-1}(P)$ (invertible). We know that $\mu = B\mu_F$, which can be written as $\mu = CQ^{-1}\mu_F$. Let $\gamma = Q^{-1}\mu_F$. Then, we have that $\mu = C\gamma$. ■

- So, you start with a model with $R^2 = 1$ and end up with a new model with $R^2 = 1$. But the coefficients are not those implied by the theory. The coefficients γ are not necessarily equal to the risk premiums on the factors P

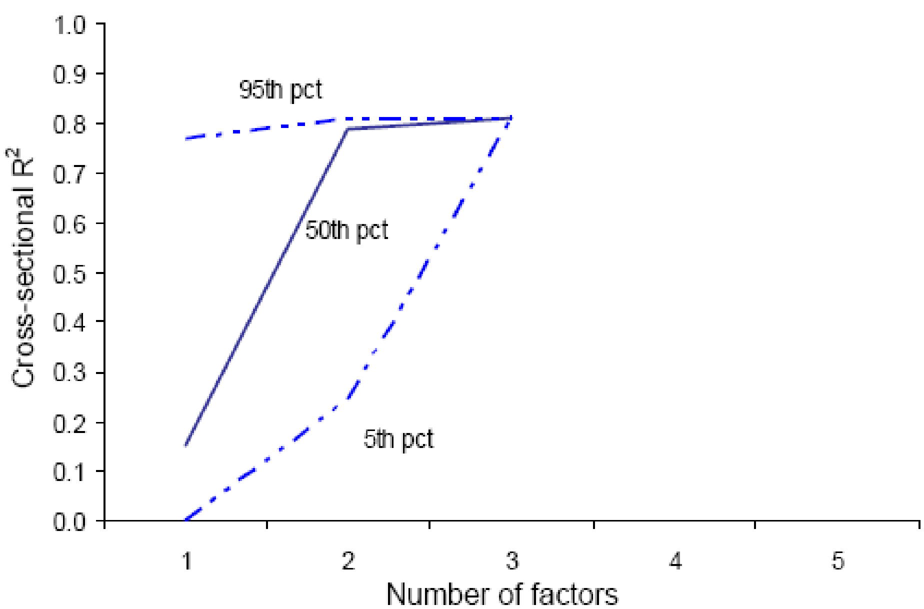
- The proof is interesting because it points out that γ is not necessarily the price for risk. Which suggests a way to test that the spanning that we get is not mechanical: test the theoretical restrictions on γ

- Notice that a crucial assumption is that the new factors P are not correlated with the residuals e. This is very likely to be the case for the FF 3 factor model where the factors explain 90% of the variation and the residuals are negligible

More general result

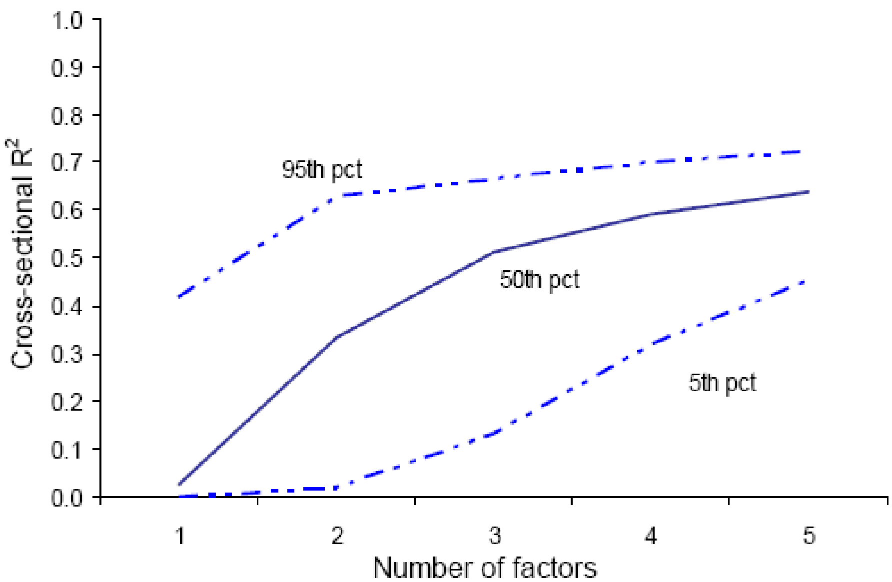
- The paper deals with the more general case where $J < K$. In this case, when the true R^2=1, the expected R^2 for the model P is J/K

- In simulations, they show that when they generate factors to be correlated with R_{mkt}, HML, and SMB, but uncorrelated with the residuals, the R^2 in cross-sectional regressions of average returns of the 25 portfolios on factor loadings increases with the number of simulated factors



- Even more compelling, they randomly generate factors that have zero mean by construction ($\mu_p =$

0). The model should have no explanatory power (because the theoretical restriction is $\gamma = \mu_P = 0$). Instead, the R^2 is positive and increasing in J



Other results

- It is worth stressing that these results concern population parameters. That is, parameters of the data generating model

- In other words, these statements abstract from sampling error because they assume that the estimated loadings are equal to the true loadings

- Once estimation error is taken into account, the problem is exacerbated. They show that the sample R^2 can end up being very different from the population R^2

- In other words, even if the true R^2 is close to zero, the sample R^2 can be very high due to sampling error. They show this result by means of simulations

- All these results similary apply to statistics such as the HJ-distance that are based on pricing errors. Indeed, pricing errors of P are zero as the $R^2 = 1$

- Similarly, using a SDF approach would not help because this representation is equivalent to the expected return-beta representation, as we know

Remedies

- They make suggestions to get around these issues:
  1. Include different sets of portfolios in the tests (e.g. industry portfolios)
  2. Impose the theoretical restrictions before estimation, or test the restrictions ex-post
  3. They suggest using GLS R^2 as opposed to OLS. We know that GLS rotates the original assets into a new set of portfolios
  4. They propose a method to provide confidence intervals for the true population R^2. This would point out that the sample R^2 can be very distant from the true one

Conclusions

- They apply these suggestions to the tests of different AP models that were successful in pricing the 25 portfolios

- They show that none of these models performs better than the CAPM in terms of the GLS R^2. Also, these models produce very large confidence intervals for the true R^2

- These conclusions are disruptive for the state of the art of the research in ‘rational’ asset pricing

- On a related note, see Daniel and Titman, 2005, “Testing factor model explanations of market anomalies”

- On the positive side, this paper develops more stringent criteria for testing AP models

- So, going back to Cochrane’s point of view, is it really true, as he suggests, that the statistical aspects of the AP tests have secondary importance?

- You can give your own answer