Static Equilibrium Asset Pricing

Markus Leippold

Spring 2019

University of Zurich
Department of Banking and Finance

Table of Contents

In this lecture, we discuss two classic theories which are widely used in asset pricing as they impose some **testable restrictions on asset returns**:

1. The **Capital Asset Pricing Model (CAPM)** uses an equilibrium argument to identify the mean-variance efficient portfolio of risky assets.
2. The **Arbitrage Pricing Theory (APT)** assumes a factor structure in the variance-covariance matrix of risky returns, effectively reducing the dimension of the portfolio choice problem from N to K, where K is the number of factors.

The CAPM

Assumptions

Universität Zürich[UZH]

Assumptions underlying the CAPM:

A1 All investors are **price takers**.

A2 The investors are only **mean-variance optimizers of single-period returns**.

A3 The investors have **common beliefs** about the means, variances and covariances of returns.

A4 There are **no nontraded assets**, **taxes**, or **transactions costs**.

A5 For the basic version of the model due to Sharpe et al. (1964) and Lintner (1965), that **investors can borrow or lend at a given riskfree interest rate**.

Given A1-A5, Tobin's separation theorem holds, all investors hold the **tangency portfolio** and risky assets in the same proportions.

- In equilibrium (**demand=supply**), the portfolio weights of the mean-variance efficient risky portfolio must be those of the market portfolio/value-weighted index.
- In principle, an investor can "free-ride" on the analyses of other investors reflected in asset prices and use the market portfolio as the optimal mutual fund of risky assets.
- The optimal capital allocation line (CAL) is just the capital market line (CML) connecting the riskfree asset to the market portfolio.
- Relaxing A5, we would still get a two-fund separation theorem and the market portfolio is a combination of these two funds.

Outline

The equilibrium result that the **market portfolio is mean-variance efficient** has some implications for asset returns:

- Consider an increase in the weight of asset i in portfolio p, w_i, financed by a decrease in the weight on the riskless asset. Then,

$$\frac{d\mu_p}{dw_i} = \mu_i - R_f \ ,$$

$$\frac{d\mathrm{Var}[R_p]}{dw_i} = 2\mathrm{Cov}[R_i,R_p] \ .$$

- The ratio of the effect on mean to the effect on variance is

$$\frac{d\mu_p/dw_i}{d\mathrm{Var}[R_p]/dw_i} = \frac{\mu_i - R_f}{2\mathrm{Cov}[R_i,R_p]} \ .$$

- If portfolio p is mean-variance efficient, this ratio **must be the same** for all individual assets $i,j \in \{1,\ldots,N\}$:

$$\frac{\mu_i - R_f}{2\text{Cov}[R_i,R_p]} = \frac{\mu_j - R_f}{2\text{Cov}[R_j,R_p]} = \frac{\mu_p - R_f}{2\text{Var}[R_p]} \; .$$

- Therefore, we can write the following as a regression for any mean-variance efficient portfolio p:

$$\mu_i - R_f = \frac{\text{Cov}[R_i,R_p]}{\text{Var}[R_p]}(\mu_p - R_f) = \beta_{ip}(\mu_p - R_f) \; .$$

- The CAPM imposes the restriction that the market portfolio m is mean-variance efficient

$$\mu_i - R_f = \frac{\text{Cov}[R_i,R_m]}{\text{Var}[R_m]}(\mu_m - R_f) = \beta_i(\mu_m - R_f) \; .$$

- Thus, if we consider the regression of excess returns on the market excess return

$$R_{it} - R_f = \alpha_i + \beta_i(R_{mt} - R_f) + \epsilon_{it} \ ,$$

  where the intercept (**Jensen's alpha**) $\alpha_i = \mu_i - R_f - \beta_i(\mu_m - R_f)$ should be zero for all assets.

- Hence, the Jensen's alpha measures the misspricing, i.e., the deviation from the **security market line (SML)**: $\mu_i = R_f + \beta_i(\mu_m - R_f)$.

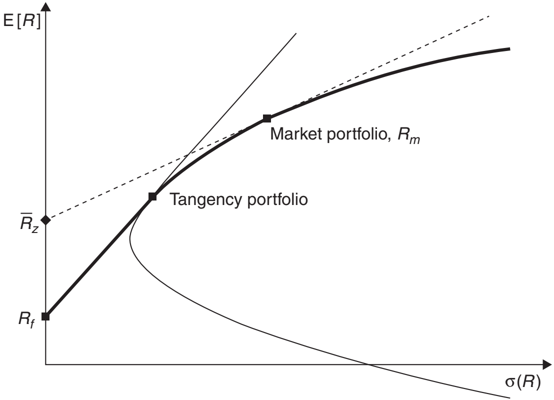
Outline

What if we relax the assumption that investors can borrow at the riskfree interest rate?

- Conservative investors combine the tangency portfolio with cash as in the basic Sharpe-Lintner model.
- Aggressive investors choose different risky portfolios with higher expected returns.
- All these portfolios will still be mean-variance efficient, and so Tobin's mutual fund theorem (with two risky mutual funds) still describes the set of risky portfolios chosen by all investors.
- When μ_z is the mean return of the zero-beta portfolio (uncorrelated with the market) we have

$$\mu_i - \mu_z = \beta_i(\mu_m - \mu_z) \ .$$

- Since all investors hold either the tangency portfolio or a riskier portfolio with a higher expected return, the market portfolio lies on the portion of the mean-variance frontier above the tangency portfolio.



What is wrong with this picture?

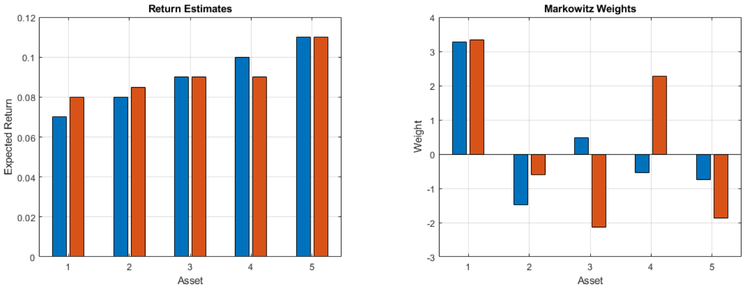
Outline

- The Black and Litterman (1992) model takes the Markowitz Model one step further.
- It incorporates an investor's own views in determining asset allocations.
- There are four crucial steps:
    1. Find implied returns.
    2. Formulate investor views.
    3. Determine what the expected returns are.
    4. Find the asset allocation for the optimal portfolio.
- You will find tons of resources and variations on this model on black-litterman.org.
- A good overview is given in Walters et al. (2014).

Black-Litterman Model (cont.)

Mean-variance analysis often produces unreasonable portfolios:

- Let the investment universe consist of 5 stocks.
- Based on historical returns, portfolio weights are determined via MV.
- Return adjustments: +1%pts for Asset 1, +0.5%pts for Asset 2, -1%pts for Asset 4.



- Problems: Communication of results, (re-) allocation in real portfolios, acceptance of method.

Black-Litterman Model (cont.)

Universität Zürich[UZH]

- Black and Litterman (1992) assume that the investor has views about returns that may differ from those implied by the CAPM.
- The Black-Litterman approach resembles Bayesian techniques in that it attempts to improve estimates of mean returns by combining information from data with the investor's priors.
- However, the information from data employed is not the historical sample average returns. Instead, starting point are returns **implicit** in market allocations.
- More specifically, it starts with an intuitive prior, the CAPM equilibrium market portfolio by reverse engineering.
- Hence, in contrast to using an uninformative prior or shrinking towards the global min-var portfolio (e.g., Frost and Savarino (1986) and Jorion (1986)), it establishes a prior with a direct connection to the market.

- Black and Litterman (1992) use the CAPM in a different way to assist portfolio construction in the context of asset allocation:

$$\boldsymbol{\mu} = \lambda \Sigma \boldsymbol{w}_m \ ,$$

  where $\boldsymbol{\mu}$ is the expected excess return vector implied by the CAPM, $\boldsymbol{w}_m$ is the market portfolio weight vector, Σ is the $N \times N$ variance-covariance matrix of the excess returns, and λ is a scale factor reflecting the average risk aversion of investors.
- Multiply both sides with $\boldsymbol{w}_m$ and solve for λ:

$$\lambda = \mu_m / \sigma_m^2,$$

  where μ_m is the market's excess return and σ_m^2 its variance.
- The term λ is often calibrated to the historical Sharpe ratio, hence $\lambda \approx 3$ for equity markets.
- Note: Here, Σ is the conditional variance $\text{Var}(R|\hat{\boldsymbol{\mu}})$. Some authors prefer the unconditional variance.

Universität Zürich[UZH]

Prior Distribution

- Since the equilibrium returns are not actually estimated, the estimation error cannot be directly derived.
- A reasonable assumption: Estimation error of the means of returns should be less than the covariance of the returns.
- A practical approach is to use an estimation error which is proportional (by a scalar τ) to the covariance matrix of returns.
- A scalar τ less than 1 is used to scale down the covariance matrix Σ of the returns. Some say that $\tau = 0.3$ is plausible.
- Given the above assumption, the prior distribution for the expected returns is $N(\boldsymbol{\mu}, \tau\hat{\Sigma})$, where $\hat{\Sigma}$ is the estimate of the covariance matrix Σ.

Universität Zürich[UZH]

Specifying Views

- By construction, each view is uncorrelated with other views. Improves stability and simplifies the problem!
- Either the sum of weights in a view is zero (**relative view**) or is one (an **absolute view**).
- We do not require a view on all assets. Views may even conflict, the mixing process will merge the views based on the **confidence in the views (priors)**.

Forming Our Views

- Our view can be formalized as:

$$\boldsymbol{q} = P\boldsymbol{\mu} + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(0,\Omega), \quad P \in \mathbb{R}^{k \times n},$$

  where $\boldsymbol{\mu}$ are the expected equilibrium returns, $\Omega \in \mathbb{R}^{k \times k}$ is diagonal, i.e., no cross-information on views is assumed.
- In $P\boldsymbol{\mu}$, each row of P represents a weight vector of n assets, so each row is a portfolio of assets (the view portfolio).
- Hence, $P\boldsymbol{\mu}$ means that we express our views with k view portfolios.
- For a relative view the sum of the weights will be 0, for an absolute view the sum of the weights will be 1.
- The number of outperforming and underperforming assets in a relative view need not coincide.

Universität Zürich[UZH]

Example

- Consider four assets and two views:
  - First, a relative view in which the investor believes that Asset 1 will outperform Asset 3 by 2% with confidence ω_{11}.
  - Second, an absolute view in which the investor believes that Asset 2 will return 3% with confidence ω_{22}.
- No views are specified on Asset 4.
- These views are specified as follows:

$$P = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}; \quad \boldsymbol{q} = \begin{bmatrix} 0.02 \\ 0.03 \end{bmatrix}; \quad \Omega = \begin{bmatrix} \omega_{11} & 0 \\ 0 & \omega_{22} \end{bmatrix}$$

Combining Equilibrium and View

- We start with $\hat{\boldsymbol{\mu}} = \boldsymbol{\mu} + \boldsymbol{\epsilon}, \boldsymbol{\epsilon} \sim \mathcal{N}(0,\tau\hat{\Sigma}), \boldsymbol{q} = P\boldsymbol{\mu} + \boldsymbol{\eta}, \boldsymbol{\eta} \sim \mathcal{N}(0,\Omega)$.
- Let $\boldsymbol{y} = [\hat{\boldsymbol{\mu}}, q]^\top$, $X = [I, P^\top]^\top \in \mathbb{R}^{(k+n)\times n}$, and $\boldsymbol{u} = [\boldsymbol{\epsilon}, \boldsymbol{\eta}]^\top$, where I is the $n \times n$ identity matrix. Then $\boldsymbol{u} \sim \mathcal{N}(0,\Psi)$ with

$$\Psi = \begin{bmatrix} \tau\hat{\Sigma} & 0 \\ 0 & \Omega \end{bmatrix}.$$

- With the above definitions, we get the regression equation $\boldsymbol{y} = X\boldsymbol{\mu}+\boldsymbol{u}$, which leads to the Generalized Least-Square Estimator

$$\boldsymbol{\mu}^c = (X^\top\Psi^{-1}X)^{-1}X^\top\Psi^{-1}\boldsymbol{y}.$$

- With the corresponding definitions, we get the Black-Litterman Master Formula.

Explicit calculation...

$$
\begin{aligned}
\boldsymbol{\mu}^c &= (X^\top \Psi^{-1} X)^{-1} X^\top \Psi^{-1} \boldsymbol{y} \\
&= \left(\begin{bmatrix} I & P^\top \end{bmatrix} \begin{bmatrix} \tau^{-1}\hat{\Sigma}^{-1} & 0 \\ 0 & \Omega^{-1} \end{bmatrix} \begin{bmatrix} I \\ P^\top \end{bmatrix} \right)^{-1} \\
&\quad \times \begin{bmatrix} I & P^\top \end{bmatrix} \begin{bmatrix} \tau^{-1}\hat{\Sigma}^{-1} & 0 \\ 0 & \Omega^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \boldsymbol{q} \end{bmatrix} \\
&= \left(\begin{bmatrix} I & P^\top \end{bmatrix} \begin{bmatrix} \tau^{-1}\hat{\Sigma}^{-1} & 0 \\ 0 & \Omega^{-1} P \end{bmatrix} \right)^{-1} \begin{bmatrix} I & P^\top \end{bmatrix} \begin{bmatrix} \tau^{-1}\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}} \\ \Omega^{-1}\boldsymbol{q} \end{bmatrix} \\
&= \left(\left(\tau\hat{\Sigma}\right)^{-1} + P^\top \Omega^{-1} P \right)^{-1} \left(\left(\tau\hat{\Sigma}\right)^{-1} \hat{\boldsymbol{\mu}} + P^\top \Omega^{-1} \boldsymbol{q} \right).
\end{aligned}
$$

Interpretation:

- First factor serves as normalization. Second factor balances between equilibrium returns $\hat{\boldsymbol{\mu}}$ and views P:
  - $\left(\tau\hat{\Sigma}\right)^{-1}$ serves as precision factor for the equilibrium returns.
  - $P^\top\Omega^{-1}$ serves as weighting factor or precision of the views.
- Limiting case "no view" $P = 0$, then $\boldsymbol{\mu}^c = \hat{\boldsymbol{\mu}}$.
- Limiting case "absolute confidence" $\Omega \to 0$, then $\boldsymbol{\mu}^c = P^{-1}\boldsymbol{q}$.
- Limiting case "no confidence" $\Omega \to \infty$, then $\boldsymbol{\mu}^c = \hat{\boldsymbol{\mu}}$.
- Limiting case "no estimation error" $\tau \to 0$, then $\boldsymbol{\mu}^c = \hat{\boldsymbol{\mu}}$.
- Limiting case "infinite estimation error" $\tau \to \infty$, then $\boldsymbol{\mu}^c = P^{-1}\boldsymbol{q}$.

- By using $\boldsymbol{q} = P\boldsymbol{\mu} + \boldsymbol{\eta}$, we have that the investor's view alone imply an OLS estimate of expected return views $\bar{\boldsymbol{\mu}}$ of

$$\bar{\boldsymbol{\mu}} = (P^\top P)^{-1} P^\top \boldsymbol{q}.$$

- Since $P(P^\top P)^{-1} P^\top = I$, we can rewrite the BL master formula as a weighted linear combination of market equilibrium $\hat{\boldsymbol{\mu}}$ and the expected return $\bar{\boldsymbol{\mu}}$ implied by the investor's views:

$$\mu^c = W_M \hat{\boldsymbol{\mu}} + W_I \bar{\boldsymbol{\mu}},$$

with weights matrices $W_M + W_I = I$ given by

$$\begin{aligned} W_M &= \left(\left(\tau\hat{\Sigma}\right)^{-1} + P^\top \Omega^{-1} P\right)^{-1} \left(\tau\hat{\Sigma}\right)^{-1}, \\ W_I &= \left(\left(\tau\hat{\Sigma}\right)^{-1} + P^\top \Omega^{-1} P\right)^{-1} P^\top \Omega^{-1} P. \end{aligned}$$

Some Remarks

- The above derivation is an application of Theil (1971)'s mixed estimation model.
- This model was created for the purpose of estimating parameters from a mixture of complete prior data and partial conditional data.
- It lends itself to our problem as it allows us to express views on only a subset of the asset returns.
- Note, the views can also be expressed on a single asset, or on arbitrary combinations of the assets.
- The views do not even need to be consistent, the estimation model will take each into account based on the investors confidence.

Some Remarks (cont.)

- The Black-Litterman model can be derived as a solution to the following optimization problem:

$$\boldsymbol{\mu}^c = \arg\min_{\boldsymbol{\mu}} \left\{ (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^\top \Sigma^{-1} (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) + \tau (\boldsymbol{q} - P\boldsymbol{\mu})^\top \Omega^{-1} (\boldsymbol{q} - P\boldsymbol{\mu}) \right\}.$$

- From this formulation we see that $\boldsymbol{\mu}^c$ is chosen such that it is simultaneously
  1. as close to $\hat{\boldsymbol{\mu}}$,
  2. and $P\boldsymbol{\mu}$ is as close to q as possible.
- The distances are scaled by Σ^{-1} and Ω^{-1}.
- Furthermore, the relative importance of the equilibrium versus the views is determined by τ.

Universität Zürich[UZH]

Mean Variance Optimization

- Decision theory (e.g., Chapter 2 of Robert (2007)) teaches us that the optimal decision is the one maximizing posterior expected utility.
- Hence, in a Markowitz framework, the last step is optimization with adjusted mean and the given covariance structure.
- To get the weights for the market portfolio with views, we can plug in

$$
\begin{aligned}
w^c &= \frac{1}{\lambda}\hat{\Sigma}^{-1}\boldsymbol{\mu}^c \\
&= \frac{1}{\lambda}\hat{\Sigma}^{-1}\left(\hat{\boldsymbol{\mu}} + \tau\hat{\Sigma}P^T\left(P\tau\hat{\Sigma}P^\top + \Omega\right)^{-1}(q - P\hat{\boldsymbol{\mu}})\right) \\
&= \boldsymbol{w} + P^T\left(P\hat{\Sigma}P^\top + \frac{1}{\tau}\Omega\right)^{-1}\left(\frac{1}{\lambda}\boldsymbol{q} - P\hat{\Sigma}\boldsymbol{w}\right).
\end{aligned}
$$

- Often, Ω is set to $\Omega = \mathrm{diag}\left[P\hat{\Sigma}P^\top\right]\tau$ to make w^c independent of τ.

Arbitrage Pricing and Multifactor Models

Outline

- Arbitrage Pricing Theory (APT) was developed primarily by Ross (1976).
- It is a one-period model in which every investor believes that the stochastic properties of returns of capital assets are consistent with a **factor structure**.
- Ross argues that if equilibrium prices offer no arbitrage opportunities over static portfolios of the assets, then the **expected returns on the assets are approximately linearly related to the factor loadings**.
- Ross' (1976) heuristic argument for the theory is based on the **preclusion of arbitrage**.
- Ross' formal proof shows that the **linear pricing relation is a necessary condition** for equilibrium in a market where agents maximize certain types of utility.

Beta pricing relationships can be derived using **asymptotic no arbitrage arguments**, opening the path to multi-factor models.

- We start with a regression model

$$R^e_{it} = \alpha_i + \beta_i R^e_{mt} + \epsilon_{it} \ ,$$

  where R^e_{it} denotes the excess return on asset i over the riskless interest rate, and assume $\mathbb{E}[\epsilon_{it}\epsilon_{jt}] = 0$ for $i \neq j$.

- Consider a portfolio

$$R^e_{pt} = \alpha_p + \beta_p R^e_{mt} + \epsilon_{pt} \ ,$$

  where $\alpha_p = \sum_{j=1}^N w_j \alpha_j$, $\beta_p = \sum_{j=1}^N w_j \beta_j$ and $\epsilon_{pt} = \sum_{j=1}^N w_j \epsilon_{jt}$.

- For an **equal-weighted portfolio** and when all stocks have the same idiosyncratic variance,

$$\mathrm{Var}[\epsilon_{pt}] = \sum_{j=1}^{N} w_j^2 \mathrm{Var}[\epsilon_{jt}] = \frac{\sigma^2}{N} \; ,$$

  which will shrink rapidly with N.

- Therefore, it would be possible to construct a "**well diversified**" portfolio with $R^e_{pt} = \alpha_p + \beta_p R^e_{mt}$. The portfolio's return contains only factor variance, but then we must have $\alpha_p = 0$ to avoid arbitrage opportunities.

- This argument lies at the heart of Ross (1976): $\alpha_p = 0$ for all well-diversified portfolios implies that "**almost all**" individual assets have α_i very close to zero.

There are different formal proofs for the APT. They rely on some structure on the covariance matrix $\mathbb{E}[\epsilon\epsilon^\top]$:

- Huberman (1982) requires $\mathbb{E}[\epsilon\epsilon^\top]$ to be diagonal and uniformly bounded.
- Ingersoll (1984) does not require diagonality.
- Using utility-based arguments, we can show that the exact APT holds iff $\mathbb{E}[M\epsilon] = 0$, where M is the stochastic discount factor (SDF). This holds for all linear SDFs. For nonlinear SDFs, we require also $\mathbb{E}[\epsilon | R^e_{mt}] = 0$.
- Connor (1984) shows that if the market portfolio is well diversified, every investor holds a well-diversified portfolio ($k+1$ separation theorem). Then, the first order condition of any investor implies exact arbitrage pricing in a competitive equilibrium.

Outline

The assumption of a single-factor APT having uncorrelated residual risk from the market model is counterfactual.

- If there are K "**factor portfolios**" with excess return R^e_{kt} capturing the common influence of K underlying sources of risk,

$$R^e_{it} = \alpha_i + \sum_{k=1}^{K} \beta_{ik} R^e_{kt} + \epsilon_{it}, \quad \mathbb{E}[\epsilon_{it}\epsilon_{jt}] = 0.$$

- Alternatively, we can measure the factors directly as mean-zero shocks

$$R_{it} = \mu_i + \sum_{k=1}^{K} \beta_{ik} f_{kt} + \epsilon_{it} \ ,$$

with $\mu_i = \alpha_i + \lambda_0 + \sum_{k=1}^{K} \beta_{ik}\lambda_k$, where λ_0 is the zero-beta or riskfree rate, and λ_k is the price of risk of the k'th factor.

Universität Zürich[UZH]

Some implications for portfolio selection:

1. For an exact K-factor model ($\alpha_i = 0 \ \forall \ i$), the full mean-variance-efficient frontier can be constructed from the K factor portfolios and the zero-beta portfolio.
2. We can construct "**factor-mimicking portfolios**" that have the highest correlation with a given factor among traded portfolios.
3. Any stock-level characteristic that predicts returns must be associated with some common risk factor.

Some weaknesses of APT:

1. The model predicts that **nonzero alphas are rare in the limit** of a large cross-section, and it is not clear how this prediction can be falsified in a finite cross-section of asset returns.
2. We need some confidence that the K-factor model holds ex ante as well as ex post. In other words, **we need theoretical reasons** to believe that a K-factor model is structural.
3. APT does **not determine the signs or magnitudes** of the risk premia.
4. It is **not clear how the APT can be falsified** in a finite cross-section of asset returns.
5. Ex-post, we can easily fit a K-factor model, but we do not know why it should also hold ex-ante. **For that, we need some general equilibrium argument.**

- The APT was developed as a generalization of the CAPM.
- The CAPM says that the market portfolio is mean-variance efficient.

✎ If the factors can be identified with traded assets then exact arbitrage pricing implies that a portfolio of these factors is mean-variance efficient.

- Jobson and Korkie (1982, 1985) and Jobson (1982) note the relation between the APT and mean-variance efficiency, proposing likelihood ratio tests of the joint hypothesis that
  a) a given set of random variables are indeed factors
  b) and that exact arbitrage pricing obtains.

Universität Zürich[UZH]

✎ We can derive an unconditional multifactor model from a conditional version of the CAPM (Hansen and Richard, 1987).

- If the CAPM holds conditionally, then $\mathbb{E}_t[R^e_{i,t+1}] = \beta_{imt}\mathbb{E}_t[R^e_{m,t+1}]$, where β_{imt} is the **conditional beta**.
- Taking unconditional expectations, we get

$$\begin{aligned}
\mathbb{E}[R^e_{i,t+1}] &= \mathbb{E}[\beta_{imt}]\mathbb{E}[R^e_{m,t+1}] + \text{Cov}(\beta_{imt}, \mathbb{E}_t[R^e_{m,t+1}]) \\
&= \beta_{im}\mathbb{E}[R^e_{m,t+1}] + \left(\mathbb{E}[\beta_{imt}] - \beta_{im}\right)\mathbb{E}[R^e_{m,t+1}] \\
&\quad + \text{Cov}(\beta_{imt}, \mathbb{E}_t[R^e_{m,t+1}]),
\end{aligned}$$

  with the **unconditional beta** β_{im}.
- Lewellen and Nagel (2006) and Boguth et al. (2011) show that

$$\mathbb{E}[\beta_{imt}] - \beta_{im} \approx -\frac{\text{Cov}(\beta_{imt}, \sigma^2_{mt})}{\sigma^2_m}.$$

Hence, we can interpret the conditional CAPM as the unconditional CAPM with two additional terms:

- The term $\mathbb{E}[\beta_{imt}] - \beta_{im}$ a **volatility timing** effect:

  If an asset has high conditional beta when market volatility is low, the time-series average of its conditional beta will be higher than its unconditional beta, **resulting in a higher unconditional average return** than predicted by the unconditional CAPM.

- The term $Cov(\beta_{imt}, \mathbb{E}_t[R^e_{m,t+1}])$ is a **equity premium timing** effect:

  An asset with a beta that tends to be high when the market risk premium is high, **it delivers market risk exposure at times when market risk is highly rewarded**, and this increases its unconditional CAPM-implied average return.

To empirically test a conditional CAPM, we can parameterize the evolution of the betas:

- We can use a state variable z_t and write

$$\beta_{imt} = \beta_{i0} + \beta_{i1} z_t.$$

- The conditional model is then

$$\mathbb{E}_t[R^e_{i,t+1}] = \beta_{i0}\mathbb{E}_t[R^e_{m,t+1}] + \beta_{i1}\mathbb{E}_t[z_t R^e_{m,t+1}]$$

- Taking unconditional expectations, we get

$$\mathbb{E}[R_{i,t+1}] = \beta_{i0}\mathbb{E}[R^e_{m,t+1}] + \beta_{i1}\mathbb{E}[z_t R^e_{m,t+1}],$$

  which is a multifactor model, with the excess market return and the excess market return scaled by the state variable z_t as factors.

- We can interpret second factor as the return on a **dynamic investment strategy that incorporates volatility and equity premium timing**.

Empirical Evidence

Outline

Universität Zürich[UZH]

Miller (2000): "I still remember the teasing we financial economists, Harry Markowitz, William Sharpe, and I, had to put up with from the physicists and chemists in Stockholm when we conceded that the basic unit of our research, the expected rate of return, was not actually observable. I tried to tease back by reminding them of their neutrino –a particle with no mass whose presence was inferred only as a missing residual from the interactions of other particles. But that was eight years ago. In the meantime, the neutrino has been detected."

- The CAPM relationship is expressed in terms of expected values, which are **not observable**.
- Even if the CAPM is a correct model, it will **never exactly describe sample moments in a finite sample**.
- We need a statistical test to tell whether sample deviations from the model (mean-variance inefficiency of the market portfolio, or equivalently nonzero alphas) are statistically significant, i.e., our Null is

$$H_0 \ : \ \alpha_i = 0, \forall i.$$

- To **jointly** test for zero alphas, there are two leading approaches, in historical order, **time-series** and **cross-sectional**.

- To obtain a model with observable quantities, we write

$$R^e_{it} = \alpha_i + \beta_i R^e_{mt} + \epsilon_{it}, \quad \mathrm{Cov}(\epsilon_t) = \Omega \ .$$

- The time-series approach tests the null hypothesis $\alpha_i = 0, \forall i = 1, \ldots, N$.
- **Assumption:** ϵ_{it} are iid normal and independent of R^e_{mt}.
- The joint density of excess returns over $t \in [1, \ldots T]$ is

$$\begin{aligned} f(R^e_{1:T} | R^e_{m,1:T}) &= \prod_{t=1}^{T} f(R^e_t | R^e_{m,t}) \\ &= \frac{e^{\left(-\frac{1}{2}\sum_{t=1}^{T}\left(R^e_t - \alpha - \beta R^e_{m,t}\right)^\top \Omega^{-1}\left(R^e_t - \alpha - \beta R^e_{m,t}\right)\right)}}{(2\pi)^{NT/2}|\Omega|^{T/2}}. \end{aligned}$$

- To estimate the unknown parameters, α, β, and Ω of this density, we maximize the log-likelihood function, i.e., the log of the joint density viewed as a function of the unknown parameters.
- From the MLE, we get that the parameters are unbiased and normally distributed.
- The alphas have the following covariance matrix:

$$
\begin{aligned}
\mathrm{Cov}(\hat{\alpha},\hat{\alpha}) &= \mathrm{Cov}(\bar{R}^e - \hat{\beta}\bar{R}^e_m) \\
&= \mathrm{Cov}\left[\frac{1}{T}\sum_{t=1}^{T} R^e_t - \bar{R}^e_m \sum_{t=1}^{T} \frac{(R^e_{m,t} - \bar{R}^e_m)R^e_t}{T\hat{\sigma}^2_m}\right] \\
&= \frac{1}{T^2\hat{\sigma}^4_m}\sum_{t=1}^{T}\left[\hat{\sigma}^2_m - \bar{R}^e_m(R^e_{m,t} - \bar{R}^e_m)^2\right]^2 \mathrm{Cov}(R_t) \\
&= \frac{1}{T^2\hat{\sigma}^4_m} T\left[\hat{\sigma}^4_m + (\bar{R}^e_m)^2\hat{\sigma}^2_m\right]\hat{\Omega} = \frac{1}{T}\left(1 + \frac{(\bar{R}^e_m)^2}{\hat{\sigma}^2_m}\right)\hat{\Omega}.
\end{aligned}
$$

- Hence, the (asymptotic) **Wald test** for testing the null hypothesis (for a refresher, see here):

$$\hat{\alpha}\mathrm{Cov}(\hat{\alpha},\hat{\alpha})^{-1}\hat{\alpha} = T\left[1+\left(\frac{\bar{R}^e_m}{\hat{\sigma}_m}\right)^2\right]^{-1}\hat{\alpha}\hat{\Omega}^{-1}\hat{\alpha} \overset{\text{asy}}{\sim} \chi^2_N \; .$$

- As an alternative test, we can use a **Likelihood Ratio (LR) Test**, which is based on the comparison between the log–likelihood values of the unconstrained (L_1) and the constrained model (L_0).
- Hence, along the same lines as above, we need to first compute the MLE under the Null $\alpha = 0$, which gives us the Likelihood L_0.
- The LR test statistic is given by

$$\mathcal{LR} = -2\left(\log L_0 - \log L_1\right) = T\left(\log|\hat{\Omega}_0| - \log|\hat{\Omega}_1|\right) \overset{\text{asy}}{\sim} \chi^2(N).$$

- A **finite-sample test** makes a further correction for the fact that the variance-covariance matrix Ω must be estimated.
- Since in our case, $\hat{\Omega}$ has a Wishart distribution, and since the quadratic form of a Wishart with normal variates is F-distributed, we can derive:

$$J = \left(\frac{T-N-1}{N}\right)\left[1 + \left(\frac{\bar{R}^e_m}{\hat{\sigma}(R^e_{mt})}\right)^2\right]^{-1} \hat{\alpha}\hat{\Omega}^{-1}\hat{\alpha} \overset{\text{asy}}{\sim} F_{N,T-N-1} \; .$$

- Gibbons et al. (1989) provide a geometric interpretation that relates the J-statistic to mean-variance analysis:
  - Define $\hat{SR}^2_m$ as Sharpe ratio of the portfolio used as a market proxy and $\hat{SR}^2_T$ as the squared Sharpe ratio of the ex-post efficient tangency portfolio.
  - Then, the J-statistic is equivalent to

$$J = \left(\frac{T-N-1}{N}\right)\left(\frac{\hat{SR}^2_T - \hat{SR}^2_m}{1+\hat{SR}^2_m}\right) \; .$$

- This relation uncovers what we are actually testing: We test whether our market proxy is so far away from the ex post efficient portfolio that we are not willing to believe that it is the population tangency portfolio, where the distance is measured in terms of the Sharpe Ratio.

- To prove the above relation, we have to show that $\hat{SR}_T^2 - \hat{SR}_m^2 = \alpha'\hat{\Omega}^{-1}\alpha$. Let $\tilde{R}^e = [\bar{R}_m^e, (\bar{R}^e)']$ having sample covariance matrix

$$V = \begin{bmatrix} \hat{\sigma_m}^2 & \hat{\sigma_m}^2\hat{\beta}' \\ \hat{\sigma_m}^2\hat{\beta}' & \hat{\Omega} + \hat{\sigma_m}^2\hat{\beta}\hat{\beta}' \end{bmatrix}.$$

We know that the efficient portfolio using the N assets has weights $w = V^{-1}\tilde{R}^e/(\mathbf{1}'V^{-1}\mathbf{1})$. Hence, $\hat{SR}_T^2 = (\tilde{R}^e)'V^{-1}\tilde{R}^e$. Also, we can verify that

$$V^{-1} = \begin{bmatrix} \hat{\sigma_m}^{-2} + \hat{\beta}'\hat{\Omega}^{-1}\hat{\beta} & -\hat{\beta}'\hat{\Omega}^{-1} \\ -\hat{\beta}'\hat{\Omega}^{-1} & \hat{\Omega}^{-1} \end{bmatrix}.$$

- Plugging V^{-1} into $\hat{SR}^2_T$ and recalling $\hat{\alpha} = \bar{R}^e - \hat{\beta}\bar{R}^e_m$, we get the desired result.
- For the zero-beta CAPM version, we can use a similar test methodology. For details, see, e.g., Shanken (1986).
- For a test of the conditional CAPM, see, e.g., Jagannathan and Wang (1998).
- Remark:
  The above statistics, and especially the GRS test, are derived from the classical assumptions that returns are independent over time, uncorrelated across assets, and (for finite sample results) normally distributed. There are more modern formulas that account for correlation of asset returns over time, conditional heteroskedasticity (returns are more volatile at some times than others) and non-normality. Mostly, they are all instances of GMM (Cochrane, 2009).

Universität Zürich[UZH]

- Another idea to test CAPM and APT is built on the models main idea: understand average returns *across* assets.
- The cross-sectional approach first estimates betas from a time-series regression.
- Then, we run a cross-sectional regression

$$\bar{R}_i^e = \lambda \hat{\beta}_i + a_i \ ,$$

with no intercept, where λ is the cross-sectional reward for bearing market risk.

- As the sample size increases $T \to \infty$,

$$\bar{R}_i^e \to \mathbb{E}[R_{it}^e], \quad \hat{\beta}_i \to \beta_i, \quad \lambda \to \mathbb{E}[R_{mt}^e], \quad a_i \to \alpha_i .$$

- Hence, the cross-sectional residuals are consistent estimates of the true alphas and we need to test whether the residuals of our regression are close to zero.
- Since the cross-sectional residuals tend to be highly correlated, we prefer to run a Generalized Least Squares (GLS):

$$\hat{\lambda}_{GLS} = (\hat{\beta}'\hat{\Omega}^{-1}\hat{\beta})^{-1}\hat{\beta}'\hat{\Omega}^{-1}\bar{R}^e \ ,$$
$$\hat{a}_{GLS} = \bar{R}^e - \hat{\lambda}_{GLS}\hat{\beta} \ ,$$

  where $\hat{\beta} = [\hat{\beta}_1, \ldots, \hat{\beta}_N]'$.

- Correcting for the fact that the betas are not known but are estimated from a prior time-series regression, the test statistics becomes

$$T\left[1 + \left(\frac{\hat{\lambda}_{GLS}}{\hat{\sigma}(R^e_{mt})}\right)^2\right]^{-1} \hat{a}'_{GLS}\hat{\Omega}^{-1}\hat{a}_{GLS} \overset{\text{asy}}{\sim} \chi^2_{N-1} \ .$$

- The most common cross-sectional regression is based on Fama and MacBeth (1973) (FMB).
- Again, we run a time-series regression to obtain the betas. Then, at each time period, we run a cross-sectional regression of N excess returns on estimated betas

$$R^e_{it} = \lambda_t \hat{\beta}_i + a_{it} \ .$$

- The coefficients $\hat{\lambda}_t$ and $\hat{a}_{it}$ residuals are then averaged over time to estimate the average reward for beta exposure $\hat{\lambda}_{FM} := (1/T) \sum \hat{\lambda}_t$ and the average residuals $\hat{a}_{i,FM} := (1/T) \sum \hat{a}_{it}$.

- Under the standard assumption that all returns are iid over time the standard errors of these averages are

$$\hat{\sigma}^2(\hat{a}_{i,FM}) = \frac{1}{T}\hat{\sigma}^2(\hat{a}_{it}) = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{a}_{it} - \hat{a}_{i,FM})^2 \ ,$$

  and similarly for $\hat{\lambda}_{FM}$.

- The FMB method for calculating standard errors does not adjust for the fact that stocks' betas are not known but must be estimated from time-series regressions.
- See, e.g., Gow et al. (2010) for an interesting review of the adjustment of cross-sectional regressions for serial and cross-sectional dependence for FMB methods, typically used in the accounting literature.

Fama-MacBeth Approach (cont.)

- One advantage of FMB: it's easy to do models in which the betas change over time. That's harder to incorporate in time-series and cross-sectional regression.
- Other applications of FMB: anytime you have a big cross-section, which may be correlated with each other. (If the errors are uncorrelated across time and space, just stack it up to a huge OLS regression)
  - E.g., in corporate finance, one often has something like:

    $$investment_{it} = a + b \times Book/Market_{it} + c \times profits_{it} + \epsilon_{it}$$

    If we stack all into one OLS, the estimates are ok — but you can't use the standard errors, because the ϵ's are correlated with each other. FMB is one way to correct the standard errors.
  - In a recent survey, Petersen (2009) finds that more than half of all published papers in the Journal of Finance ignored the fact that errors are cross-correlated, and that standard errors are typically off by a factor of 10. **Half the (old) stuff in the JF is wrong!**

We briefly summarize an enormous empirical literature on the cross-section of stock returns, with particular attention to characteristics of stocks that appear to predict returns in a way that cannot be explained by the CAPM (for a recent contribution in this direction, see Hou et al. (2018)):

1. **Beta**:
   Very early studies found a positive and close to linear relationship between subsequent betas of beta-sorted portfolios and their average returns (Black et al., 1972; Fama and MacBeth, 1973). Subsequent studies of the beta-return relationship have found a weak relationship.
2. **Size:**
   Banz et al. (1981) finds that firms with low market capitalization ("small" firms) tend to have higher average returns than their betas justify.

Universität Zürich[UZH]

3. **Value versus growth**:
   Value investors seek to buy stocks with low market capitalization relative to accounting measures. This approach to investing was popularized several decades before the formulation of the CAPM (Graham and Dodd, 1934). Since Fama and French (1993), one uses book-market ratio as "value."
4. **Momentum**:
   Going long past winners and going short past losers.
5. **Post-Event Drift**:
   Post-event drift (e.g., earning announcements) is similar to price momentum in that higher initial returns are associated with higher subsequent returns, but different in that initial returns are measured conditional on the occurrence of an event.

6. **Turnover and Volatility**:
   Both turnover and idiosyncratic volatility appear to be negatively associated with returns at intermediate 3–12 month horizons, although turnover may have a shorter-term positive effect. These phenomena may be related to a negative effect of disagreement among equity analysts (see, e.g., Leippold and Lohre (2014)), since investors who disagree about a firm's prospects are likely to trade more aggressively with one another and may create volatility in doing so.
7. **Insider Trading**:
   There is some evidence that purchases and sales by corporate insiders predict future stock returns (Seyhun 1988), with stronger evidence for purchases (Jeng, Metrick, and Zeckhauser 2003) and for sales that are not part of a regular sale program (Cohen, Malloy, and Pomorski 2012).

8. **Growth of the Firm**:
   Variables that capture growth of a firm's assets, such as capital expenditures and total asset growth, are negatively associated with subsequent returns. The same is true for variables that capture growth of a firm's liabilities, including measures of equity and debt issuance and the percentage change in shares outstanding.
9. **Earnings Quality**:
   A large literature in accounting argues that different components of current earnings have different degrees of relevance for future earnings and hence may predict returns if investors do not fully appreciate these distinctions.

10. **Profitability**:
    Firms with higher ratios of gross profitability to book assets have higher subsequent returns. Gross profitability differs from current accounting earnings in that it adds back a number of corporate expenses that may enhance the long-run prospects of the firm, including advertising, sales commissions, and R&D. For new evidence, see Feng et al. (2019).

By the way, when is a factor a factor? Consult, e.g., Pukthuanthong et al. (2018).

Evidence that characteristics predict stock returns presents a challenge to the CAPM and has generated a variety of responses.

1. **Data mining:** The evidence may have been found by a process of experimentation with many candidate explanatory variables. Of every 100 variables considered, 5 will appear to predict returns at the 5% significance level purely by chance (p-hacking).
2. **Anomaly elimination:** Even if return predictability is the genuine result of investor behavior in the period during which it is measured, it is possible that publicizing the result leads investors to alter their behavior to weaken or even eliminate the phenomenon.
3. **Roll critique:** CAPM tests are flawed because we do not know the composition of the true market portfolio, and so we use a broad stock index as an imperfect proxy. A rejection of the CAPM just tells us that our proxy is inadequate.

4. **Illiquidity:** Some economists accept deviations from the CAPM as genuine but downplay their importance on the grounds that they are concentrated in small, illiquid stocks.
5. **Conditional CAPM:** Standard tests of the CAPM assume that stocks' betas and the market risk premium are constant over time.
6. **A theoretical or equilibrium multifactor risk model:** Under quite weak conditions we should expect to be able to find a multifactor risk model that describes the cross-section of stock returns.
7. **Behavioral finance:** Behavioral finance economists believe that investors with irrational beliefs (and/or highly nonstandard preferences) have an important influence on market equilibrium.

A Closer Look on Momentum

- Momentum is one of the most puzzling observations in financial time series. Two main strategies:
  - Price momentum: past winning stocks continue to deliver superior returns, past losing stocks continue to disappoint.
  - Earnings momentum: momentum in stock prices follow the direction of analyst's earning forecast revision.
- Price and earnings momentum are defying market efficiency:
  - Jegadesh and Titman (1993): Price momentum.
  - Chan, Jegadesh, and Lakonishok (1996): Earnings momentum.
  - Rouwenhorst (1998), Griffith, Ji, and Martin (2003, 2005): International evidence.
- But can we christen momentum returns as an "anomaly?"

If we find the momentum anomaly to be robust, we may ask the following questions:

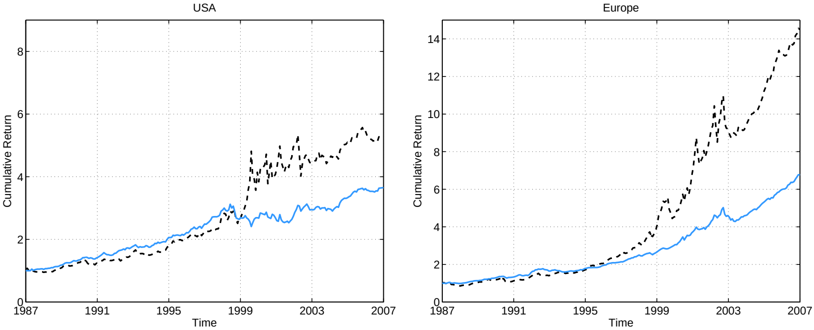
1. Is price momentum subsumed by earnings momentum or vice versa, i.e., do investors **underreact** to fundamental news represented by earnings revision or do they **overreact** to price changes?
2. What is the economic rationale for the existence of momentum? In particular:
   a) Is there a relation between macroeconomy and momentum, i.e., does momentum proxy macroeconomic risk?
   b) Is momentum related to information uncertainty, i.e., do investors underreact when the signal about fundamentals is noisier?
3. Why is momentum not arbitraged away and what causes limits to arbitrage?

Momentum Portfolios: Risk and Return

		Momentum Ranking					Hedge Strategy
Country		Lowest	2	3	4	Highest	
Price Momentum							
USA	Return	0.93	1.15	1.21	1.27	1.72	0.79 (2.80)
	Volatility	6.48	4.41	3.98	4.17	5.98	4.40
	Beta	1.20	0.82	0.72	0.76	1.07	-0.14
	Size	19.77	20.29	20.46	20.49	20.21	2.80
Europe	Return	0.56	0.88	1.10	1.25	1.75	1.19 (5.00)
	Volatility	5.76	4.24	3.93	4.02	4.77	3.69
	Beta	1.24	0.94	0.87	0.89	1.03	-0.21
	Size	20.32	20.92	21.16	21.29	21.15	5.00
Earnings Momentum							
USA	Return	1.27	1.16	1.10	1.43	1.85	0.58 (4.11)
	Volatility	5.50	4.40	3.82	4.21	4.91	2.17
	Beta	1.15	0.90	0.74	0.81	0.99	-0.04
	Size	19.47	20.17	20.61	20.60	20.04	4.11
Europe	Return	0.91	0.98	1.06	1.28	1.74	0.83 (7.52)
	Volatility	4.82	4.28	3.82	3.74	3.99	1.71
	Beta	1.18	1.05	0.92	0.89	0.96	-0.14
	Size	19.96	21.02	21.43	21.43	20.63	7.52

- Extreme quintiles: large volatility and beta, size bias.
- Momentum portfolios: negative betas!
- Europe: 12 out of 16 with significant alphas.

Cumulative Returns for the U.S. and Europe



- Price momentum (dashed), earnings momentum (solid).
- Price momentum: Higher returns at a higher volatility.
- Earnings momentum higher risk-adjusted performance.

$\gg$ Not explained by Fama-French factors, i.e., compensation for bearing risk?

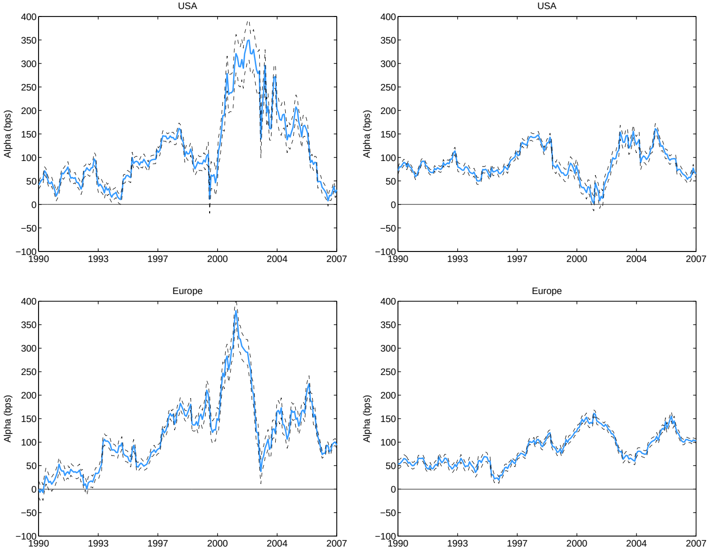
Universität Zürich[UZH]

$$R_{Lt} - R_{St} = \alpha + \beta(R_{Mt} - R_{Ft}) + \gamma R_{SMBt} + \delta R_{HMLt} + \varepsilon_t$$

		Fama-French Model								
		α	β	γ	δ	$t(\alpha)$	$t(\beta)$	$t(\gamma)$	$t(\delta)$	Adj. R^2
Price Momentum										
USA	1	-0.90	1.00	0.34	0.08	-5.29	19.38	5.43	1.28	84.5
	5	0.11	0.82	0.34	-0.30	0.63	15.64	5.45	-5.06	81.0
	5-1	1.01	-0.18	0.01	-0.38	3.57	-2.07	0.08	-3.88	7.0
Europe	1	-0.41	0.76	0.41	0.21	-2.54	8.23	5.55	2.76	84.6
	5	1.05	0.52	0.45	-0.20	7.82	6.80	7.32	-3.19	84.7
	5-1	1.46	-0.24	0.04	-0.41	5.84	-1.68	0.33	-3.49	9.4
Earnings Momentum										
USA	1	-0.63	1.00	0.22	0.12	-6.10	27.17	5.30	3.38	92.5
	5	0.22	0.75	0.33	-0.01	1.80	17.50	6.85	-0.16	87.3
	5-1	0.85	-0.25	0.11	-0.12	6.15	-5.12	2.01	-2.67	14.5
Europe	1	-0.15	0.72	0.38	0.14	-1.59	12.73	8.56	3.06	92.5
	5	0.89	0.47	0.42	0.03	10.14	9.13	10.25	0.82	90.9
	5-1	1.05	-0.25	0.04	-0.10	9.68	-3.94	0.70	-2.07	24.3

- Risk factors explain most of the variation in extreme quintiles.
- Very low explanatory power for momentum portfolios.
- Statistically and economically significant alphas for 14 countries (above 90bps).

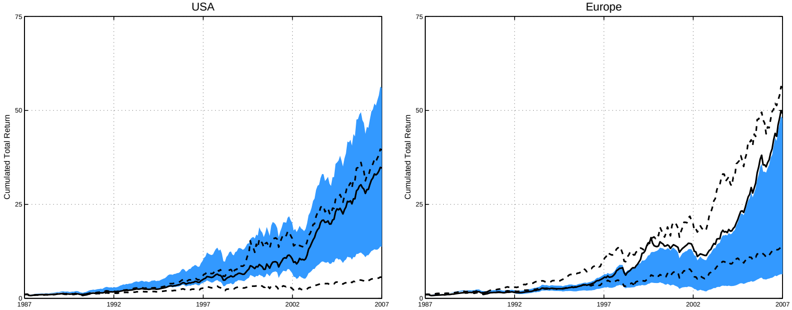
Trailing Alphas for the U.S. and Europe



Is Momentum due to Data Snooping?

Country	Price Momentum θ_s	StepM c_l	StepM rej	FDP-StepM c_l	FDP-StepM rej	Earnings Momentum θ_s	StepM c_l	StepM rej	FDP-StepM c_l	FDP-StepM rej
USA	0.0101	0.0046	1	0.0067	1	0.0085	0.0054	1	0.0067	1
Europe	0.0146	0.0082	1	0.0106	1	0.0105	0.0079	1	0.0090	1
UK	0.0090	0.0037	1	0.0057	1	0.0080	0.0052	1	0.0063	1
Ireland	0.0040	-0.0041	0	-0.0011	0	0.0145	0.0030	1	0.0076	1
Germany	0.0128	0.0060	1	0.0086	1	0.0087	0.0049	1	0.0064	1
Austria	0.0032	-0.0036	0	-0.0010	0	0.0089	0.0030	1	0.0054	1
Switzerland	0.0093	0.0025	1	0.0051	1	0.0081	0.0035	1	0.0054	1
France	0.0116	0.0063	1	0.0083	1	0.0100	0.0063	1	0.0078	1
Italy	0.0119	0.0056	1	0.0080	1	0.0042	-0.0003	0	0.0015	1
Greece	0.0217	0.0120	1	0.0156	1	0.0045	-0.0031	0	0.0000	0
Spain	0.0066	0.0008	1	0.0030	1	0.0103	0.0042	1	0.0067	1
Portugal	0.0102	0.0006	1	0.0042	1	0.0106	0.0031	1	0.0061	1
Netherlands	0.0113	0.0057	1	0.0078	1	0.0108	0.0052	1	0.0074	1
Belgium	0.0118	0.0052	1	0.0077	1	0.0088	0.0045	1	0.0062	1
Sweden	0.0122	0.0055	1	0.0080	1	0.0093	0.0035	1	0.0058	1
Norway	0.0106	0.0025	1	0.0056	1	0.0071	0.0003	1	0.0030	1
Denmark	0.0134	0.0077	1	0.0099	1	0.0123	0.0055	1	0.0082	1
Finland	0.0124	0.0047	1	0.0076	1	0.0140	0.0077	1	0.0103	1
Σ			16		16			16		17

Linking Price and Earnings Momentum



- Time series correlation of price (dashed lines) and earnings (shaded) momentum strategies suggest a close relationship.
- Price and earnings momentum may be traced back to similar sources.
- Is price momentum just a noisy proxy of earnings momentum?

Does Earnings Momentum Subsume Price Momentum?

Universität Zürich[UZH]

$$R_{WMLt} = \alpha + \beta(R_{Mt} - R_{Ft}) + \gamma R_{SMBt} + \delta R_{HMLt} + \zeta R_{PMNt} + \varepsilon_t$$

		Fama-French Model			*4-Factor Model*				
		α	$t(\alpha)$	Adj R^2.	α	ζ	$t(\alpha)$	$t(\zeta)$	Adj. R^2
USA	1	-0.90	-5.29	84.5	-0.80	-0.17	-4.20	-2.02	83.7
	2	-0.24	-2.10	84.8	-0.30	0.16	-2.38	2.83	83.7
	3	-0.07	-0.63	80.3	-0.17	0.27	-1.43	5.10	80.5
	4	-0.04	-0.34	81.0	-0.15	0.32	-1.33	6.38	83.0
	5	0.11	0.63	81.0	0.01	0.30	0.04	3.74	80.9
	5-1	**1.01**	**3.57**	7.0	**0.80**	0.47	**2.65**	3.51	14.5
Europe	1	-0.41	-2.54	84.6	0.36	-0.81	2.22	-9.71	88.6
	2	0.15	1.74	92.3	0.35	-0.19	3.50	-3.66	91.7
	3	0.47	6.01	92.4	0.40	0.11	4.41	2.31	91.7
	4	0.64	7.64	91.7	0.41	0.27	4.51	5.84	91.9
	5	1.05	7.82	84.7	0.52	0.56	3.60	7.55	85.6
	5-1	**1.46**	**5.84**	9.4	**0.16**	1.37	**0.66**	11.13	42.9

- Typically high R^2 for quintile portfolios.
- U.S. 4-factor alpha still significant, contrasting Chordia and Shivakumar (2006).
- European aggregate 4-factor alpha not significant (same for other 7 countries).

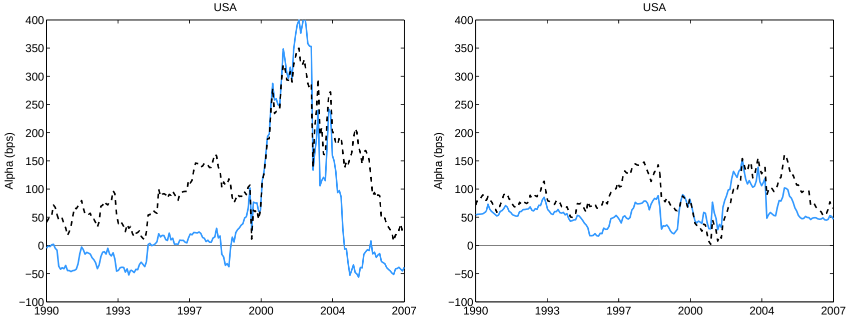
Does Price Momentum Subsume Earnings Momentum?

$$R_{PMNt} = \alpha + \beta(R_{Mt} - R_{Ft}) + \gamma R_{SMBt} + \delta R_{HMLt} + \eta R_{WMLt} + \varepsilon_t$$

		Fama-French Model			*4-Factor Model*				
		α	$t(\alpha)$	Adj. R^2	α	η	$t(\alpha)$	$t(\eta)$	Adj. R^2
USA	1	-0.63	-6.10	92.5	-0.57	-0.05	-5.35	-2.18	92.6
	2	-0.37	-3.88	89.9	-0.30	-0.06	-3.02	-2.95	90.3
	3	-0.18	-1.60	82.2	-0.11	-0.06	-0.97	-2.23	82.5
	4	0.06	0.46	82.4	0.09	-0.03	0.69	-0.92	82.4
	5	0.22	1.80	87.3	0.15	0.06	1.20	2.08	87.5
	5-1	**0.85**	**6.15**	14.5	**0.72**	0.11	**5.14**	3.51	18.5
Europe	1	-0.15	-1.59	92.5	0.10	-0.16	1.08	-6.98	93.7
	2	0.13	1.77	94.3	0.28	-0.09	3.60	-4.89	94.8
	3	0.34	4.57	92.8	0.42	-0.04	5.12	-2.23	92.9
	4	0.57	6.97	91.0	0.49	0.05	5.58	2.39	91.2
	5	0.89	10.14	90.9	0.74	0.09	8.06	4.21	91.5
	5-1	**1.05**	**9.68**	24.3	**0.64**	0.26	**6.76**	11.13	50.6

- Alphas remain large and significant, both for the U.S. and the European aggregate.
- 13/15 European countries remain significant.

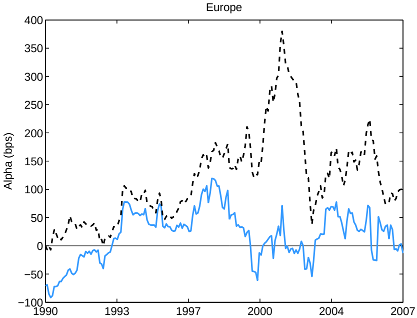
Fama-French vs. Four-Factor Alphas: U.S.

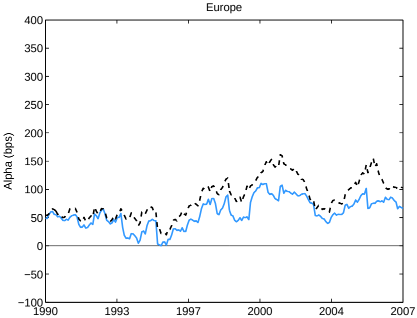


- Fama-French (dashed), 4-Factor (solid).
- Left figure: Earnings momentum subsumes price momentum except for 2000-2004.
- Right figure: Price momentum does not subsume earnings momentum.

Fama-French vs. Four-Factor Alphas: Europe

Universität Zürich[UZH]





- Fama-French (dashed), 4-Factor (solid).
- Left figure: Earnings momentum subsumes price momentum consistently throughout sample period.
- Right figure: Price momentum does not subsume earnings momentum.

What we have learned so far:

- Price momentum is mostly subsumed by earnings momentum in European equity markets.
- The market frenzy at the end of the nineties caused a temporary decoupling of this relationship in the U.S..
- Nevertheless, given this close relation suggests that both momentum phenomena may be traced back to similar sources.

What causes momentum?

Two strands of research:

- Risk-based explanation, see Chordia and Shivakumar (2002, 2006): Momentum may reflect future economic activity or mispricing of certain macroeconomic variables.
- Behavioral-based explanation, see Zhang (2006): Investors tend to underreact to fundamental news.

Universität Zürich[UZH]

Conjecture and Test Procedure:

- Momentum may reflect future economic activity or mispricing of certain macroeconomic variables.
- Regression of future GDP growth (and other variables) on lagged values of FF factors and one momentum factor.

Results:

- Market factor tends to be a leading indicator of future economic growth.
- Negative relation: small cap or value stocks suffer(thrive) prior to periods of economic growth(slowdown).
- Earnings momentum as a macroeconomic risk factor?
- Rather weak link between future GDP growth (and other variables) and both momentum factors.

Conjecture:

- Theoretical models suggest investors' underreaction to cause the momentum effects, see Hong and Stein (1999).
- If momentum is due to investors' underreaction, momentum should be stronger in more opaque information environments for which information diffusion is slowest.

Test Procedure:

- Analyze winner and loser portfolios limited to different degrees of information uncertainty as measured by:
  - Analyst coverage.
  - Company size.
  - (Total volatility.)
  - Idiosyncratic volatility.

Momentum and Information Uncertainty

Country	*Analyst coverage* Low	Mid	High	*Size* Low	Mid	High	*Idiosyncratic Volatility* Low	Mid	High
Price Momentum									
USA	1.38	1.04	0.87	1.53	0.91	0.48	0.97	0.98	1.41
	5.77	4.15	2.69	6.78	3.24	1.37	3.01	3.79	5.49
Europe	1.66	1.69	1.16	2.10	1.41	1.04	1.47	1.22	1.71
	6.77	7.21	4.83	7.54	5.46	3.87	5.50	5.23	6.74
# max	4	6	1	9	2	0	4	0	7
ranking	1.82	1.45	2.73	1.27	1.82	2.91	1.82	2.82	1.36
Earnings Momentum									
USA	1.04	0.60	0.16	1.08	0.58	0.24	0.50	0.63	0.70
	7.41	3.78	0.75	7.02	3.55	1.22	3.56	4.28	4.26
Europe	0.92	1.00	0.58	1.20	0.71	0.70	0.87	0.89	0.89
	8.42	7.35	3.48	8.09	6.06	4.42	7.62	7.97	5.99
# max	4	6	1	8	3	0	3	3	6
ranking	1.91	1.55	2.55	1.45	1.91	2.64	2.18	2.00	1.73

- U.S./most European countries: price and earnings momentum more pronounced when limited to high information uncertainty stocks.
- Hence, high arbitrage costs (high idiosyncratic risk) may deter investors from exploiting the momentum anomalies.

- Lesmond, Schill, and Zhou (2004) and Korajczyk and Sadka (2004) evidence that exploiting U.S. price momentum is costly.
- The trading costs basically derive from frequent trading in mostly illiquid stocks.
- Consequently, Sadka (2006) documents a close relation between liquidity risk and U.S. momentum strategies.
- Liu (2006)'s liquidity-augmented two-factor asset pricing model almost completely subsumes the U.S. price momentum alpha.
- Hence, we expect liquidity to also play a crucial role in inhibiting profitable execution of the European momentum strategies.

- Arbitrage costs may also be linked to liquidity.
- Liquidity metrics:
  1. Quantity dimension: volume and turnover.
  2. Price impact dimension: Amihud's (2002) *ILLIQ* measure.
  3. Including trading speed as in Liu (2006):

$$\text{Liu Measure} = \ \# \text{ No-Trading Days} + \frac{1/\mathit{Turnover}}{1{,}000{,}000}.$$

- Stocks are sorted into five quintiles based on past returns or earnings revisions.
- For each quintile, the stocks are further sorted into three terciles based on one of the liquidity measures.
- We exclude the six smallest countries from the analysis, i.e., Austria, Finland, Greece, Ireland, Norway, and Portugal.

Country	*Dollar Volume*			*Share Turnover*			*ILLIQ*			*Liu Measure*		
	High	Mid	Low	High	Mid	Low	Low	Mid	High	Low	Mid	High
USA	0.72	0.92	1.15	1.08	0.70	0.76	0.63	0.97	1.29	0.94	0.56	1.34
	2.06	3.42	5.61	3.55	2.80	3.57	1.81	3.56	5.66	3.21	2.08	5.96
Europe	1.23	1.51	1.45	1.63	1.18	1.19	1.20	1.50	1.50	1.41	1.33	1.49
	4.76	5.66	7.08	6.00	5.14	6.02	4.65	5.56	6.85	5.18	5.40	7.47
UK	0.91	1.19	1.29	1.18	1.00	1.19	0.88	1.16	1.33	1.02	1.02	1.29
	3.12	4.23	4.59	4.15	3.66	4.22	3.04	3.98	4.85	3.70	3.60	4.58
Germany	1.14	1.15	1.06	1.21	1.03	0.87	1.08	1.07	1.16	1.07	1.17	1.11
	3.33	3.82	4.29	3.74	3.61	3.79	3.50	3.16	4.17	3.23	4.04	4.51
Switzerland	1.51	0.85	1.15	1.34	0.97	1.17	1.33	1.01	1.18	1.29	1.17	1.19
	4.20	2.75	3.82	3.78	3.29	4.56	3.79	3.07	3.99	3.78	3.67	4.11
France	0.66	1.36	1.22	1.08	1.25	0.95	0.71	1.38	1.14	1.06	1.22	1.06
	1.94	4.52	4.42	3.30	4.25	3.50	2.09	4.37	4.29	3.26	3.98	3.71
Italy	1.39	1.30	0.65	1.16	0.82	0.61	1.16	1.17	0.88	1.34	1.34	0.72
	3.18	3.09	1.65	2.56	2.15	1.57	2.79	2.76	2.29	2.80	3.84	1.53
Spain	0.35	0.33	0.98	0.78	0.54	0.16	0.23	0.34	0.93	0.69	0.08	0.51
	0.87	0.87	1.80	1.87	1.31	0.43	0.54	0.86	2.02	1.72	0.21	1.09
Netherlands	0.67	0.79	1.24	0.75	0.89	1.15	0.73	0.95	0.89	0.80	1.23	0.60
	1.69	2.14	3.70	1.98	2.85	3.43	1.80	2.78	2.87	2.09	3.52	1.97
Sweden	1.02	1.52	0.27	1.47	0.92	-0.18	1.10	0.92	0.40	1.25	0.84	0.58
	2.50	3.25	0.58	3.42	2.26	-0.45	2.63	2.06	0.94	2.94	2.03	1.44
Denmark	1.16	0.95	0.92	1.08	0.76	1.34	1.32	0.97	0.72	1.26	1.18	0.79
	3.57	3.06	2.42	3.12	2.44	3.61	4.14	3.14	2.09	3.60	3.71	1.97
# max	3	4	4	7	1	3	3	4	5	5	4	3
ranking	2.18	1.82	2.00	1.45	2.36	2.18	2.27	1.82	1.82	1.64	2.00	2.09

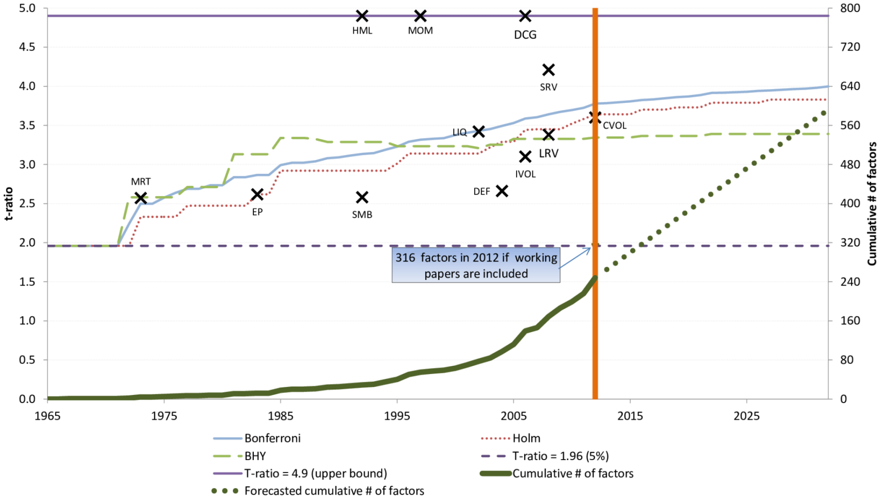
Country	Dollar Volume			Share Turnover			ILLIQ			Liu Measure		
	High	Mid	Low	High	Mid	Low	Low	Mid	High	Low	Mid	High
USA	0.27	0.45	0.97	0.31	0.57	0.79	0.23	0.54	1.01	0.4	0.42	1.02
	1.38	2.62	6.96	1.55	3.66	5.94	1.19	3.16	6.97	2.11	2.58	8.18
Europe	0.78	0.81	0.98	0.91	0.74	0.91	0.81	0.79	0.95	0.89	0.83	0.95
	5.07	6.41	8.63	6.14	5.56	8.59	5.41	6.35	8.04	5.67	6.73	9.68
UK	0.89	0.87	1	1.07	0.65	0.97	0.88	0.98	0.92	0.88	0.78	1.03
	5.11	5.18	5.51	6.38	4.08	5.14	5.42	5.7	5.04	5.38	4.7	5.61
Germany	0.56	0.83	0.98	0.8	0.68	0.91	0.54	0.8	0.82	0.6	0.79	0.8
	2.78	4.76	4.05	3.75	4.07	4.57	2.87	4.52	3.39	3.11	4.76	3.25
Switzerland	0.71	0.48	0.64	0.78	0.43	0.66	0.85	0.23	0.54	0.72	0.47	0.73
	2.42	1.76	2.38	2.59	1.67	2.65	2.95	0.75	2.07	2.41	1.82	2.76
France	0.38	0.93	0.49	0.65	0.78	0.73	0.69	0.74	0.53	0.7	1.01	0.39
	1.32	4.3	2.11	2.48	3.42	3.32	2.69	3.16	2.38	2.58	4.83	1.75
Italy	0.86	0.25	-0.05	0.93	0.07	0.13	0.8	0.21	-0.08	0.62	0.13	0.13
	3.09	0.87	-0.15	2.62	0.25	0.46	3.02	0.76	-0.24	2.02	0.47	0.45
Spain	1	0.88	0.92	1.11	0.78	0.88	0.85	0.89	1.05	0.85	0.9	0.99
	1.96	2.05	2.34	2.37	1.91	2.38	2.06	2.17	2.7	1.88	2.37	2.47
Netherlands	0.4	0.77	1.37	0.42	0.5	1.42	0.56	1.07	0.99	0.57	0.97	1.04
	0.99	2.15	5.22	1.23	1.68	4.89	1.35	3.35	3.69	1.64	2.69	4.03
Sweden	0.39	0.65	0.97	0.64	0.94	0.84	0.54	0.63	1.15	0.59	1.01	0.87
	0.91	1.92	2.6	1.56	2.43	2.4	1.28	1.87	3.12	1.33	3.34	2.33
Denmark	0.97	1.66	2.36	1.15	1.25	2.26	1.25	1.46	1.85	1.43	1.65	1.77
	2.13	3.76	2.58	2.62	2.26	2.75	2.56	2.91	2.84	2.2	4.07	2.47
# max	3	1	7	5	2	5	2	3	6	1	2	8
ranking	2.36	2.18	1.45	2.00	2.36	1.55	2.45	1.91	1.64	2.45	2.09	1.36

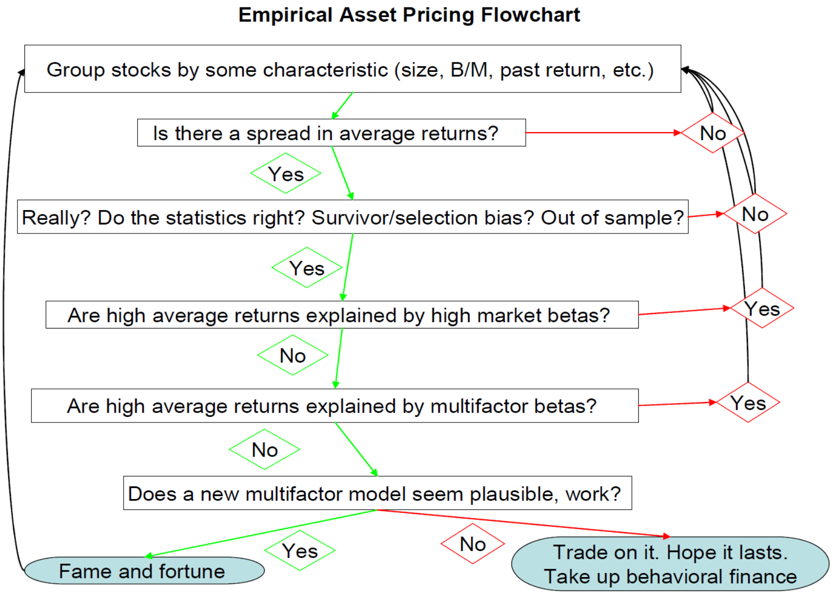
- International momentum effects are robust when controlling for data snooping biases.
- Price momentum is mostly earnings momentum in disguise, with some decoupling in the U.S. after the burst of the tech bubble.
- The evidence of momentum is due to investors' underreaction to fundamental news, not macroeconomic risk.
- Liquidity turns out to be a crucial driver in governing the momentum effects.
- The U.S. momentum effects are clearly most pronounced among illiquid winner and loser stocks.
- However, some European markets exhibit very profitable price momentum strategies even for highly liquid stocks.
- Liquidity appears to be a more severe impediment to implementing earnings momentum strategies as opposed to price momentum strategies.

Solve Problem 3.5 “Fama-French Portfolios” of Campbell (2017) with the software of your choice.

A final thought

If you go factor-fishing, check your t-stats.... Harvey et al. (2016):




Empirical Asset Pricing Flowchart
Group stocks by some characteristic (size, B/M, past return, etc.)
Is there a spread in average returns?
No
Yes
Really? Do the statistics right? Survivor/selection bias? Out of sample?
No
Yes
Are high average returns explained by high market betas?
Yes
No
Are high average returns explained by multifactor betas?
Yes
No
Does a new multifactor model seem plausible, work?
Yes
Fame and fortune
No
Trade on it. Hope it lasts.
Take up behavioral finance

Banz, R. W. et al. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3–18.

Black, F., Jensen, M. C., and Scholes, M. (1972). The capital asset pricing model: Some empirical tests. *in: Michael C. Jensen, Studies in the Theory of Capital Markets, Praeger Publishers Inc.*

Black, F. and Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5):28.

Boguth, O., Carlson, M., Fisher, A., and Simutin, M. (2011). Conditional risk and performance evaluation: Volatility timing, overconditioning, and new estimates of momentum alphas. *Journal of Financial Economics*, 102(2):363–389.

Campbell, J. Y. (2017). *Financial decisions and markets: a course in asset pricing*. Princeton University Press.

Cochrane, J. (2009). *Asset Pricing: (Revised Edition)*. Princeton University Press.

Connor, G. (1984). A unified beta pricing theory. *Journal of Economic Theory*, 34(1-2):13–31.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy*, 81(3):607–636.

References (cont.)

Feng, G., Giglio, S., and Xiu, D. (2019). Taming the factor zoo: A test of new factors. Technical report, National Bureau of Economic Research.

Frost, P. A. and Savarino, J. E. (1986). An empirical Bayes approach to efficient portfolio selection. *Journal of Financial and Quantitative Analysis*, 21(3):293–305.

Gibbons, M. R., Ross, S. A., and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152.

Gow, I. D., Ormazabal, G., and Taylor, D. J. (2010). Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review*, 85(2):483–512.

Graham, B. and Dodd, D. (1934). Securities analysis. *PRINCIPLES AND TECHNIQUES (4th*.

Grinblatt, M. and Titman, S. (1987). The relation between mean-variance efficiency and arbitrage pricing. *Journal of Business*, pages 97–112.

Hansen, L. P. and Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica: Journal of the Econometric Society*, pages 587–613.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). . . . and the cross-section of expected returns. *The Review of Financial Studies*, 29(1):5–68.

Hou, K., Xue, C., and Zhang, L. (2018). Replicating Anomalies.

Huberman, G. (1982). A simple approach to arbitrage pricing theory. *Journal of Economic Theory*, 28(1):183–191.

Ingersoll, J. E. (1984). Some results in the theory of arbitrage pricing. *The Journal of Finance*, 39(4):1021–1039.

Jagannathan, R. and Wang, Z. (1998). An asymptotic theory for estimating beta-pricing models using cross-sectional regression. *The Journal of Finance*, 53(4):1285–1309.

Jobson, J. D. (1982). A multivariate linear regression test for the arbitrage pricing theory. *Journal of Finance*, 37(4):1037–42.

Jobson, J. D. and Korkie, B. (1982). Potential performance and tests of portfolio efficiency. *Journal of Financial Economics*, 10(4):433–466.

Jobson, J. D. and Korkie, B. (1985). Some tests of linear asset pricing with multivariate normality. *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration*, 2(1):114–138.

Jorion, P. (1986). Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3):279–292.

Leippold, M. and Lohre, H. (2014). The dispersion effect in international stock returns. *Journal of Empirical Finance*, 29:331–342.

Lewellen, J. and Nagel, S. (2006). The conditional capm does not explain asset-pricing anomalies. *Journal of Financial Economics*, 82(2):289–314.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4):587–615.

References (cont.)

Miller, M. H. (2000). The history of finance: an eyewitness account. *Journal of Applied Corporate Finance*, 13(2):8–14.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1):435–480.

Pukthuanthong, K., Roll, R., and Subrahmanyam, A. (2018). A Protocol for Factor Identification.

Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341 – 360.

Shanken, J. (1986). Testing portfolio efficiency when the zero-beta rate is unknown: A note. *The Journal of Finance*, 41(1):269–276.

Sharpe, W. F. et al. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442.

Theil, H. (1971). Principles of econometrics. Technical report.

Walters, C. et al. (2014). The Black–Litterman model in detail. Technical report, SSRN Working Paper.