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

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

how should the world work?
are the prices "wrong"?
- trading opportunities?
- cost of capital
- non traded assets: "fair" price

positive

how does the world work?

Empirical asset pricing - Introduction (2)

Basic : Prices equal discounted $\underbrace{\text{expected}}_{\text{What probability measure?}}$ payoff

Absolute Asset Pricing

↓

exposure to "fundamental" macroeconomic risk

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

↑

Relative Asset Pricing

e.g. CAPM:

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

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

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 t

stochastic discount factor (r.v.)

payoff (r.v.)

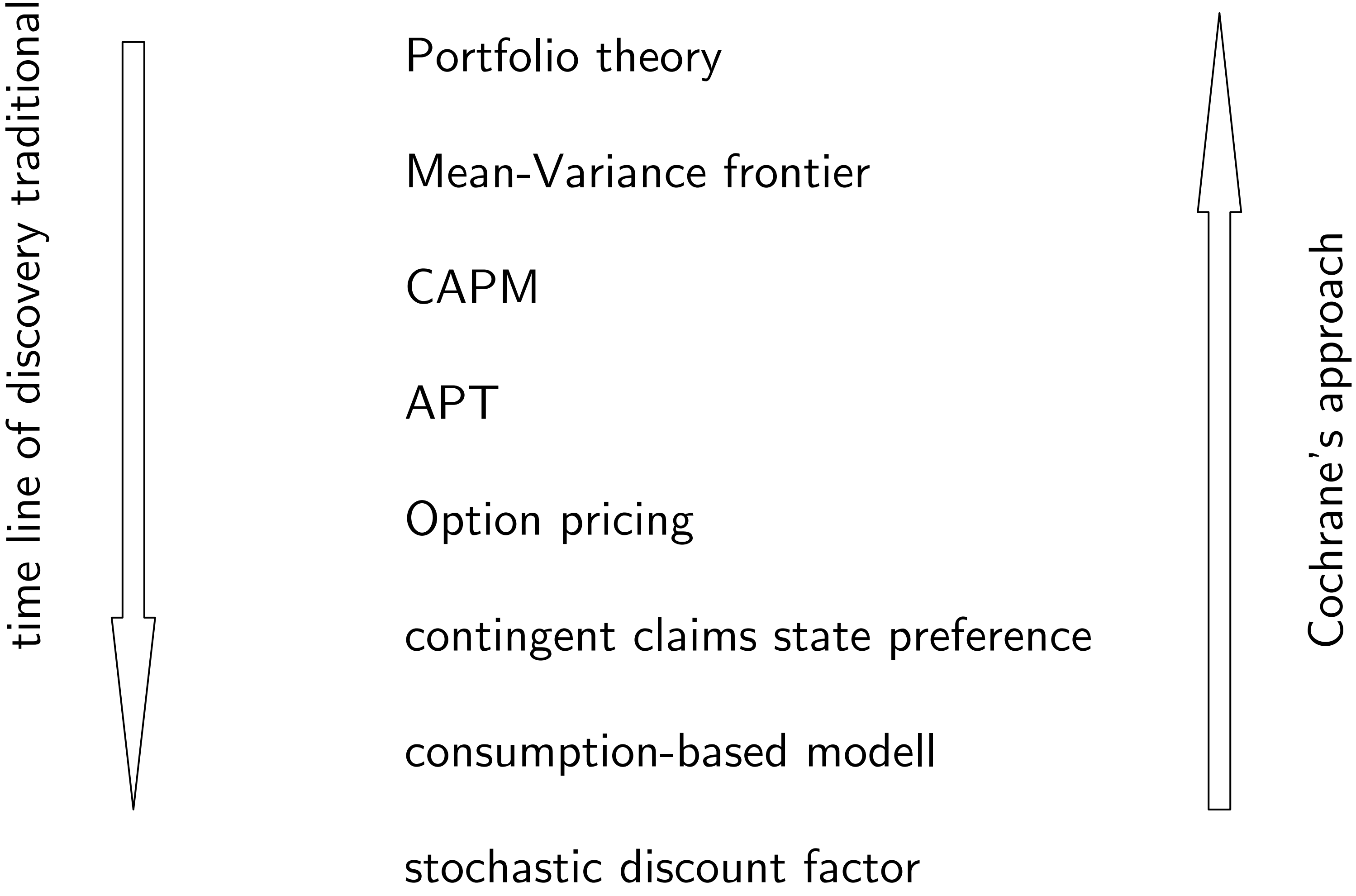
$$m_{t+1} = f(\underbrace{\text{data , parameters}}_{\text{the model}})$$

Moment condition: $\mathbb{E}_t(m_{t+1}x_{t+1}) - p_t \quad = 0$

use $\frac{1}{n}\sum \;\rightarrow \mathbb{E}()$ 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 modell

stochastic discount factor

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}

dividend

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

price of asset in t+1

Utility function

period utility function

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

expected utility

consumption

subjective discount factor

consumption level without asset purchase (other income)

$$\begin{aligned} c_t &= e_t - p_t\xi \\ c_{t+1} &= e_{t+1} + x_{t+1}\xi \end{aligned}$$

quantity of asset bought/sold

Random variables: $p_{t+1},\ d_{t+1},\ x_{t+1},\ e_{t+1},\ c_{t+1},\ u\left(c_{t+1}\right)$ $\qquad \mathbb{E}_t\left[\cdot\right] \triangleq \mathbb{E}\left[\cdot \mid \mathcal{F}_t\right]$

From an utility maximising investor`s first order conditions we obtain the basic asset pricing formula (2)

$$\max_{(\xi)} \left[U\left(c_t, c_{t+1}\right) \right] \text{ s.t.}$$

$$c_t = e_t - p_t \xi; \; c_{t+1} = e_{t+1} + x_{t+1} \xi$$

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

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

utility loss if investor buys another unit of the asset

discounted expected utility increase from extra payoff

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

Investor continues to buy or sell the asset until marginal loss equals marginal gain.

No complete solution: endogenous variables

Turning off uncertainty we are in the standard two-goods case (1)

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

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

$$\begin{aligned} p_t\cdot u'\left(c_t\right) &= x_{t+1}\cdot\beta u'\left(c_{t+1}\right)\\ p_t &= x_{t+1}\cdot\frac{\beta u'\left(c_{t+1}\right)}{u'\left(c_t\right)}\end{aligned}$$

marginal valuation of consumption in t+1 in terms of consumption in t →

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

← opportunity cost to transfer consumption from t to t+1

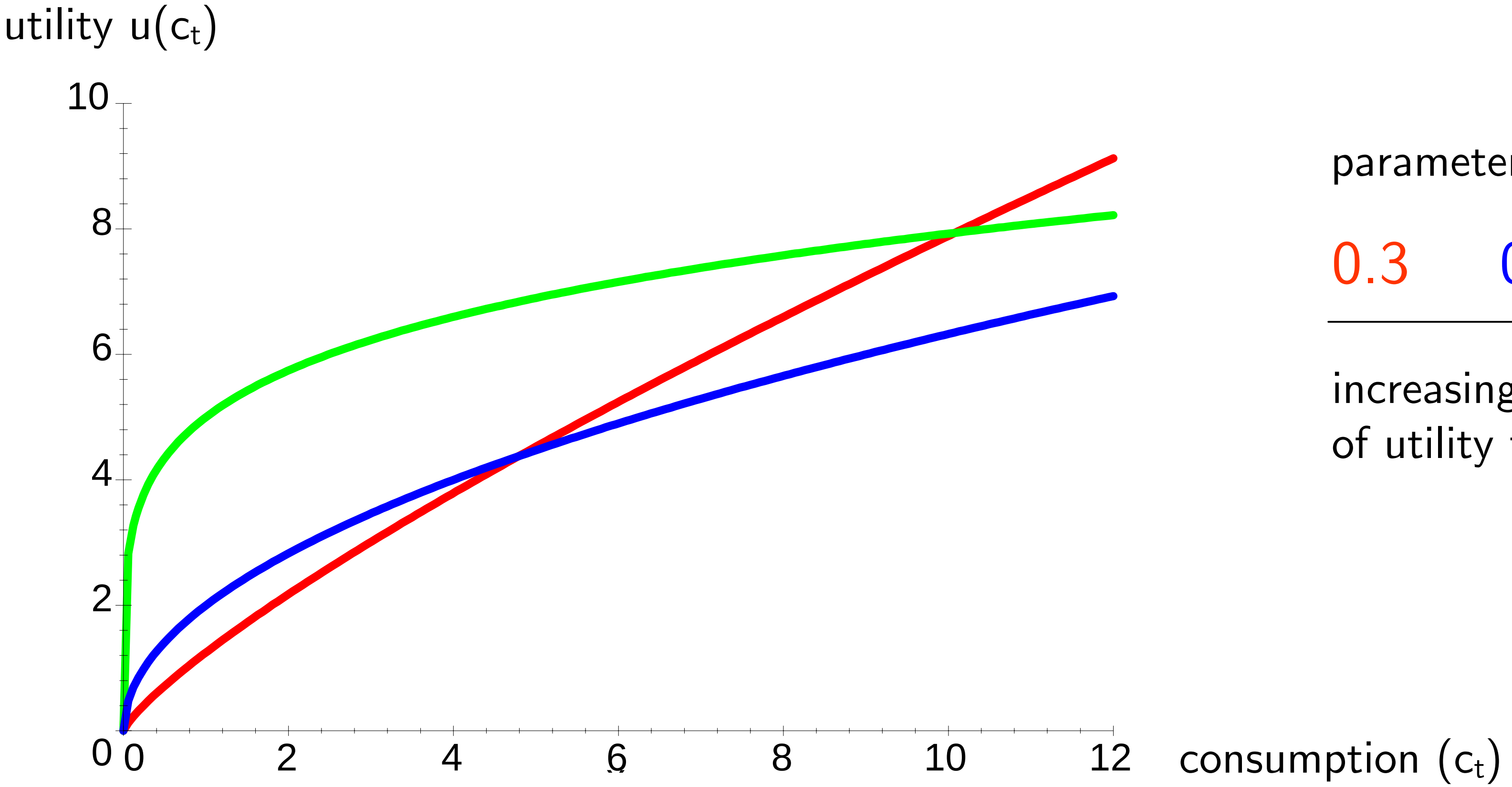
$$\begin{aligned} p_t u'\left(c_t\right) &= \mathbb{E}_t\left[\beta u'\left(c_{t+1}\right)x_{t+1}\right]\\ p_t &= \mathbb{E}_t\left[\beta\frac{u'\left(c_{t+1}\right)}{u'\left(c_t\right)}x_{t+1}\right]\end{aligned}$$

We often use a convenient power utility function (1)

$$u\left(c_t\right) = \frac{1}{1-\gamma} c_t^{1-\gamma} \qquad \lim_{\gamma \to 1} \left(\frac{1}{1-\gamma} c_t^{1-\gamma} \right) = \ln\left(c_t\right)$$

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

← marginal rate of substitution



parameter γ:

0.3 0.5 0.8

→

increasing concavity of utility function

Prices, payoffs, excess returns

	Price p_t	Payoff x_{t+1}
stock	p_t	$p_{t+1} + d_{t+1}$
return	1	R_{t+1}
excess return	0	$R^e_{t+1} = R^a_{t+1} - R^b_{t+1}$
one $ one period discount bond	p_t	1
risk-free rate	1	R^f

Payoff x_{t+1} divided by price $p_t \Rightarrow$ gross return $R_{t+1} = \frac{x_{t+1}}{p_t}$
Return: payoff with price one

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

Zero-cost portfolio:
Short selling one stock, investing proceeds in another stock
$\Rightarrow$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

$$
\begin{aligned}
cov\left(m_{t+1}, x_{t+1}\right) &= \mathbb{E}\left(m_{t+1} \cdot x_{t+1}\right) - \mathbb{E}\left(m_{t+1}\right)\mathbb{E}\left(x_{t+1}\right) \\
p_t &= \mathbb{E}\left(m_{t+1} \cdot x_{t+1}\right) \\
&= \mathbb{E}\left(m_{t+1}\right)\mathbb{E}\left(x_{t+1}\right) + cov\left(m_{t+1}, x_{t+1}\right)
\end{aligned}
$$

$$
\begin{aligned}
R^f &= \frac{1}{\mathbb{E}\left(m_{t+1}\right)} \\
p_t &= \frac{\mathbb{E}\left(x_{t+1}\right)}{R^f} + cov\left(m_{t+1}, x_{t+1}\right)
\end{aligned}
$$

$$
\begin{aligned}
p_t &= \frac{\mathbb{E}\left(x_{t+1}\right)}{R^f} + cov\left(\beta\frac{u'\left(c_{t+1}\right)}{u'\left(c_t\right)}, x_{t+1}\right) \\
p_t &= \underbrace{\frac{\mathbb{E}\left(x_{t+1}\right)}{R^f}}_{\text{price in risk-neutral world}} + \beta\underbrace{\frac{cov\left(u'\left(c_{t+1}\right), x_{t+1}\right)}{u'\left(c_t\right)}}_{\text{risk adjustment}}
\end{aligned}
$$

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

$$
\begin{aligned}
1 &= \mathbb{E}\left(m_{t+1} \cdot R^i_{t+1}\right) \\
1 &= \mathbb{E}\left(m_{t+1}\right)\mathbb{E}\left(R^i_{t+1}\right) + cov\left(m_{t+1}, R^i_{t+1}\right) \\
R^f &= \frac{1}{\mathbb{E}\left(m_{t+1}\right)};\ 1 - \frac{1}{R^f}\mathbb{E}\left(R^i_{t+1}\right) = cov\Big(m_{t+1}, R^i_{t+1} \\
\mathbb{E}\left(R^i_{t+1}\right) - R^f &= -R^f \cdot cov\left(m_{t+1}, R^i_{t+1}\right) \\
\mathbb{E}\left(R^i_{t+1}\right) - R^f &= -\frac{1}{\mathbb{E}\left(\beta\frac{u'(c_{t+1})}{u'(c_t)}\right)} \cdot cov\left(\beta\frac{u'\left(c_{t+1}\right)}{u'\left(c_t\right)}, R^i_{t+1}\right)
\end{aligned}
$$

$$
\overbrace{\mathbb{E}\left(R^i_{t+1}\right) - R^f}^{\text{excess return}} = -\frac{cov\left(u'\left(c_{t+1}\right), R^i_{t+1}\right)}{\mathbb{E}\left(u'\left(c_{t+1}\right)\right)}
$$

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

The basic pricing equation has an expected return-beta representation

$$\mathbb{E}\left(R^i_{t+1}\right) - R^f \;=\; -R^f \cdot cov\left(R^i_{t+1}, m_{t+1}\right)$$

$$\mathbb{E}\left(R^i_{t+1}\right) - R^f \;=\; -\frac{cov\left(R^i_{t+1}, m_{t+1}\right)}{Var\left(m_{t+1}\right)} \frac{Var\left(m_{t+1}\right)}{\mathbb{E}\left(m_{t+1}\right)}$$

$$\mathbb{E}\left(R^i_{t+1}\right) = R^f - \left(\frac{cov\left(R^i_{t+1}, m_{t+1}\right)}{Var\left(m_{t+1}\right)}\right) \cdot \left(\frac{Var\left(m_{t+1}\right)}{\mathbb{E}\left(m_{t+1}\right)}\right)$$

asset specific quantity of risk

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 $\frac{c_{t+1}}{c_t}$

$$\begin{aligned} \mathbb{E}\left(R^i\right) \;&=\; R^f + \beta_{R^i,\Delta c} \cdot \lambda_{\Delta c} \\ \lambda_{\Delta c} \;&\approx\; \gamma \cdot Var\left(\Delta \ln c\right) \end{aligned}$$

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:

$$p_t u'(c_t) = \mathbb{E}_t\left(\beta\left(u'(c_{t+1})\right)\overbrace{(p_{t+1} + d_t)}^{x_{t+1}}\right)$$

Market efficiency $\Leftrightarrow$ Prices follow martingales (random walks)? **NO!**

Required:

- Risk neutral investors u'( )=const. or no variation in consumption
- $\beta = 1$ $\Leftarrow$ OK short time horizon
- no dividends

Then:

$$\begin{aligned} p_t &= \mathbb{E}(p_{t+1}) \\ p_{t+1} &= p_t + \varepsilon_{t+1} \\ \text{if} \quad \sigma^2(\varepsilon_{t+1}) &= \sigma^2 \end{aligned}$$

= Random Walk

$\Rightarrow$ Returns are not predictable $\mathbb{E}\left(\frac{p_{t+1}}{p_t}\right) = 1$

Marginal utility weighted prices follow martingales (2)

With risk aversion (but no dividends) and β=1

$$
\begin{aligned}
\tilde{p}_t &= \mathbb{E}(\tilde{p}_{t+1}) \\
\tilde{p}_t &= \tilde{p}_t \cdot u'(c_t)
\end{aligned}
$$

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

Predictability in the short horizon?

$$\left.\begin{array}{l}\text{consumption}\\ \text{risk aversion}\end{array}\right\} \text{does not change day by day}$$

⇒ Random Walks successful ⇒ Predictability of asset returns (day by day)?

Technical analysis, media reports...

Some popular linear factor models

Factor pricing models

return on wealth portfolio

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 vector

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

excess return small vs.
large stocks

- Fama French model

$$m_{t+1} = b_0 + b_m R^{em}_{t+1} + b_{SMB} SMB_{t+1} + b_{HML} HML_{t+1}$$

excess return value
stocks vs. growth stocks
(high book-to-markt – low
book-to-market)

Market excess return

´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 CONDTIONAL moment restrictions

$$\begin{aligned} p_t &= \mathbb{E}_t\left(m_{t+1}x_{t+1}\right) \\ &= \mathbb{E}\left(m_{t+1}x_{t+1} \mid I_t\right) \end{aligned}$$

$\{m_t\}$ 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

$$\begin{aligned} p_t &= \mathbb{E}_t\left(m_{t+1}x_{t+1}\right) \\ \mathbb{E}(p_t) &= \mathbb{E}\left(\mathbb{E}_t\left(m_{t+1}x_{t+1}\right)\right) \qquad \text{I.i.e.} \\ &= \mathbb{E}\left(m_{t+1}x_{t+1}\right) \end{aligned}$$

Estimation and evaluation of asset pricing models (Basics)

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)

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

$$\uparrow f(data, parameters)$$

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

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

Assume model correct: "Best" choice for β, γ?
Best "fit", smallest (average) pricing errors

Estimation and evaluation of asset pricing models. The basic idea.

Estimates $\widehat{b}$ from data, distribution of $\widehat{b}$?

Average pricing errors:

$$\text{sample mean } \underbrace{\text{(observed price - predicted price)}}_{\text{should be close to zero}} = \alpha$$

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

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

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

$$\underbrace{\frac{1}{T}\sum_{t=1}^{T} p_t - \frac{1}{T}\sum_{i=1}^{T} m_{t+1}(b)x_{t+1}}_{\alpha} \approx 0$$

GMM basic idea(first step):
choose $\widehat{b}$ to minimize α^2 (squared average pricing error) among set of test assets.

The two asset, two parameter case

$$\begin{aligned}
\mathbb{E}\left(m_{t+1}(\beta,\gamma)\, x^1_{t+1} - p^1_t\right) &= 0 \\
\mathbb{E}\left(m_{t+1}(\beta,\gamma)\, x^2_{t+1} - p^2_t\right) &= 0
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left(m_{t+1}(\beta,\gamma)\, R^1_{t+1} - 1\right) &= 0 \\
\mathbb{E}\left(m_{t+1}(\beta,\gamma)\, R^2_{t+1} - 1\right) &= 0
\end{aligned}$$

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} m_{t+1}(\beta,\gamma)\, R^1_{t+1} - 1 &= 0 \\
\frac{1}{T}\sum_{t=1}^{T} m_{t+1}(\beta,\gamma)\, R^2_{t+1} - 1 &= 0
\end{aligned}$$

solve equations for $\beta,\ \gamma\ \Rightarrow\ \widehat{\beta},\ \widehat{\gamma} \Rightarrow$

To apply GMM data have to be generated by stationary (and ergodic) processes (not necessarily i.i.d.)

Problem: WLLN works for **stationary data**:

(Weakly) stationary process: $\{Y_t\}_{t=-\infty}^{\infty}$

$\{\ldots, \mathrm{y}_0, y_1, \ldots, y_5, \ldots\}$

$\mathbb{E}(Y_t) = u$

$var(Y_t) = \sigma^2$

$cov(Y_t, Y_{t-j}) = \gamma_j$

Solution: $\Rightarrow$ We use:

$$1 = \mathbb{E}\left(m_{t+1}(b) \cdot R_{t+1}\right) \qquad \text{instead of} \qquad \mathbb{E}(p_t) = \mathbb{E}\left(m_{t+1}(b) \cdot x_{t+1}\right)$$

$$0 = \mathbb{E}\left(m_{t+1}(b) \cdot R_{t+1} - 1\right)$$

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
${R_t}^1, {R_t}^2, \cdots, {R_t}^N \quad t = 1, \cdots, T$

$$\begin{bmatrix} \mathbb{E}_T[u_t^1(b)] \\ \mathbb{E}_T[u_t^2(b)] \\ \vdots \\ \vdots \\ \mathbb{E}_T[u_t^N(b)] \end{bmatrix} = g_T(b) \quad N \times 1 \quad \text{vector}$$

If $\sharp$ assets $= \sharp$ parameters b can be chosen such that average pricing errors are zero usually $\sharp$ assets $> \sharp$ parameters.

GMM objective function

$$\widehat{b} = \underset{\{b\}}{\text{argmin}} \quad g_T'(b) \cdot I_N \cdot g_T(b) \qquad \text{first step } \textbf{GMM estimate}$$

$$= \underset{\{b\}}{\text{argmin}} \quad \left[\mathbb{E}_T[u_{t+1}^1(b)]\right]^2 + \left[\mathbb{E}_T[u_{t+1}^2(b)]\right]^2 + \ldots + \left[\mathbb{E}_T[u_{t+1}^N(b)]\right]^2$$

$\Rightarrow$ minimize sum of squared average (pricing)errors

equal weight for all test assets $1, \ldots, N$

Alternatively other weight matrix

$$\widehat{b} = \underset{\{b\}}{\text{argmin}} \quad g_T'(b)\, W\, g_T(b) \quad \text{e. g. } W = \begin{bmatrix} 1 & 0 & & \\ 0 & 2 & & \\ & & 100 & \ddots \\ 0 & & & \end{bmatrix}$$

Under mild assumptions (stationarity) GMM estimators have desirable properties

GMM estimators consistent:
Bias and variance of estimator go to zero asymptotically $\widehat{b} \underset{p}{\rightarrow} b$

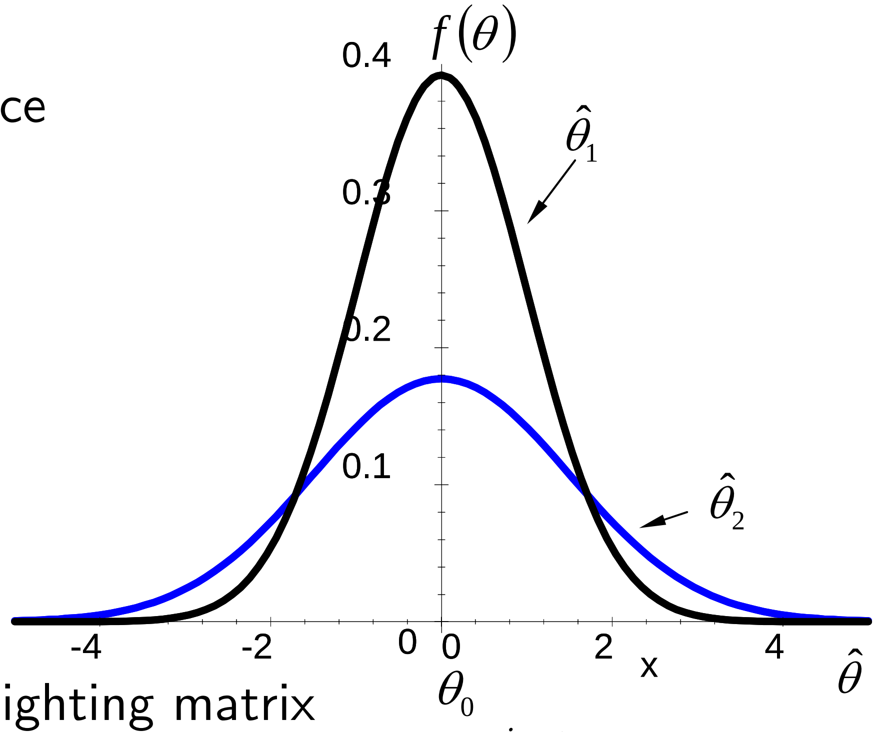
GMM estimators asymptotically normal. Required for inference:

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

To conduct $t-$test: $\frac{\widehat{b}_k}{\widehat{\sigma} k} \overset{a}{\sim} N(0,1)$

Efficient estimates obtained by using the optimal weighting matrix

Efficiency: Smallest asymptotic variance
among GMM esimators



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

$$S = var(g_T(b)) = \mathbb{E}(g_T(b)g_T(b)') = \mathbb{E}(u_t(b)u_t(b)') \quad \text{resp.} \quad = \underbrace{\sum_{j=-\infty}^{j=+\infty} \mathbb{E}(u_t(b)u_{t-j}(b)')}_{\text{with serial correlation in moment conditions}}$$

variance-covariance
matrix average pricing errors

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

There exists an optimal weighting matrix

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

$$\widehat{b} = \underset{\{b\}}{argmin} \quad g_T(b)' \, \widehat{S}^{-1} \, g_T(b)$$

write $u_t(b) = \begin{pmatrix} u_t^1(b) \\ \vdots \\ u_t^N(b) \end{pmatrix} \qquad \Big(\underset{i=assets}{u_t^i(b) = m_{t+1}(b) x_{t+1}^i - p_t^i} \Big)$

Recall: $\mathbb{E}(u_t^i) = 0 \quad \Rightarrow \quad \mathbb{E}(u_t(b)) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

The optimal weighing matrix takes into account variances and covariances of pricing errors across assets

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

With no serial correlation in pricing errors!

$S=$ variance covariance matrix of pricing errors

$$= \begin{bmatrix} var\left(u_t^1(b)\right) \cdots & & \\ cov\left(u_t^1(b) u_t^2(b)\right) & var\left(u_t^2(b)\right) & \ddots \\ \vdots & & \\ & & var\left(u_t^N(b)\right) \end{bmatrix}$$

Estimate $\widehat{S}$: Replace $\mathbb{E}$ by $\frac{1}{N}\sum$ using $\widehat{b}$ obtained with weighting matrix $I_N \Rightarrow \widehat{S}$.

Steps of iterated GMM estimation

1) $\widehat{b}^1 = \underset{\{b\}}{argmin} \quad g_T(b)' \, I_N \, g_T(b) \Rightarrow$

2) $\widehat{S} \Rightarrow$

3) $\widehat{b}^2 = \underset{\{b\}}{argmin} \quad \mathrm{g}_T(b)'\widehat{S}^{-1} \, g_T(b)$

...repeat... ...

Intuition behind optimal weighting matrix (1)

Intuition behind GMM weighting matrix

Example

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

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

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

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

$$\underset{\{b\}}{argmin}\, \mathbb{E}_T\left[u_t^1(b)\right]^2 \cdot W_1 + \mathbb{E}_T\left[u^2(b)\right]^2 \cdot W_2$$

Example

$W_1 :\ 0.1 \Rightarrow var\left(u_t^1(b)\right) = 10$

$W_2 :\ 10 \Rightarrow var\left(u_t^2(b)\right) = 0.1$

$\Rightarrow$ 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.

$$\text{Example } S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.999 \\ 0 & 0.999 & 1 \end{pmatrix} \qquad cov(u_t^2, u_t^3) = 0.999$$

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

$$\underset{\{b\}}{argmin} \left[\mathbb{E}_T(u_t^1(b)),\ \mathbb{E}_T(u_t^2(b)), \mathbb{E}_T(u_t^3(b)) \right] \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.99 \\ 0 & 0.99 & 1 \end{bmatrix}^{-1} \times$$

$$\begin{bmatrix} \mathbb{E}_T(u_t^1(b)) \\ \mathbb{E}_T(u_t^2(b)) \\ \mathbb{E}_T(u_t^3(b)) \end{bmatrix}$$

Some more intuition behind optimal weighting matrix: Correlations across pricing errors (2)

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

$$\underset{\{b\}}{argmin}\; g_T(b)'\, S^{-1} g_T(b) =$$

$$\Big[\mathbb{E}_T\left(u_t^1(b)\right), \mathbb{E}_T\left(u_t^2(b)\right)\cdot 500.25 - \mathbb{E}_T\left(u_t^3(b)\right)\cdot 499.75,$$

$$\mathbb{E}_T\left(u_t^3(b)\right)\cdot 500.75 - \mathbb{E}_T\left(u_t^2(b)\right)\cdot 499.75\Big] \times \begin{bmatrix} \mathbb{E}_T(u_t^1(b)) \\ \mathbb{E}_T(u_t^2(b)) \\ \mathbb{E}_T(u_t^3(b)) \end{bmatrix}$$

Some more intuition behind optimal weighting matrix: Correlations of pricing errors (3)

$$\underset{\{b\}}{argmin}\; g_T(b)'\, S^{-1} g_T(b) =$$

$$\mathbb{E}_T\left(u_t^1(b)\right)^2 + \mathbb{E}_T\left(u_t^2(b)\right)^2 \cdot 500.25 + \mathbb{E}_T\left(u_t^3(b)\right)^2 \cdot 500.25 -$$
$$2 \cdot \mathbb{E}_T\left(u_t^2(b)\right) \mathbb{E}_T\left(u_t^3(b)\right) \cdot 499.75$$

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

since

$$\mathrm{E}_T\left(u_t^2(b)\right) \approx \mathbb{E}_T\left(u_t^3(b)\right)$$

To test hypotheses we need the distribution of the GMM estimates

Standard errors of GMM estimates

We want:

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

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

$$t = \frac{\widehat{b}_k - 0}{\sqrt{var(\widehat{b}_k)}} \overset{a}{\sim} N(0, 1) \text{ under } H_0 : \; b_k = 0$$

Asyptotic distribution of GMM estimates when using optimal weighting matrix

References for nonlinear GMM results: Hayashi (2000) Econometrics, Chapter 6, Hall (2005)

$$\sqrt{T}(\widehat{b}-b) \underset{d}{\rightarrow} N(0,(d' S^{-1} d)$$

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

consistently estimated by $\widehat{d} = \frac{\partial g_T(b)}{\partial b}|\widehat{b}$

t- and Wald tests use $\widehat{var(\widehat{b})} = \frac{\widehat{d}'\widehat{S}^{-1}\widehat{d}}{T}$

Details

Some more details:

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

b) Recall

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

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

Details

$$\frac{\partial g_T(b)}{\partial b} = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} & \frac{\partial m_t(b)}{\partial b_1} R_t, \frac{\cdots}{\partial b_2} \cdots \\ \downarrow & \longrightarrow \\ & \text{Parameters} \\ N & \end{bmatrix}$$

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

$var(\widehat{b})$ used for testing hypotheses:

$H_0: \quad b_k = 0$

t-statistic: $\frac{\widehat{b}_k - 0}{\sqrt{var(\widehat{b}_k)}} \overset{a}{\sim} N(0,1) \,\widehat{=}$ Standard t-test.

joint significance:

$$H_0 : \underbrace{(b_{j1} = b_{j2} = \quad = b_{jN}}_{\text{some subset of } b} = \; 0) \text{ or } \underset{J\times 1}{b_J} = 0$$

$$\left.\widehat{b}_j' \left[\underbrace{var(\widehat{b})_J}\right]^{-1} \widehat{b}_j \overset{a}{\sim} \chi^2(J)\right] \widehat{=}$$ Standard Wald test use to test *Rb*=*r*

appropriate subset of $var(\widehat{b})$

Nonlinear restrictions testable applying delta method => EVIEWS example

Testing the validity of the model (moment conditions) by J-test

$\{R_t, \Delta c_t, \ldots\}$ data is a random sample $\Rightarrow \widehat{b}$ is a random variable $\Rightarrow$

$u_t(b)$ is a random variable $\Rightarrow \mathbb{E}_T\left(u_t(b)\right) = \frac{1}{N}\sum\cdots$ is a random variable

pricing errors too large to be explained by random sampling?

$\Leftrightarrow$ Is the model in correct?

$$T \cdot J_T = T \cdot \underbrace{\left[g_T(\widehat{b})'\, \widehat{S}^{-1}\, g_T(\widehat{b})\right]}_{\text{objective function at minimum using optimal weighting matrix estimate}} \overset{a}{\sim} \chi^2\left(\begin{array}{l}\text{no. moment conditions}\\ \text{- no. of parameters.}\end{array}\right)$$

$\Rightarrow$ Reject or non-reject model (i.e. moment conditions) at given significance level
Example: no. of moment conditions: 10, no. parameters: 2,

$$T J_T = 7.9,\ \chi^2_{95}(1) = 2.73 \Rightarrow$$

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)

Chooese W to be positive semi-definite and symmetric

$$\widehat{b} = \underset{\{b\}}{\arg\min}\; g_T(b)' \overset{\swarrow}{W} g_T(b)$$

$$\underbrace{\frac{\partial g_T(b)'}{\partial b} W}_{\text{K linear combinations set to zero}} \qquad \times g_T(b) = 0$$

$$N \times 1 \text{ vector of r.v. with } \underbrace{g_T(\widehat{b})}_{} \; K \text{ linear dependencies}$$

General GMM results (Hayashi Ch. 6)

$$\sqrt{T}(\widehat{b}-b) \underset{d}{\rightarrow} N(0,(d'Wd)^{-1}d'WSWd(d'Wd)^{-1})$$

For *t*- and Wald-tests use

$$\widehat{var(\widehat{b})} = \frac{(\widehat{d}'W\widehat{d})^{-1}\widehat{d}'W\widehat{S}W\widehat{d}(\widehat{d}'W\widehat{d})^{-1}}{T}$$

General GMM results (Hayashi Ch. 6)

$$\sqrt{T} g_T(\widehat{b}) \underset{d}{\rightarrow} N(0, Avar(g_T(\widehat{b})))$$

$$Avar(g_T(\widehat{b})) = (I - d(d'Wd)^{-1}d'W)S(I - d(d'Wd)^{-1}d'W$$

General form of J-statistic

Pseudo inverse, linear dependencies in g by construction, V-C matrix singular

$$Tg_T(\widehat{b})'[\widehat{Avar(g_T(\widehat{b}))}]^{+} g_T(\widehat{b}) \underset{d}{\rightarrow} \chi(N-K)$$

Performance comparison (1)

Problems using J-statistic

Popular measure

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

From

$$
\begin{aligned}
1 =& \ \mathbb{E}(mR) \\
1 =& \ \mathbb{E}(m)\mathbb{E}(R) + cov(m, R) \\
\mathbb{E}(R) =& \ \tfrac{1}{\mathbb{E}(m)} - \tfrac{cov(m,R)}{\mathbb{E}(m)}
\end{aligned}
$$

Use as predictor

$$
\widehat{\mathbb{E}(R)} = \frac{1}{\frac{1}{T}\sum\limits_{t=1}^{T} m_t} - \frac{\frac{1}{T}\sum\limits_{t=1}^{T} m_t R_t - \frac{1}{T}\sum\limits_{t=1}^{T} m_t \frac{1}{T}\sum\limits_{t=1}^{T} R_t}{\frac{1}{T}\sum\limits_{t=1}^{T} m_t}
$$

Performance comparison (2)

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

Similarly using excess returns as test assets

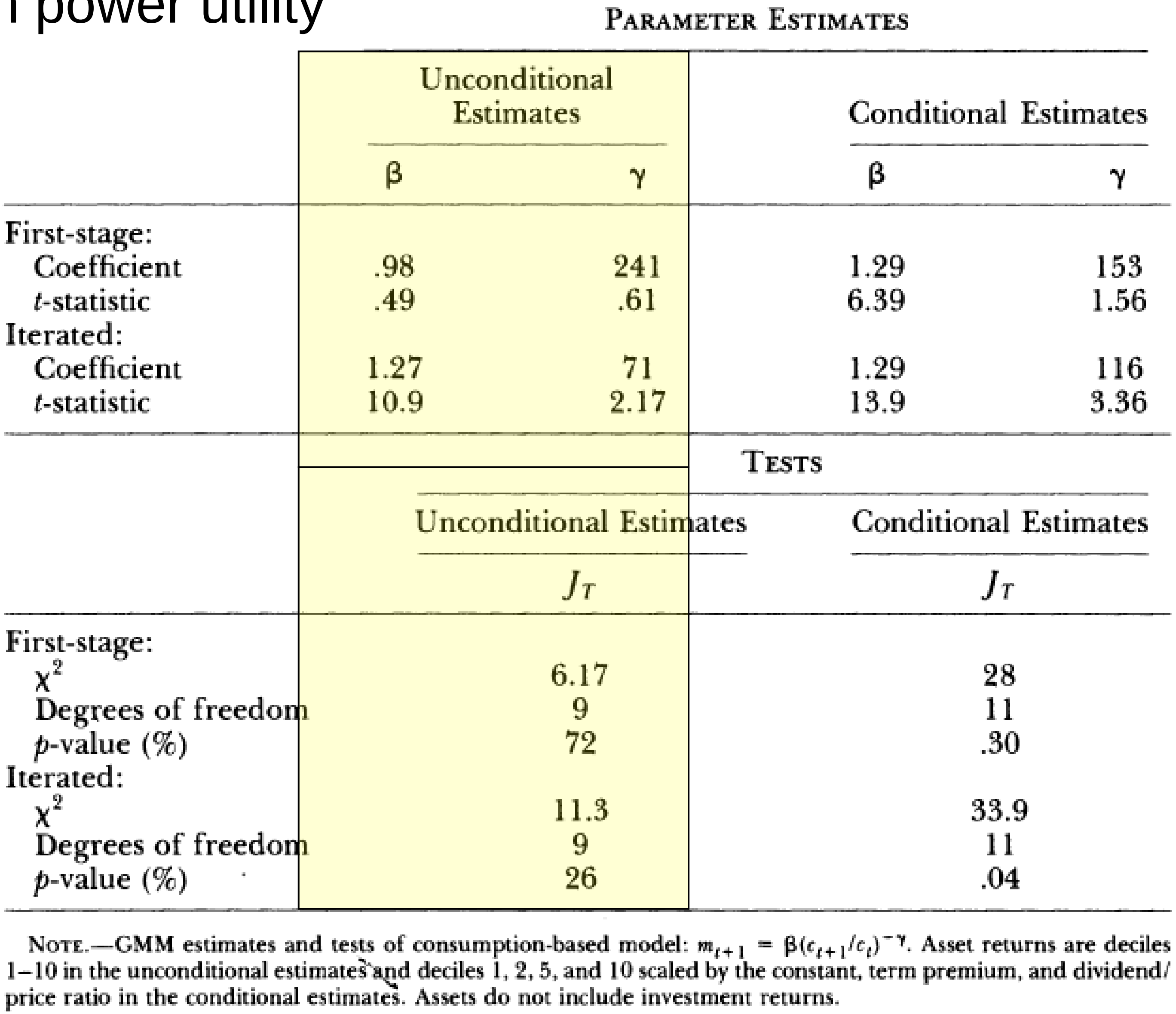
$$
\begin{aligned}
\text{From} \qquad 0 =&\ \mathbb{E}(mR^e) \\
0 =&\ \mathbb{E}(m)\mathbb{E}(R^e) + cov(m, R^e) \\
\mathbb{E}(R^e) =&\ -\tfrac{cov(m,R^e)}{\mathbb{E}(m)}
\end{aligned}
$$

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

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

RMSE $= \sqrt{\sum\limits_{j=1}^{N}\left[\widehat{\mathbb{E}(R^j)} - \bar{R}^j\right]^2}$ or $= \sqrt{\sum\limits_{j=1}^{N}\left[\widehat{\mathbb{E}(R^{ej})} - \bar{R}^{ej}\right]^2}$ used to rank and compare alternative models

Cochrane‘s (1996) estimation results for the consumption based model with power utility



PARAMETER ESTIMATES

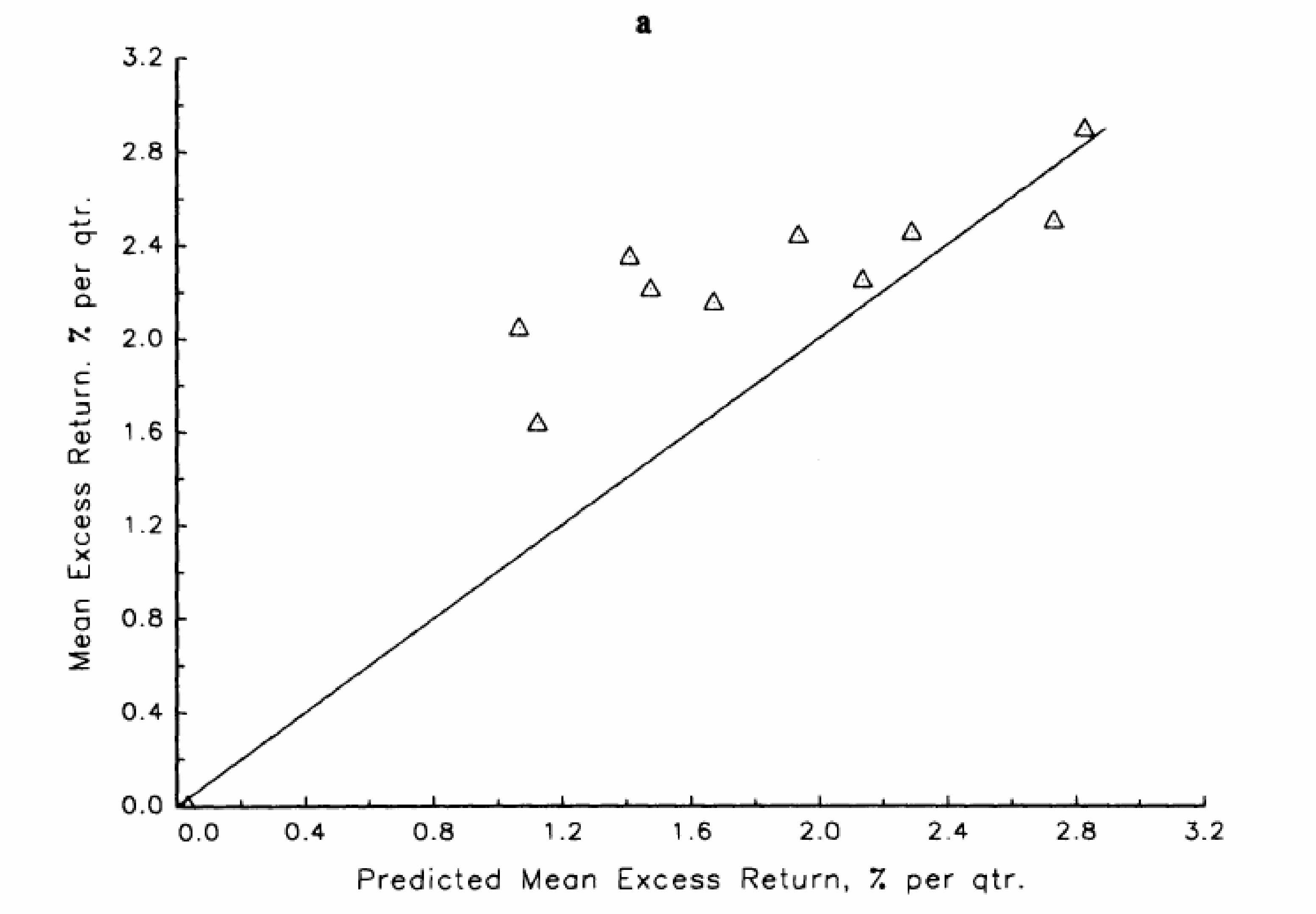
	Unconditional Estimates		Conditional Estimates	
	β	γ	β	γ
First-stage:				
Coefficient	.98	241	1.29	153
t-statistic	.49	.61	6.39	1.56
Iterated:				
Coefficient	1.27	71	1.29	116
t-statistic	10.9	2.17	13.9	3.36

TESTS

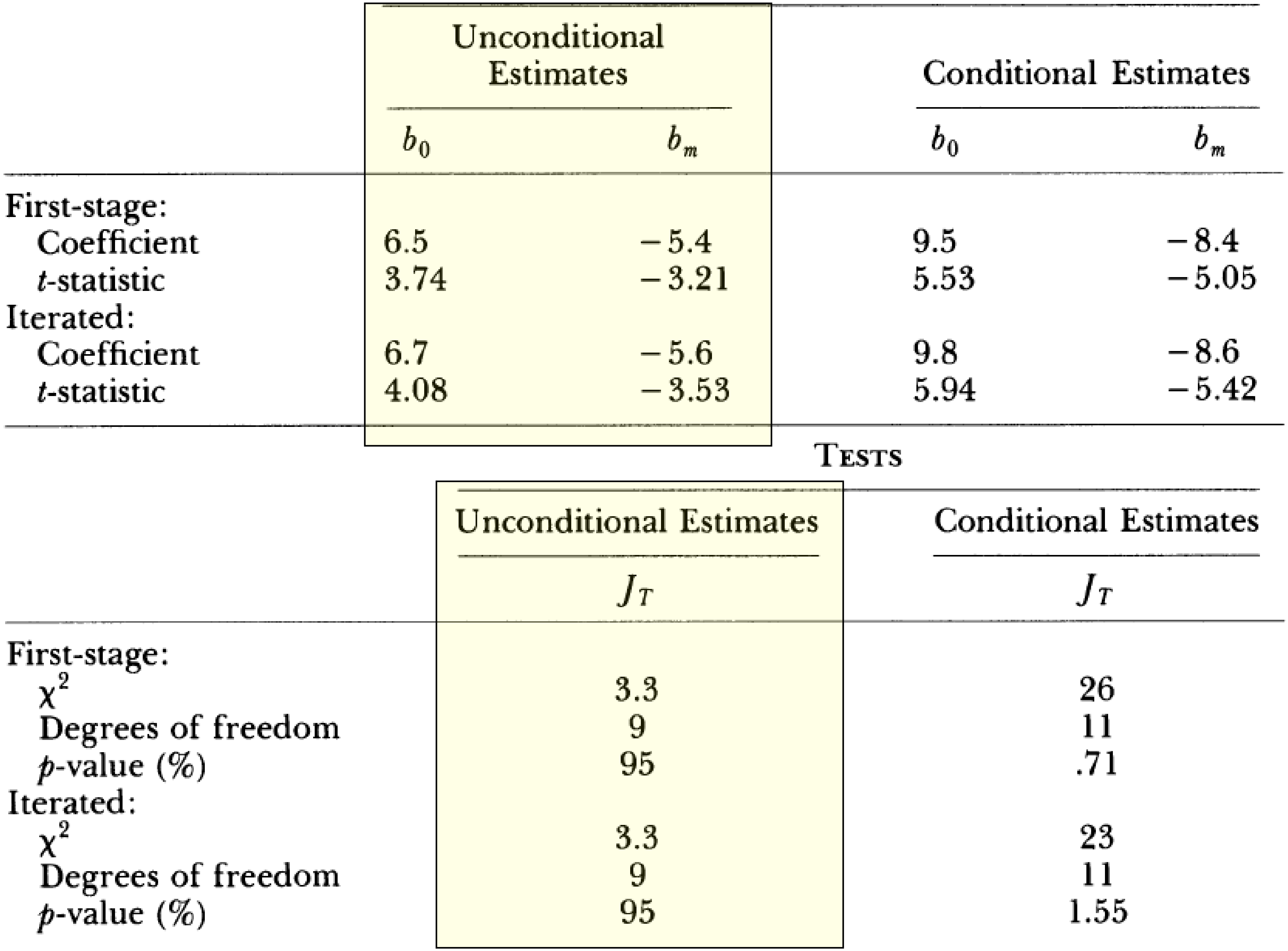
	Unconditional Estimates	Conditional Estimates
	J_T	J_T
First-stage:		
χ^2	6.17	28
Degrees of freedom	9	11
p-value (%)	72	.30
Iterated:		
χ^2	11.3	33.9
Degrees of freedom	9	11
p-value (%)	26	.04

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.

Non-rejection doesn‘t mean a thing



Cochrane‘s (1996) results for unconditional estimation of CAPM

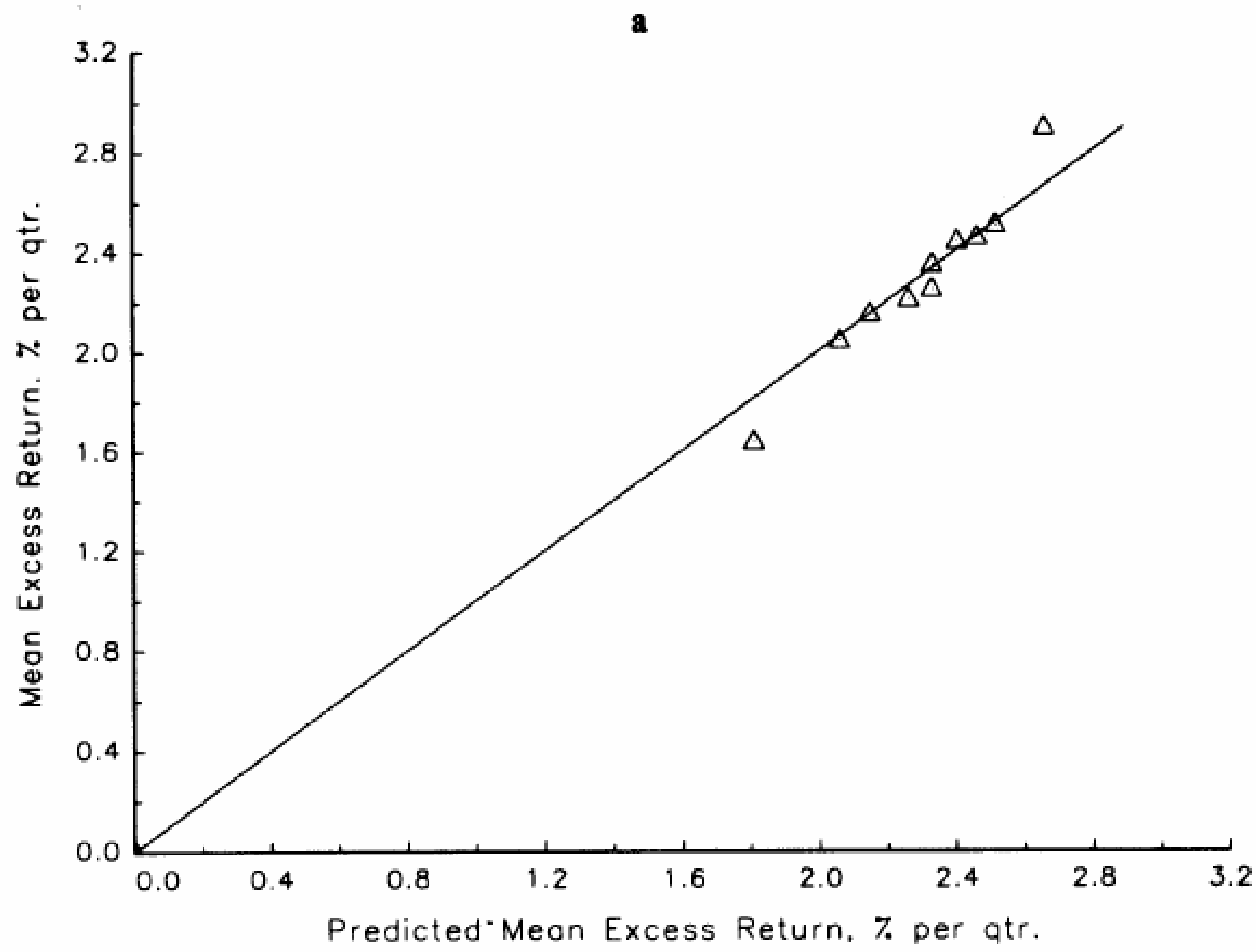


	Unconditional Estimates		Conditional Estimates	
	b_0	b_m	b_0	b_m
First-stage:				
Coefficient	6.5	−5.4	9.5	−8.4
t-statistic	3.74	−3.21	5.53	−5.05
Iterated:				
Coefficient	6.7	−5.6	9.8	−8.6
t-statistic	4.08	−3.53	5.94	−5.42

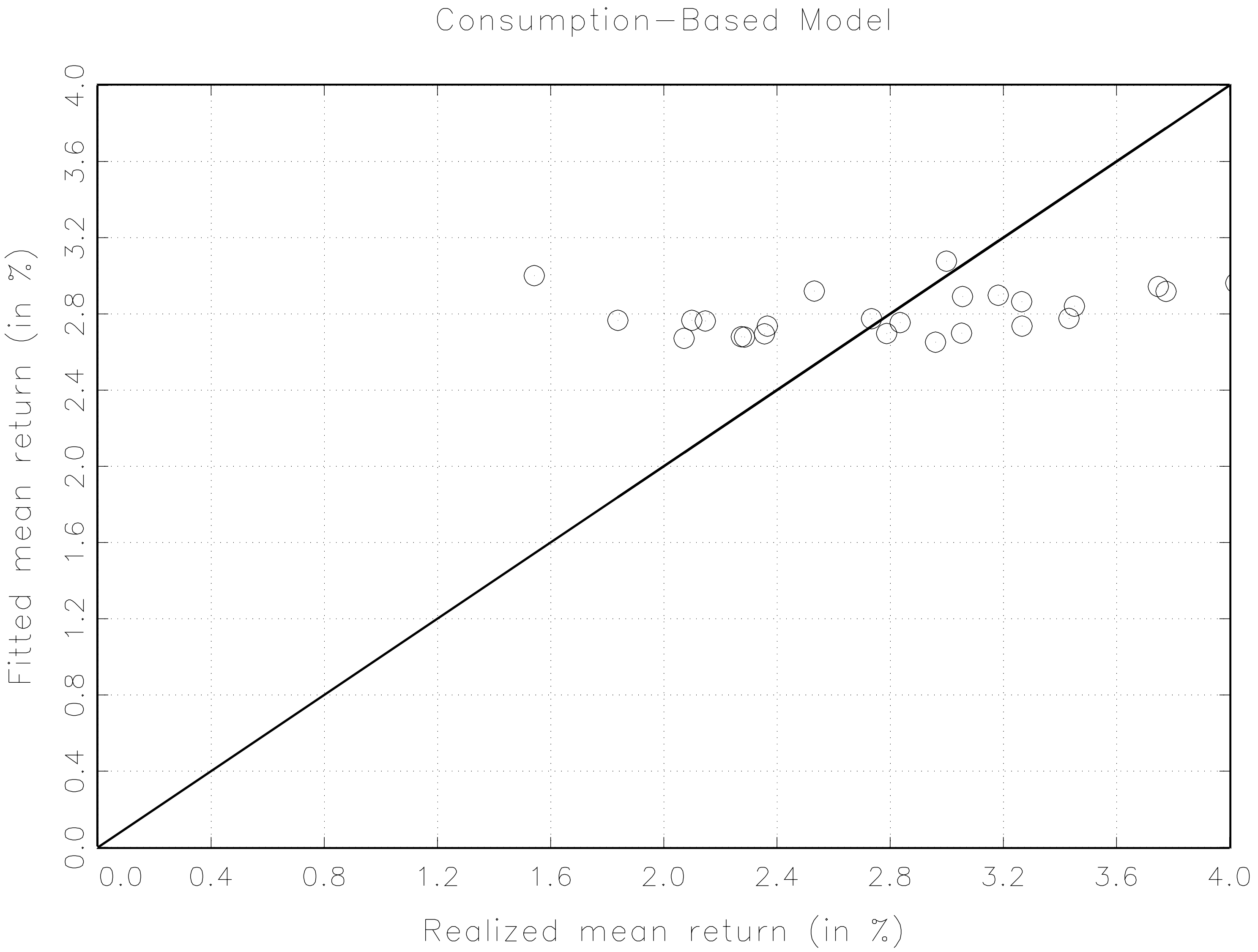
TESTS

	Unconditional Estimates	Conditional Estimates
	J_T	J_T
First-stage:		
χ^2	3.3	26
Degrees of freedom	9	11
p-value (%)	95	.71
Iterated:		
χ^2	3.3	23
Degrees of freedom	9	11
p-value (%)	95	1.55

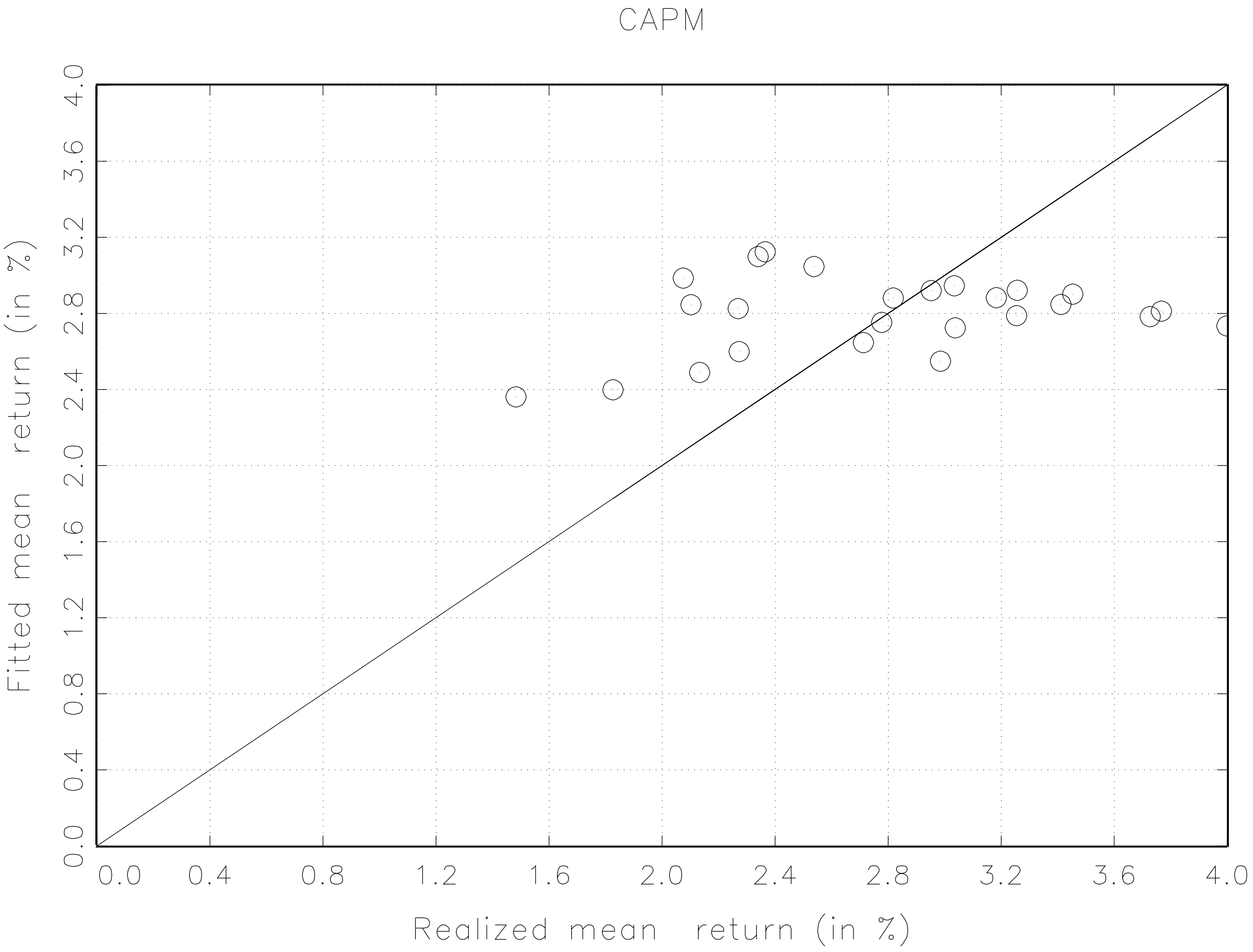
Cochrane‘s (1996) results for unconditional estimation of CAPM



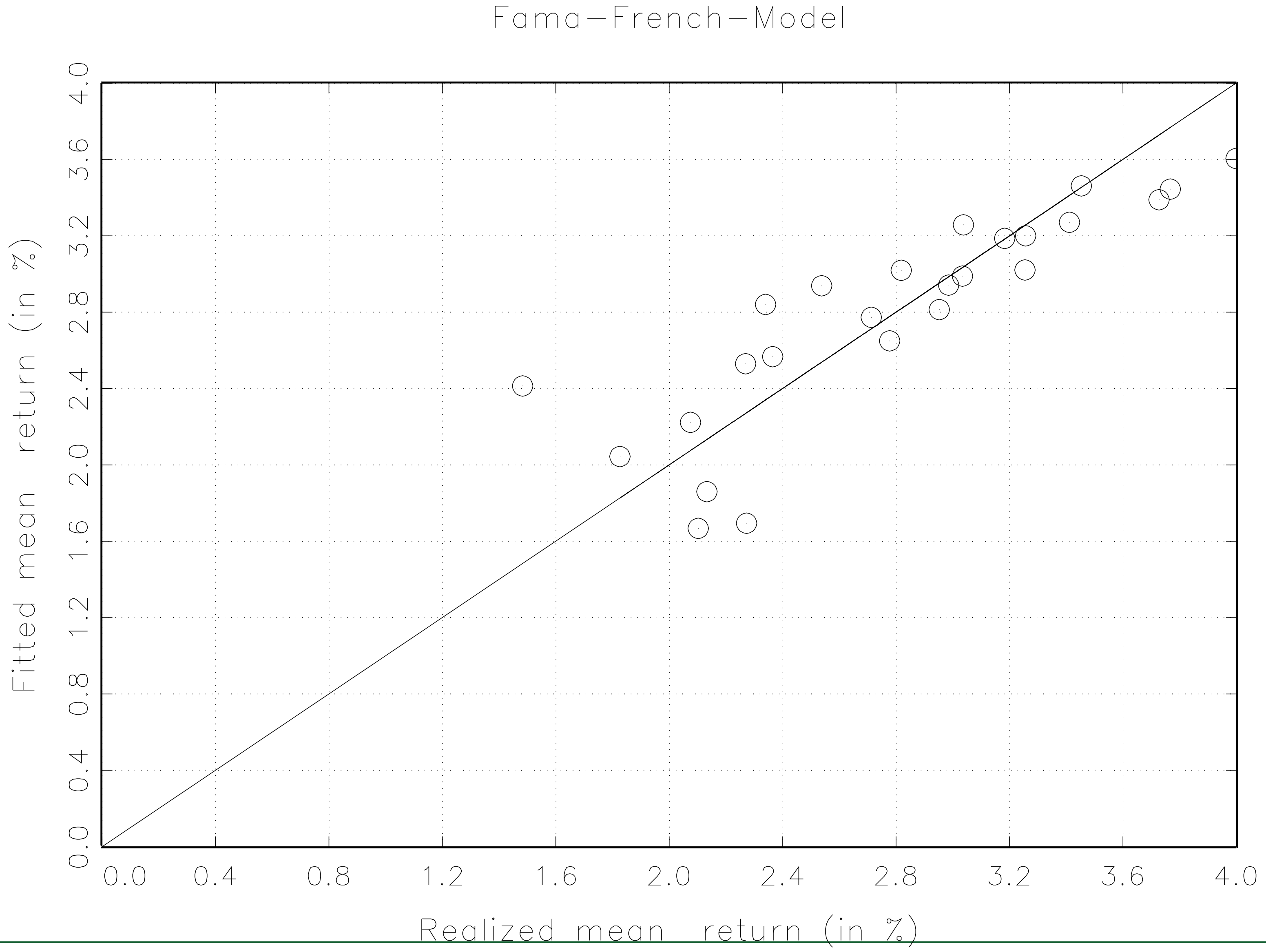
Performance comparison. Example: Consumption-Based Model estimated on 25 Fama-French portfolios



Performance comparison. Example: CAPM estimated on 25 Fama-French portfolios



Performance comparison. Example: Fama-French two factor model estimated on 25 Fama-French portfolios

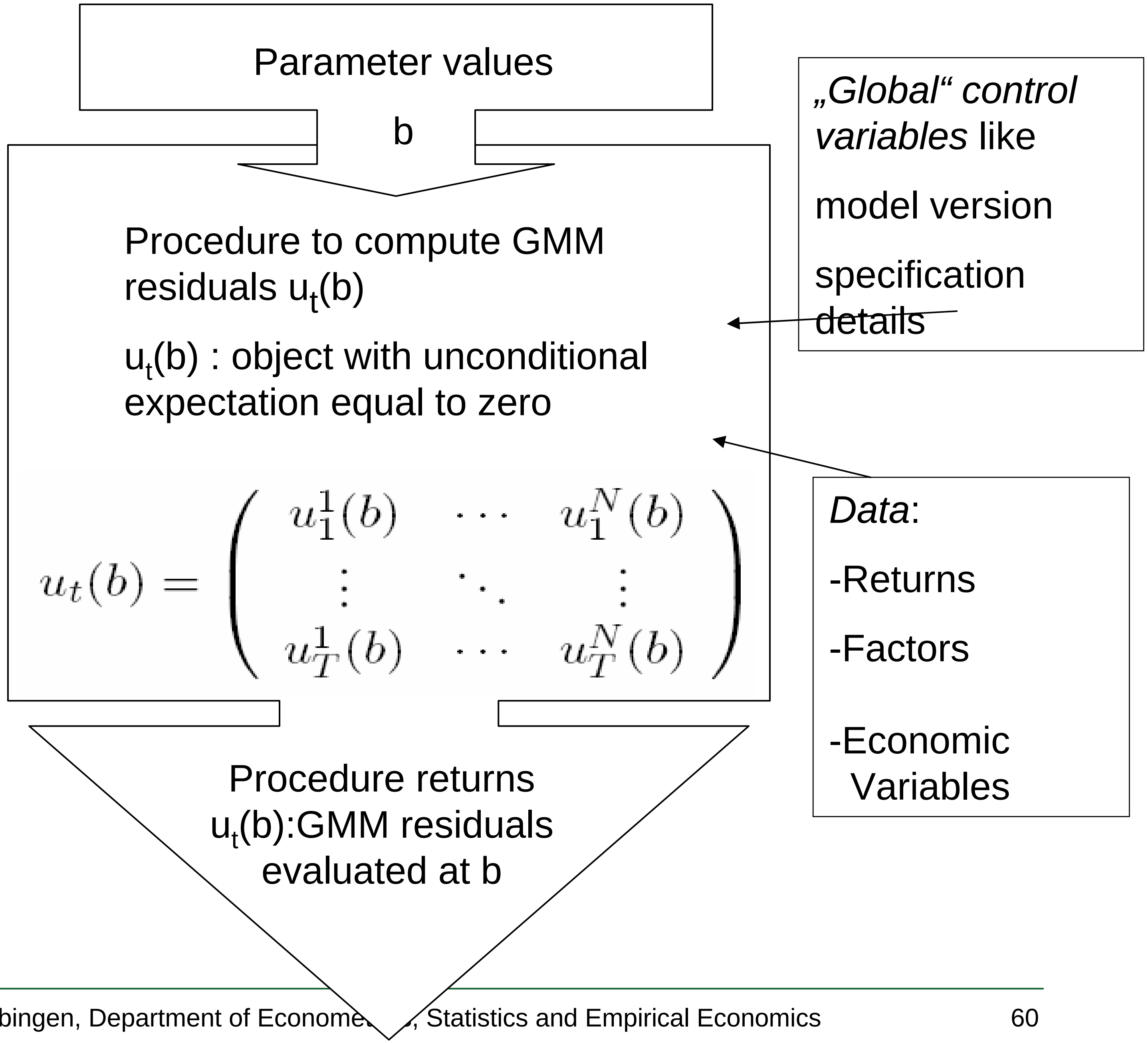


GMM estimation using the Gauss library: Ingredients and recipe

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

   iteratively calls procedure to compute GMM residuals $u_t(b)$

5. Check parameter estimates and test statistics



Parameter values

b

Procedure to compute GMM residuals $u_t(b)$

$u_t(b)$: object with unconditional expectation equal to zero

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

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

„Global“ control variables like

model version

specification details

Data:

-Returns

-Factors

-Economic Variables

The canoncical 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 γ and β using EXCEL solver.

Input: Time series of returns and consumption growth.

$$\begin{bmatrix} R_1^1 & \cdots & R_1^{10} & R_1^f & dc_1 \\ \vdots & & \vdots & & \vdots \\ R_T^1 & & R_1^{10} & R_1^f & dc_T \end{bmatrix}$$

3. Recent approaches

Readings: Lettau and Ludvigson (2001), Garcia, Renault and Semonov (2002), Yogo (2006)

Newer models consumption based model and habit formation

Garcia et al. (2003)

Period utility function

habit level (external)

$$u(c_t/H_t, H_t) = \frac{\left(\frac{c_t}{H_t}\right)^{1-\gamma} H_t^{1-\psi} - 1}{1-\gamma}$$

Marginal utility

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

Stochastic discount factor

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

$$\mathbb{E}_t \left[\delta \left(\frac{c_{t+1}}{c_t}\right)^{-\gamma} \left(\frac{H_{t+1}}{H_t}\right)^{\gamma-\psi} R^i_{t+1}\right] = 1$$

Modelling the habit level (1)

$$\begin{aligned}
H_{t+1} &= \mathbb{E}(c_{t+1}|c_t, c_{t-1}, \ldots) \\
\Delta H_{t+1} &= \lambda(c_t - H_t) \qquad 0 \leq \lambda \leq 1 \\
H_{t+1} &= a + \lambda c_t + (1-\lambda)H_t \\
H_{t+1} &= \frac{a}{\lambda} + \lambda \sum_{i=0}^{\infty} (1-\lambda)^i c_{t-i}
\end{aligned}$$

using

$$\begin{aligned}
c_{t+1} &= \frac{a}{\lambda} + \lambda \sum_{i=0}^{\infty} (1-\lambda)^i c_{t-i} + \varepsilon_{t+1} \\
c_{t+1} &= \frac{a}{\lambda} + \lambda c_t + \lambda(1-\lambda)c_{t-1} + \lambda(1-\lambda)^2 c_{t-2} + \ldots + \varepsilon_{t+1} \\
(1-\lambda)c_t &= \frac{a}{\lambda}(1-\lambda) + \lambda(1-\lambda)c_{t-1} + \ldots + (1-\lambda)\varepsilon_t
\end{aligned}$$

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 λ 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)

log return market portfolio

$$\begin{aligned}\Delta h_{t+1} &= \ln(H_{t+1}) - ln(H_t) \\ \Delta h_{t+1} &= a_0 + \sum_{i=1}^{n} a_i \cdot \Delta \ln c_{t+1-i} + b \cdot r^m_{t+1}\end{aligned}$$

with

orthogonal forecast error

$$\begin{aligned}\Delta h_{t+1} &= \mathbb{E}(\Delta \ln c_{t+1} | \Delta \ln c_t, \Delta \ln c_{t-1}, \ldots) \\ \Delta \ln c_{t+1} &= a_0 + \sum_{i=1}^{n} a_i \cdot \Delta \ln c_{t+1-i} + b \cdot r^m_{t+1} + \varepsilon_{t+1}\end{aligned}$$

$a_0, a_1, \ldots, b$ 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

$$\begin{aligned} \mathbb{E}(m_{t+1}R^i_{t+1} - 1) &= 0 \\ &\vdots \\ \mathbb{E}(m_{t+1}R^N_{t+1} - 1) &= 0 \end{aligned}$$

along with

$$\begin{aligned} \mathbb{E}(\varepsilon_{t+1}r^m_{t+1}) &= 0 \\ \mathbb{E}(\varepsilon_{t+1}\Delta \ln c_t) &= 0 \\ &\vdots \end{aligned}$$

An alternative model for the habit process (3)

Habit growth is then

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

exp(a_0)

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

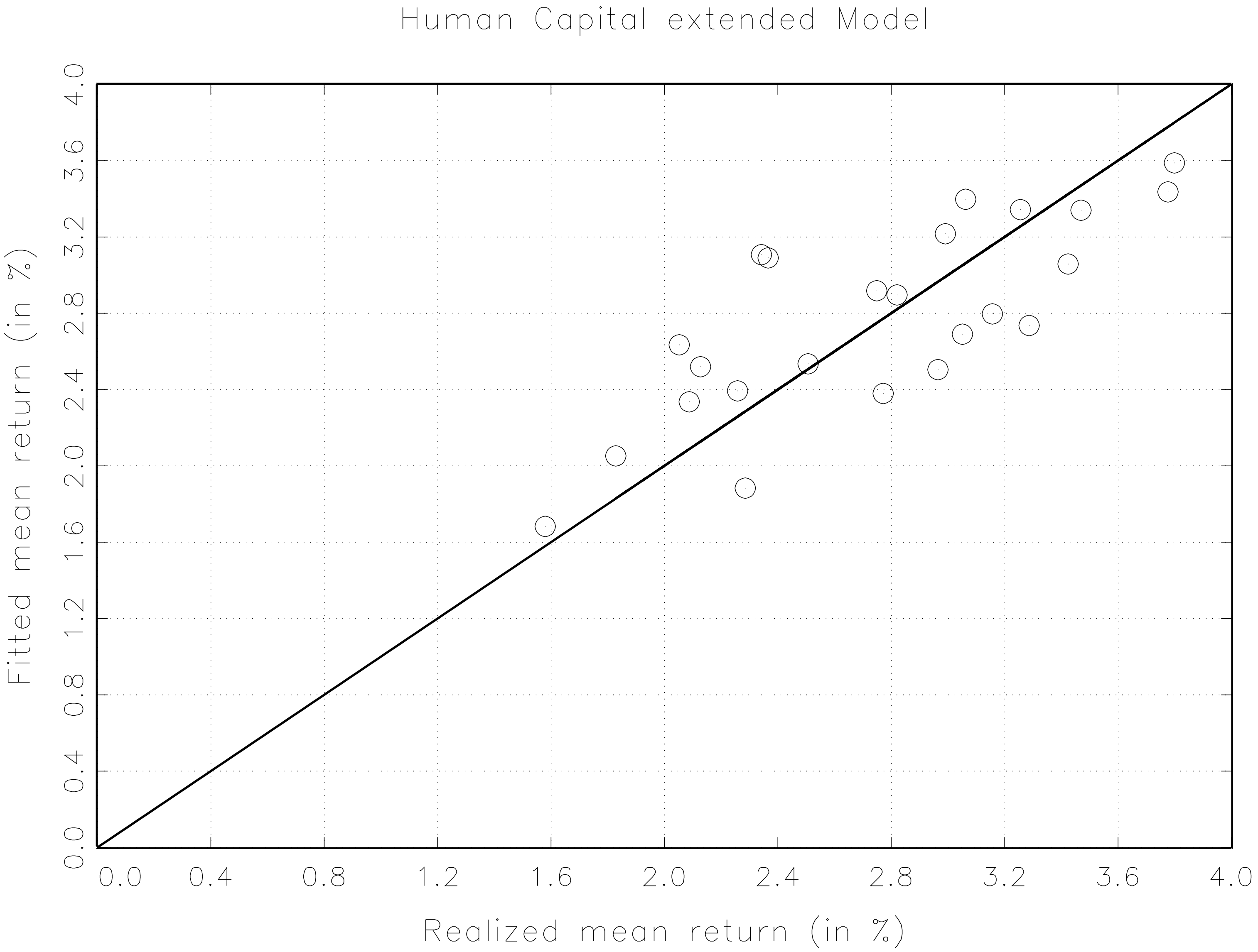
$$m_{t+1} = \delta^* \left[\frac{c_{t+1}}{c_t}\right]^{-\gamma} \prod_{i=0}^{n} \left[\frac{c_{t+1-i}}{c_{t-i}}\right]^{\frac{a_i \cdot \kappa}{b}} \left(R_{t+1}^m\right)^{\kappa}$$

We estimate using

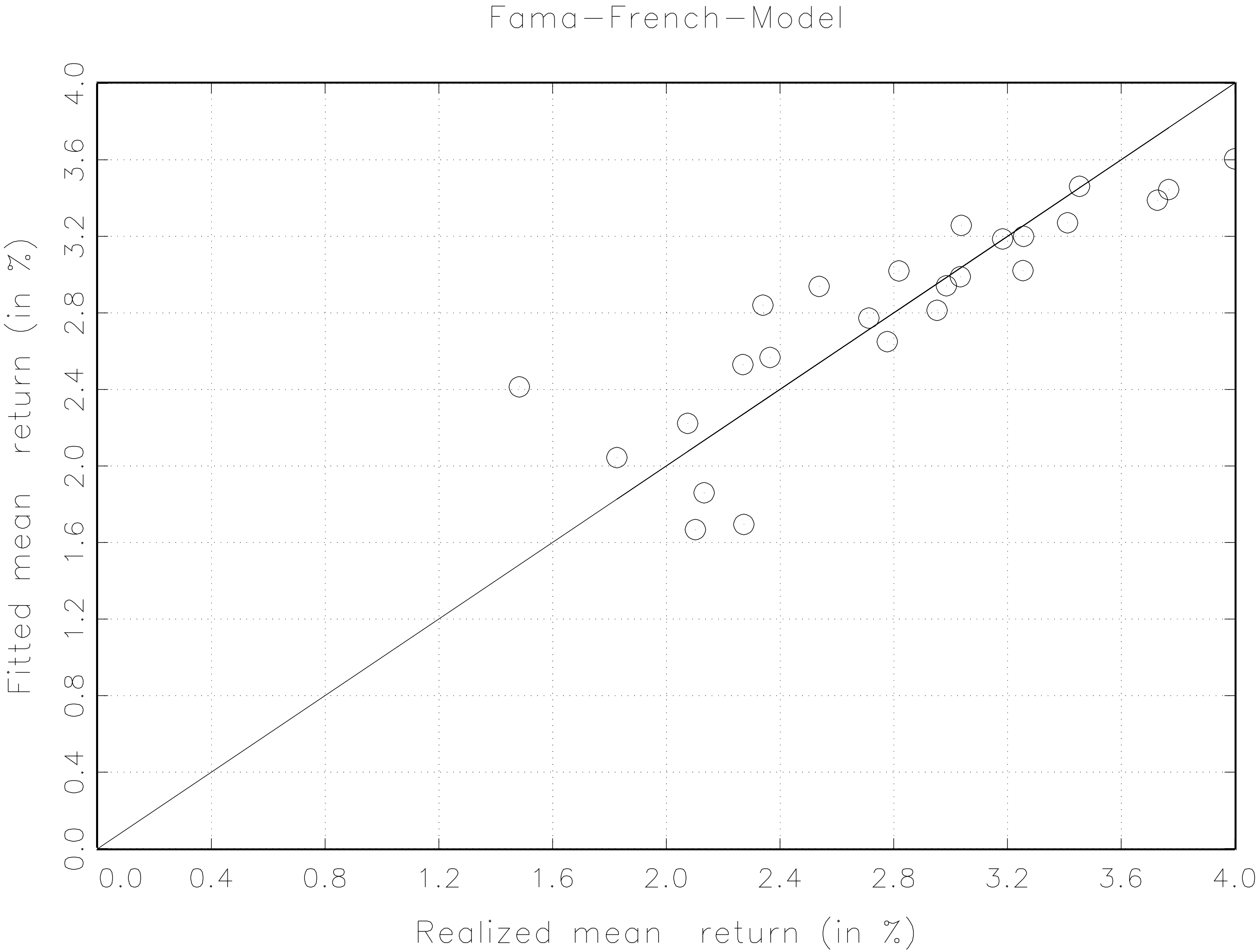
$n = 0$ "Epstein-Zin SDF"

$n = 1$

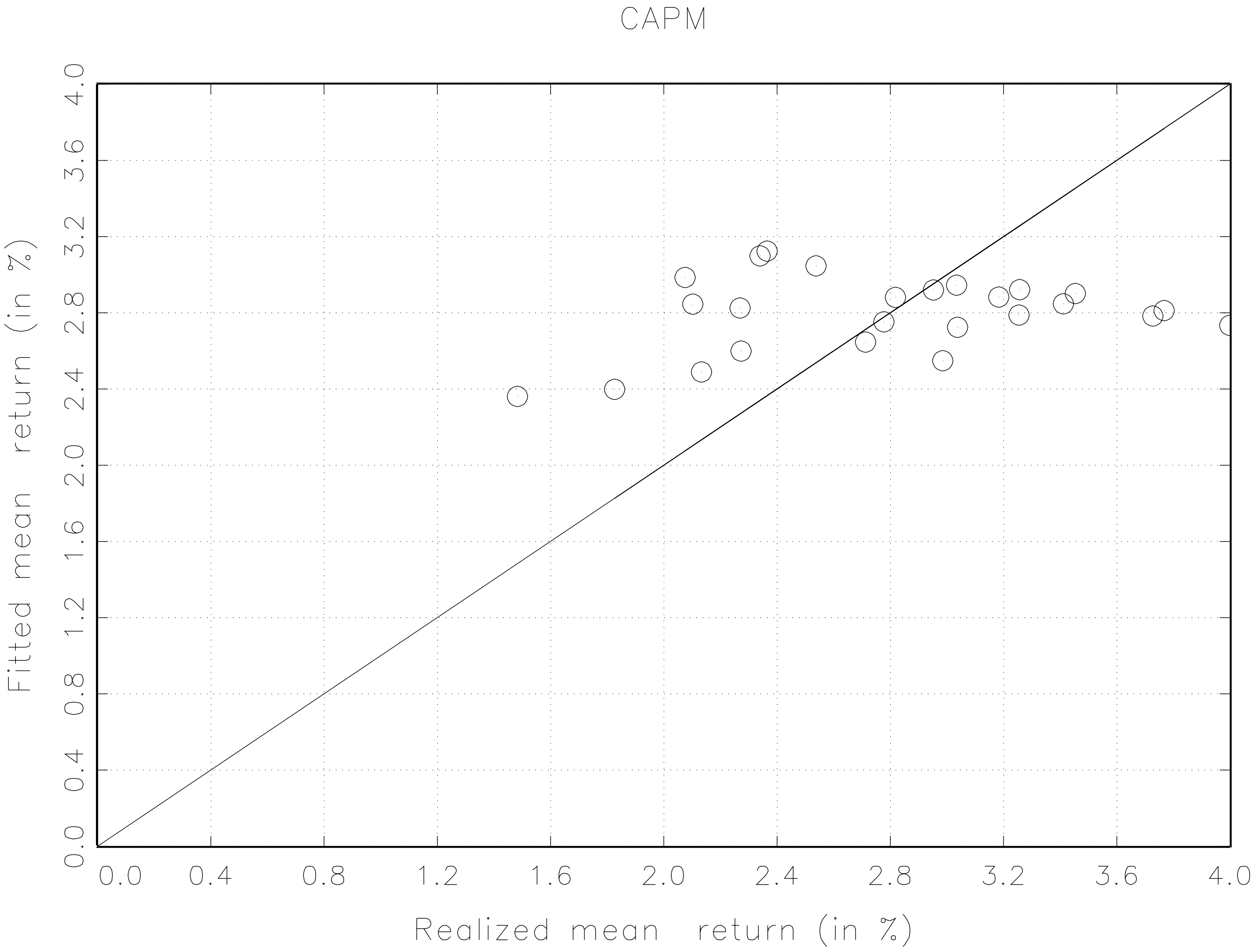
Performance comparison. Example: Habit model Grammig/Schrimpf (2005) estimated on 25 Fama-French portfolios



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

Depreciation rate

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

Stock of durable goods

Expenditures durable goods

Nondurable goods

Investment in assets

$$\sum_{i=0}^{N} B_t^i = W_t - C_t - P_t E_t$$

Wealth

$$W_{t+1} = \sum_{i=0}^{N} B_t^i R_{t+1}^i$$

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

$$\alpha \in (0,1) \qquad \rho \geq 0$$

The household‘s intertemporal utility is specified by a recursive function that disentangles EIS and RRA

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

Idea of recursive utility function: Epstein/Zin (Econometrica 1989), (JPE 1991)

Intertemporal elasticity of substitution (EIS)

$$\kappa = (1-\gamma)/(1-1/\sigma) \qquad \beta \in (0,1) \qquad \sigma \geq 0 \qquad \gamma > 0$$

Subjective discount factor

relative risk aversion coefficient

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

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

Dunn/Singleton (1986), Eichenbaum and Hansen (1990), Ogaki/Reinhard (1998)

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

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

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

Use as usual for

$$\mathbb{E}_t(m_{t+1}R_{t+1}^{i} = 1) \quad \mathbb{E}_t(m_{t+1}R_{t+1}^{ei}) = 0$$

An additional moment restriction for the „investment“ in the durable good is added

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

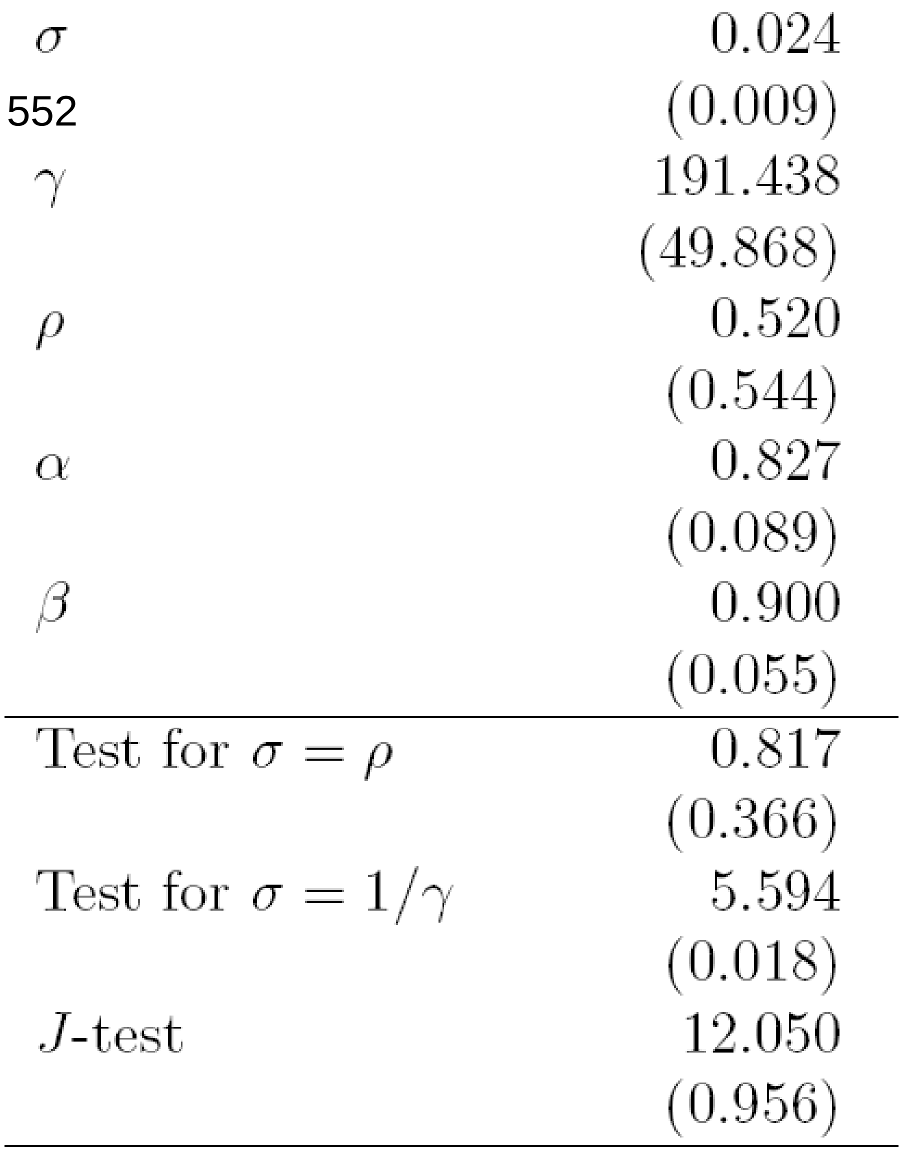
$$\mathbb{E}\left[1 - \frac{\alpha}{1-\alpha}(D_t/C_t)^{-1/\rho}\frac{1}{P_t} - (1-\delta)m_{t+1}\frac{P_{t+1}}{P_t}\right] = 0$$

Yogo‘s (2006) estimation results for Fama-French portfolios

Source: Yogo (2006) p. 552

standard errors in parentheses

p-values in parentheses



σ	0.024
	(0.009)
γ	191.438
	(49.868)
ρ	0.520
	(0.544)
α	0.827
	(0.089)
β	0.900
	(0.055)
Test for $\sigma = \rho$	0.817
	(0.366)
Test for $\sigma = 1/\gamma$	5.594
	(0.018)
J-test	12.050
	(0.956)

EIS estimate small

Risk aversion estimate high

elasticity of subsitution reasonable

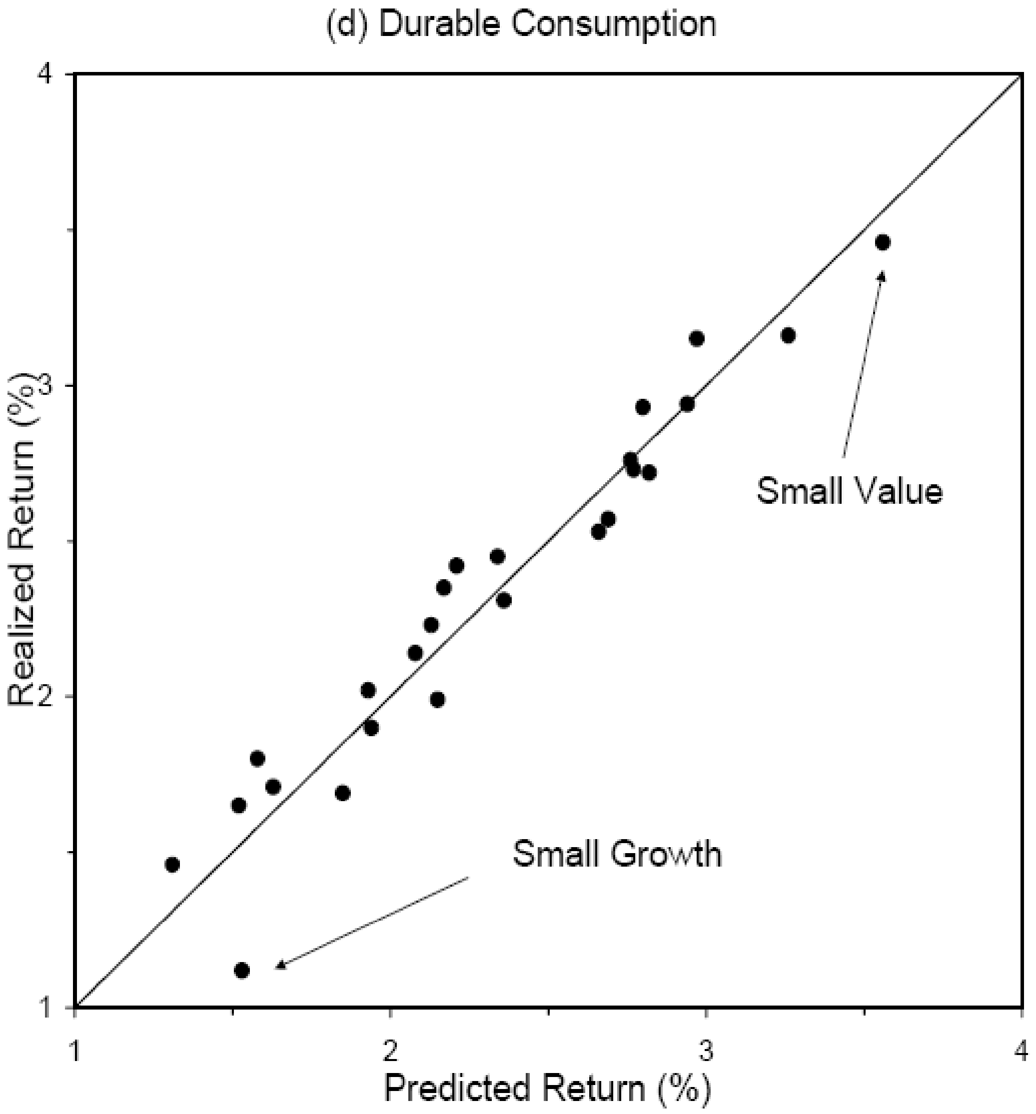
subjective discount factor < 1

Epstein/Zin (1991) non-rejected

Eichenbaum/Hansen (1987) rejected

Durable model not rejected

The fit of the durable consumption model is good (Fama French portfolios)



Source: Yogo (2006), p. 558

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

  $m_{t+1} = b_0 + b_m R^m_{t+1}$

- Scaled CAPM by Lettau and Ludvigson (2001)

  $m_{t+1} = b_0 + b_{cay} cay_t + b_m R^m_{t+1} + b_{caym} cay_t R^m_{t+1}$

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

$p_t = \mathbb{E}\left(m_{t+1}(b) \cdot x_{t+1} | I_t\right)$ or $1 = \mathbb{E}\left(m_{t+1}(b) \cdot R_{t+1} | I_t\right)$
or $0 = \mathbb{E}\left(m_{t+1}(b) \cdot R^e_{t+1} | I_t\right)$

I.i.e "integrates out" conditional implications, let us focus on unconditional implications of asset pricing model (model for S.D.F.):
$\mathbb{E}\left(m_{t+1}(b) \cdot R_{t+1} - 1\right) = 0$

To test conditional implications write
$\mathbb{E}\left(Y_{t+1} | I_t\right) = 0$ where $Y_{t+1} = \left(m_{t+1}(b) \cdot R_{t+1} - 1\right)$ or ...
$\{Y_{t+1}\}$ a martingale difference sequence.

Properties of m.d.s include:
$cov\left(Y_{t+1}, z_t\right) = 0 \quad \forall \quad z_t \in I_t$
$\mathbb{E}\left(Y_{t+1} z_t\right) = 0$ since $1 \in I_t$
Testable restrictions therefore: $\mathbb{E}\left[(m_{t+1}(b) \cdot R_{t+1} - 1) z_t\right] = 0 \ \forall \ \ z_t \in I_t$

The use of instruments has an economic interpretation: Can the model price “managed portfolios“?

$\tilde{x}_{t+1} = x^i_{t+1} z_t$ 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 wtih price $z_t p_t$
If model correct, it prices any asset, also mgt. portfolios.

$$\underbrace{z_t p_t}_{p(\tilde{x}_{t+1})} = \mathbb{E}_t(m_{t+1}(b) \cdot \underbrace{x_{t+1} z_t}_{\tilde{x}_{t+1}}) \text{ or } z_t = \mathbb{E}_t \Big(m_{t+1}(b) \cdot R_{t+1} z_t \Big)$$

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 $z_t = (1, z_t^1)'$

$$\mathbb{E}\begin{bmatrix} m_{t+1}(b)\, R_{t+1}^a - 1 \\ m_{t+1}(b)\, R_{t+1}^b - 1 \\ (m_{t+1}(b)\, R_{t+1}^a - 1)z_t^1 \\ (m_{t+1}(b)\, R_{t+1}^b - 1)z_t^1 \end{bmatrix} = 0$$

or, emphasizing the managed portfolio interpretation

$$\mathbb{E}(m_{t+1}(b)\, \underbrace{R_{t+1} \otimes z_t}_{\text{payoff}} - \underbrace{1 \otimes z_t}_{\text{price}}) = 0$$

$$\mathbb{E}(m_{t+1}(b)\, \underbrace{x_{t+1} \otimes z_t}_{\text{payoff}} - \underbrace{p_t \otimes z_t}_{\text{price}}) = 0$$

You should include economically meaningful instruments (managed portfolios)

- $p = \mathbb{E}(mx)$ 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

$$p_t z_t = \mathbb{E}_t(m_{t+1} x_{t+1} z_t) \quad \text{resp.} \quad z_t = \mathbb{E}_t(m_{t+1} R_{t+1} z_t)$$

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

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

then in

$$1 z_t = z_t \underbrace{\mathbb{E}_t(R_{t+1})}_{\uparrow} \underbrace{\mathbb{E}_t(m_{t+1})}_{\downarrow \ \text{or}} + z_t \underbrace{cov_t(m_{t+1} R_{t+1})}_{\downarrow}$$

Is a conditional asset pricing model testable at all?

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

$$1 = \mathbb{E}\left(m_{t+1}(b_t) \cdot R_{t+1} | I_t\right)$$

e.g. CAPM $m_{t+1} = a_t - b_t R^W_{t+1}$.

Parameters of factor pricing model vary over time.
⇒unconditioning via l.i.e. no longer possible:

$$1 = \mathbb{E}\left(m_{t+1}(b_t) \cdot R_{t+1} | I_t\right)$$

does NOT imply

$$1 = \mathbb{E}\left(m_{t+1}(b) \cdot R_{t+1}\right)$$

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

$$m_{t+1} = b_t' \underbrace{f_{t+1}}_{K\times 1}$$

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

$$m_{t+1} = b' \underbrace{(f_{t+1} \otimes z_t)}_{KM\times 1}$$

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:

$$m_{t+1} = b_t' \underbrace{f_{t+1}}_{K\times 1}$$

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

$$m_{t+1} = b(z_t)' f_{t+1} \quad \text{with} \quad b(z_t) = \underbrace{B}_{K\times M} z_t$$

Equivalent to $m_{t+1} = \tilde{b}' \underbrace{(f_{t+1} \otimes z_t)}_{K\times N}$ where $\tilde{b} = vec(B)$

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

$$\mathbb{E}_t\left(\underbrace{\left(\tilde{b}'(f_{t+1}\otimes z_t)\right)}_{m_{t+1}} R_{t+1}\right) = 1 \quad \text{l.i.e.} \Rightarrow \quad \underbrace{\mathbb{E}\left(\left(\tilde{b}'(f_{t+1}\otimes z_t)\right) R_{t+1} - 1\right) = 0}_{\text{unconditional moment restrictions}}$$

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

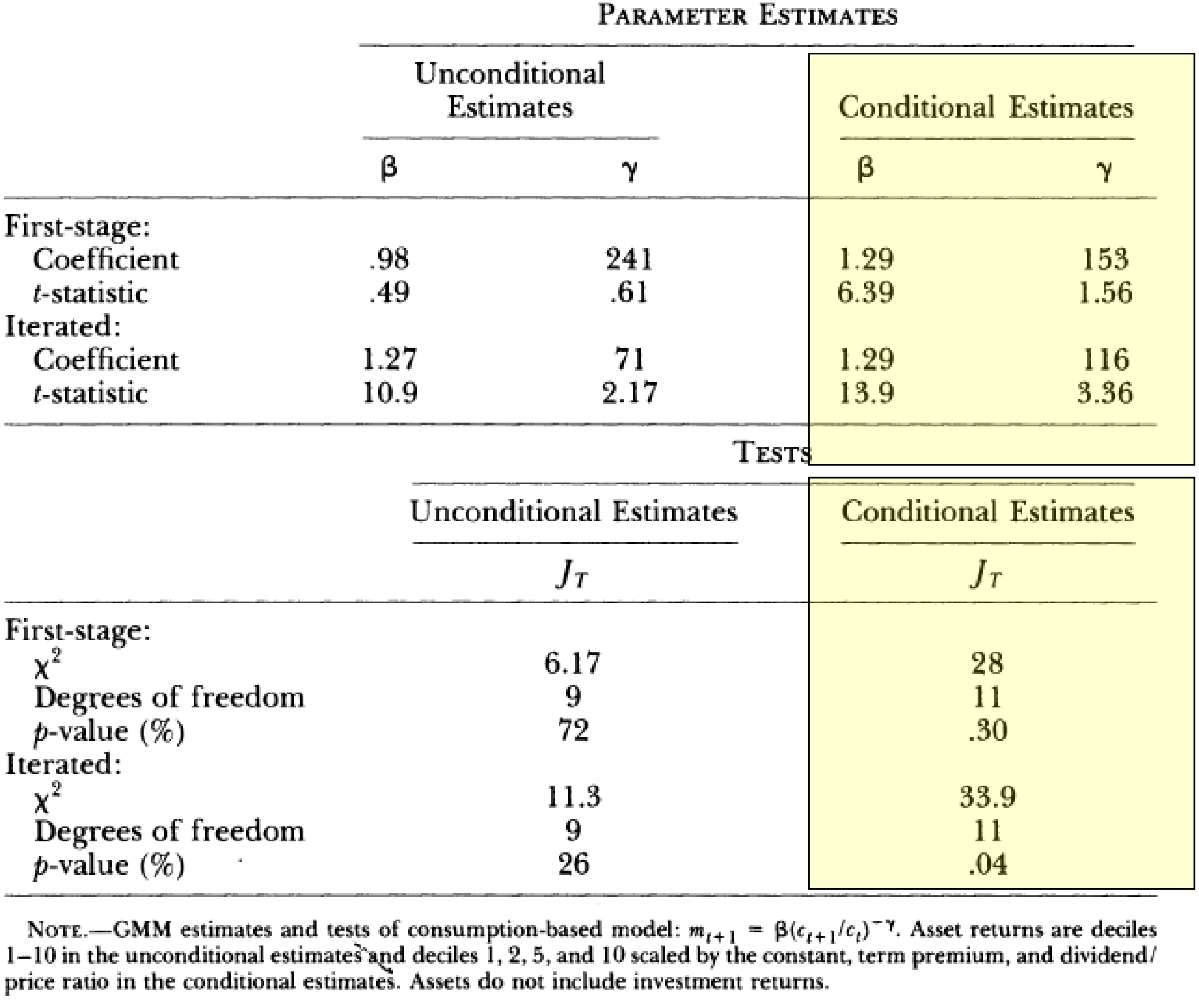
$$f = \begin{pmatrix} 1 \\ R^W \end{pmatrix} z_t = \begin{pmatrix} 1 \\ \frac{P}{D} \\ term \end{pmatrix} B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$$

$$f \otimes z = \begin{pmatrix} 1 \\ R^W \\ \frac{P}{D} \\ R^W \cdot \frac{P}{D} \\ term \\ R^W \cdot term \end{pmatrix} \tilde{b} = \begin{pmatrix} b_{11}, b_{21}, b_{12}, b_{22}, b_{13}, b_{23} \end{pmatrix}'$$

$$m = \tilde{b}'(f \otimes z) = b_{11} + b_{12}\frac{P}{D} + b_{13} term + b_{21} R^W + b_{22} R^W \cdot \frac{P}{D} + b_{23} R^W \cdot term$$

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

Cochrane‘s (JPE 1996) estimation results for the consumption based model with power utility



PARAMETER ESTIMATES

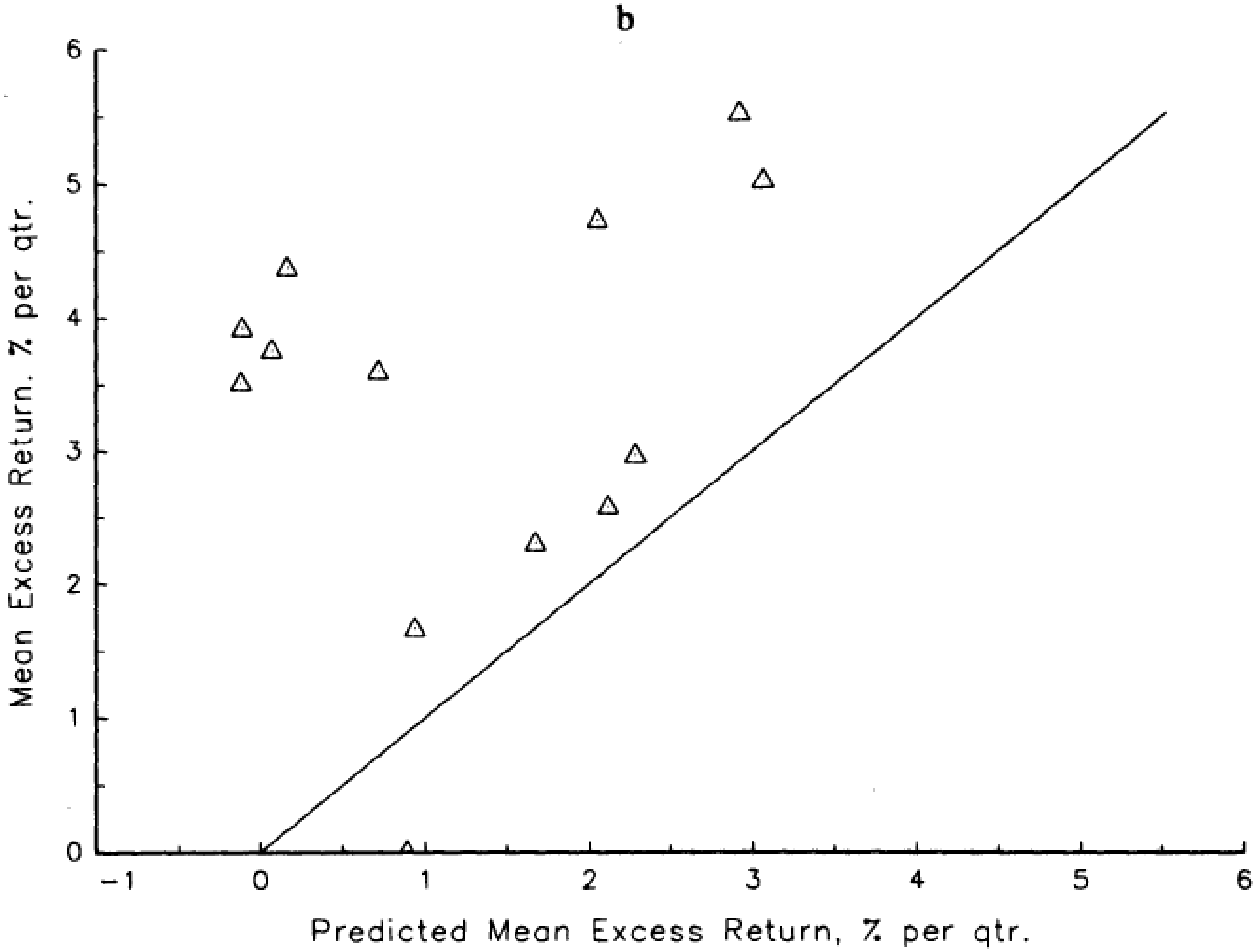
	Unconditional Estimates		Conditional Estimates	
	β	γ	β	γ
First-stage:				
Coefficient	.98	241	1.29	153
t-statistic	.49	.61	6.39	1.56
Iterated:				
Coefficient	1.27	71	1.29	116
t-statistic	10.9	2.17	13.9	3.36

TESTS

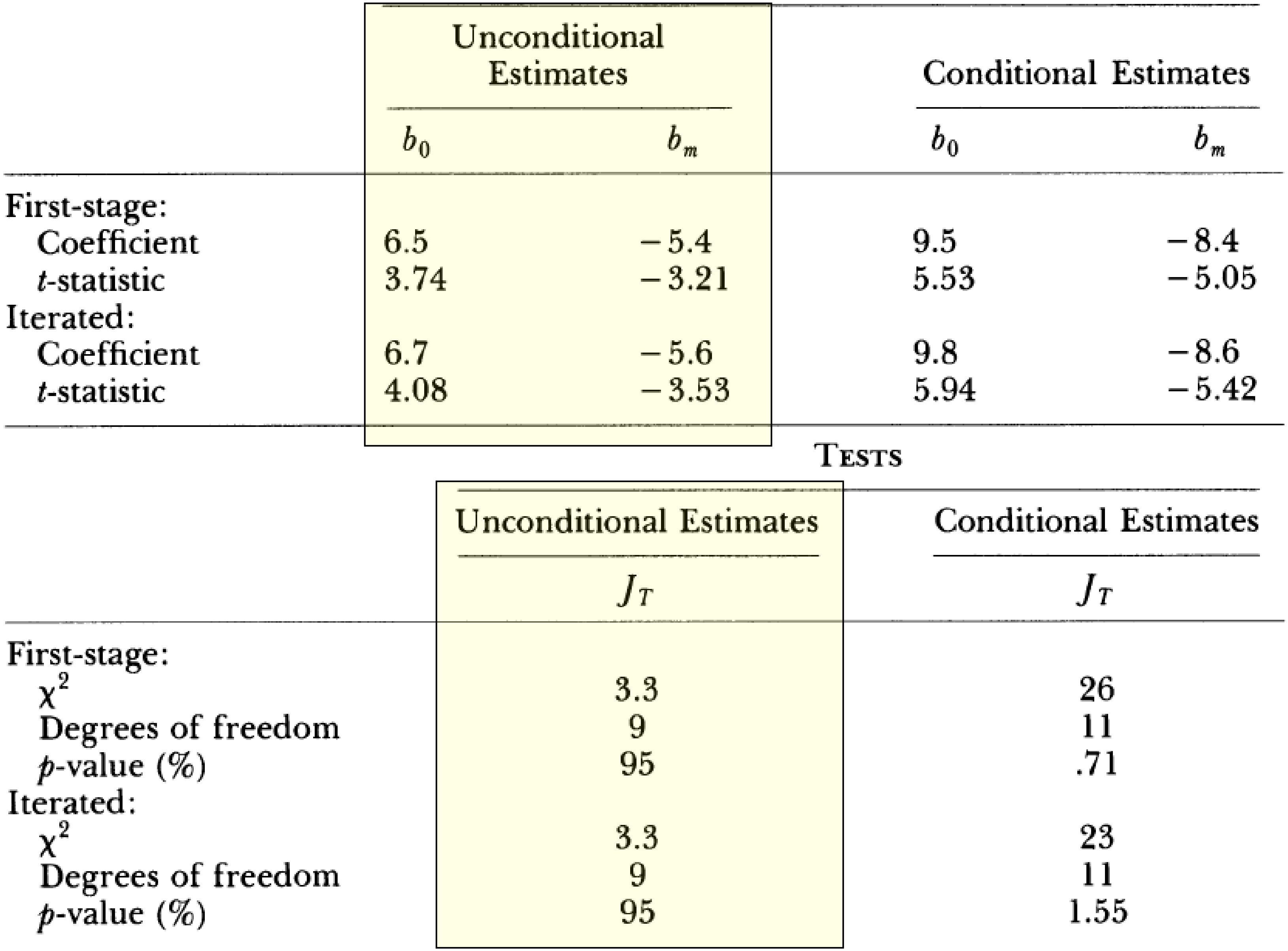
	Unconditional Estimates	Conditional Estimates
	J_T	J_T
First-stage:		
χ^2	6.17	28
Degrees of freedom	9	11
p-value (%)	72	.30
Iterated:		
χ^2	11.3	33.9
Degrees of freedom	9	11
p-value (%)	26	.04

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

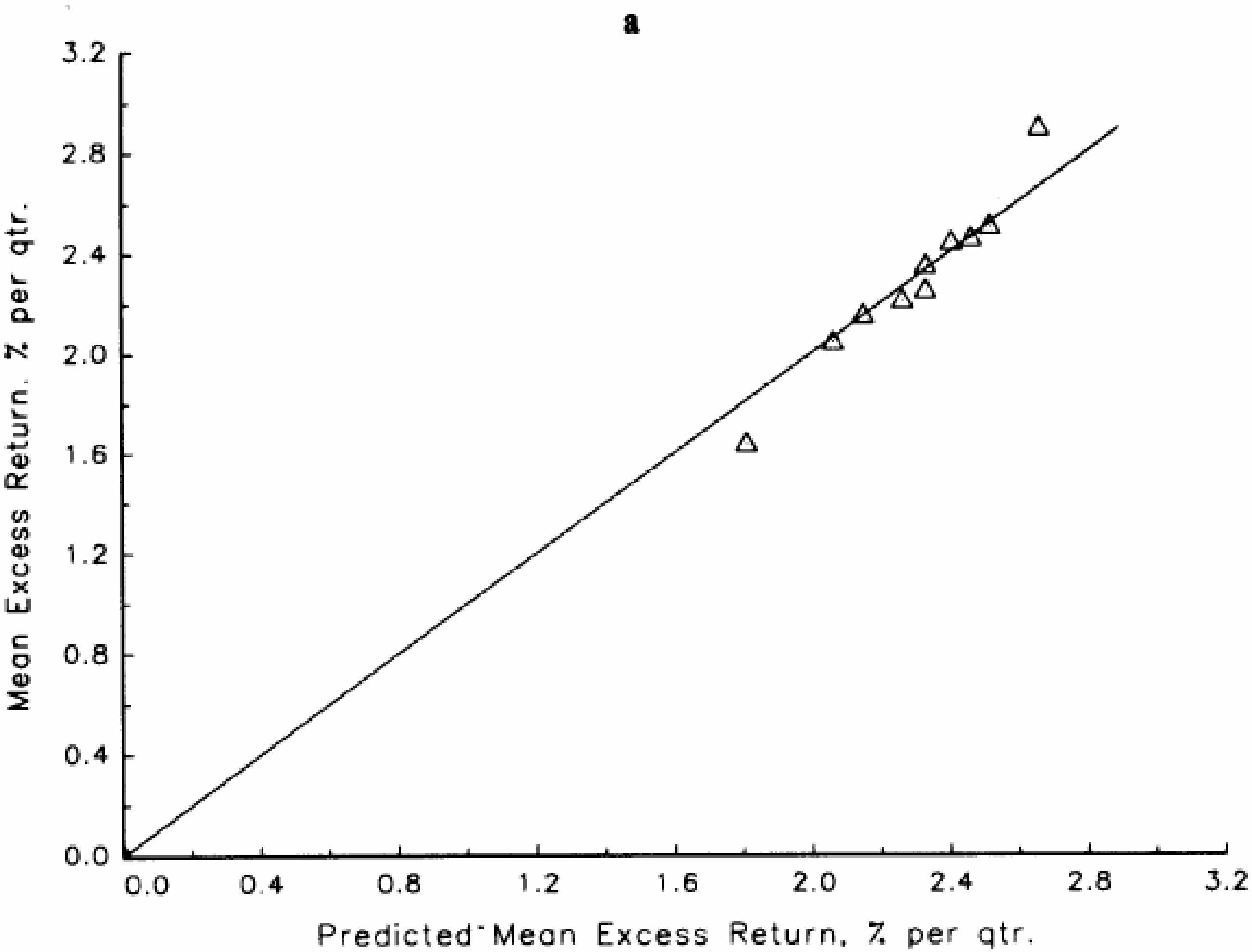


	Unconditional Estimates		Conditional Estimates	
	b_0	b_m	b_0	b_m
First-stage:				
Coefficient	6.5	−5.4	9.5	−8.4
t-statistic	3.74	−3.21	5.53	−5.05
Iterated:				
Coefficient	6.7	−5.6	9.8	−8.6
t-statistic	4.08	−3.53	5.94	−5.42

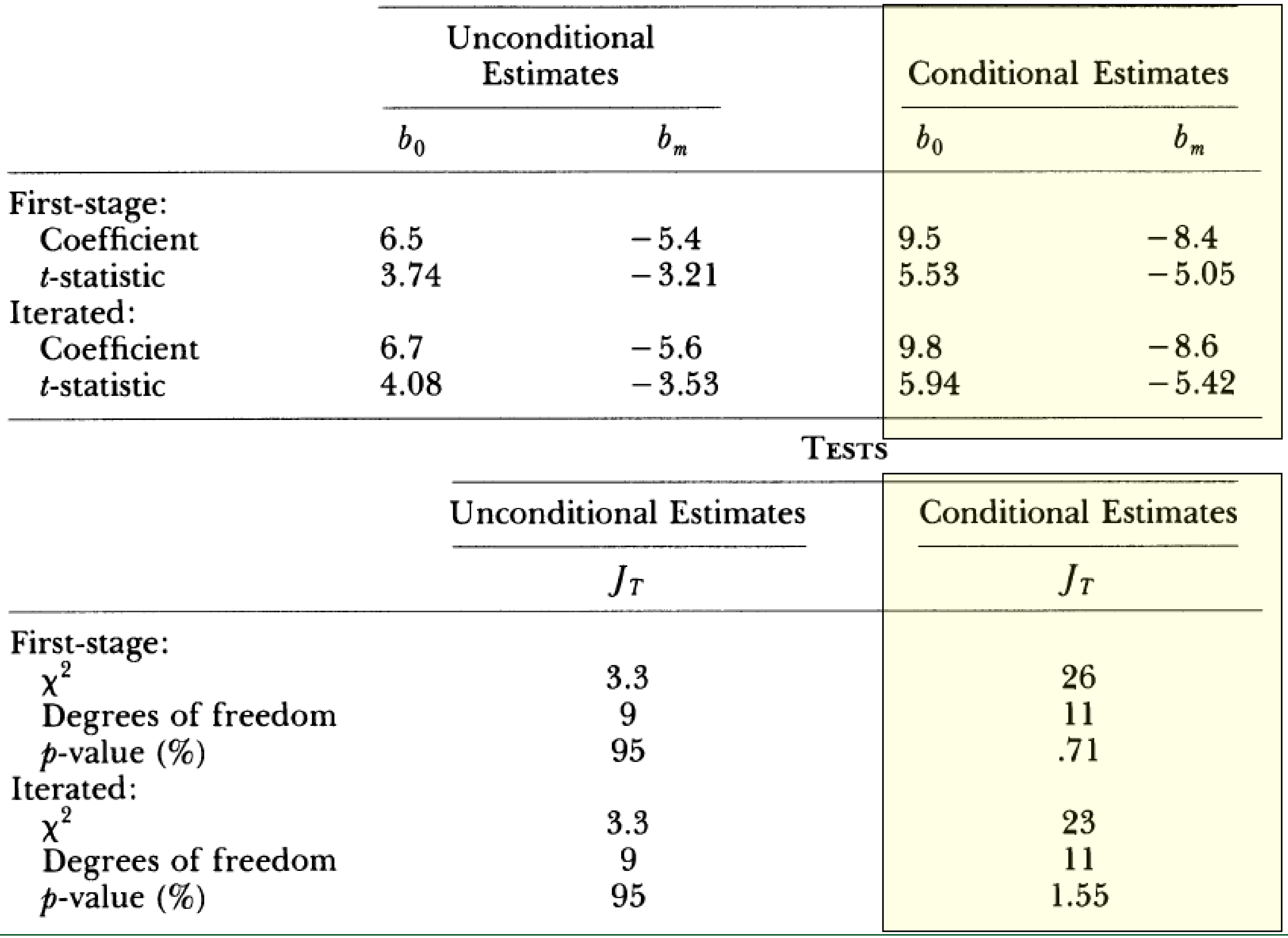
TESTS

	Unconditional Estimates	Conditional Estimates
	J_T	J_T
First-stage:		
χ^2	3.3	26
Degrees of freedom	9	11
p-value (%)	95	.71
Iterated:		
χ^2	3.3	23
Degrees of freedom	9	11
p-value (%)	95	1.55

Cochrane‘s (1996) results for unconditional estimation of CAPM



Cochrane‘s (1996) results for conditional estimation of CAPM

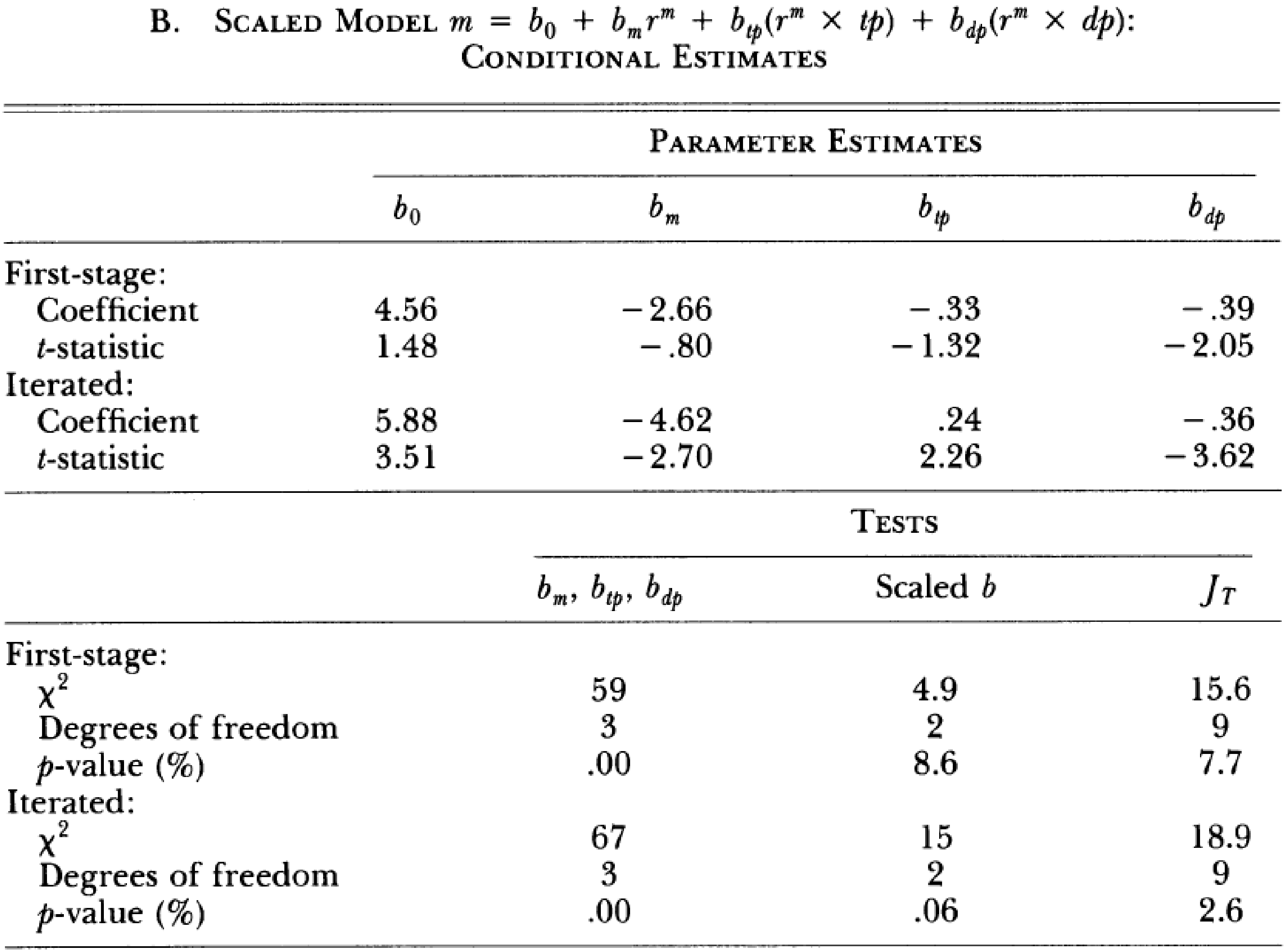


	Unconditional Estimates		Conditional Estimates	
	b_0	b_m	b_0	b_m
First-stage:				
Coefficient	6.5	−5.4	9.5	−8.4
t-statistic	3.74	−3.21	5.53	−5.05
Iterated:				
Coefficient	6.7	−5.6	9.8	−8.6
t-statistic	4.08	−3.53	5.94	−5.42

TESTS

	Unconditional Estimates	Conditional Estimates
	J_T	J_T
First-stage:		
χ^2	3.3	26
Degrees of freedom	9	11
p-value (%)	95	.71
Iterated:		
χ^2	3.3	23
Degrees of freedom	9	11
p-value (%)	95	1.55

Cochrane‘s (1996) results for conditional estimation of CAPM

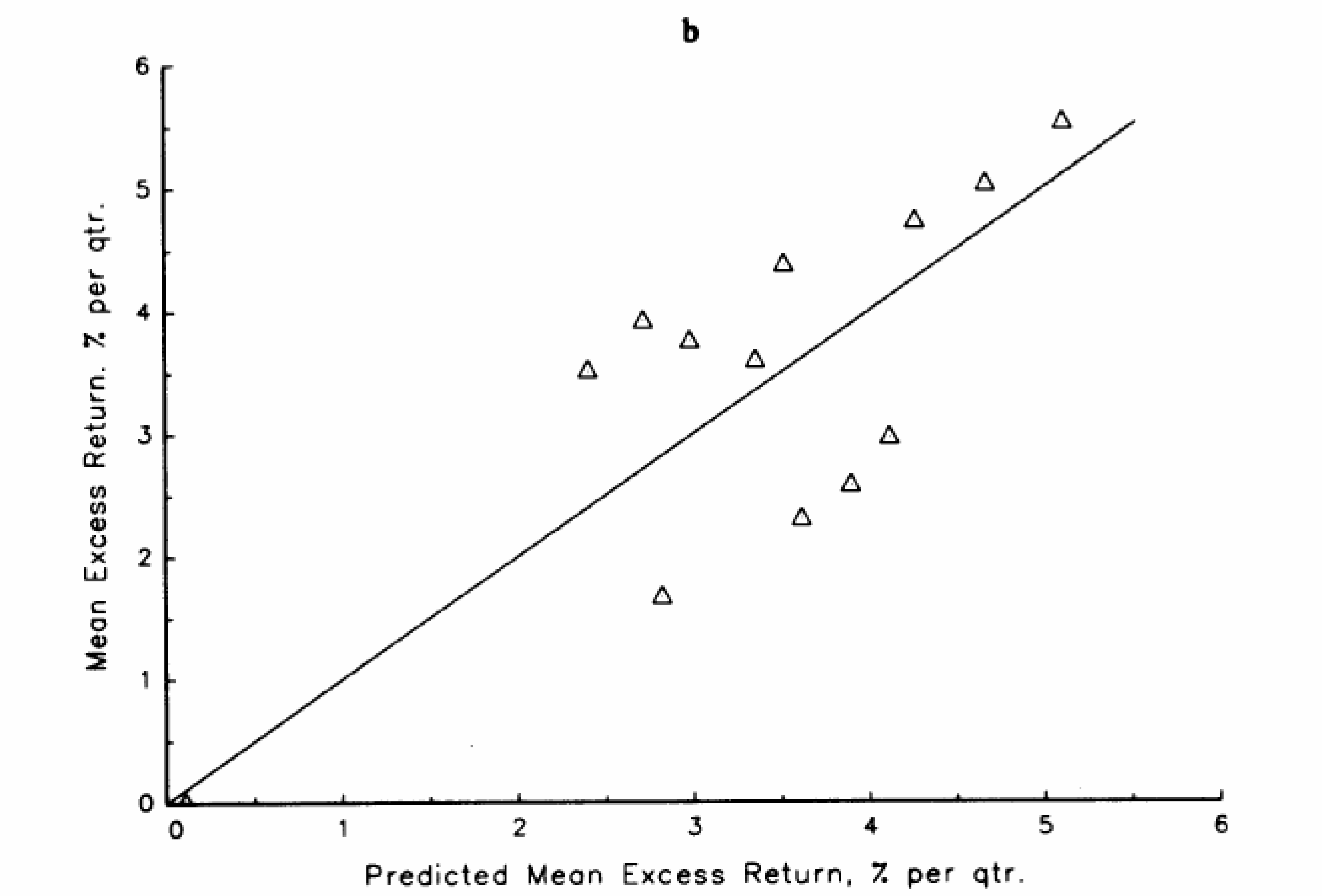


B. SCALED MODEL $m = b_0 + b_m r^m + b_{tp}(r^m \times tp) + b_{dp}(r^m \times dp)$:
CONDITIONAL ESTIMATES

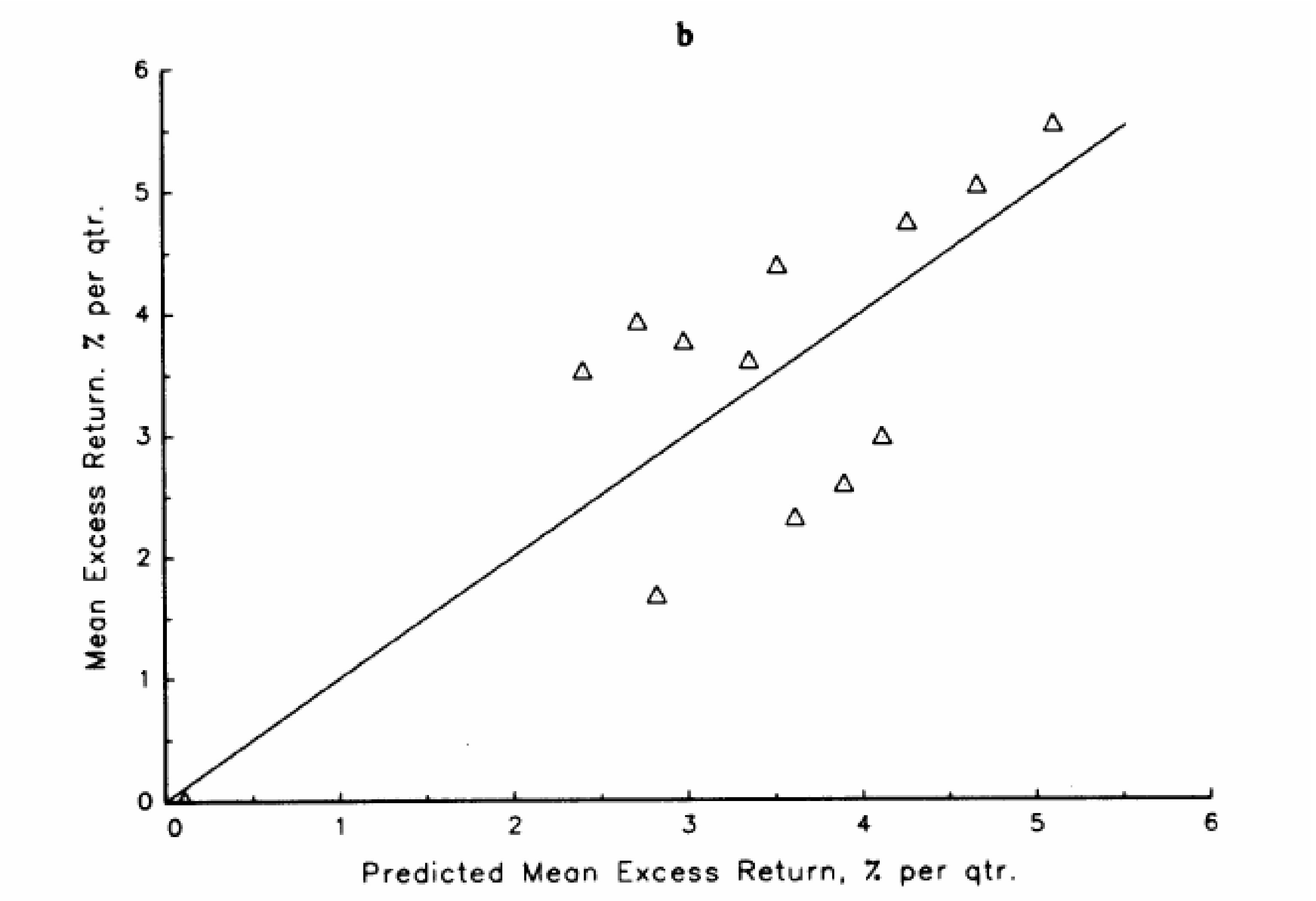
	PARAMETER ESTIMATES			
	b_0	b_m	b_{tp}	b_{dp}
First-stage:				
Coefficient	4.56	−2.66	−.33	−.39
t-statistic	1.48	−.80	−1.32	−2.05
Iterated:				
Coefficient	5.88	−4.62	.24	−.36
t-statistic	3.51	−2.70	2.26	−3.62

	TESTS		
	b_m, b_{tp}, b_{dp}	Scaled b	J_T
First-stage:			
χ^2	59	4.9	15.6
Degrees of freedom	3	2	9
p-value (%)	.00	8.6	7.7
Iterated:			
χ^2	67	15	18.9
Degrees of freedom	3	2	9
p-value (%)	.00	.06	2.6

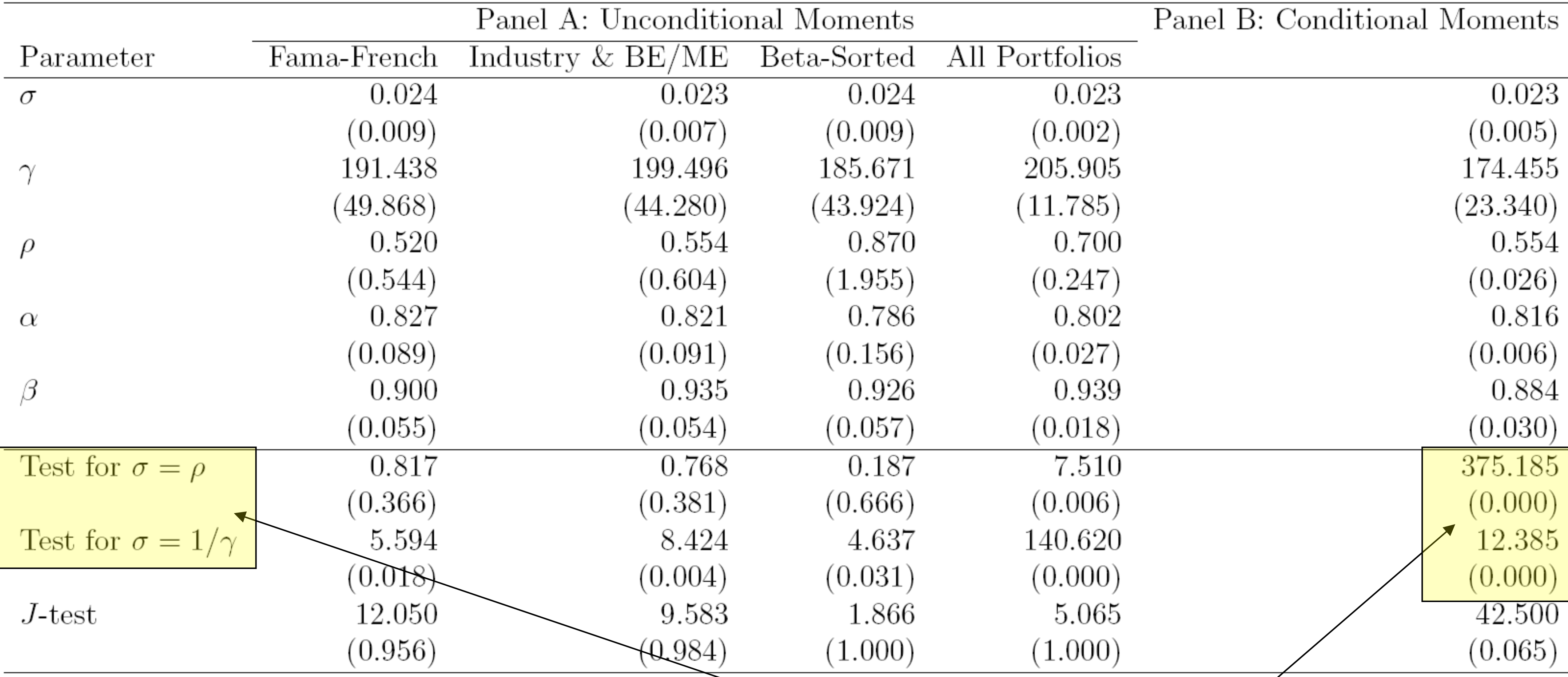
Cochrane‘s (1996) results for conditional estimation of scaled CAPM



Cochrane‘s (1996) results for conditional estimation of scaled CAPM



Yogo‘s (2006) cross section estimation results



	Panel A: Unconditional Moments				Panel B: Conditional Moments
Parameter	Fama-French	Industry & BE/ME	Beta-Sorted	All Portfolios	
σ	0.024	0.023	0.024	0.023	0.023
	(0.009)	(0.007)	(0.009)	(0.002)	(0.005)
γ	191.438	199.496	185.671	205.905	174.455
	(49.868)	(44.280)	(43.924)	(11.785)	(23.340)
ρ	0.520	0.554	0.870	0.700	0.554
	(0.544)	(0.604)	(1.955)	(0.247)	(0.026)
α	0.827	0.821	0.786	0.802	0.816
	(0.089)	(0.091)	(0.156)	(0.027)	(0.006)
β	0.900	0.935	0.926	0.939	0.884
	(0.055)	(0.054)	(0.057)	(0.018)	(0.030)
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)

Source: Yogo (2006) p. 552

Both restrictions rejected

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

$$m_{t+1} = a_0 + a_1 cay_t + a_2 r^m_{t+1} + a_3 cay_t r^m_{t+1}$$

log wealth

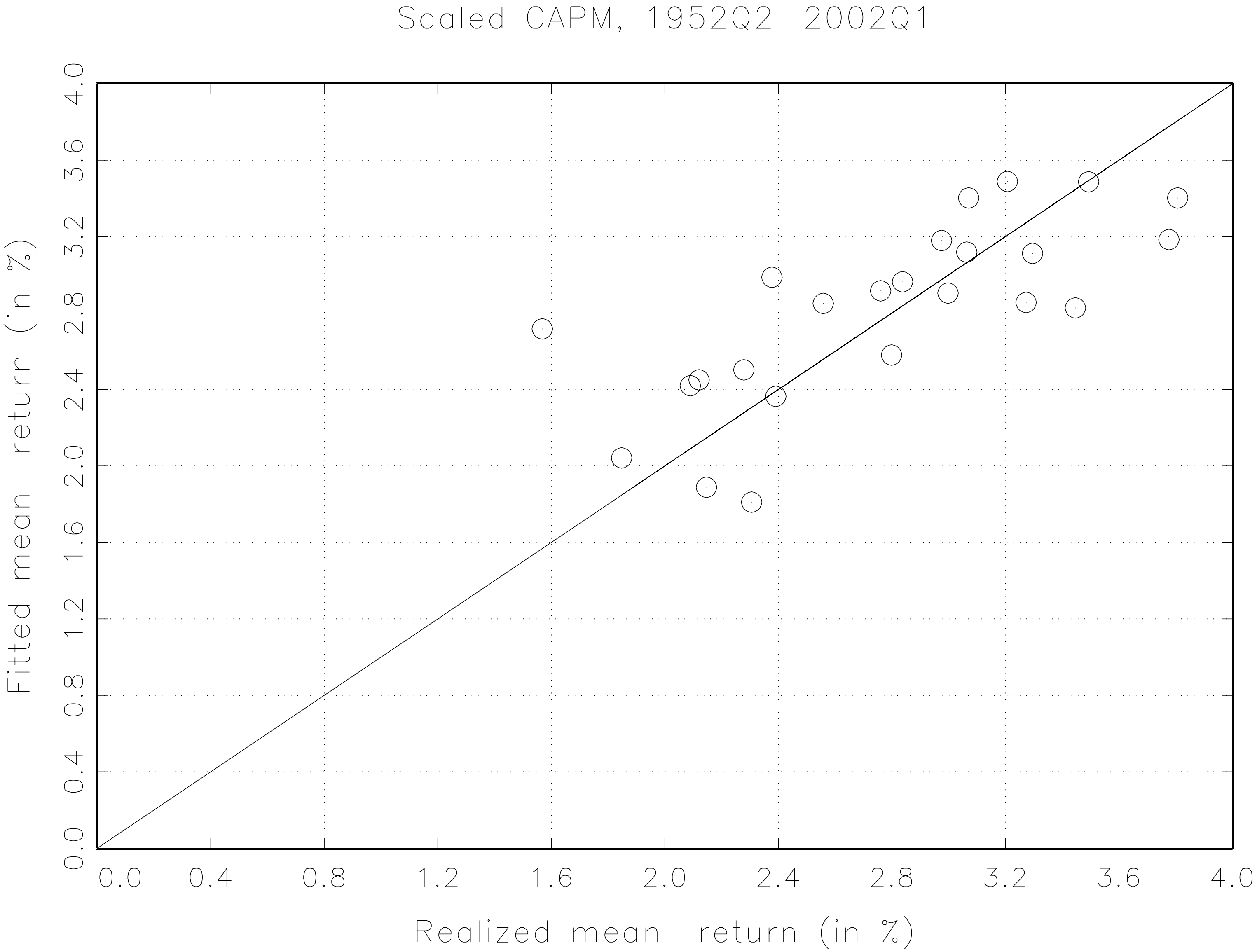
$$c_t - w_t \approx \mathbb{E}_t \sum_{i=1}^{\infty} \rho_w^i (r^m_{t+i} - \Delta c_{t+i})$$

log asset wealth

log labor income

$$cay_t = c_t - \omega a_t - (1 - \omega) y_t$$

Performance comparison. Example: Lettau/Ludvigson model estimated on 25 Fama-French portfolios



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 assests
- 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: MRS

$$b = (b_1, b_2, \ldots, b_K)' \quad f = (f_1, f_2, \ldots, f_K)'$$

often: factors are returns of portfolios, e.g market or wealth portfolio

$m = a - b \cdot R^m$: 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

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

$$\lambda = \underbrace{(\lambda_1, \ldots, \lambda_K)'}_{\text{"Price of factor k" or factor risk premium}} \qquad \beta_i = \underbrace{(\beta_{i1}, \ldots, \beta_{iK})'}_{\text{Exposure of asset i to factor k}}$$

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

Compare to linear regression:

$$y_i = a + b'x_i + \underbrace{u_i}_{\mathbb{E}(u_i)=0}$$

$$\mathbb{E}(y_i) = a + b'\mathbb{E}(x_i)$$

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 $\mathbb{E}(f_i f_j) = cov(f_i, f_j) = 0$ for $i \neq j$ (orthogonal factors)

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

Example:

$$m = a - bR^m \quad \Leftrightarrow \quad \mathbb{E}(R^i) = R^f + \beta_i(\mathbb{E}(R^m) - R^f)$$

where $R^f \widehat{=} \gamma$, $\beta_i \widehat{=}$ riskiness of asset i and $\mathbb{E}(R^m) - R^f \widehat{=} \lambda \widehat{=}$ 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: $1 = \mathbb{E}(mR) \quad \Leftrightarrow \quad \mathbb{E}(R) = \gamma + \lambda'\beta$:

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

convenient: demean factors: ”fold” means of factors into constant a

$$\begin{aligned} \tilde{f} &= \text{ factor with } \quad (\tilde{f}) \neq 0 \\ f &= \tilde{f} - \mathbb{E}(\tilde{f}) = \text{demeaned factor with } \quad \mathbb{E}(f) = 0 \end{aligned}$$

$$m = a + b'f \quad \text{where} \quad a = \tilde{a} + b'\mathbb{E}(\tilde{f})$$

$\Rightarrow \quad \mathbb{E}(m) = a$

We want to show the equivalence of the two representations (2)

Rewrite

$$\begin{aligned} 1 &= \mathbb{E}(mR) \\ &= cov(m,R) + \mathbb{E}(m)\cdot\mathbb{E}(R) \\ \Rightarrow \mathbb{E}(R) &= \frac{1}{\mathbb{E}(m)} - \frac{cov(m,R)}{\mathbb{E}(m)} \\ &= \frac{1}{a} - \frac{cov\Big((a+bf),R\Big)}{a} \end{aligned}$$

$$\begin{aligned} cov\Big((a+bf),R\Big) &= \mathbb{E}\Big[(a+bf-a)(R-\mathbb{E}(R))\Big] \\ &= \mathbb{E}(bfR) - \underbrace{\mathbb{E}(bf)\cdot\mathbb{E}(R)}_{=0 \text{ as } \mathbb{E}(f)=0} \end{aligned}$$

We want to show the equivalence of the two representations (3)

$$\mathbb{E}(R) = \frac{1}{a} - \frac{b\mathbb{E}(Rf)}{a} \quad \Big| \quad \text{we want betas}$$

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

Define

$$\gamma \equiv \frac{1}{a} = \frac{1}{\mathbb{E}(m)} = R^f \quad \text{(if traded)}$$

$$\beta \equiv \frac{cov(f,R)}{var(f)}$$

$$\lambda \equiv -\frac{bvar(f)}{a}$$

$$\Rightarrow \quad \mathbb{E}(R^i) = \gamma + \beta_i\lambda$$

λ in the expeced return- beta representation can be interpreted as the price of the risk factor

We want to interpret λ as price of risk factor

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

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

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

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

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

$$p\Big(\underbrace{\mathbb{E}(\tilde{f})}_{\substack{\text{constant}\\ \text{payoff}\\ \text{in t+1}}}\Big) = \mathbb{E}\Big(m \cdot \mathbb{E}(\tilde{f})\Big) = \mathbb{E}(m) \cdot \mathbb{E}(\tilde{f}) = \frac{\mathbb{E}(\tilde{f})}{\gamma}$$

If the factor is a return, λ has the interpretation of an expected excess return, or factor risk premium

$$\lambda = -\gamma\left(1-\frac{\mathbb{E}(\tilde{f})}{\gamma}\right) = \mathbb{E}(\tilde{f}) - \gamma \;\Bigg|\; \overbrace{\mathbb{E}(\tilde{f}) - R^f}^{\text{expected excess return}} \qquad \text{: factor risk premium}$$

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

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

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

with $\gamma = \frac{1}{\mathbb{E}(m)} = R^f$ $\tilde{f} = R^m \Rightarrow$ CAPM

Equivalence in the multifactor case (1)

In a multifactor model with k factors

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

2. $\lambda = \underbrace{\mathbb{E}(\tilde{f})}_{K \times 1} - \gamma$

$$\underbrace{\beta_i}_{K \times 1} = \Big[\mathbb{E}[f f']\Big]^{-1} \mathbb{E}[f R^i] \text{ with } \mathbb{E}(f) = 0 \text{ (demeaned factors)}$$

Equivalence in the multifactor case (2)

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

where $cov(f, R^i) = \left[\begin{array}{ccc} cov(f_1, R^i) & cov(f_2, R^i) & \cdots \end{array}\right]$

and $cov(f) = \left[\begin{array}{cccc} var(f_1) & cov(f_1, f_2) & \cdots & cov(f_1, f_K) \\ cov(f_1, f_2) & var(f_2) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ cov(f_1, f_K) & \ldots & \ldots & var(f_K) \end{array}\right]$

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

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