Empirical asset pricing:
The Stochastic Discount Factor approach

Joachim Grammig
University of Tübingen
Department of Econometrics, Statistics and Empirical Economics
Course outline

**Empirical asset pricing: The Stochastic Discount Factor approach**

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Theoretical Background</td>
<td>3</td>
</tr>
<tr>
<td>2. Stochastic Discount Factors and GMM estimation</td>
<td>20</td>
</tr>
<tr>
<td>3. Recent models</td>
<td>62</td>
</tr>
<tr>
<td>4. Testing conditional predictions of asset pricing models:</td>
<td></td>
</tr>
<tr>
<td>Managed portfolios and scaled factors</td>
<td>82</td>
</tr>
<tr>
<td>5. Linear factor models and the basic pricing equation</td>
<td>106</td>
</tr>
</tbody>
</table>
1. Theoretical background

Readings:
Cochrane (2005), Chapters 1 (without 1.5), 3 (3.1 and 3.2), 4 (4.1 and 4.2)
Empirical asset pricing - Introduction (1)

Asset pricing (Valuation of financial assets)

- delay of payoff
- account for risk of payoff

⇒ risk correction

50 years US stocks: 9% average return (real) p.a.
1% real interest rate p.a. (treasury bills)

8% premium earned for holding risk
What is the risk that is priced?

Asset pricing

- normative
- positive

how should the world work?
how does the world work?

- trading opportunities?
- cost of capital
- non traded assets: "fair" price

Prof. Joachim Grammig, University of Tübingen, Department of Econometrics, Statistics and Empirical Economics
Empirical asset pricing - Introduction (2)

Basic: Prices equal discounted expected payoff

What probability measure?

Absolute Asset Pricing

exposure to "fundamental" macroeconomic risk

Asset priced given other asset prices (e.g. option pricing)

Relative Asset Pricing

e.g. CAPM:

$$E(R^i) = R^f + \beta_i \left( E(R^m) - R^f \right)$$

$$\beta_i = \frac{cov(R^i, R^m)}{var(R^m)}$$

Market price of risk (factor) risk premium not explained
Empirical asset pricing - Introduction (3)

Basic pricing equation

\[ p_t = \mathbb{E}_t(m_{t+1} x_{t+1}) \]

- Asset price at time \( t \)
- Stochastic discount factor (r.v.)
- Payoff (r.v.)

\[ m_{t+1} = f(\text{data}, \text{parameters}) \]

The model

Moment condition:

\[ \mathbb{E}_t(m_{t+1} x_{t+1}) - p_t = 0 \]

Use \( \frac{1}{n} \sum \rightarrow \mathbb{E}(\cdot) \) WLLN

Generalized Method of Moments (GMM) to estimate parameters
Empirical asset pricing - Introduction (4)

- Portfolio theory
- Mean-Variance frontier
- CAPM
- APT
- Option pricing
- Contingent claims state preference
- Consumption-based model
- Stochastic discount factor

Time line of discovery traditional

Cochrane's approach
From an utility maximising investor's first order conditions we obtain the basic asset pricing formula (1)

Basic objective: find $p_t$, the present value of stream of uncertain payoff $x_{t+1}$

$$x_{t+1} = p_{t+1} + d_{t+1}$$

Utility function

$$U(c_t, c_{t+1}) = u(c_t) + \beta \mathbb{E}_t [u(c_{t+1})]$$

Random variables: $p_{t+1}, d_{t+1}, x_{t+1}, e_{t+1}, c_{t+1}, u(c_{t+1})$, $\mathbb{E}_t [\cdot] \triangleq \mathbb{E} [\cdot | \mathcal{F}_t]$
From an utility maximising investor’s first order conditions we obtain the basic asset pricing formula (2)

\[
\max_{\xi} \left[ U(c_t, c_{t+1}) \right] \text{ s.t.} \\
\begin{align*}
    c_t &= e_t - p_t \xi; \quad c_{t+1} = e_{t+1} + x_{t+1} \xi \\
\end{align*}
\]

\[
\max_{\xi} \left\{ u(e_t - p_t \xi) + \beta \mathbb{E}_t [u(e_{t+1} + x_{t+1} \xi)] \right\}
\]

\[
-p_t \cdot u'(c_t) + \beta \cdot \mathbb{E}_t [u'(c_{t+1}) \cdot x_{t+1}] = 0
\]

utility loss if investor buys another unit of the asset

discounted expected utility increase from extra payoff

\[
p_t \cdot u'(c_t) = \mathbb{E}_t \left[ \beta u'(c_{t+1}) x_{t+1} \right]
\]

\[
p_t = \mathbb{E}_t \left[ \frac{\beta u'(c_{t+1})}{u'(c_t)} x_{t+1} \right]
\]

No complete solution: endogenous variables

Investor continues to buy or sell the asset until marginal loss equals marginal gain.
Turning off uncertainty we are in the standard two-goods case (1)

$$\max [u(c_t) + \beta u(c_{t+1})] \text{ s.t. } c_t = e_t - p_t \cdot \xi, c_{t+1} = e_{t+1} + x_{t+1} \cdot \xi$$

$$\frac{\partial U(c_t, c_{t+1})}{\partial \xi} = -p_t \cdot \frac{\partial u(c_t)}{\partial c_t} + \beta \cdot x_{t+1} \cdot \frac{\partial u(c_{t+1})}{\partial c_{t+1}} = 0$$

$$p_t \cdot u'(c_t) = x_{t+1} \cdot \beta u'(c_{t+1})$$

$$p_t = x_{t+1} \cdot \frac{\beta u'(c_{t+1})}{u'(c_t)}$$

$$-\frac{dc_t}{dc_{t+1}} = \frac{\beta \cdot u'(c_{t+1})}{u'(c_t)} = \frac{p_t}{x_{t+1}}$$

Opportunity cost to transfer consumption from $t$ to $t+1$

Marginal valuation of consumption in $t+1$ in terms of consumption in $t$

$$p_t u'(c_t) = \mathbb{E}_t \left[ \beta u'(c_{t+1}) x_{t+1} \right]$$

$$p_t = \mathbb{E}_t \left[ \frac{\beta u'(c_{t+1})}{u'(c_t)} x_{t+1} \right]$$
We often use a convenient power utility function (1)

\[ u(c_t) = \frac{1}{1 - \gamma} c_t^{1-\gamma} \]

\[ \lim_{\gamma \to 1} \left( \frac{1}{1 - \gamma} c_t^{1-\gamma} \right) = \ln(c_t) \]

\[ u'(c_t) = c_t^{-\gamma} \]

\[ \frac{dc_t}{dc_{t+1}} = \frac{\beta u'(c_{t+1})}{u'(c_t)} = \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \]

utility \( u(c_t) \)

parameter \( \gamma \):

0.3 0.5 0.8

increasing concavity of utility function
Prices, payoffs, excess returns

<table>
<thead>
<tr>
<th></th>
<th>Price $p_t$</th>
<th>Payoff $x_{t+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>stock</td>
<td>$p_t$</td>
<td>$p_{t+1} + d_{t+1}$</td>
</tr>
<tr>
<td>return</td>
<td>1</td>
<td>$R_{t+1}$</td>
</tr>
<tr>
<td>excess return</td>
<td>0</td>
<td>$R^e_{t+1} = R^a_{t+1} - R^b_{t+1}$</td>
</tr>
<tr>
<td>one $$ one period discount bond</td>
<td>$p_t$</td>
<td>1</td>
</tr>
<tr>
<td>risk-free rate</td>
<td>1</td>
<td>$R^f$</td>
</tr>
</tbody>
</table>

Payoff $x_{t+1}$ divided by price $p_t$ ⇒ gross return $R_{t+1} = \frac{x_{t+1}}{p_t}$

Return: payoff with price one

$$1 = \mathbb{E}_t (m_{t+1} \cdot R_{t+1})$$

Zero-cost portfolio:
Short selling one stock, investing proceeds in another stock
⇒ excess return $R^e$

Example: Borrow 1$ at $R^f$, invest it in risky asset with return $R$.
Pay no money out of the pocket today → get payoff $R^e = R - R^f$.

Zero price does not imply zero payoff.
The covariance of the payoff with the discount factor rather than its variance determines the risk-adjustment

\[ cov(m_{t+1}, x_{t+1}) = \mathbb{E}(m_{t+1} \cdot x_{t+1}) - \mathbb{E}(m_{t+1}) \mathbb{E}(x_{t+1}) \]

\[ p_t = \frac{\mathbb{E}(m_{t+1}) \mathbb{E}(x_{t+1})}{R_f} + cov(m_{t+1}, x_{t+1}) \]

\[ R_f = \frac{1}{\mathbb{E}(m_{t+1})} \]

Marginal utility declines as consumption rises.

Price is lowered if payoff covaries positively with consumption. (makes consumption stream more volatile)

Price is increased if payoff covaries negatively with consumption. (smoothens consumption) Insurance!

Investor does not care about volatility of an individual asset, if he can keep a steady consumption.
All assets have an expected return equal to the risk-free rate, plus risk adjustment:

\[
1 = \mathbb{E} \left( m_{t+1} \cdot R_{t+1}^i \right)
\]

\[
1 = \mathbb{E} (m_{t+1}) \mathbb{E} \left( R_{t+1}^i \right) + \text{cov} \left( m_{t+1}, R_{t+1}^i \right)
\]

\[
R^f = \frac{1}{\mathbb{E} (m_{t+1})}; \quad 1 - \frac{1}{R^f} \mathbb{E} \left( R_{t+1}^i \right) = \text{cov} \left( m_{t+1}, R_{t+1}^i \right)
\]

\[
\mathbb{E} \left( R_{t+1}^i \right) - R^f = -R^f \cdot \text{cov} \left( m_{t+1}, R_{t+1}^i \right)
\]

\[
\mathbb{E} \left( R_{t+1}^i \right) - R^f = -\frac{1}{\mathbb{E} \left( \beta u'(c_{t+1}) u'(c_t) \right)} \cdot \text{cov} \left( \beta u'(c_{t+1}) u'(c_t), R_{t+1}^i \right)
\]

Excess return:

\[
\mathbb{E} \left( R_{t+1}^i \right) - R^f = -\frac{\text{cov} \left( u'(c_{t+1}), R_{t+1}^i \right)}{\mathbb{E} \left( u'(c_{t+1}) \right)}
\]

Investors demand higher excess returns for assets that covary positively with consumption. Investors may accept expected returns below the risk-free rate. Insurance!

---

Prof. Joachim Grammig, University of Tübingen, Department of Econometrics, Statistics and Empirical Economics
The basic pricing equation has an expected return-beta representation

\[
\mathbb{E}\left(R_{t+1}^i\right) - R^f = -R^f \cdot \text{cov}\left(R_{t+1}^i, m_{t+1}\right)
\]

\[
\mathbb{E}\left(R_{t+1}^i\right) - R^f = -\frac{\text{cov}\left(R_{t+1}^i, m_{t+1}\right)}{\text{Var}(m_{t+1})} \cdot \frac{\text{Var}(m_{t+1})}{\mathbb{E}(m_{t+1})}
\]

\[
\mathbb{E}\left(R_{t+1}^i\right) = R^f - \left(\frac{\text{cov}\left(R_{t+1}^i, m_{t+1}\right)}{\text{Var}(m_{t+1})}\right) \cdot \left(\frac{\text{Var}(m_{t+1})}{\mathbb{E}(m_{t+1})}\right)
\]

asset specific quantity of risk \[\uparrow\] \[\downarrow\] price of risk for all assets

Beta-pricing model:

\[
\mathbb{E}\left(R^i\right) = R^f + \beta_{R^i,m} \cdot \lambda_m
\]

With \( m = \beta \left(\frac{c_{t+1}}{c_t}\right)^{-\gamma} \) and lognormal consumption growth \( \frac{c_{t+1}}{c_t} \)

\[
\mathbb{E}\left(R^i\right) = R^f + \beta_{R^i,\Delta c} \cdot \lambda_{\Delta c}
\]

\[\lambda_{\Delta c} \approx \gamma \cdot \text{Var}\left(\Delta \ln c\right)\]

The more risk averse the investors or the riskier the environment, the larger the expected return premium for risky (high-beta) assets.
Marginal utility weighted prices follow martingales (1)

Basic first order condition:

\[ p_t u'(c_t) = \mathbb{E}_t \left( \beta \left( u'(c_{t+1}) \right) \left( p_{t+1} + d_t \right) \right) \]

Market efficiency ⇔ Prices follow martingales (random walks)? NO!

Risk neutral investors \( u'(\cdot) = \text{const.} \)
or no variation in consumption

\[ \beta = 1 \leftarrow \text{OK short time horizon} \]

no dividends

Then:

\[ p_t = \mathbb{E}(p_{t+1}) \]

\[ p_{t+1} = p_t + \varepsilon_{t+1} \]

if

\[ \sigma^2(\varepsilon_{t+1}) = \sigma^2 = \text{Random Walk} \]

\[ \Rightarrow \text{Returns are not predictable} \quad \mathbb{E} \left( \frac{p_{t+1}}{p_t} \right) = 1 \]
Marginal utility weighted prices follow martingales (2)

With risk aversion (but no dividends) and $\beta=1$

$$\tilde{p}_t = \mathbb{E}(\tilde{p}_{t+1})$$

$$\tilde{p}_t = \tilde{p}_t \cdot u'(c_t)$$

Scale prices by marginal utility, correct for dividends and apply risk neutral valuation formulas.

Predictability in the short horizon?
consumption risk aversion does not change day by day

$\Rightarrow$ Random Walks successful $\Rightarrow$ Predictability of asset returns (day by day)?

Technical analysis, media reports...
Some popular linear factor models

Factor pricing models

CAPM: \[ m_{t+1} = a + bR_{w,t+1} \]

Free parameters

Compatible with utility maximisation?

ICAPM: \[ m_{t+1} = a + b'f_{t+1} \]

Parameter factors

Factors (macro, term spread, price-earnings ratio help forecast conditional distribution of future asset returns)

APT: similar,

but factors determined by principal component analysis of payoff covariance matrix

Practice: just test \( m = b'f \) and don't worry about derivations
The benchmark model: Fama/French (1993,1996) three factor model

- Fama French model

\[ m_{t+1} = b_0 + b_m R_{t+1}^{em} + b_{SMB} SMB_{t+1} + b_{HML} HML_{t+1} \]

Market excess return

Excess return small vs. large stocks
Excess return value stocks vs. growth stocks (high book-to-market – low book-to-market)
2. Stochastic discount factors and GMM estimation

Readings:
Cochrane (2005), Chapters 7, 10, 11
Hamilton (1994), Chapter 14
Hayashi (2000), Chapter 7
Hall (2005) (new GMM textbook)
The basic pricing equation implies a set of CONDITIONAL moment restrictions

\[ p_t = \mathbb{E}_t \left( m_{t+1} x_{t+1} \right) \]
\[ = \mathbb{E} \left( m_{t+1} x_{t+1} \mid I_t \right) \]

\{m_t\} and \{x_t\} non i.i.d. \Rightarrow \mathbb{E}_t (\cdot) \neq \mathbb{E} (\cdot)

Information set (partially) not observed, conditional density not known, conditional expectation cannot be computed

Conditioning down to coarser information set

\[ p_t = \mathbb{E}_t \left( m_{t+1} x_{t+1} \right) \]
\[ = \mathbb{E} \left( \mathbb{E}_t \left( m_{t+1} x_{t+1} \right) \right) \quad \text{i.i.e.} \]
\[ = \mathbb{E} \left( m_{t+1} x_{t+1} \right) \]
Models contain **free parameters**

\[ p_t = \mathbb{E}_t \left( \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} x_{t+1} \right) \]

- Estimation from data
- Testing hypotheses about parameters
- How good is the model?
Estimation and evaluation of asset pricing models (CBM)

\[ p_t = \mathbb{E}_t(m_{t+1} x_{t+1}) \quad \text{or} \quad 1 = \mathbb{E}_t(m_{t+1} R_{t+1}) \]

\[ \uparrow f(\text{data, parameters}) \]

e.g. CBM with \( u(c) = \frac{1}{1-\gamma} c^{1-\gamma} \Rightarrow m_{t+1} = \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \)

\( \frac{c_{t+1}}{c_t} : \text{data (random variables)} \)
\( b = (\beta, \gamma)' : \text{free parameters} \)

Assume model correct: "Best” choice for \( \beta, \gamma \)?
Best ”fit”, smallest (average) pricing errors
Estimation and evaluation of asset pricing models. The basic idea.

Estimates \( \hat{b} \) from data, distribution of \( \hat{b} \)?

Average pricing errors:

\[
\text{sample mean } \left( \text{observed price - predicted price} \right) = \alpha
\]

should be close to zero

\[
p_t = \mathbb{E}_t \left( m_{t+1}(b) \cdot x_{t+1} \right) = \mathbb{E} \left( m_{t+1}(b) \cdot x_{t+1} | I_t \right)
\]

\[
\mathbb{E}(p_t) = \mathbb{E}[\mathbb{E}_t \left( m_{t+1}(b) \cdot x_{t+1} \right)] = \mathbb{E}[m_{t+1}(b) \cdot x_{t+1}]
\]

Unconditional expectation:

\[
\mathbb{E}[m_{t+1}(b)x_{t+1} - p_t] = 0
\]

Equivalently using returns:

\[
1 = \mathbb{E}_t \left( m_{t+1}(b)R_{t+1} \right) \Rightarrow 0 = \mathbb{E} \left( m_{t+1}(b)R_{t+1} - 1 \right)
\]
Generalized Methods of Moments estimation is based on the WLLN

\[
WLLN : \frac{1}{N} \sum_{i=1}^{N} y_i \quad \overrightarrow{p} \quad \mathbb{E}(Y)
\]

sample average consistent estimate for population moment

\[
\frac{1}{T} \sum_{t=1}^{T} p_t - \frac{1}{T} \sum_{i=1}^{T} m_{t+1}(b)x_{t+1} \approx 0
\]

GMM basic idea (first step):
choose \( \hat{b} \) to minimize \( \alpha^2 \) (squared average pricing error) among set of test assets.
The two asset, two parameter case

\[ \mathbb{E}(m_{t+1}(\beta, \gamma) x_{t+1}^1 - p_t^1) = 0 \]
\[ \mathbb{E}(m_{t+1}(\beta, \gamma) x_{t+1}^2 - p_t^2) = 0 \]
\[ \mathbb{E}(m_{t+1}(\beta, \gamma) R_{t+1}^1 - 1) = 0 \]
\[ \mathbb{E}(m_{t+1}(\beta, \gamma) R_{t+1}^2 - 1) = 0 \]

\[ \frac{1}{T} \sum_{t=1}^{T} m_{t+1}(\beta, \gamma) R_{t+1}^1 - 1 = 0 \]
\[ \frac{1}{T} \sum_{t=1}^{T} m_{t+1}(\beta, \gamma) R_{t+1}^2 - 1 = 0 \]

solve equations for \( \beta, \gamma \) \( \Rightarrow \) \( \hat{\beta}, \hat{\gamma} \Rightarrow \)
To apply GMM data have to be generated by stationary (and ergodic) processes (not necessarily i.i.d.)

Problem: WLLN works for \textbf{stationary data}:

(Weakly) stationary process: \{\(Y_t\)\}_{t=-\infty}^{\infty}
\{\ldots,y_0,y_1,\ldots, y_5,\ldots\}
\(E(Y_t) = u\)
\(var(Y_t) = \sigma^2\)
\(cov(Y_t, Y_{t-j}) = \gamma_j\)

Solution: \(\Rightarrow\) We use:

\[1 = E \left( m_{t+1}(b) \cdot R_{t+1} \right) \quad \text{instead of} \quad E(p_t) = E \left( m_{t+1}(b) \cdot x_{t+1} \right)\]
\[0 = E \left( m_{t+1}(b) \cdot R_{t+1} - 1 \right)\]
Define the GMM residual or “pricing error“

Define GMM residual: object whose mean should be zero

\[ u_{t+1}(b) = m_{t+1}(b) R_{t+1} - 1 \]

\[ \mathbb{E}(u_{t+1}(b)) = 0 \]

\[ \mathbb{E}_T[u_t(b)] = \frac{1}{T} \sum_{t=1}^{T} u_t(b) \approx 0 \]

Notational convenience (Hansen’s notation, sometimes causing confusion)

\[ \mathbb{E}_T(\cdot) = \frac{1}{T} \sum_{t=1}^{T} (\cdot) \]
We have more assets than unknown model parameters

For GMM parameter estimation: Select $N$ test assets $R_t^1, R_t^2, \ldots, R_t^N$ $t = 1, \ldots, T$

$$\begin{bmatrix}
E_T[u_1^1(t)(b)] \\
E_T[u_2^2(t)(b)] \\
\vdots \\
E_T[u_N^N(t)(b)]
\end{bmatrix} = g_T(b) \quad N \times 1 \text{ vector}$$

If $\# \text{ assets} = \# \text{ parameters} b$ can be chosen such that average pricing errors are zero usually $\# \text{ assets} > \# \text{ parameters}$. 
GMM objective function

\[ \hat{b} = \arg\min_{\{b\}} g_T'(b) \cdot I_N \cdot g_T(b) \quad \text{first step GMM estimate} \]

\[ = \arg\min_{\{b\}} \left[ \mathbb{E}_T[u_{t+1}^1(b)] \right]^2 + \left[ \mathbb{E}_T[u_{t+1}^2(b)] \right]^2 + \ldots + \left[ \mathbb{E}_T[u_{t+1}^N(b)] \right]^2 \]

\[ \Rightarrow \text{minimize sum of squared average (pricing) errors} \]

\[ \text{equal weight for all test assets } 1, \ldots, N \]

Alternatively other weight matrix

\[ \hat{b} = \arg\min_{\{b\}} g_T'(b) W g_T(b) \quad \text{e. g. } W = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \]
Under mild assumptions (stationarity) GMM estimators have desirable properties

GMM estimators consistent:
Bias and variance of estimator go to zero asymptotically $\hat{b} \xrightarrow{p} b$

GMM estimators asymptotically normal. Required for inference:

$$
\text{var}(\hat{b}) = 
\begin{pmatrix}
\text{var}(\hat{b}_1) & \cdots & \\
\text{cov}(\hat{b}_1, \hat{b}_2) & \text{var}(\hat{b}_2) & \\
\vdots & \ddots & \\
\text{cov}(\hat{b}_1, \hat{b}_k) & \cdots & \text{var}(\hat{b}_k)
\end{pmatrix}
$$

To conduct $t$-test: $\frac{\hat{b}_k}{\sigma_k} \sim N(0, 1)$
Efficient estimates obtained by using the optimal weighting matrix

Efficiency: Smallest asymptotic variance among GMM estimators

Efficient estimator: employ $S^{-1}$ as weighting matrix

\[ S = \text{var}(g_T(b)) = \mathbb{E}(g_T(b)g_T(b)' - 1) = \mathbb{E}(u_t(b)u_t(b)') \]

\[ \text{resp. } \sum_{j=-\infty}^{j=+\infty} \mathbb{E}(u_t(b)u_{t-j}(b)') \]

variance-covariance matrix of pricing errors (no serial correlation)

with serial correlation in moment conditions
There exists an optimal weighting matrix

**Optimal weighting matrix (and GMM parameter standard errors):** use consistent estimate \( \hat{S} \) of \( S \) in minimization:

\[
\hat{b} = \arg\min_{\{b\}} g_T(b)' \hat{S}^{-1} g_T(b)
\]

write \( u_t(b) = \begin{pmatrix} u_1^1(b) \\ \vdots \\ u_N^N(b) \end{pmatrix} \)

\( u_i^i(b) = m_{t+1}(b)x_{t+1}^i - p_i^i \)

**Recall:** \( \mathbb{E}(u_t^i) = 0 \) \( \Rightarrow \) \( \mathbb{E}(u_t(b)) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \)
The optimal weighing matrix takes into account variances and covariances of pricing errors across assets

\[ S = \mathbb{E} \left[ u_t(b) \cdot u'_t(b) \right] = \begin{bmatrix} \mathbb{E} \left( [u^1_t(b)]^2 \right) & \cdots & \mathbb{E} \left( [u^N_t(b)]^2 \right) \\ \vdots & \ddots & \vdots \\ \mathbb{E} \left( u^1_t(b) u^2_t(b) \right) & \cdots & \mathbb{E} \left( u^1_t(b) u^N_t(b) \right) \end{bmatrix} \]

With no serial correlation in pricing errors!

\[ S = \text{variance covariance matrix of pricing errors} \]

\[ = \begin{bmatrix} \text{var} \left( u^1_t(b) \right) & \cdots & \text{var} \left( u^N_t(b) \right) \\ \text{cov} \left( u^1_t(b) u^2_t(b) \right) \text{var} \left( u^2_t(b) \right) & \cdots \\ \vdots & \ddots & \vdots \\ \end{bmatrix} \]

Estimate \( \hat{S} \): Replace \( \mathbb{E} \) by \( \frac{1}{N} \sum \) using \( \hat{b} \) obtained with weighting matrix \( I_N \Rightarrow \hat{S} \).
Steps of iterated GMM estimation

1) $\hat{b}^1 = \operatorname{argmin}_{\{b\}} g_T(b)' I_N g_T(b) \Rightarrow$

2) $\hat{S} \Rightarrow$

3) $\hat{b}^2 = \operatorname{argmin}_{\{b\}} g_T(b)' \hat{S}^{-1} g_T(b)$
   \[ \text{...repeat...} \ldots \]
Intuition behind optimal weighting matrix (1)

Intuition behind GMM weighting matrix

Example

\( N = 2, \quad \text{cov}(u_t^1(b), u_t^2(b)) = 0 \) [zero covariance of pricing errors]

\[
S = \begin{bmatrix}
\text{var}[u_t^1(b)] & 0 \\
0 & \text{var}[u_t^2(b)]
\end{bmatrix}
\]

\[
S^{-1} = \begin{bmatrix}
\frac{1}{\text{var}[u_t^1(b)]} & 0 \\
0 & \frac{1}{\text{var}[u_t^2(b)]}
\end{bmatrix} = \begin{bmatrix}
W_1 & 0 \\
0 & W_2
\end{bmatrix}
\]

Example \( S = \begin{pmatrix}
10 & 0 \\
0 & 0.1
\end{pmatrix} \)
Intuition behind optimal weighting matrix (2)

GMM objective \( g_T(b)'S^{-1}g_T(b) \) becomes

\[
\arg\min_{\{b\}} \mathbb{E}_T \left[ u^1_t(b) \right]^2 \cdot W_1 + \mathbb{E}_T \left[ u^2(b) \right]^2 \cdot W_2
\]

Example

\( W_1 : 0.1 \implies \text{var} \left( u^1_t(b) \right) = 10 \)
\( W_2 : 10 \implies \text{var} \left( u^2_t(b) \right) = 0.1 \)

\( \implies \) Asset (1) gets less weight in minimization

"Model imprecise" for asset 1, more precise for asset 2.
Some more intuition behind optimal weighting matrix: Correlations across pricing errors (1)

Another example: Correlations between asset returns: Two "similar" assets (high correlation of pricing errors) are downweighted. Count more like one asset.

Example \( S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.999 \\ 0 & 0.999 & 1 \end{pmatrix} \)

\[ \text{cov}(u_t^2, u_t^3) = 0.999 \]

\[ \text{corr}(u_t^2, u_t^3) \approx 1 = \frac{0.999}{\sqrt{1} \sqrt{1}} \]

\[ \arg\min_{\{b\}} \left[ \mathbb{E}_T(u_t^1(b)), \mathbb{E}_T(u_t^2(b)), \mathbb{E}_T(u_t^3(b)) \right] \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.99 \\ 0 & 0.99 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} \mathbb{E}_T(u_t^1(b)) \\ \mathbb{E}_T(u_t^2(b)) \\ \mathbb{E}_T(u_t^3(b)) \end{bmatrix} \]
Some more intuition behind optimal weighting matrix: Correlations across pricing errors (2)

\[
S^{-1} = \begin{bmatrix}
1 & 0 & 0 \\
0 & 500.25 & -499.75 \\
0 & -499.75 & 500.25 
\end{bmatrix}
\]

\[
\text{argmin}_{\{b\}} g_T(b)' S^{-1} g_T(b) = \\
\left[ \mathbb{E}_T \left( u_1^1(b) \right), \mathbb{E}_T \left( u_2^2(b) \right) \cdot 500.25 - \mathbb{E}_T \left( u_3^3(b) \right) \cdot 499.75, \\
\mathbb{E}_T \left( u_3^3(b) \right) \cdot 500.75 - \mathbb{E}_T \left( u_2^2(b) \right) \cdot 499.75 \right] \times \begin{bmatrix}
\mathbb{E}_T(u_1^1(b)) \\
\mathbb{E}_T(u_2^2(b)) \\
\mathbb{E}_T(u_3^3(b))
\end{bmatrix}
\]
Some more intuition behind optimal weighting matrix: Correlations of pricing errors (3)

\[
\arg\min_{\{b\}} g_T(b)' S^{-1} g_T(b) = \\
\mathbb{E}_T \left( u_t^1(b) \right)^2 + \mathbb{E}_T \left( u_t^2(b) \right)^2 \cdot 500.25 + \mathbb{E}_T \left( u_t^3(b) \right)^2 \cdot 500.25 - \\
2 \cdot \mathbb{E}_T \left( u_t^2(b) \right) \mathbb{E}_T \left( u_t^3(b) \right) \cdot 499.75
\]

\[
\approx \mathbb{E}_T \left( u_t^1(b) \right)^2 + 0.5 \mathbb{E}_T \left( u_t^2(b) \right)^2 + 0.5 \mathbb{E}_T \left( u_t^3(b) \right)^2
\]

since

\[
\mathbb{E}_T \left( u_t^2(b) \right) \approx \mathbb{E}_T \left( u_t^3(b) \right)
\]
To test hypotheses we need the distribution of the GMM estimates

Standard errors of GMM estimates

We want:

$$\text{var}(\hat{b}) = \begin{pmatrix}
\text{var}(\hat{b}_1) & \text{cov}(\hat{b}_1, \hat{b}_2) & \cdots & \text{cov}(\hat{b}_1, \hat{b}_k) \\
\text{cov}(\hat{b}_1, \hat{b}_2) & \text{var}(\hat{b}_2) & & \\
\text{cov}(\hat{b}_1, \hat{b}_k) & \cdots & \text{var}(\hat{b}_k)
\end{pmatrix} (K \times K)$$

$$b = (b_0, b_1, \cdots, b_k)$$

$$t = \frac{\hat{b}_k - 0}{\sqrt{\text{var}(\hat{b}_k)}} \sim N(0, 1) \text{ under } H_0 : b_k = 0$$
Asyptotic distribution of GMM estimates when using optimal weighting matrix


\[
\sqrt{T} (\hat{b} - b) \xrightarrow{d} N(0, (d' S^{-1} d))
\]

\[
d = \mathbb{E} \left( \frac{\partial u_t(b)}{\partial b} \right)
\]

consistently estimated by

\[
\hat{d} = \left. \frac{\partial g_T(b)}{\partial b} \right|_b
\]

t- and Wald tests use

\[
\text{var}(\hat{b}) = \frac{\hat{d}' S^{-1} \hat{d}}{T}
\]
Details

Some more details:

a) In application: replace $S^{-1}$ by consistent estimate $\hat{S}^{-1}$
b) Recall

$$g_T(b) = \begin{bmatrix} \frac{1}{T} \sum u_t^1(b) \\ \vdots \\ \frac{1}{T} \sum u_t^N(b) \end{bmatrix} = \begin{bmatrix} \frac{1}{T} \sum m_t(b) R_t^1 - 1 \\ \vdots \\ \frac{1}{T} \sum m_t(b) R_t^N - 1 \end{bmatrix}$$

$$\frac{\partial g_T(b)}{\partial b} = \begin{bmatrix} \frac{1}{T} \sum \frac{\partial u_t^1(b)}{\partial b_1} \\ \frac{1}{T} \sum \frac{\partial u_t^1(b)}{\partial b_2} \\ \vdots \\ \frac{1}{T} \sum \frac{\partial u_t^1(b)}{\partial b_k} \\ \frac{1}{T} \sum \frac{\partial u_t^N(b)}{\partial b_1} \\ \frac{1}{T} \sum \frac{\partial u_t^N(b)}{\partial b_2} \\ \vdots \\ \frac{1}{T} \sum \frac{\partial u_t^N(b)}{\partial b_k} \end{bmatrix}_{[N \times k]}$$
Details

$$\frac{\partial g_T(b)}{\partial b} = \begin{bmatrix} \frac{1}{T} \sum_{t=1}^{T} & \frac{\partial m_t(b)}{\partial b_1} R_t, \ldots \ldots \ldots \downarrow \downarrow N \end{bmatrix}$$

Parameters

For power utility

$$m_{t+1}(b) = \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma}$$

$$b = \beta, \gamma$$

Linear factor models  $$m_{t+1} = b' f_{t+1} \quad b \neq 0$$

Risk factor?

$$\frac{\partial m_{t+1}(b)}{\partial b_1} = ?$$
We employ the estimated variance covariance matrix to test hypotheses

\[ \text{var}(\hat{b}) \text{ used for testing hypotheses:} \]

\[ H_0 : \quad b_k = 0 \]

\[ t\text{-statistic:} \quad \frac{\hat{b}_k - 0}{\sqrt{\text{var}(\hat{b}_k)}} \sim N(0, 1) \triangleq \text{Standard } t\text{-test.} \]

joint significance:

\[ H_0 : \quad (b_{j1} = b_{j2} = \ldots = b_{jN} = 0) \text{ or } b_J = 0 \]

some subset of \( b \)

\[ \hat{b}_j' \left[ \sqrt{\text{var}(\hat{b})}_J \right]^{-1} \hat{b}_j \sim \chi^2(J) \subseteq \text{Standard Wald test use to test } Rb=r \]

appropriate subset of \( \text{var}(\hat{b}) \)

Nonlinear restrictions testable applying delta method => EVIEWS example

Prof. Joachim Grammig, University of Tübingen, Department of Econometrics, Statistics and Empirical Economics 45
Testing the validity of the model (moment conditions) by J-test

\[ \{ R_t, \Delta c_t, \ldots \} \quad \text{data is a random sample} \quad \Rightarrow \hat{b} \quad \text{is a random variable} \quad \Rightarrow \]

\[ u_t(b) \quad \text{is a random variable} \quad \Rightarrow \mathbb{E}_T(u_t(b)) = \frac{1}{N} \sum \cdots \quad \text{is a random variable} \]

pricing errors too large to be explained by random sampling?

\[ \Leftrightarrow \text{Is the model in correct?} \]

\[ T \cdot J_T = T \cdot \left[ g_T(\hat{b})' \tilde{S}^{-1} g_T(\hat{b}) \right] \sim \chi^2 \left[ \begin{array}{c} \text{no. moment conditions} \\ \text{no. of parameters.} \end{array} \right] \]

objective function at minimum using optimal weighting matrix estimate

\[ \Rightarrow \text{Reject or non-reject model (i.e. moment conditions) at given significance level} \]

Example: no. of moment conditions: 10, no. parameters: 2,

\[ T J_T = 7.9, \quad \chi^2_{95} (1) = 2.73 \Rightarrow \]
Remarks

Inference is different if other weighting matrix than optimal weighting matrix is used

- different formula for parameter standard errors

- different formula for J-statistic. Watch out when using EVIEWS!

When comparing alternative models (e.g. parameter restrictions) use the same weighting matrix (weighting matrix depends on unknown parameters)
General GMM results (Hayashi Ch. 6)

Choose $W$ to be positive semi-definite and symmetric

$$\hat{b} = \arg \min_{\{b\}} g_T(b)' W g_T(b)$$

$$\frac{\partial g_T(b)'}{\partial b} W \times g_T(b) = 0$$

K linear combinations set to zero

$$N \times 1 \text{ vector of r.v. with } K \text{ linear dependencies}$$
General GMM results (Hayashi Ch. 6)

\[ \sqrt{T}(\hat{b} - b) \xrightarrow{d} N(0, (d'Wd)^{-1}d'WSWd(d'Wd)^{-1}) \]

For \( t \)- and Wald-tests use

\[ \text{var}(\hat{b}) = \frac{(\hat{d}'W\hat{d})^{-1}\hat{d}'W\hat{S}W\hat{d}(\hat{d}'W\hat{d})^{-1}}{T} \]
General GMM results (Hayashi Ch. 6)

\[ \sqrt{T} g_T(\hat{b}) \rightarrow_d N(0, Avar(g_T(\hat{b}))) \]

\[ Avar(g_T(\hat{b})) = (I - d(d'Wd)^{-1}d'W)S(I - d(d'Wd)^{-1}d'W) \]

General form of J-statistic

\[ T g_T(\hat{b})' [Avar(g_T(\hat{b}))]^+ g_T(\hat{b}) \rightarrow_d \chi(N - K) \]

Pseudo inverse, linear dependencies in \( g \) by construction, V-C matrix singular
Performance comparison (1)

Problems using J-statistic

Popular measure

Compare observed average return with $\mathbb{E}(R)$ predicted by model

From

$$1 = \mathbb{E}(mR)$$

$$1 = \mathbb{E}(m)\mathbb{E}(R) + \text{cov}(m, R)$$

$$\mathbb{E}(R) = \frac{1}{\mathbb{E}(m)} - \frac{\text{cov}(m, R)}{\mathbb{E}(m)}$$

Use as predictor

$$\widehat{\mathbb{E}(R)} = \frac{1}{\frac{1}{T} \sum_{t=1}^{T} m_t} \left( \frac{1}{T} \sum_{t=1}^{T} m_t R_t - \frac{1}{T} \sum_{t=1}^{T} m_t \frac{1}{T} \sum_{t=1}^{T} R_t \right)$$
Performance comparison (2)

Plot $\mathbb{E}(R)$ vs. $\frac{1}{T} \sum_{t=1}^{T} R_t = \bar{R}$

Similarly using excess returns as test assets

From $0 = \mathbb{E}(mR^e)$

$0 = \mathbb{E}(m)\mathbb{E}(R^e) + \text{cov}(m, R^e)$

$\mathbb{E}(R^e) = -\frac{\text{cov}(m, R^e)}{\mathbb{E}(m)}$

Again: replace $\mathbb{E}()$ by $\frac{1}{T} \sum()$ to obtain $\widehat{\mathbb{E}(R^e)}$

Plot $\widehat{\mathbb{E}(R^e)}$ against $\bar{R}^e$

$\text{RMSE} = \sqrt{\sum_{j=1}^{N} \left[ \mathbb{E}(R^j) - \bar{R}^j \right]^2}$ or $\sqrt{\sum_{j=1}^{N} \left[ \mathbb{E}(R^{ej}) - \bar{R}^{ej} \right]^2}$ used to rank and compare alternative models
Cochrane's (1996) estimation results for the consumption based model with power utility

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Estimates</th>
<th>Conditional Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>γ</td>
</tr>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>.98</td>
<td>241</td>
</tr>
<tr>
<td>t-statistic</td>
<td>.49</td>
<td>.61</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>1.27</td>
<td>71</td>
</tr>
<tr>
<td>t-statistic</td>
<td>10.9</td>
<td>2.17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Estimates</th>
<th>Conditional Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tests</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>J&lt;sub&gt;T&lt;/sub&gt;</td>
<td>J&lt;sub&gt;T&lt;/sub&gt;</td>
</tr>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>6.17</td>
<td>28</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>72</td>
<td>.30</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>11.3</td>
<td>33.9</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>26</td>
<td>.04</td>
</tr>
</tbody>
</table>

NOTE.—GMM estimates and tests of consumption-based model: $m_{t+1} = \beta (e_{t+1}/e_t)^{-\gamma}$. Asset returns are deciles 1–10 in the unconditional estimates and deciles 1, 2, 5, and 10 scaled by the constant, term premium, and dividend/price ratio in the conditional estimates. Assets do not include investment returns.
Non-rejection doesn’t mean a thing
Cochrane’s (1996) results for unconditional estimation of CAPM

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Estimates</th>
<th>Conditional Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b_0$</td>
<td>$b_m$</td>
</tr>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>6.5</td>
<td>-5.4</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>3.74</td>
<td>-3.21</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>6.7</td>
<td>-5.6</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>4.08</td>
<td>-3.53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Estimates</th>
<th>Conditional Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$J_T$</td>
<td>$J_T$</td>
</tr>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>3.3</td>
<td>26</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>95</td>
<td>.71</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>3.3</td>
<td>23</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>95</td>
<td>1.55</td>
</tr>
</tbody>
</table>
Cochrane's (1996) results for unconditional estimation of CAPM

![Graph showing the relationship between predicted mean excess return and mean excess return.](image)
Performance comparison. Example: Consumption-Based Model estimated on 25 Fama-French portfolios
Performance comparison. Example: CAPM estimated on 25 Fama-French portfolios

![Graph showing CAPM performance comparison](image-url)
Performance comparison. Example: Fama-French two factor model estimated on 25 Fama-French portfolios
GMM estimation using the Gauss library: Ingredients and recipe

1. Supply data
2. Provide GMM/optimization settings (number of iterations, weighting matrix)
3. Supply initial parameter values
4. Call GMM minimization procedure

Procedure to compute GMM residuals $u_t(b)$

$$u_t(b) = \begin{pmatrix} u_1^1(b) & \cdots & u_1^N(b) \\ \vdots & \ddots & \vdots \\ u_T^1(b) & \cdots & u_T^N(b) \end{pmatrix}$$

Parameter values $b$

Data:
- Returns
- Factors
- Economic Variables

„Global“ control variables like model version specification details

Procedure returns $u_t(b)$: GMM residuals evaluated at $b$

5. Check parameter estimates and test statistics
The canonical example: Estimate the CBM by GMM

For consumption based model with power utility

$$\mathbb{E}_T(u_t(b)) = \frac{1}{T} \sum_{t=1}^{T} \beta \left( \frac{c_{t+1}}{c_t} \right)^{\gamma} \cdot R_t^i - 1 = 0$$

Exercise: 10 test assets (NYSR decile portfolios)
Perform GMM estimation of $\gamma$ and $\beta$ using EXCEL solver.

Input: Time series of returns and consumption growth.

$$\begin{bmatrix}
R_1^1 & \cdots & R_1^{10} & R_1^f & dc_1 \\
\vdots & \ddots & \vdots & \vdots \\
R_T^1 & R_T^{10} & R_T^f & dc_T
\end{bmatrix}$$
3. Recent approaches

Newer models consumption based model and habit formation

Garcia et al. (2003)

Period utility function

\[ u(c_t/H_t, H_t) = \frac{(c_t/H_t)^{1-\gamma} H_t^{1-\psi}}{1 - \gamma} - 1 \]

Marginal utility

\[ u'(c_t) = c_t^{-\gamma} H_t^{\gamma - \psi} \]

Stochastic discount factor

\[ m_{t+1} = \delta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \left( \frac{H_{t+1}}{H_t} \right)^{\gamma - \psi} \]

\[ E_t \left[ \delta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \left( \frac{H_{t+1}}{H_t} \right)^{\gamma - \psi} R_{t+1}^i \right] = 1 \]
Modelling the habit level (1)

\[
H_{t+1} = \mathbb{E}(c_{t+1}|c_t, c_{t-1}, \ldots)
\]

\[
\Delta H_{t+1} = \lambda(c_t - H_t) \quad 0 \leq \lambda \leq 1
\]

\[
H_{t+1} = a + \lambda c_t + (1 - \lambda) H_t
\]

\[
H_{t+1} = \frac{a}{\lambda} + \lambda \sum_{i=0}^{\infty} (1 - \lambda)^i c_{t-i}
\]

using

\[
c_{t+1} = \frac{a}{\lambda} + \lambda \sum_{i=0}^{\infty} (1 - \lambda)^i c_{t-i} + \varepsilon_{t+1}
\]

\[
c_{t+1} = \frac{a}{\lambda} + \lambda c_t + \lambda(1 - \lambda) c_{t-1} + \lambda(1 - \lambda)^2 c_{t-2} + \ldots + \varepsilon_{t+1}
\]

\[
(1 - \lambda) c_t = \frac{a}{\lambda} (1 - \lambda) + \lambda(1 - \lambda) c_{t-1} + \ldots + (1 - \lambda) \varepsilon_t
\]
Modelling the habit level (2)

Subtracting two previous equations

\[ c_{t+1} - (1 - \lambda)c_t = a + \lambda c_t + \ldots + \varepsilon_{t+1} - (1 - \lambda)\varepsilon_t \]

\[ \Delta c_{t+1} = a - (1 - \lambda)\varepsilon_t + \varepsilon_{t+1} \]

ARIMA(0,1,1) model - Estimation by Maximum Likelihood

Use parameter estimates of \(a\) and \(\lambda\) to iterate on

\[ H_{t+1} = a + \lambda c_t + (1 - \lambda)H_t. \]

to estimate habit level

Plug in GMM objective function
An alternative model for the habit process (1)

Log habit growth (unobservable)

\[ \Delta h_{t+1} = \ln(H_{t+1}) - \ln(H_t) \]
\[ \Delta h_{t+1} = a_0 + \sum_{i=1}^{n} a_i \cdot \Delta \ln c_{t+1-i} + b \cdot r_{m}^m \]

with

\[ \Delta h_{t+1} = \mathbb{E}(\Delta \ln c_{t+1} | \Delta \ln c_t, \Delta \ln c_{t-1}, \ldots) \]
\[ \Delta \ln c_{t+1} = a_0 + \sum_{i=1}^{n} a_i \cdot \Delta \ln c_{t+1-i} + b \cdot r_{m}^m + \epsilon_{t+1} \]

\( a_0, a_1, \ldots, b \) can be estimated by GMM additional moment restrictions
An alternative model for the habit process (2)

Estimation
Add to usual moment conditions additional moment restrictions from habit equation:

use

\[ \mathbb{E}(m_{t+1}R^i_{t+1} - 1) = 0 \]
\[ \vdots \]
\[ \mathbb{E}(m_{t+1}R^N_{t+1} - 1) = 0 \]

along with

\[ \mathbb{E}(\varepsilon_{t+1}r^m_{t+1}) = 0 \]
\[ \mathbb{E}(\varepsilon_{t+1} \Delta \ln c_t) = 0 \]
\[ \vdots \]
An alternative model for the habit process (3)

Habit growth is then

\[
\frac{H_{t+1}}{H_t} = A \prod_{i=0}^{n} \left[ \frac{c_{t+1-i}}{c_{t-i}} \right]^{a_i} \left( R_{t+1}^m \right)^b
\]

Stochastic discount factor

\[
m_{t+1} = \delta A^{\gamma-\psi} \left[ \frac{c_{t+1}}{c_t} \right]^{-\gamma} \prod_{i=0}^{n} \left[ \frac{c_{t+1-i}}{c_{t-i}} \right]^{a_i(\gamma-\psi)} \left( R_{t+1}^m \right)^{b(\gamma-\psi)}
\]

Used for estimation

\[
m_{t+1} = \delta^* \left[ \frac{c_{t+1}}{c_t} \right]^{-\gamma} \prod_{i=0}^{n} \left[ \frac{c_{t+1-i}}{c_{t-i}} \right]^{a_i \cdot \kappa \cdot \frac{1}{b}} \left( R_{t+1}^m \right)^{\kappa}
\]

We estimate using

\[n = 0 \quad \text{"Epstein-Zin SDF"}
\]
\[n = 1\]
Performance comparison. Example: Habit model
Grammig/Schrömpf (2005) estimated on 25 Fama-French portfolios

Human Capital extended Model

Prof. Joachim Grammig, University of Tübingen, Department of Econometrics, Statistics and Empirical Economics
Performance comparison. Example: Fama-French two factor model estimated on 25 Fama-French portfolios
Performance comparison. Example: CAPM estimated on 25 Fama-French portfolios
Yogo’s durable consumption model (JF, 2006) includes durable and nondurables in investor utility function.

\[
D_t = (1 - \delta)D_{t-1} + E_t \\
\delta \in (0, 1)
\]

- **Stock of durable goods**
- **Depreciation rate**
- **Expenditures durable goods**
- **Nondurable goods**
- **Investment in assets**
- **Wealth**

\[
\sum_{i=0}^{N} B_t^i = W_t - C_t - P_tE_t
\]

\[
W_{t+1} = \sum_{i=0}^{N} B_t^i R_{i+1}^i
\]
The intra-period CES utility function contains durables and nondurables

$$u(C, D) = [(1 - \alpha)C^{1-1/\rho} + \alpha D^{1-1/\rho}]^{(1-1/\rho)}$$

Elasticity of substitution between durables and nondurables

$$\alpha \in (0, 1) \quad \rho \geq 0$$
The household’s intertemporal utility is specified by a recursive function that disentangles EIS and RRA

\[ u_t = \left\{ (1 - \beta)u(C_t, D_t)^{1 - 1/\sigma} + \beta(\mathbb{E}_t[u_{t+1}^{1-\gamma}]^{1/\kappa}) \right\}^{1/(1 - 1/\sigma)} \]


**Intertemporal elasticity of substitution (EIS)**

\[ \kappa = (1 - \gamma)/(1 - 1/\sigma) \quad \beta \in (0, 1) \quad \sigma \geq 0 \quad \gamma > 0 \]

Special cases: Epstein/Zin (1991) \( \sigma = \rho \)

Dunn/Singleton (1986) nonsperable expected utility model \( \sigma = 1/\gamma \)

Additive separable Model \( \sigma = 1/\gamma = \rho \)
Special case I $\sigma = \rho$

\[ u_t = \{ (1 - \beta)[(1 - \alpha)C_t^{1-1/\sigma} + \alpha D_t^{1-1/\sigma}] + \beta(\mathbb{E}_t[u_{t+1}^{1-\gamma}])^{1/\kappa} \}^{1/(1-1/\sigma)} \]

Additively seperable model by Epstein/Zin 1989, 1991
Special case II $\sigma=1/\gamma$ : additively separable utility model

$$u_t^{1-\gamma} = (1 - \beta)E_t \sum_{s=0}^{\infty} \beta^s u(C_{t+s}, D_{t+s})^{1-\gamma}$$

Solving the intertemporal asset allocation problem Yogo (2006) obtains the following SDF

\[
m_{t+1} = \left[ \beta \left( \frac{C_{t+1}}{C_t} \right)^{-1/\sigma} \left( \frac{v(D_{t+1}/C_{t+1})}{v(D_t/C_t)} \right)^{1/\rho-1/\sigma} R^W_{t+1} (1-1/\kappa) \right]^\kappa
\]

\[
v\left( \frac{D}{C} \right) = \left[ 1 - \alpha + \alpha \left( \frac{D}{C} \right)^{1-1/\rho} \right]^{1/(1-1/\rho)} \quad \text{with} \quad u(C, D) = C v(D/C)
\]

Use as usual for

\[
\mathbb{E}_t (m_{t+1} R^i_{t+1} = 1) \quad \mathbb{E}_t (m_{t+1} R^{ei}_{t+1}) = 0
\]
An additional moment restriction for the „investment“ in the durable good is added

\[
\frac{u_{Dt}}{u_{Ct}} = P_t - (1 - \delta) \mathbb{E}_t[m_{t+1}P_{t+1}] = \frac{\alpha}{1 - \alpha} \left( \frac{D_t}{C_t} \right)^{-1/\rho}
\]

\[
\mathbb{E} \left[ 1 - \frac{\alpha}{1 - \alpha} \left( \frac{D_t}{C_t} \right)^{-1/\rho} \frac{1}{P_t} - (1 - \delta)m_{t+1} \frac{P_{t+1}}{P_t} \right] = 0
\]
Yogo’s (2006) estimation results for Fama-French portfolios

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma )</td>
<td>0.024</td>
<td>EIS estimate small</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td></td>
</tr>
<tr>
<td>( \gamma )</td>
<td>191.438</td>
<td>Risk aversion estimate high</td>
</tr>
<tr>
<td></td>
<td>(49.868)</td>
<td></td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.520</td>
<td>elasticity of substitution reasonable</td>
</tr>
<tr>
<td></td>
<td>(0.544)</td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.827</td>
<td>subjective discount factor &lt; 1</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td></td>
</tr>
</tbody>
</table>

Eichenbaum/Hansen (1987) rejected
Epstein/Zin (1991) non-rejected

Test for \( \sigma = \rho \) 0.817  (0.366)
Test for \( \sigma = 1/\gamma \) 5.594  (0.018)

J-test 12.050  (0.956)
The fit of the durable consumption model is good (Fama French portfolios)

Some more models

- Linearized consumption based model

\[ m_{t+1} = b_0 + b_\Delta c \Delta \ln c_{t+1} \]

Taylor approximation of \( \frac{u'(c_{t+1})}{u'(c_t)} \)

- CAPM

\[ m_{t+1} = b_0 + b_m R_{t+1}^m \]

- Scaled CAPM by Lettau and Ludvigson (2001)

\[ m_{t+1} = b_0 + b_{cay} cay_t + b_m R_{t+1}^m + b_{caym} cay_t R_{t+1}^m \]
4. Testing conditional predictions of asset pricing models: Managed portfolios and scaled factors

Readings: Cochrane (2002), Ch. 8, 10, Cochrane (1996), Lettau and Ludvigson (2001 (JPE))
We use instruments to test the conditional predictions of asset pricing models

\[ p_t = \mathbb{E}(m_{t+1}(b) \cdot x_{t+1}|I_t) \text{ or } 1 = \mathbb{E}(m_{t+1}(b) \cdot R_{t+1}|I_t) \]

or \( 0 = \mathbb{E}(m_{t+1}(b) \cdot R^{e}_{t+1}|I_t) \)

l.i.e "integrates out" conditional implications, let us focus on unconditional implications of asset pricing model (model for S.D.F.):

\[ \mathbb{E}(m_{t+1}(b) \cdot R_{t+1} - 1) = 0 \]

To test conditional implications write

\[ \mathbb{E}(Y_{t+1}|I_t) = 0 \] where \( Y_{t+1} = (m_{t+1}(b) \cdot R_{t+1} - 1) \) or ...

\( \{Y_{t+1}\} \) a martingale difference sequence.

Properties of m.d.s include:

\( \text{cov}(Y_{t+1}, z_t) = 0 \quad \forall \quad z_t \in I_t \)

\[ \mathbb{E}(Y_{t+1} z_t) = 0 \text{ since } 1 \in I_t \]

Testable restrictions therefore: \( \mathbb{E}[(m_{t+1}(b) \cdot R_{t+1} - 1) z_t] = 0 \quad \forall \quad z_t \in I_t \)
The use of instruments has an economic interpretation: Can the model price “managed portfolios“?

\[ \tilde{x}_{t+1} = x_{t+1}^i z_t \] conceived as (payoff of) managed portfolios, i.e. artificial assets.

Example: \( z_t = \frac{d_t}{p_t} \) invest if \( z_t \uparrow \)

\[ \tilde{x}_{t+1} \] conceived as another payoff with price \( z_t p_t \)

If model correct, it prices any asset, also mgt. portfolios.

\[ \frac{z_t p_t}{p(\tilde{x}_{t+1})} = \mathbb{E}_t (m_{t+1}(b) \cdot x_{t+1} z_t) \quad \text{or} \quad z_t = \mathbb{E}_t \left( m_{t+1}(b) \cdot R_{t+1} z_t \right) \]

i.e.

\[ \mathbb{E}(z_t) = \mathbb{E}(m_{t+1} R_{t+1} z_t) \quad \text{or} \quad \mathbb{E}[ (m_{t+1} R_{t+1} - 1) z_t ] = 0 \]
To test the conditional implications you simply “blow up“ the number of assets by including meaningful managed portfolios and proceed as before.

Practice: \( N \) assets, \( M \) instruments
\( M \) moment restrictions

\[
\mathbb{E} \left( \left[ m_{t+1} (b) R_{t+1} - 1 \right] \otimes z_t \right) = 0
\]

With two assets and two instruments \( z_t = (1, z_1^1)' \)

\[
\mathbb{E} \left[ \begin{array}{c}
    m_{t+1} (b) R_{t+1}^a - 1 \\
    m_{t+1} (b) R_{t+1}^b - 1 \\
    (m_{t+1} (b) R_{t+1}^a - 1) z_1^1 \\
    (m_{t+1} (b) R_{t+1}^b - 1) z_1^1 \\
\end{array} \right] = 0
\]

or, emphasizing the managed portfolio interpretation

\[
\mathbb{E}(m_{t+1} (b) \overbrace{R_{t+1}}^{\text{payoff}} \otimes z_t - \overbrace{1 \otimes z_t}^{\text{price}}) = 0
\]

\[
\mathbb{E}(m_{t+1} (b) \overbrace{x_{t+1}}^{\text{payoff}} \otimes z_t - \overbrace{p_t \otimes z_t}^{\text{price}}) = 0
\]
You should include economically meaningful instruments (managed portfolios)

- $p = \mathbb{E}(m.x)$ should price any asset, also managed portfolios
- if model prices all managed portfolios, conditional asset pricing model true.
- select few selected instruments (we also select few assets from millions available). New managed funds example
- Select meaningful instruments: Those affecting conditional distribution of returns
- Any $z_t \in I_t$ qualifies as an instruments, but if $corr((m_{t+1}R_{t+1}), z_t) = 0$ but $corr(R_{t+1}, z_t)$ small: weak instrument
- danger of using weak instruments (Hamilton, 1994, p. 426 for references)
Some more details and intuition on the choice of instruments

\[ p_t z_t = \mathbb{E}_t(m_{t+1}x_{t+1}z_t) \] resp. \[ z_t = \mathbb{E}_t(m_{t+1}R_{t+1}z_t) \]

holds true trivially if \( \text{corr}((m_{t+1}R_{t+1} - 1), z_t) = 0 \)
but an interesting instrument implies \( \text{corr}(R_{t+1}, z_t) \neq 0 \) and/or \( \text{corr}(m_{t+1}, z_t) \neq 0 \)

if \( \mathbb{E}_t(R_{t+1}) \uparrow \) when \( z_t \uparrow \)

then in

\[ 1z_t = z_t \underbrace{\mathbb{E}_t(R_{t+1})}_{\uparrow} \underbrace{\mathbb{E}_t(m_{t+1})}_{\downarrow} + z_t \underbrace{\text{cov}_t(m_{t+1}R_{t+1})}_{\downarrow} \]
Is a conditional asset pricing model testable at all?

Most asset pricing models imply **conditional** moment restrictions

\[ 1 = \mathbb{E}(m_{t+1}(b_t) \cdot R_{t+1}|I_t) \]

e.g. CAPM \( m_{t+1} = a_t - b_t R_{t+1}^W \).

Parameters of factor pricing model vary over time. 
\( \Rightarrow \) unconditioning via i.i.e. no longer possible:

\[ 1 = \mathbb{E}(m_{t+1}(b_t) \cdot R_{t+1}|I_t) \]

does NOT imply

\[ 1 = \mathbb{E}(m_{t+1}(b) \cdot R_{t+1}) \]

this is not repaired by using scaled returns. GMM estimation no possible.

Hansen and Richard critique: CAPM (or other factor model) is not testable.
Scaled factors are a partial solution to the problem

With linear factor model

$$ m_{t+1} = b'_t \left( f_{t+1} \right) $$

use of "scaled factors" a partial solution:

"Blow up" number of factors by scaling factors with \((M \times 1)\) instruments vector \(z_t\) observable at \(t\)

$$ m_{t+1} = b'_t \left( f_{t+1} \otimes z_t \right) $$

Unconditioning via l.i.e. and GMM procedure as above.
Time varying parameters lead to scaled factors (single factor case)

Motivation
Consider linear one factor model \( m_{t+1} = a_t + b_t f_{t+1} \) (\( f_{t+1} \) scalar)
Assume Parameters vary with \( M \times 1 \) instruments vector \( z_t \).

\[
m_{t+1} = a(z_t) + b(z_t) f_{t+1}
\]

With linear functions

\[
a(z_t) = a' z_t \quad \text{and} \quad b(z_t) = b' z_t
\]

\[
\Rightarrow m_{t+1} = a' z_t + (b' z_t) f_{t+1}
\]

Mathematically equivalent to

\[
m_{t+1} = \tilde{b}' (\tilde{f}_{t+1} \otimes z_t)
\]

where \( \tilde{b} = \begin{bmatrix} a \\ b \end{bmatrix} \), \( \tilde{f}_{t+1} = \begin{bmatrix} 1 \\ f_{t+1} \end{bmatrix} \)

Number of parameters to estimate \( 2 \cdot M \)
Time varying parameters lead to scaled factors (multi factor case)

Multi-factor case:

\[ m_{t+1} = b_t^l f_{t+1}^{K \times 1} \]

Again: Time varying parameters linear functions of \( M \times 1 \) vector of observables \( z_t \).

\[ m_{t+1} = b(z_t)^l f_{t+1} \quad \text{with} \quad b(z_t) = B_{K \times M} z_t \]

Equivalent to \( m_{t+1} = \tilde{b}' (f_{t+1} \otimes z_t) \) where \( \tilde{b} = \text{vec}(B) \)

In practical application some elements of \( B \) may be set to zero.
Using scaled factors we can condition down and apply GMM

Conditioning down and GMM estimation possible

\[
\mathbb{E}_t \left( \begin{pmatrix} \tilde{b}' (f_{t+1} \otimes z_t) \\ m_{t+1} \end{pmatrix} R_{t+1} \right) = 1
\]  

\[ \Rightarrow \mathbb{E} \left( \left( \tilde{b}' (f_{t+1} \otimes z_t) \right) R_{t+1} - 1 \right) = 0\]

Scaled factors and managed portfolios can be combined. ($z_t$ might be the same).

\[ \Rightarrow \mathbb{E}(\tilde{b}' (f_{t+1} \otimes z_t) R_{t+1} - 1) \otimes z_t = 0 \]

- Inclusion of conditioning information as managed portfolios (scaled returns, increases number of test assets.

- Scaled factors increase number of unknown parameters
Cochranes (1996) CAPM with scaled factors

\[ f = \begin{pmatrix} 1 \\ RW \end{pmatrix} z_t = \begin{pmatrix} 1 \\ \frac{P}{D} \text{term} \end{pmatrix} B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \]

\[ f \otimes z = \begin{pmatrix} 1 \\ RW \\ \frac{P}{D} \\ RW \cdot \frac{P}{D} \\ \text{term} \\ RW \cdot \text{term} \end{pmatrix} \tilde{b} = (b_{11}, b_{21}, b_{12}, b_{22}, b_{13}, b_{23})' \]

\[ m = \tilde{b}'(f \otimes z) = b_{11} + b_{12} \frac{P}{D} + b_{13} \text{term} + b_{21} RW + b_{22} RW \cdot \frac{P}{D} + b_{23} RW \cdot \text{term} \]

In application Cochrane (1996) restricts \( b_{12} \) and \( b_{13} \) to zero

Prof. Joachim Grammig, University of Tübingen, Department of Econometrics, Statistics and Empirical Economics
Cochrane's (JPE 1996) estimation results for the consumption based model with power utility

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional Estimates</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>First-stage:</td>
</tr>
<tr>
<td>Coefficient</td>
</tr>
<tr>
<td>t-statistic</td>
</tr>
<tr>
<td>Iterated:</td>
</tr>
<tr>
<td>Coefficient</td>
</tr>
<tr>
<td>t-statistic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional Estimates</td>
</tr>
<tr>
<td>( J_T )</td>
</tr>
<tr>
<td>First-stage:</td>
</tr>
<tr>
<td>( \chi^2 )</td>
</tr>
<tr>
<td>Degrees of freedom</td>
</tr>
<tr>
<td>( p )-value (%)</td>
</tr>
<tr>
<td>Iterated:</td>
</tr>
<tr>
<td>( \chi^2 )</td>
</tr>
<tr>
<td>Degrees of freedom</td>
</tr>
<tr>
<td>( p )-value (%)</td>
</tr>
</tbody>
</table>

**Note.**—GMM estimates and tests of consumption-based model: \( m_{t+1} = \beta(c_{t+1}/c_t)^{-\gamma} \). Asset returns are deciles 1–10 in the unconditional estimates and deciles 1, 2, 5, and 10 scaled by the constant, term premium, and dividend/price ratio in the conditional estimates. Assets do not include investment returns.
Conditional estimation yields a poor performance of the consumption based model (Cochrane (1996))
Cochrane's (1996) results for unconditional estimation of CAPM

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Estimates</th>
<th>Conditional Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b_0$</td>
<td>$b_m$</td>
</tr>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>6.5</td>
<td>-5.4</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>3.74</td>
<td>-3.21</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>6.7</td>
<td>-5.6</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>4.08</td>
<td>-3.53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Estimates</th>
<th>Conditional Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$J_T$</td>
<td>$J_T$</td>
</tr>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>3.3</td>
<td>26</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>95</td>
<td>.71</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>3.3</td>
<td>23</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>95</td>
<td>1.55</td>
</tr>
</tbody>
</table>
Cochrane’s (1996) results for unconditional estimation of CAPM
Cochrane's (1996) results for conditional estimation of CAPM

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Estimates</th>
<th>Conditional Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b_0$</td>
<td>$b_m$</td>
</tr>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>6.5</td>
<td>-5.4</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>3.74</td>
<td>-3.21</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>6.7</td>
<td>-5.6</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>4.08</td>
<td>-3.53</td>
</tr>
</tbody>
</table>

**Tests**

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Estimates</th>
<th>Conditional Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$J_T$</td>
<td></td>
</tr>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>95</td>
<td></td>
</tr>
</tbody>
</table>
Cochrane's (1996) results for conditional estimation of CAPM

B. Scaled Model $m = b_0 + b_m r^n + b_{tp} (r^n \times \tau_p) + b_{dp} (r^n \times \tau_d)$:

**Conditional Estimates**

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
<th>$b_0$</th>
<th>$b_m$</th>
<th>$b_{tp}$</th>
<th>$b_{dp}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>4.56</td>
<td>-2.66</td>
<td>-.33</td>
<td>-.39</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>1.48</td>
<td>-.80</td>
<td>-1.32</td>
<td>-2.05</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>5.88</td>
<td>-4.62</td>
<td>.24</td>
<td>-.36</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>3.51</td>
<td>-2.70</td>
<td>2.26</td>
<td>-3.62</td>
</tr>
</tbody>
</table>

**Tests**

<table>
<thead>
<tr>
<th></th>
<th>$b_m$, $b_{tp}$, $b_{dp}$</th>
<th>Scaled $b$</th>
<th>$J_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First-stage:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>59</td>
<td>4.9</td>
<td>15.6</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>3</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>.00</td>
<td>8.6</td>
<td>7.7</td>
</tr>
<tr>
<td><strong>Iterated:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>67</td>
<td>15</td>
<td>18.9</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>3</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>$p$-value (%)</td>
<td>.00</td>
<td>.06</td>
<td>2.6</td>
</tr>
</tbody>
</table>
Cochrane’s (1996) results for conditional estimation of scaled CAPM
Cochrane’s (1996) results for conditional estimation of scaled CAPM
Yogo’s (2006) cross section estimation results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Panel A: Unconditional Moments</th>
<th>Panel B: Conditional Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fama-French</td>
<td>Industry &amp; BE/ME</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.024</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>191.438</td>
<td>199.496</td>
</tr>
<tr>
<td></td>
<td>(49.868)</td>
<td>(44.280)</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.520</td>
<td>0.554</td>
</tr>
<tr>
<td></td>
<td>(0.544)</td>
<td>(0.604)</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.827</td>
<td>0.821</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.091)</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.900</td>
<td>0.935</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.054)</td>
</tr>
</tbody>
</table>

| Test for \( \sigma = \rho \) | 0.817 | 0.768 | 0.187 | 7.510 | 375.185 |
|                               | (0.366) | (0.381) | (0.666) | (0.006) | (0.000) |

| Test for \( \sigma = 1/\gamma \) | 5.594 | 8.424 | 4.637 | 140.620 | 12.385 |
|                                   | (0.018) | (0.004) | (0.031) | (0.000) | (0.000) |

| J-test                            | 12.050 | 9.583 | 1.866 | 5.065 | 42.500 |
|                                   | (0.956) | (0.984) | (1.000) | (1.000) | (0.065) |

Resurrection of the C(CAPM) by Lettau and Ludvigson (2001)

Scaled CCAPM

\[ m_{t+1} = b_0 + b_1 cay_t + b_2 \Delta \ln c_{t+1} + b_3 cay_t \Delta \ln c_{t+1} \]

Scaled CAPM

\[ m_{t+1} = a_0 + a_1 cay_t + a_2 r^m_{t+1} + a_3 cay_t r^m_{t+1} \]

\[ c_t - w_t \approx \mathbb{E}_t \sum_{i=1}^{\infty} \rho^i_w (r^m_{t+i} - \Delta c_{t+i}) \]

\[ cay_t = c_t - \omega a_t - (1 - \omega) y_t \]
Performance comparison. Example: Lettau/Ludvigson model estimated on 25 Fama-French portfolios

Scaled CAPM, 1952Q2–2002Q1
Model comparison (practical exercise)

- 10 decile portfolios and t-bill rate (Cochrane 1996)
- 25 size/book-to-market portfolios and t-bill rate
- Excess returns or gross returns as test assets
- Estimation using GMM (alternatives ⇒ course 1)
- J-test
- RMSE comparisons (plots)

Models:
* Consumption Based Model (CBM), CAPM, Scaled (LL) CBM, Scaled (LL) CAPM, various habit model variants
5. Linear factor model and the basic pricing equation

Readings: Cochrane (2005), Ch. 6
Linear factor model dominate the empirical work because they have been easier to estimate.

Linear factor models

\[ p = \mathbb{E}(mx) \quad \text{or} \quad 1 = \mathbb{E}(mR) \quad \text{or} \quad 0 = \mathbb{E}(mR^e) \]

linear models for discount factor \( m = a + b'f \)

\( m = a - b \cdot R^m \): single factor model

What qualifies as a factor? Anything that affects investors MRS!
Linear factor models are equivalent to the more familiar expected return-beta representation

\[ m = a + b'f \quad \Leftrightarrow \quad \mathbb{E}(R^i) = \gamma + \lambda'\beta_i \quad \text{resp.} \quad \mathbb{E}(R^{ei}) = \lambda'\beta_i \]

\[ \lambda = \left(\lambda_1, \ldots, \lambda_K\right)' \quad \text{"Price of factor k" or factor risk premium} \]

\[ \beta_i = \left(\beta_{i1}, \ldots, \beta_{iK}\right)' \quad \text{Exposure of asset i to factor k} \]

\[ \gamma = \frac{1}{\mathbb{E}(m)} = R^f \]

Compare to linear regression:
\[ y_i = a + b'x_i + u_i \quad \mathbb{E}(u_i) = 0 \]
\[ \mathbb{E}(y_i) = a + b'\mathbb{E}(x_i) \]
If the factors have certain properties, the betas are given by the ratio of a covariance and a variance

Special cases:

if \( \mathbb{E}(f) = 0 \) (demeaned factors)
and \( \mathbb{E}(f_i f_j) = \text{cov}(f_i, f_j) = 0 \) for \( i \neq j \) (orthogonal factors)

\[ \Rightarrow \quad \beta_{ik} = \frac{\text{cov}(f_k, R_i)}{\text{var}(f_k)} \]

Example:

\[ m = a - bR^m \quad \Rightarrow \quad \mathbb{E}(R^i) = R^f + \beta_i(\mathbb{E}(R^m) - R^f) \]

where \( R^f \equiv \gamma \), \( \beta_i \equiv \text{riskiness of asset } i \) and \( \mathbb{E}(R^m) - R^f \equiv \lambda \equiv \text{market risk premium} \)
How can one estimate linear factor models?

Estimation and testing:

a) Use GMM \( (1 = \mathbb{E}(mR)) \)

b) linear regression - time series or cross section - Fama/McBeth

General problem for linear factor models: "fishing for factors"
We want to show the equivalence of the two representations (1)

We want to show: \(1 = \mathbb{E}(mR) \iff \mathbb{E}(R) = \gamma + \lambda' \beta:\)

single factor case: if \(m = \tilde{a} + b' \tilde{f}\)

convenient: demean factors: ”fold” means of factors into constant \(a\)

\[
\begin{align*}
\tilde{f} &= \text{factor with } (\tilde{f}) \neq 0 \\
 f &= \tilde{f} - \mathbb{E}(\tilde{f}) = \text{demeaned factor with } \mathbb{E}(f) = 0 \\
 m &= a + b' f \quad \text{where} \quad a = \tilde{a} + b' \mathbb{E}(\tilde{f}) \\
 \Rightarrow \quad \mathbb{E}(m) &= a
\end{align*}
\]
We want to show the equivalence of the two representations (2)

Rewrite

\[ 1 = \mathbb{E}(mR) \]
\[ = \text{cov}(m, R) + \mathbb{E}(m) \cdot \mathbb{E}(R) \]
\[ \Rightarrow \mathbb{E}(R) = \frac{1}{\mathbb{E}(m)} - \frac{\text{cov}(m, R)}{\mathbb{E}(m)} \]
\[ = \frac{1}{a} - \frac{\text{cov}((a + bf), R)}{a} \]

\[ \text{cov}((a + bf), R) = \mathbb{E}[(a + bf - a)(R - \mathbb{E}(R))] \]
\[ = \mathbb{E}(bfR) - \mathbb{E}(bf) \cdot \mathbb{E}(R) \]
\[ = 0 \text{ as } \mathbb{E}(f) = 0 \]
We want to show the equivalence of the two representations (3)

\[ \mathbb{E}(R) = \frac{1}{a} - \frac{b \mathbb{E}(Rf)}{a} \]

we want betas

\[ = \frac{1}{a} - \frac{\text{cov}(f, R)}{\text{var}(f)} \cdot \frac{\text{bvar}(f)}{a} \]

Define

\[ \gamma \equiv \frac{1}{a} = \frac{1}{\mathbb{E}(m)} = R^f \] (if traded)

\[ \beta \equiv \frac{\text{cov}(f, R)}{\text{var}(f)} \]

\[ \lambda \equiv -\frac{\text{bvar}(f)}{a} \]

\[ \Rightarrow \mathbb{E}(R^i) = \gamma + \beta_i \lambda \]
We want to interpret $\lambda$ as price of risk factor

$$
\lambda = -\frac{b\mathbb{E}(f^2)}{a} = -\frac{\mathbb{E}((a + bf) \cdot f)}{a} \quad \text{note: } \mathbb{E}(af) = a\mathbb{E}(f) = 0
$$

$$
= -\frac{\mathbb{E}(m \cdot f)}{a} = -\frac{p(f)}{a} = -\gamma \cdot p(f)
$$

if $\tilde{f}$ (non-demeaned factor) is a return, e.g. $R^m$

$$
-\gamma \cdot p(f) = -\gamma p(\tilde{f} - \mathbb{E}(\tilde{f})) = -\gamma (p(\tilde{f}) - p(\mathbb{E}(\tilde{f}))) \quad \text{since expectation operator is linear}
$$

$p(\tilde{f}) = 1$ if $\tilde{f}$ is a return

$$
p\left(\mathbb{E}(\tilde{f})\right) = \mathbb{E}(m \cdot \mathbb{E}(\tilde{f})) = \mathbb{E}(m) \cdot \mathbb{E}(\tilde{f}) = \frac{\mathbb{E}(\tilde{f})}{\gamma}
$$

constant payoff in $t+1$
If the factor is a return, $\lambda$ has the interpretation of an expected excess return, or factor risk premium

$$
\lambda = -\gamma \left(1 - \frac{\mathbb{E}(\tilde{f})}{\gamma}\right) = \mathbb{E}(\tilde{f}) - \gamma \mathbb{E}(\tilde{f}) - R^\tilde{f} \quad : \text{factor risk premium}
$$

$$
\Rightarrow \quad 1 = \mathbb{E}(mR)
$$

with $m = a + b \cdot f$ and $f = \tilde{f} - \mathbb{E}(\tilde{f})$ and $\tilde{f}$ is a return

$$
\Leftrightarrow \quad \mathbb{E}(R) = \gamma + \beta \left(\mathbb{E}(\tilde{f}) - \gamma\right)
$$

with $\gamma = \frac{1}{\mathbb{E}(m)} = R^f \Rightarrow \tilde{f} = R^m \Rightarrow \text{CAPM}$
In a multifactor model with k factors

1. \( \mathbb{E}(R^i) = \gamma + \lambda' \beta_i \)

2. \( \lambda = \mathbb{E}(\tilde{f}) - \gamma \)

\( \beta_i = \left[ \mathbb{E}(f f') \right]^{-1} \mathbb{E}(f R^i) \) with \( \mathbb{E}(f) = 0 \) (demeaned factors)
Equivalence in the multifactor case (2)

\[ \Rightarrow \beta_i = \text{cov}(f, R^i) \cdot [\text{cov}(f)]^{-1} \]

where \( \text{cov}(f, R^i) = \begin{bmatrix} \text{cov}(f_1, R^i) & \text{cov}(f_2, R^i) & \cdots \\ \end{bmatrix} \)

and \( \text{cov}(f) = \begin{bmatrix} \text{var}(f_1) & \text{cov}(f_1, f_2) & \cdots & \text{cov}(f_1, f_K) \\ \text{cov}(f_1, f_2) & \text{var}(f_2) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(f_1, f_K) & \cdots & \cdots & \text{var}(f_K) \end{bmatrix} \)

if demeaned factors orthogonal: \( \mathbb{E}(f_i f_j) = 0 \) for \( i \neq j \)

\[ \beta_{ik} = \frac{\text{cov}(f_k, R^i)}{\text{var}(f_k)} \]