15.472: Cross-Sectional Asset Pricing

Daniel Greenwald

Fall 2020

Cross-Sectional Asset Pricing

- Key research questions:
  1. Why do some stocks have higher returns than others?
  2. What can this tell us about investors' preferences and the risks they face?
- Fundamental equation(s) of finance:

$$E_t \left[M_{t+1} R_{i,t+1} - 1 \right] = 0 \qquad E_t \left[M_{t+1} R^e_{i,t+1} \right] = 0$$

- Unconditional equivalents

$$E \left[\left(M_{t+1} R_{i,t+1} - 1 \right) z_t \right] = 0 \qquad E \left[M_{t+1} R^e_{i,t+1} z_t \right] = 0$$

- Challenge: estimate M_{t+1} as a function of observable factors.

Linear SDF Approach

- Linear specification for SDF: $M_t = b' f_t$.
  - Can drop constant WLOG by redefining $f_t' = (1, \tilde{f}_t')$.
- Linear GMM moment conditions:

$$E\left[\underbrace{Z_t'}_{m\times n}\left(\underbrace{R_{t+1}}_{n\times 1}\underbrace{f_{t+1}'}_{1\times k}\underbrace{b}_{k\times 1} - 1\right)\right] = 0 \qquad E_t\left[\underbrace{Z_t'}_{m\times n}\underbrace{R_{t+1}^e}_{n\times 1}\underbrace{f_{t+1}'}_{1\times k}\underbrace{b}_{k\times 1}\right] = 0$$

- Why not estimate $E_t\left[M_{t+1}x_{t+1} - p_t\right] = 0$?
- Note: for excess return version, need to normalize.

Linear SDF Approach

- Linear specification for SDF: $M_t = b' f_t$.
  - Can drop constant WLOG by redefining $f_t' = (1, \tilde{f}_t')$.
- Linear GMM moment conditions:

$$E\left[\underbrace{Z_t'}_{m\times n}\left(\underbrace{R_{t+1}}_{n\times 1}\underbrace{f_{t+1}'}_{1\times k}\underbrace{b}_{k\times 1} - 1\right)\right] = 0 \qquad E_t\left[\underbrace{Z_t'}_{m\times n}\underbrace{R_{t+1}^e}_{n\times 1}\underbrace{f_{t+1}'}_{1\times k}\underbrace{b}_{k\times 1}\right] = 0$$

- Why not estimate $E_t\left[M_{t+1}x_{t+1} - p_t\right] = 0$? Need GMM data to be stationary.
- Note: for excess return version, need to normalize.

Warm-Up: Single Factor is Excess Return

- Simplest case: single factor f_t which is an excess return, $M_{t+1} = \gamma_0 + \gamma_1 f_{t+1}$.
- Recall: $E\left(R^e_{i,t+1}\right) = -\text{Cov}\left(R^e_{i,t+1}, M_{t+1}\right) E(M_{t+1})^{-1}$
- Now use some algebra and use the fact that f_t is itself an excess return.

$$E\left(R^e_{i,t+1}\right) = -\beta_i \text{Var}(f_{t+1})\gamma_1 E\left(M_{t+1}\right)^{-1}, \qquad \beta_i = \frac{\text{Cov}(R_{i,t+1}, f_{t+1})}{\text{Var}(f_{t+1})}$$

$$E\left(f_{t+1}\right) = -\text{Var}(f_{t+1})\gamma_1 \left(M_{t+1}\right)^{-1}$$

- Putting it all together: $E\left(R^e_{i,t}\right) = \beta_i E\left(f_t\right)$
- Implementation: regress $R^e_{i,t} = \alpha_i + \beta_i f_t + \varepsilon_{i,t}$ and then jointly test $\alpha_i = 0$.

Testing $\alpha = 0$

- Could state as "DM" test:

$$T\left(g_{R,t}(\hat{b}_R)' S_U^{-1} g_{R,t}(\hat{b}_R) - g_{U,t}(\hat{b}_U)' S_U^{-1} g_{U,t}(\hat{b}_U)\right) \xrightarrow{d} \chi^2(\#\text{restrictions})$$

- But can also just do Wald test, which requires only unrestricted estimate

$$Tr(\hat{b}_U)' \left[R(\hat{b}_U)' \hat{V}_U R(\hat{b}_U)\right]^{-1} r(\hat{b}_U) \xrightarrow{d} \chi^2(\#\text{restrictions})$$

where restriction is $r(b) = 0$ and $R(b) = \nabla r(b)$, and $\hat{V} = \text{acov}(\hat{b})$ under efficient GMM.

- In this case:

$$T\alpha' V_{11}^{-1} \alpha \xrightarrow{d} \chi^2(n)$$

where V_{11} is top left block of $\text{acov}(b)$ for $b' = (\alpha', \beta')$, and $n = \#\text{assets}$.

Testing $\alpha = 0$: Special Case

- "Recall" that for OLS with homoskedastic, serially uncorrelated errors:

$$V_{OLS} = E[x_t x_t']^{-1} \otimes E[\varepsilon_t \varepsilon_t']$$

- Here $x_t' = (1, f_t)$, so

$$V_{OLS} = \begin{bmatrix} 1 & E(f_t) \\ E(f_t) & E(f_t^2) \end{bmatrix}^{-1} \otimes \Sigma = \text{Var}(f_t)^{-1} \begin{bmatrix} E(f_t^2) & -E(f_t) \\ -E(f_t) & 1 \end{bmatrix} \otimes \Sigma.$$

- Top left block:

$$V_{11} = \text{Var}(f_t)^{-1} E(f_t^2) \Sigma = \left(1 + \frac{E(f_t)^2}{\text{Var}(f_t)} \right) \Sigma$$

- GMM can easily handle heteroskedasticity and autocorrelation.

General Factor Structure

- General structure: multiple factors, not excess returns. $M_{t+1} = \gamma_0 + \gamma_1' f_{t+1}$.
  - Assume that $\text{Cov}_t(f_{t+1}, f_{t+1}), \text{Cov}_t(f_{t+1}, R_{t+1})$ are constant over time (constant beta).
- Now have

$$E_t(R^e_{t+1}) = -B\text{Cov}(f_{t+1})\gamma_1 R_{f,t} = B\lambda_t \quad (1)$$

$$E\left(R^e_{t+1}\right) = B\lambda \quad (2)$$

where B is the OLS coefficient matrix on $R^e_t = a + Bf_t + \varepsilon_t$.

- Goal: test whether (2) holds while correcting for fact that B is estimated.
  - Note that we are losing information by going from (1) to (2).

When Factor $\neq$ Excess Return

- Need a different approach this time.
  - Before, $E(f_t) = \lambda$ means
$$E(R^e_{i,t}) = \beta_i \lambda = \alpha_i + \beta_i E(f_t) \implies \alpha_i = 0.$$
  - Now, $E(f_t) \neq \lambda$:
$$E(R^e_{i,t}) = B_i \lambda = a_i + B_i E(f_t) \implies R^e_{i,t} = \underbrace{B_i \left(\lambda - E(f_t)\right)}_{a_i} + B_i f_t + \varepsilon_{i,t}$$
    so we need to know λ to test this.
- Previously, were getting k restrictions from theory (definition of excess return).
  - Now, need to estimate λ using at least k new moment conditions.
  - Many possible moments to add, which should we use?

Special Case: I.I.D. Return

- Ideal approach: WWMLD ("what would maximum likelihood do?").
- If returns (errors) are jointly i.i.d. normal:

$$L = \text{const} - \sum_{t=1}^{T} \frac{1}{2} (R_t^e - B\lambda)' S^{-1} (R_t^e - B\lambda)$$

$$\frac{\partial L}{\partial \lambda} = \sum_{t=1}^{T} (R_t^e - B\lambda)' S^{-1} B = 0$$

$$\hat{\lambda}_{ML} = (B' S^{-1} B)^{-1} B' S^{-1} \bar{R}^e$$

- This is the GLS estimator of the regression $\bar{R}^e = B\lambda + \alpha_i$
- Can use our moment condition to target this solution.

Efficient GMM Approach

- Can impose something like this in GMM.
- System of equations:

$$E\begin{bmatrix} R_t^e - a - F_t'\beta \\ F_t\left(R_t^e - a - F_t'\beta\right) \\ R_t^e - \Lambda'\beta \end{bmatrix} = 0$$

where $F_t = (F_t \otimes I_n)$, $\Lambda = (\lambda \otimes I_n)$.

- Connection to MLE? Imagine estimating last moment by itself for known B:

$$g_T = \bar{R}^e - B\lambda \qquad \hat{\lambda} = (B'S^{-1}B)^{-1}B'S^{-1}\bar{R}^e$$

- Note that we still estimate β using OLS. (Why?)

Efficient GMM Approach

- Sample moment condition:

$$g_T = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} R_t^e - a - F_t'\beta \\ F_t\left(R_t^e - a - F_t'\beta\right) \\ R_t^e - \Lambda'\beta \end{bmatrix}$$

where $\bar{R}^e = E_T\left(R_t^e\right)$.

- Derivative matrix for $b' = (a', \beta', \lambda')$:

$$d = -E\begin{bmatrix} I & F_t & 0 \\ F_t & F_tF_t' & 0 \\ 0 & \Lambda' & B \end{bmatrix} = -E\left(\begin{bmatrix} 1 & f_t \\ f_t & f_tf_t' \\ 0 & \lambda' \end{bmatrix} \otimes I_n, \quad \begin{bmatrix} 0 \\ 0 \\ B \end{bmatrix}\right)$$

- Sample equivalent:

$$d_T = -\frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} I & F_t & 0 \\ F_t & F_tF_t' & 0 \\ 0 & \Lambda' & B \end{bmatrix}$$

Three-Pass Regression

- Two-pass regression recovers λ values if all factors are included, but can be biased (in both stages) if factors are omitted.
  - Giglio and Xiu (2019): use PCA to span common sources of variation in returns.
- Assume that you want to price a factor g_t and you observe a vector of returns r_t with

$$
\begin{aligned}
r_t &= \beta\gamma + \beta v_t + u_t \\
g_t &= \delta + \eta v_t + z_t
\end{aligned}
$$

- **Pass 1:** Compute first p PCs of r_t. Denote components $\hat{v}_t$, loadings as $\hat{\beta}$.
- **Pass 2:** Regress average returns $\bar{r}$ on $\hat{\beta}$ to obtain risk prices $\hat{\gamma}$.
- **Pass 3:** Regress g_t on $\hat{v}_t$ and compute expected return as $\hat{\gamma}_g = \hat{\eta}\hat{\gamma}$.

Fama-MacBeth

- Historically important procedure useful for understanding GMM estimate.

  1. Estimate betas using

  $$R^e_{i,t} = a_i + \beta_i' f_t + \varepsilon_{i,t}$$

  2. For each t, estimate λ_t using cross-sectional estimate

  $$R^e_{i,t} = \lambda_t' \beta_i + \alpha_{i,t}$$

  3. Estimate $\hat{\lambda}$, $\hat{\alpha}$, and asymptotic covariances using

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T} \hat{\lambda}_t \qquad \hat{\alpha} = \frac{1}{T}\sum_{t=1}^{T} \hat{\alpha}_t$$

$$V(\hat{\lambda}) = \frac{1}{T}\sum_{t=1}^{T} (\hat{\lambda}_t - \hat{\lambda})^2 \qquad V(\hat{\alpha}) = \frac{1}{T}\sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})^2$$

Fama-MacBeth

- Totally different approach (regress for fixed t then average). But delivers similar result because β_i terms are constant across time.
- Stacking $R_t^e = B\lambda + \alpha_t$ implies $\hat{\lambda}_t = (B'B)^{-1}B'R_t^e$.
- Sample expectation of this object:

$$E_T(\hat{\lambda}_T) = (B'B)^{-1}B'\bar{R}^e$$

  identical to cross-sectional OLS estimator on averaged data: $\bar{R}^e = B\lambda + \bar{\alpha}$.
- Sample covariance assuming α_t independent across time:

$$\begin{aligned} \text{Cov}_T(\hat{\lambda}_t) &= (B'B)^{-1}B'\text{Cov}_T(R_t^e)B(B'B)^{-1} \\ &= (B'B)^{-1}B'\text{Cov}_T(\hat{\alpha}_t)B(B'B)^{-1} \\ &= T^{-1}(B'B)^{-1}B'\text{Cov}_T(\bar{\alpha})B(B'B)^{-1} \end{aligned}$$

  which is averaged OLS, corrected for X-Eqn corr. (no serial corr., known, not estimated B).

Time-Varying SDF

- Specification $M_{t+1} = a + b' f_{t+1}$ implies that risk premia and risk free rates should be constant over time. If they aren't, this can lead to poor performance even with correct factors.
- Instead, could use $M_{t+1} = a_t + b_t' f_{t+1}$. Unrestricted problem hard to estimate.
- More parsimonious approach:

$$a_t = \gamma_0 + \gamma_1 z_t$$
$$b_t = \eta_0 + \eta_1 z_t.$$

- Write in factor form using

$$\mathbf{f}_{t+1} = \begin{bmatrix} 1 \\ z_t \\ f_{t+1} \\ z_t f_{t+1} \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \eta_0 \\ \eta_1 \end{bmatrix}$$

so that $M_{t+1} = \mathbf{b}' \mathbf{f}_{t+1}$. Now use existing tools.

Lettau and Ludvigson (2001)

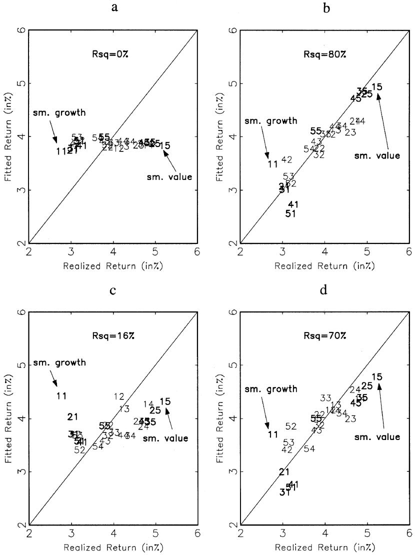
- Use $f_{t+1} = \Delta c_{t+1}$ as in traditional C-CAPM.
- But also use $z_t = cay_t$.
  - This is the residual from a cointegrating relationship inspired by the budget constraint.
  - Good empirical predictor of stock returns.
- Estimates equivalent to two stage procedure

$$R^e_{i,t+1} = a_i + \beta_{i,z} z_t + \beta_{i,f} f_{t+1} + \beta_{i,f,z} z_t f_{t+1} + \varepsilon^i_{t+1}$$
$$E[R^e_{i,t+1}] = \beta_{i,z}\lambda_z + \beta_{i,f}\lambda_f + \beta_{i,f,z}\lambda_{f,z}.$$

  allowing for testing of the z-specific parameters.
- LL find strong explanatory power, rivaling Fama-French when labor income included as an additional factor.

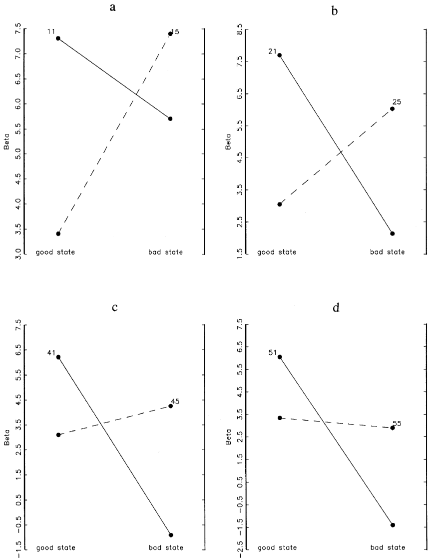
Lettau and Ludvigson (2001)

- Figures:
  a. CAPM.
  b. Fama-French
  c. Consumption CAPM
  d. Scaled Consumption CAPM



Lettau and Ludvigson (2001)

- Good state: high *cay* (low risk premia).
- Intuition: different portfolios can have same average betas, but what matters is if β is high when risk premia (λ_t) are high.
- Question: what is cay_t and why does it proxy for risk premia?



Derivation of *cay*

- Complete markets rep. agent economy.
- Denote W_t as aggregate wealth (human capital plus asset holdings), C_t as consumption, and $R_{w,t+1}$ as net return on aggregate wealth.
- Accumulation equation for aggregate wealth:

$$W_{t+1} = R_{w,t+1}(W_t - C_t).$$

- Rearranging the budget constraint and taking log-linear approximation:

$$\Delta w_{t+1} = k + r_{w,t+1} + (1 - \rho^{-1})(c_t - w_t).$$

where lowercase letters denote log variables, $\rho = (W - C)/W$.

Tool: Lag Polynomial

- Lag operator L defined by $L^k x_t = x_{t-k}$.
- Geometric sum formula: $\left(\sum_{j=0}^{\infty} \rho^j\right) x = (1-\rho)^{-1} x$
- Lag polynomial versions:

$$\sum_{j=0}^{\infty} \rho^j x_{t-j} = \sum_{j=0}^{\infty} \rho^j L^j x_t = (1-\rho L)^{-1} x_t, \qquad \sum_{j=0}^{\infty} \rho^j x_{t+j} = \sum_{j=0}^{\infty} \rho^j L^{-j} x_t = \left(1-\rho L^{-1}\right)^{-1} x_t$$

- Denote $cw_t = c_t - w_t$. Then:

$$cw_t - \rho cw_{t+1} = \left(1-\rho L^{-1}\right) wc_t = \rho\left(k + r_{w,t+1} - \Delta c_{t+1}\right)$$

Log-Linear Approximation

- Solving forward and imposing the transversality condition $\lim_{k \to \infty} \rho^k (c_{t+k} - w_{t+k}) = 0$:

$$c_t - w_t = \text{const} + \sum_{j=1}^{\infty} \rho^j (r_{w,t+1} - \Delta c_{t+1}).$$

- This is an ex post relation, but it most also hold ex ante:

$$c_t - w_t = \text{const} + E_t \sum_{j=1}^{\infty} \rho^j (r_{w,t+1} - \Delta c_{t+1}).$$

- Conclusion: wealth-consumption ratio should contain predictable information on future consumption growth and wealth returns.

Further Approximations

- Challenge #1: can't observe human capital component of wealth.

  1. Take log-linear approximation

  $$w_t \simeq \omega a_t + (1-\omega) h_t$$
  $$r_{w,t} \simeq \omega r_{a,t} + (1-\omega) r_{h,t}$$

  2. Assume that

  $$h_t = \kappa + y_t + z_t$$

  where y_t is labor income, and z_t is stationary with mean zero.

- Challenge #2: can't observe service flows from durables.
  - Approach: assume that total consumption proportional to nondurables/services: $c_t = \lambda c_{n,t}$.

Putting it All Together

- Putting it all together

$$\lambda c_{n,t} - \omega a_t - (1-\omega) y_t = E_t \sum_{i=1}^{\infty} \rho^i \left\{ [\omega r_{a,t+i} + (1-\omega) r_{h,t+i}] - \Delta c_{t+i} \right\} + (1-\omega) z_t.$$

- Scale the LHS to define

$$cay_t \equiv \text{const} + c_{n,t} - \beta_a a_t - \beta_y y_t$$

  where $\beta_a = \omega / \lambda$, $\beta_y = (1-\omega)/\lambda$.

- Note that cay_t is stationary, even though (c, a, y) all appear to contain unit roots.
  - Estimate using cointegration.

Estimation of Cointegration Parameters

- Estimate β_a, β_y using the dynamic least squares (DLS) method of Stock and Watson (1993).
- DLS applied to this model specifies a single OLS regression equation

$$c_{n,t} = \alpha + \beta_a a_t + \beta_y y_t + \sum_{i=-k}^{k} b_{a,i} \Delta a_{t-i} + \sum_{i=-k}^{k} b_{y,i} \Delta y_{t-i} + \epsilon_t \tag{3}$$

- Point estimates are: $c_{n,t} = 0.61 + 0.31 a_t + 0.59 y_t$
- Adjusting for $\lambda = 1.1$ implies $\sim 2/3$ of wealth in human capital.

Dynamic Least Squares

- To see why (3) works, define $x_t = (a_t, y_t)$ and note that $(c_{n,t}, a_t, y_t)$ being individually I(1) and cointegrated implies the triangular representation

$$\Delta x_t = \mu_1 + u_t^1 \quad (4)$$
$$c_{n,t} = \mu_2 + \beta' x_t + u_t^2. \quad (5)$$

- The obstacle is that u_t^2 and x_t may be correlated. To orthogonalize them, project u_t^2 onto $\{u_t^1\}$ and use (4) to obtain

$$E[u_t^2 | \{u_t^1\}] = E[u_t^2 | \{\Delta x_t\}] = \mu_u + d(L)\Delta x_t$$

  where $d(L)$ is an unknown two-sided lag polynomial.
- Substituting into (5) now yields

$$c_{n,t} = \mu + \beta' x_t + d(L)\Delta x_t + v_t^2$$

  where $v_t^2 \perp x_t$.

DLS In Practice

- To apply the DLS estimator, assume $d(L) = \sum_{i=-k}^{k} d_i L^i$. LL use $k = 8$.
- Stock (1987) establishes that parameter estimates are superconsistent, in that $T(\beta - \hat{\beta}) \xrightarrow{p} 0$ instead of the usual $\sqrt{T}(\beta - \hat{\beta}) \xrightarrow{p} 0$.
- Intuition: sharp disparity between stationary (finite cov) and nonstationary (infinite cov) distributions allows for faster convergance.
- Superconsistency allows us to use the estimated $\widehat{cay}_t$ as if it were the true cay_t (i.e. no adjustment for generated regressors).

Lewellan and Nagel (2006)

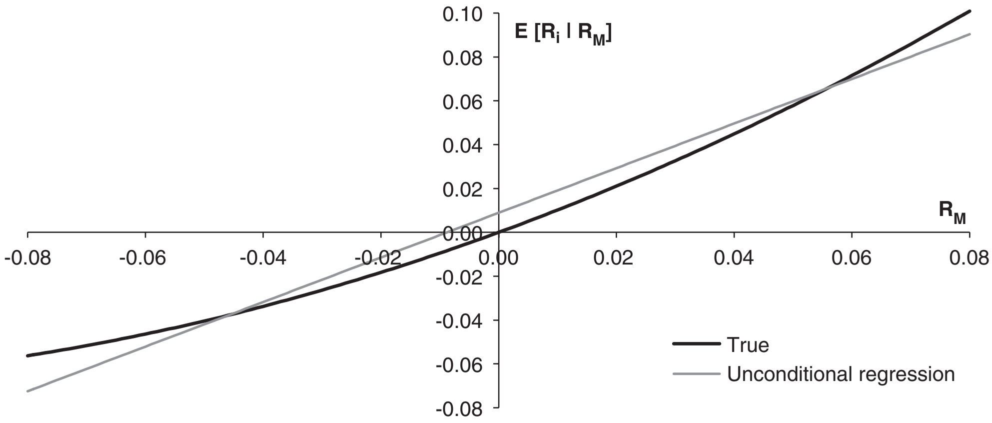
- Many conditional CAPM papers seek to reproduce return properties of Fama-French portfolios using time-varying SDFs and a single traditional factor ($R_{m,t}$ or Δc_t).
- LN argue that this approach cannot explain observed asset pricing "anomalies."
- Two-part argument:
  1. Existing studies ignore theoretical relations when freely estimating λ.
  2. Directly estimating conditional CAPM yields poor performance.

Lewellan and Nagel (2006)

- Goal: see if reasonable data generating processes can produce large unconditional alphas observed on some portfolios:
$$\alpha_i^u = E\left(R_{i,t+1}^e\right) - \beta_i^u \lambda$$
- Conditional relation for single factor (market excess return):
$$E_t\left(R_{i,t+1}^e\right) = \beta_{i,t}\lambda_t \qquad\qquad \lambda_t = E_t(R_{m,t+1}^e)$$
- Taking unconditional expectations (defining $\beta_i \equiv E\left(\beta_{i,t}\right)$, $\lambda \equiv E\left(\lambda_t\right)$):
$$E\left(R_{i,t+1}^e\right) = \beta_i\lambda + \text{Cov}(\beta_{i,t}, \lambda_t)$$
- Rewrite unconditional alpha as
$$\alpha^u = \lambda(\beta_i - \beta^u) + \text{Cov}(\beta_{i,t}, \lambda_t)$$
where β^u (from unconditional regression) is not necessarily the same as β!

Unconditional β: Intuition

- Example: β_t and λ_t are positively correlated.



Unconditional Beta of a Stock

- Assume CAPM holds, so that: $R^e_{i,t+1} = \beta_{i,t} R^e_{m,t+1} + \varepsilon_{i,t+1}$.
- Define $\sigma^2_{m,t} \equiv \text{Var}_t\left(R^e_{m,t}\right)$, $\sigma^2_m \equiv \text{Var}\left(R^e_{m,t}\right)$, and also define $\eta_{i,t} \equiv \beta_{i,t} - \beta_i$. Then:

$$\begin{aligned}
\text{Cov}(R^e_{i,t+1}, R^e_{m,t+1}) &= \text{Cov}\left(\beta_{i,t} R^e_{m,t+1}, R^e_{m,t+1}\right) \\
&= \beta_i \sigma^2_m + E\left[\eta_{i,t}\left(R^e_{m,t+1}\right)^2\right] - E\left(\eta_{i,t} R^e_{m,t+1}\right) E(R^e_{m,t+1}) \\
&= \beta_i \sigma^2_m + E\left[\eta_{i,t}\left(\lambda_t^2 + \sigma^2_{m,t}\right)\right] - \lambda E\left(\eta_{i,t}\lambda_t\right) \\
&= \beta_i \sigma^2_m + \text{Cov}\left(\beta_{i,t}, \lambda_t^2\right) + \text{Cov}\left(\beta_{i,t}, \sigma^2_{m,t}\right) - \lambda \text{Cov}\left(\beta_{i,t}, \lambda_t\right) \\
&= \beta_i \sigma^2_m + \text{Cov}\left(\beta_{i,t}, (\lambda_t - \lambda)^2\right) + \text{Cov}\left(\beta_{i,t}, \sigma^2_{m,t}\right) + \lambda \text{Cov}\left(\beta_{i,t}, \lambda_t\right)
\end{aligned}$$

- Unconditional beta:

$$\beta_i^u = \beta_i + \sigma_m^{-2}\left[\text{Cov}\left(\beta_{i,t}, (\lambda_t - \lambda)^2\right) + \text{Cov}\left(\beta_{i,t}, \sigma^2_{m,t}\right) - \lambda \text{Cov}\left(\beta_{i,t}, \lambda_t\right)\right]$$

Unconditional Beta of a Stock

- Putting it all together:

$$\alpha_i^u = \left(1 - \lambda^2 \sigma_m^{-2}\right) \text{Cov}\left(\beta_{i,t}, \lambda_t\right) - \lambda \sigma_m^{-2} \text{Cov}\left(\beta_{i,t}, (\lambda_t - \lambda)^2\right) - \lambda \sigma_m^{-2} \text{Cov}\left(\beta_{i,t}, \sigma_{m,t}^2\right)$$

- Removing quantitatively small terms λ^2 / σ_m^2 and $\text{Cov}\left(\beta_{i,t}, (\lambda_t - \lambda)^2\right)$ yields

$$\alpha_i^u \simeq \text{Cov}\left(\beta_{i,t}, \lambda_t\right) - \lambda \sigma_m^{-2} \text{Cov}\left(\beta_{i,t}, \sigma_{m,t}^2\right)$$

- Let's look for an upper bound. Ignore second term for now, so that

$$\alpha_i^u \simeq \text{Cov}(\beta_{i,t} \lambda_t) = \rho \sigma_\beta \sigma_\lambda$$

- Large alphas require extremely volatile betas. Do these show up in the data?

Estimating Conditional Betas

- Conditional CAPM approaches generate $\beta_{i,t}$ series but depend on correctly specified model.
- LN's approach: directly estimate $\beta_{i,t}$ using high-frequency data.
- Key idea: assume $\beta_{i,t}$ is stable within e.g., one quarter: $\beta_{i,t} = \beta_{i,q}$. Then run daily regression

$$R^e_{i,t} = \alpha_{i,q} + \beta_{i,q}(L) R^e_{m,t} + \varepsilon_{i,t}$$

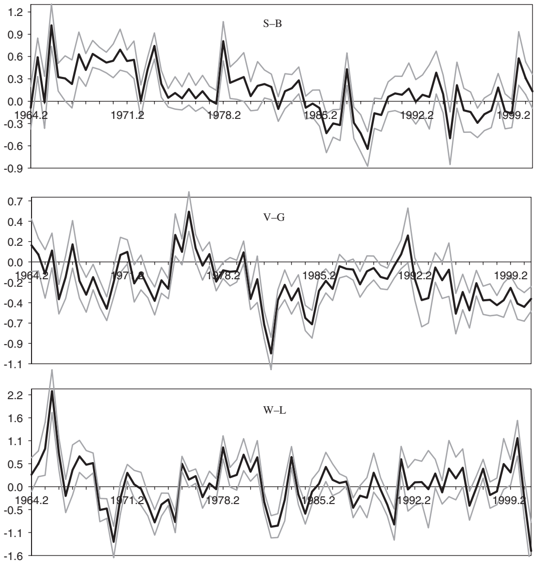
- Lags are useful for allowing some stocks (esp. small stocks) to have delayed reaction to market return. Approach follows Dimson (1979)

$$R^e_{i,t} = \alpha_{i,q} + \beta_{i,q,0} R^e_{m,t} + \beta_{i,q,1} R^e_{m,t-1} + \beta_{i,q,2} \left[\left(R^e_{m,t-2} + R^e_{m,t-3} + R^e_{m,t-4} \right) / 3 \right] + \varepsilon_{i,t}$$

- If conditional CAPM is correct, then conditional alphas should be close to zero.
  - Also produce estimates of $\beta_{i,q}$ that can be used to evaluate theory.

Conditional Betas

- Betas do move around over time.
- Vary systematically with relevant state variables (risk-free rate, dividend yield, term spread, etc.).
- But not enough to overturn anomalies.
- Conditional alphas large and close to unconditional versions.



Implied Alphas

- Examples: book-market portfolio earns 0.59% monthly on $\sigma_\beta = 0.25$, momentum portfolio earns 1.01% monthly on $\sigma_\beta = 0.60$.

	σ_β			σ_β		
σ_γ	0.3	0.5	0.7	0.3	0.5	0.7
	$\rho =$ **0.6**			$\rho =$ **1.0**		
0.1	0.02	0.03	0.04	0.03	0.05	0.07
0.2	0.04	0.06	0.08	0.06	0.10	0.14
0.3	0.05	0.09	0.12	0.09	0.15	0.21
0.4	0.07	0.12	0.17	0.12	0.20	0.28
0.5	0.09	0.15	0.21	0.15	0.25	0.35

What About C-CAPM?

- Don't have high frequency consumption data, so hard to estimate conditional betas directly.
- But LL theory implies that

$$R^e_{i,t+1} = \underbrace{a_i + \beta_{i,z} z_t}_{a_{i,t}} + \underbrace{\left(\beta_{i,f} + \beta_{i,f,z} z_t\right)}_{\beta_{i,t}} f_{t+1}$$

$$E[R^e_{i,t}] = \beta_i \lambda + \mathrm{Cov}(\beta_{i,t}, \lambda_t) = \beta_i \lambda + \beta_{i,f,z} \mathrm{Cov}(z_t, \lambda_t) = \beta_i \lambda + \beta_{i,f,z} \cdot \rho_{z,\lambda} \sigma_z \sigma_\lambda$$

- LL implies $\mathrm{Cov}(z_t, \lambda_t) \simeq 0.07\%$. Since $\sigma_z \simeq 0.019$, so $\sigma_\lambda \geq 3.2\%$ quarterly.
  - Average λ is small (-0.02% to 0.22% quarterly), need highly volatile (and skewed) price of risk.
- So what's the point? Does it matter if cay_t is factor or scaling variable?
- General warning: be careful explaining portfolios with strong factor structure.

Pitfalls of Cross-Sectional Asset Pricing Research

- Typical approach: run XSAP regs, declare victory if p-value on long-short return < 0.05.
- Many problems with this approach (Harvey, 2017).
  - Many possible factors, unsuccessful ones not reported (publication bias).
  - Many possible specifications for each factor (p-hacking).
  - Base rate $p(H)$ is very important for $p(H|\text{data})$. Very low base rate implies many false positives.
- How can you avoid this trap?
  - Do not consider any $t < 3$ to be strong unilateral evidence (Harvey, Liu, Zhu, 2016 RFS).
  - Use **Minimum Bayes Factor** (Harvey, 2017). Weighs prior on null against strongest possible Bayesian evidence against the null (taken over all priors on alternative hypothesis).
  - For large n tests (e.g., alphas) False Discovery Rate control (Benjamini, Hochberg, 1995; Giglio, Liao, Xiu 2020)
  - Bring theory and other supporting evidence to bear.

Recap: Cross-Sectional Asset Pricing

- Framework based on beta representations implied by theory.
- Estimating risk premia/risk prices uses generated regressors, can easily perform inference using GMM.
  - Fama-MacBeth is special case not correcting for generated regressors.
- Adding additional variables helps, but need to use theory to determine if these are factors or changes in risk prices.
- Tools:
  1. Cointegration/dynamic least squares
  2. Lag polynomial