Bayesian Analysis of Stochastic Betas

Jostova and Philipov (2005, JFQA)

Authors

Gergana Jostova
George Washington

Alexander Philipov
George Mason

Abstract

- A Mean-Reverting Stochastic Process for the Market β
- Precise β Estimates than
  - GARCH βs, Firm-level Conditioned βs and Rolling-regression βs
- Stronger Support for the Conditional CAPM
- Resolve asset pricing anomalies
  - Size, B/M and Idiosyncratic Volatility Effects

We propose a mean-reverting stochastic process for the market beta. In a simulation study, the proposed model generates significantly more precise beta estimates than GARCH betas, betas conditioned on aggregate or firm-level variables, and rolling regression betas, even when the true betas are generated based on these competing specifications. Our model significantly improves out-of-sample hedging effectiveness. In asset pricing tests, our model provides substantially stronger support for the conditional CAPM relative to competing beta models and helps resolve asset pricing anomalies such as the size, book-to-market, and idiosyncratic volatility effects in the cross section of stock returns.

I. Introduction

- Capital Asset Pricing Model: Normative
  - Sharpe (1964, JF), Merton (1973, EMA): β Prices Capital Assets
  - Ang and Liu (2004, JF): β Matters in Explaining Long-run Cash-flows
  - Zhang (2004, JF), Petkova and Zhang (2005, JFE): β and Business Cycles
  - Avramov and Chordia (2006, RFS): Time-varying β Causes Anomalies
- β as a General Mean-reverting Stochastic Process: More Precise
  - Embraces Existing β Models as Special Cases
  - Allows Both Time (Recent Models) and Stochastic (Past) Variation
  - Justifies Excess Kurtosis (Empirically Observed)
  - Explains High β Persistence
    - Berk, Green and Naik (1999, JF): Persistent β and Momentum Effects
    - Ang and Chen (2003, JEF): Persistent β and B/M Effects

I. Introduction (Cont’d)

- Why Mean-reverting β?
  - Theoretically
    - Gomes, Kogan and Zhang (2003, JPE) : Ensure Return Stationarity
    - Kaldor (1961) : Ensure Non-exploding Growth Rate of Output
  - Empirically
    - Ang and Chen (2003, JEF): Time-varying βs with Slow Mean-reversion
    - Campbell and Vuolteenaho (2004, AER): Discount-rate βs Reverts
    - Zhang (2004, JF), Pekova and Zhang (2005, JFE)
- **Why Bayesian?**
  - Bayesian Prior-Posterior Relations: Investor’s Learning Process
  - Inherent Uncertainty about Model Parameters
  - MCMC: Exact Finite Sample Inference (Unlike MLE and GMM)
    - Ang and Chen (2003, JEF): Asymptotic Theory Can Be Misleading

I. Introduction (Cont'd)

- SBETA (Stochastic β) Estimates...
  - Highly Precise Relative to
    - Multivariate GARCH βs
    - βs Scaled by Aggregate and Firm-level Variables
    - Rolling Regression βs
  - Even When the True βs Are Generated Based on These Competing Models
  - When βs Are Persistent, OLS βs Fail to Capture the Time-variation
    - Lewellen and Nagel (2006, JFE): Disagree with the Results of This Paper
  - Empirically, βs Exhibit Both Mean-reversion and Stochastic Behavior
  - Generates Superior β Forecasts: Effective Hedging
  - Provides Strong Support for the Conditional CAPM
  - Helps Resolve Previously Documented Asset Pricing Anomalies
    - Size, B/M and Idiosyncratic Risk Effects
    - Gomes, Kogan and Zhang (2003, JPE): βs Are Correlated with Those Variables

II. The Model

$$
\begin{array}{llcll}
(1) & r_{pt} & = & \beta_{pt} r_{mt} + \sigma_p \varepsilon_{pt}, & \varepsilon_{pt} \sim N(0,1) \\
& \beta_{pt} & = & \alpha_p + \delta_p (\beta_{p,t-1} - \alpha_p) + \sigma_{\beta_p} \nu_{pt}, & \nu_{pt} \sim N(0,1) \\
& \left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right) & \sim & p\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right), &
\end{array}
$$

- r_{pt}≡Portfolio Excess Return, r_{mt}≡Market Excess Return
- β_{pt}≡Portfolio p's Sensitivity to Market Movements
- α_p≡Unconditional Mean of β_{pt}
- δ_p≡Persistence of β_{pt}
- $\sigma_{\beta p}{}^2$≡Conditional Volatility of β_{pt}
- $\sigma_p{}^2$≡Portfolio p's Idiosyncratic Return Volatility
- $p(\alpha_p,\delta_p,\sigma_{\beta p}{}^2,\sigma_p{}^2)$≡Joint Distribution of the Model Parameters
- ε_{pt}≡Stochastic Components of r_{pt}, ν_{pt}≡Stochastic Components of β_{pt}

II. The Model (Cont'd)

$$
\begin{aligned}
(1) \qquad r_{pt} &= \beta_{pt} r_{mt} + \sigma_p \varepsilon_{pt}, & \varepsilon_{pt} &\sim N(0,1) \\
\beta_{pt} &= \alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p) + \sigma_{\beta_p} \nu_{pt}, & \nu_{pt} &\sim N(0,1) \\
\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right) &\sim p\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right),
\end{aligned}
$$

- High δ_p: Captures Discrete and Rare One-time Shocks to β_p
- Low δ_p: Captures More Frequent and Dynamic Events
  - Reversion to the Long-term Mean over Shorter Periods of Time

$$
(2) \qquad E(\beta_p) = \frac{\alpha_p(1-\delta_p)}{1-\delta_p} = \alpha_p
$$

$$
(3) \qquad \mathrm{var}(\beta_p) = \frac{\sigma^2_{\beta_p}}{1-\delta^2_p}.
$$

- $\delta_p \downarrow \rightarrow \mathrm{var}(\beta_p) \uparrow$: Slow Reversion Implies Volatile Unconditional βs

II. The Model (Cont'd)

$$
\begin{array}{lrcll}
(1) & r_{pt} & = & \beta_{pt} r_{mt} + \sigma_p \varepsilon_{pt}, & \varepsilon_{pt} \sim N(0,1) \\
& \beta_{pt} & = & \alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p) + \sigma_{\beta_p} \nu_{pt}, & \nu_{pt} \sim N(0,1) \\
& \left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right) & \sim & p\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right), &
\end{array}
$$

- **Generality**
  - $\delta_p=0 \rightarrow$ "Noisy β" Model: Chen and Lee (1982, JEB)
  - $\sigma_\beta^2=0$
    - Multivariate GARCH Type Filter: Braun, Nelson and Sunier (1995, JF)
    - Conditional β Model: Shanken (1990, Journal of Econometircs)
    - Characteristic-based β Model: Avramov and Chordia (2006, RFS)
  - $\delta_p=\sigma_\beta^2=0 \rightarrow$ Constant β Model: Sharpe (1964, JF)
- Significance of Parameters: Test the Restrictions Imposed by these Embedded Models

III. Estimation Methodology

$$
\begin{aligned}
(1) \qquad r_{pt} &= \beta_{pt} r_{mt} + \sigma_p \varepsilon_{pt}, & \varepsilon_{pt} &\sim N(0,1) \\
\beta_{pt} &= \alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p) + \sigma_{\beta_p} \nu_{pt}, & \nu_{pt} &\sim N(0,1) \\
\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right) &\sim p\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right),
\end{aligned}
$$

$$
\begin{aligned}
(4) \qquad p(\boldsymbol{\omega}_p) &\propto N(\mu_\alpha, \sigma^2_\alpha) \times N(\mu_\delta, \sigma^2_\delta) \times IG(a_1, b_1) \times IG(a_2, b_2) \\
&= N(1, 100) \times N(0, 100) \\
&\quad \times IG(0.001, 0.001) \times IG(0.001, 0.001).
\end{aligned}
$$

A. Choice of Prior Distributions

- Where $\boldsymbol{\omega}_p = (\alpha_p, \delta_p, \sigma_{\beta p}{}^2, \sigma_p{}^2)$

- $\alpha_p \sim N(1,10^2)$: Since It Is Unconditional Mean of β
- $\delta_p \sim \text{Truncated}_{(-1,1)}\text{–}N(.5,10^2)$: Ensure Stationarity of β
- $\sigma_{\beta p}{}^2 \sim IG(10^{-3},10^{-3})$: Variance of β
- $\sigma_p{}^2 \sim IG(10^{-3},10^{-3})$: Variance of r_p

III. Estimation Methodology (Cont'd)

$$
\begin{aligned}
(1) \qquad r_{pt} &= \beta_{pt} r_{mt} + \sigma_p \varepsilon_{pt}, & \varepsilon_{pt} &\sim N(0,1) \\
\beta_{pt} &= \alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p) + \sigma_{\beta_p} \nu_{pt}, & \nu_{pt} &\sim N(0,1) \\
\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right) &\sim p\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right),
\end{aligned}
$$

B. Likelihood Function

$$
(6) \qquad r_{pt} \mid \beta_{pt}, r_{mt} \sim N\left(\beta_{pt} r_{mt}, \sigma^2_p\right).
$$

$$
(5) \qquad \beta_{pt} \mid \beta_{p,t-1} \sim N\left(\alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p),\ \sigma^2_{\beta_p}\right).
$$

$$
\begin{aligned}
(7) \quad L\left(\boldsymbol{\beta}_p, \alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p \middle| \mathbf{r}_t, \mathbf{r}_m\right) &\propto \prod_{t=1}^{T} N\left(\alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p), \sigma^2_{\beta_p}\right) \\
&\quad \times \prod_{t=1}^{T} N\left(\beta_{pt} r_{mt}, \sigma^2_p\right),
\end{aligned}
$$

where $\boldsymbol{\beta}_p = (\beta_{p1}, \ldots, \beta_{pT})$, $\mathbf{r}_p = (r_{p1}, \ldots, r_{pT})$, and $\mathbf{r}_m = (r_{m1}, \ldots, r_{mT})$.

III. Estimation Methodology (Cont'd)

$$(8) \qquad p(\boldsymbol{\theta}_p \mid \mathbf{r}_p, \mathbf{r}_m) \quad \propto \quad p(\boldsymbol{\omega}_p) L(\boldsymbol{\theta}_p \mid \mathbf{r}_p, \mathbf{r}_m).$$

$$(9) \quad p(\boldsymbol{\theta}_p \mid \mathbf{r}_p, \mathbf{r}_m) \quad \propto \quad N(\mu_\alpha, \sigma_\alpha^2) \times N\left(\mu_\delta, \sigma_\delta^2\right) \times IG(a_1, b_1) \times IG(a_2, b_2)$$
$$\times \prod_{t=1}^{T} N\left(\alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p), \sigma_{\beta_p}^2\right)$$
$$\times \prod_{t=1}^{T} N\left(\beta_{pt} r_{mt}, \sigma_p^2\right).$$

C. Joint Posterior Distribution
   - Where $\boldsymbol{\theta}_p = (\boldsymbol{\beta}_p, \boldsymbol{\omega}_p)$

- Use Gibbs Sampler
   - $\beta_{pt}, \alpha_p \sim$ Normal, $\delta_p \sim$ Truncated–Normal, $\sigma_{\beta p}^2, \sigma_p^2 \sim$ Inverse χ^2
   - Appendix A Contains Posteriors

$$\text{(B-1)} \quad p(\theta_k^{(i+1)} \mid \theta_1^{(i+1)}, \ldots, \theta_{k-1}^{(i+1)}, \theta_{k+1}^{(i)}, \ldots, \theta_K^{(i)}, \mathbf{y}), \quad k = 1, \ldots, K, \quad i = 1, \ldots, I.$$

IV. Simulation Study

A. SBETA Data-Generating Process

- For Each Parameter Combination $(\alpha,\delta,\sigma_\beta,\sigma)$, **Generate**
  - 500 Sets of True β Series (Each of Length T=120)
  - 500 Samples of r_p Series Based on the True β, σ and r_m (1991–2000)
- Use the Gibbs Sampler to **Estimate** the Model Parameters
  - Iterates 3,000 Draws, Discard the First 1,000 Draws
  - Use the Last 2,000 Draws to Calculate the Posterior Means
  - Compare the Posterior Means to the True Parameters

$$\text{(10)} \qquad \text{RMSE}_{\widehat{\theta}} = \sqrt{\frac{1}{500}\sum_{i=1}^{500}\left(\widehat{\theta}_i - \theta\right)^2}, \qquad \rightarrow \text{For } \boldsymbol{\omega}_p\text{s}$$

$$\text{(11)} \qquad \text{RMSE}_{\widehat{\beta}_{\mathcal{M}}} = \frac{1}{500}\sum_{i=1}^{500}\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\widehat{\beta}_{\mathcal{M},i,t} - \beta_{i,t}\right)^2}. \qquad \rightarrow \text{For } \boldsymbol{\beta}_p\text{s}$$

Table 1: MCMC Estimates SBETA Precisely

TABLE 1

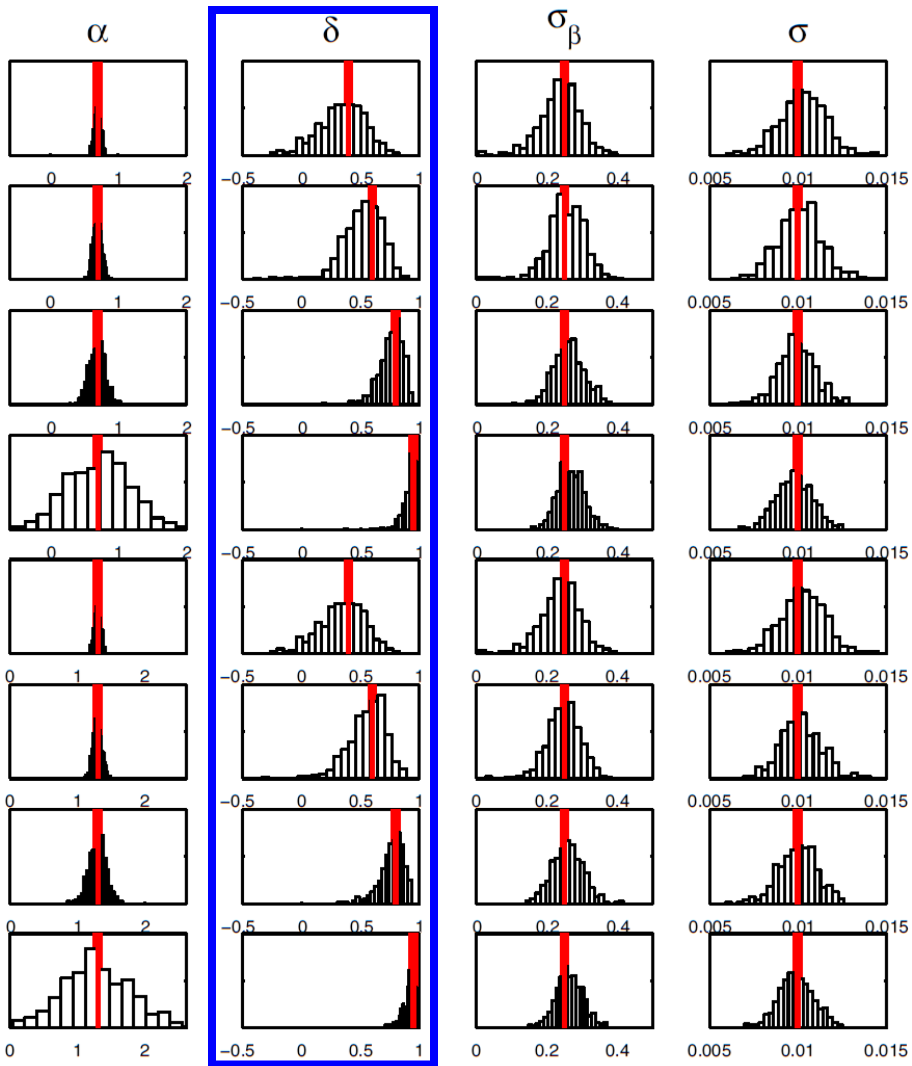
Sampling Properties of the SBETA Estimates

True α	True δ		True	$\sigma_\beta = 0.25,$	$\sigma = 0.01$
		$\hat{\alpha}$	$\hat{\delta}$	$\widehat{\sigma_\beta}$	$\hat{\sigma}$
0.7	0.40	0.6996 (0.0466)	0.3422 (0.2066)	0.2379 (0.0602)	0.0104 (0.000028)
	0.60	0.7066 (0.0656)	0.5517 (0.1716)	0.2486 (0.0510)	0.0102 (0.0012)
	0.80	0.7019 (0.1291)	0.7529 (0.1180)	0.2629 (0.0464)	0.0100 (0.0010)
	0.95	0.7141 (0.4845)	0.9118 (0.0718)	0.2697 (0.0433)	0.0099 (0.0010)
1.3	0.40	1.2977 (0.0514)	0.3422 (0.2113)	0.2398 (0.0615)	0.0103 (0.0014)
	0.60	1.3034 (0.0728)	0.5419 (0.1695)	0.2522 (0.0506)	0.0102 (0.0013)
	0.80	1.3014 (0.1223)	0.7608 (0.1062)	0.2615 (0.0449)	0.0100 (0.0011)
	0.95	1.2872 (0.5214)	0.9140 (0.0629)	0.2654 (0.0411)	0.0099 (0.0010)

Persistent≈high autocorrelation

Less accurate unconditional mean

Figure 1: α, σ_β, σ Are Unbiased, But δ Is Biased



Positive skew due to truncation

Table 2: SBETA Better Captures Time-variation

TABLE 2

Performance of the SBETA and Rolling OLS Betas

		Average	Average Bias		RMSE	
True α	True δ	True Beta	SBETA	OLS	SBETA	OLS
0.7	0.40	0.6995	−0.0003	−0.0011	0.2119	0.2771
	0.60	0.6996	0.0008	−0.0007	0.2115	0.3087
	0.80	0.6913	0.0016	−0.0001	0.2092	0.3916
	0.95	0.7101	0.0010	0.0018	0.2031	0.5630
1.3	0.40	1.2995	−0.0003	−0.0011	0.2119	0.2771
	0.60	1.2989	−0.0005	0.0004	0.2100	0.3116
	0.80	1.2976	0.0001	0.0034	0.2084	0.3948
	0.95	1.2875	0.0005	−0.0036	0.2037	0.5547

$$\mathrm{RMSE}_{\hat{\theta}} = \sqrt{\frac{1}{500}\sum_{i=1}^{500}\left(\hat{\theta}_i - \theta_i\right)^2}.$$

Both are performing well
In estimating averages of βs

But for every case, SBETA has smaller RMSE than rolling OLS

Conclusion: MCMC is Efficient, Quick and Precise

Table 3: SBETA Estimates βs More Accurately

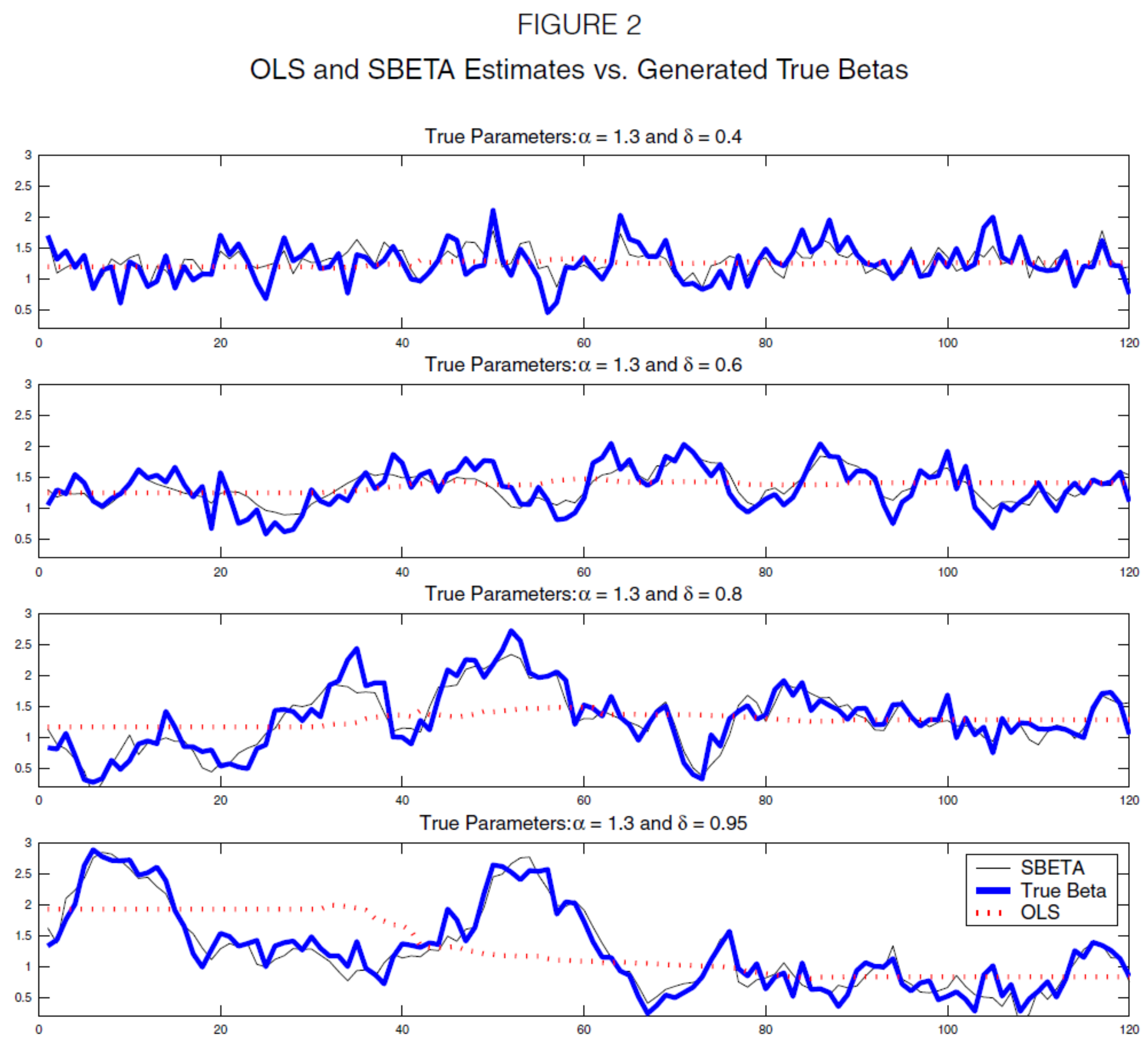
TABLE 3

True Beta vs. Estimated Confidence Bounds

		Percentage of Times True Beta Lies Outside the:			
		SBETA Confidence Bounds		OLS Confidence Bounds	
True α	True δ	$\beta < \widehat{\beta}_{0.05}$	$\beta > \widehat{\beta}_{0.95}$	$\beta < \widehat{\beta}_{0.05}$	$\beta > \widehat{\beta}_{0.95}$
0.7	0.40	6.85%	6.83%	37.55%	37.76%
	0.60	5.76%	5.91%	38.39%	38.11%
	0.80	5.45%	5.25%	39.48%	39.38%
	0.95	5.01%	4.91%	40.60%	40.27%
1.3	0.40	6.85%	6.83%	37.55%	37.76%
	0.60	5.74%	5.96%	38.52%	38.27%
	0.80	5.28%	5.42%	39.64%	39.17%
	0.95	5.04%	5.08%	40.20%	40.72%

True βs touch OLS confidence interval more frequently than SBETA confidence interval

Figure 2: SBETA Closely Follows True βs



In all cases, SBETA closely follows true β, while OLS does not follow it (especially when β is persistent)

IV. Simulation Study (Cont'd)

$$\begin{array}{lrcl} (12) & \mathbf{x}_t & \sim & N_k(0, \mathbf{H}_t) \\ & \boxed{\mathbf{H}_t} & = & \mathbf{C} + \mathbf{A}\boldsymbol{\Sigma}_{t-1} + \mathbf{B}\mathbf{H}_{t-1}, \end{array}$$

B. GARCH Data-Generating Process

$$\boxed{\mathbf{H}_t} = \mathbb{E}\begin{pmatrix} x_{1t}^2 & x_{1t}x_{2t} \\ x_{2t}x_{1t} & x_{2t}^2 \end{pmatrix}$$

$$\boldsymbol{\beta} = \frac{\mathbb{E}(x_{1t}x_{2t})}{\mathbb{E}(x_{2t}^2)}$$

- Ledoit, Santa-Clara and Wolf (2003, RES)
- Where **A**, **B** and **C** are Matrix Parameters
- H_t=Conditional Covariance Matrix
- $\Sigma_t = \mathbf{x}_t \mathbf{x}_t^T$=Cross-product of Variables Observed at Time t
- Use 2 Step Procedure for the Diagonal and Off-diagonal $\mathbf{H}_t$ Elements
- Time-varying βs Are Implied by $\mathbf{H}_t$, i.e. $Cov(x_{1t},x_{2t})/Var(x_{2t})$
- Braun, Nelson and Sunier (1995, JF): Heavy Parametrization, Often Infeasible

- Generate 300 Observations of $\mathbf{x}_t$ and $\mathbf{H}_t$ 100 Times
- Estimate **A**, **B**, **C**, $\mathbf{H}_t$ and $\boldsymbol{\beta}_{pt}$ for Each Simulation
  - Estimate SBETA and Rolling OLS βs as Well

Table 4: SBETA Works Better Under GARCH

TABLE 4

Performance of Beta Estimates when the Data is Generated Using a GARCH Model

Panel A. Predetermined True Parameters of the GARCH Model

$$\mathbf{A} = \begin{bmatrix} 0.1326 & 0.1252 \\ 0.1252 & 0.1208 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0.8486 & 0.8532 \\ 0.8532 & 0.8581 \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} 0.9321 & 0.7585 \\ 0.7585 & 0.7162 \end{bmatrix}$$

Panel B. Parameter Estimates of the GARCH Model

$$\widehat{\mathbf{A}} = \begin{bmatrix} \begin{matrix} 0.1604 \\ (0.0505) \end{matrix} & \begin{matrix} 0.1469 \\ (0.0487) \end{matrix} \\ \begin{matrix} 0.1469 \\ (0.0487) \end{matrix} & \begin{matrix} 0.1458 \\ (0.0499) \end{matrix} \end{bmatrix} \qquad \widehat{\mathbf{B}} = \begin{bmatrix} \begin{matrix} 0.7992 \\ (0.0767) \end{matrix} & \begin{matrix} 0.7814 \\ (0.1204) \end{matrix} \\ \begin{matrix} 0.7814 \\ (0.1204) \end{matrix} & \begin{matrix} 0.8074 \\ (0.0679) \end{matrix} \end{bmatrix} \qquad \widehat{\mathbf{C}} = \begin{bmatrix} \begin{matrix} 1.7282 \\ (1.6589) \end{matrix} & \begin{matrix} 1.3603 \\ (1.0143) \end{matrix} \\ \begin{matrix} 1.3603 \\ (1.0143) \end{matrix} & \begin{matrix} 1.3659 \\ (1.1003) \end{matrix} \end{bmatrix}$$

Panel C. Parameter Estimates of the SBETA Model

	$\hat{\alpha}$	$\hat{\delta}$	$\hat{\sigma}_\beta$	$\hat{\sigma}$
Mean	1.1195	0.7815	0.0416	13.6887
BCI	[0.9984 1.2433]	[0.3036 0.9475]	[0.0171 0.0864]	[11.7388 15.9084]

Panel D. RMSE of the Beta Estimates

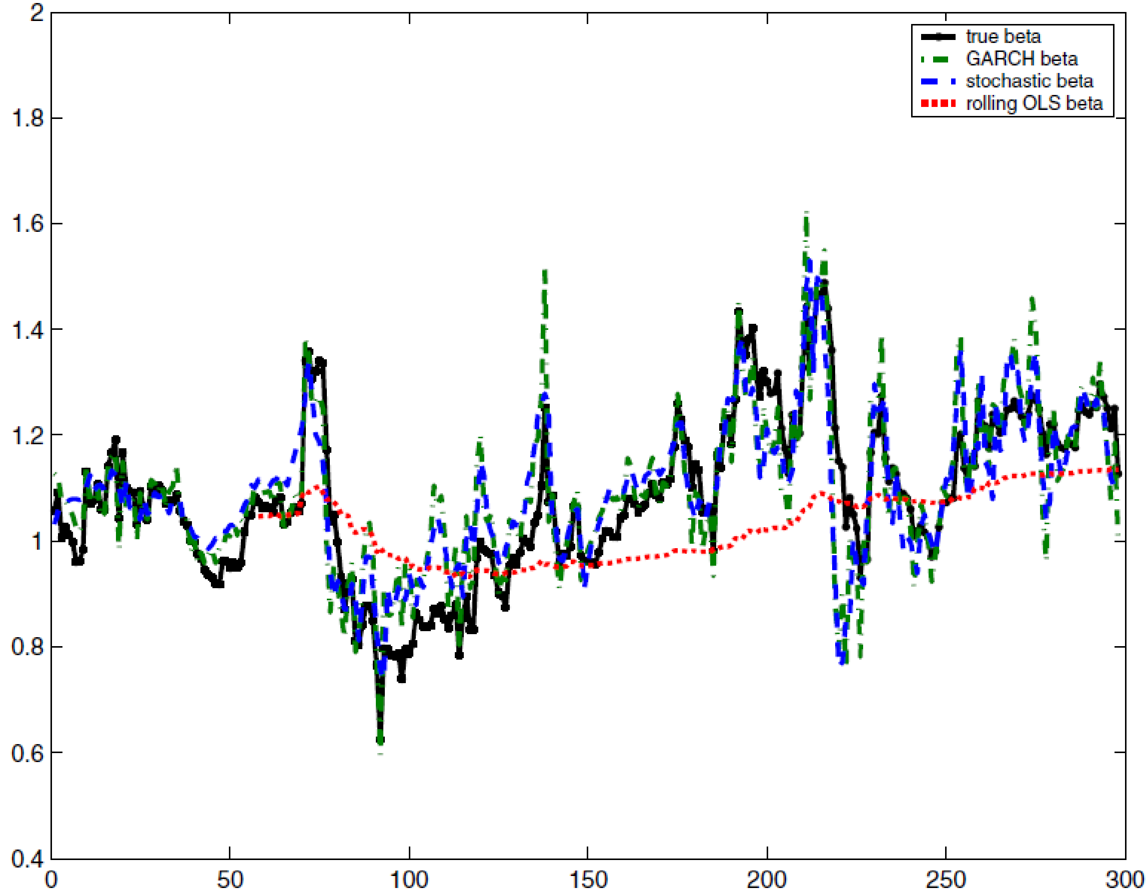
	GARCH	SBETA	Rolling OLS
RMSE	0.1180 (0.0183)	0.1051 (0.0029)	0.2104 (0.0086)

Smallest RMSE

Figure 3: SBETA Works Better Under GARCH

FIGURE 3

SBETA, GARCH, and Rolling Regression Betas Using GARCH Model Generated Data



IV. Simulation Study (Cont'd)

$$y_t = \alpha + \beta_t x_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2) \tag{13}$$

$$\beta_t = \alpha_\beta + \delta_1 z_{1,t-1} + \delta_2 z_{2,t-1}. \tag{14}$$

C. Beta Process with Conditioning Information

- Where x_t=value-weighted index from CRSP
- z_{1t}=Yield on the 3-month T-bill, z_{2t}=Dividend Yield of the Dow Jones Index
  - Shanken (1990, Journal of Econometrics), Ferson and Harvey (1993, RFS), Avramov and Chordia (2005, RFS)
- Draw 100 Data Samples of 300 Observations Each
- **Estimate the Conditional βs Using the Cross-product (Shanken)**

$$y_t = \alpha + \left(\alpha_\beta + \delta_1 z_{1,t-1} + \delta_2 z_{2,t-1}\right) x_t + \varepsilon_t$$

- Estimate based on
  - All True Instruments
  - A Subsample of the True Instruments
  - Proxy Instruments (i.e. the True Instruments Are Unknown)

Table 5: SBETA Outperforms Under Cond'l β

TABLE 5

Performance of Beta Estimates when the Data is Generated Based on Conditional Betas

Panel A. Predetermined True Parameters of the Conditional Beta Model

	α	α_β	δ_1	δ_2	σ
True parameter value	0.000	0.400	0.077	0.144	0.050

Panel B. RMSE Estimates

Model	RMSE
True model	0.101 (0.004)
True z_1 only	0.154 (0.002)
True z_2 only	0.161 (0.002)
SBETA	0.296 (0.005)
Rolling OLS	0.414 (0.006)
Single instrument (T-bill vol.)	0.423 (0.001)
Single instrument (PE ratio)	0.447 (0.002)
GARCH	0.527 (0.033)

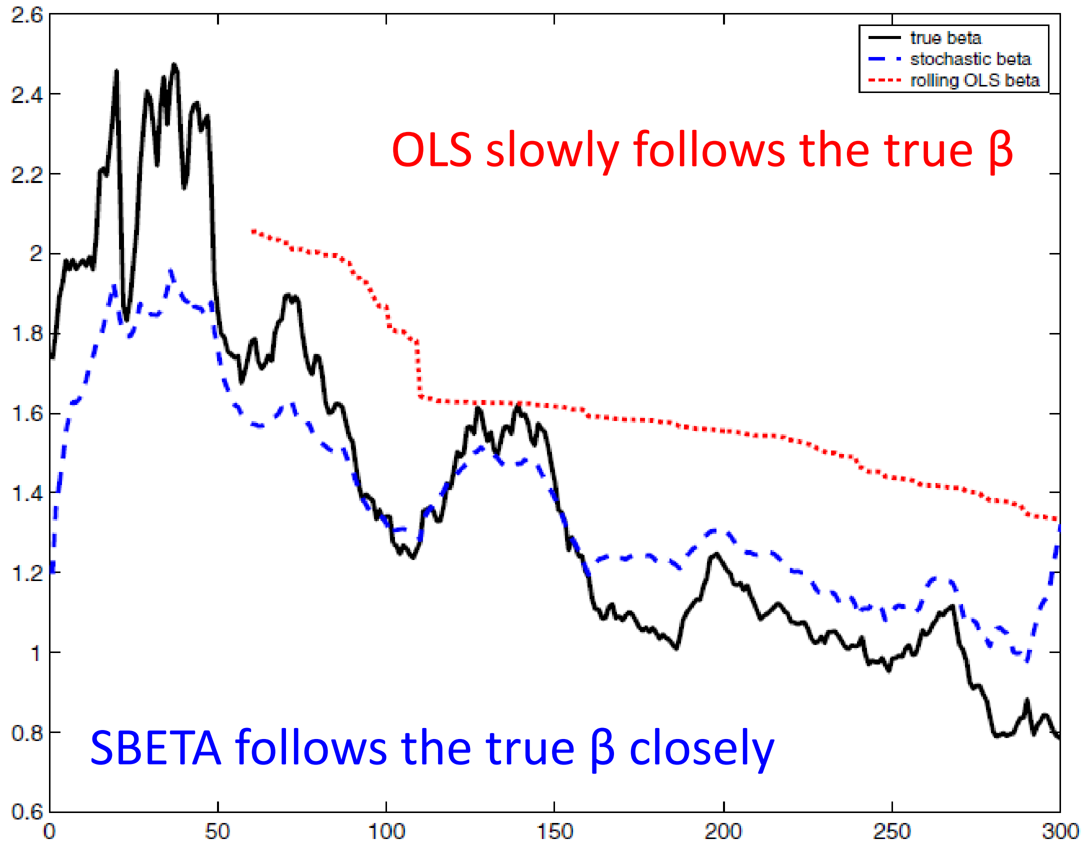
When the true z is available, the conditional β model works best

However, when the true z is unavailable, SBETA outperforms all the other β models

Figure 4: SBETA Outperforms Under Cond'l β

FIGURE 4

SBETA and Rolling Regression Beta Estimates when the True Betas are Generated Based on the Conditional Beta Model



V. Industry Betas over the Past 40 Years

- January 1962–December 2001 Period
- Fit the SBETA Model to r_{pt}s of 5 US Industry Portfolios
  - Value-weighted Across All NYSE, AMEX and NASDAQ Stocks
  - Assign to a Portfolio Based on Their 4-digit SIC Codes
- r_{mt}: Value-weighted CRSP Index Return−1-month T-bill Rate
- Geweke (1992): Numerical Standard Errors (NSE)

$$\text{NSE}(\bar{\theta}_n) = \sqrt{\frac{S_\theta(0)}{n}}. \tag{15}$$

If the Gibbs sampler draws were i.i.d. then

$$\text{s.e.}\ (\bar{\theta}_n) = \sqrt{\frac{S_\theta(0)}{n}} = \sqrt{\frac{{\sigma_\theta}^2}{n}} = \frac{\sigma_\theta}{n^{1/2}}.$$

Portfolio	Industry	SIC Code
1	Manufacturing	2000–3999
2	Utilities	4900–4999
3	Shops (wholesale, retail, and some services)	5000–5999, 7000–7999
4	Finance	6000–6999
5	Other (all remaining industries)	All others

Table 6: βs Vary Differently by Industry

Strongly mean-reverting

Implies that both r_{mt} and β_{pt} affect the portfolio returns r_{pt}

TABLE 6

Analysis of Industry Portfolio Betas

	Manufacturing	Utilities	Shops	Finance	Other
$\hat{\alpha}$					
Mean	1.0138	0.5207	1.1916	0.9983	0.9647
NSE	0.0000	0.0046	0.0005	0.0017	0.0015
BCI	[0.997 1.030]	[0.330 0.673]	[1.138 1.245]	[0.860 1.120]	[0.932 0.986]
$\hat{\delta}$					
Mean	0.1975	0.9299	0.3821	0.9348	0.0136
NSE	0.0331	0.0050	0.0092	0.0049	0.0947
BCI	[−0.386 0.840]	[0.824 0.988]	[0.040 0.673]	[0.841 0.985]	[−0.867 0.937]
$\hat{\sigma}_\beta$					
Mean	0.0602	0.0963	0.2362	0.0742	0.0430
NSE	0.0017	0.0039	0.0018	0.0029	0.0103
BCI	[0.025 0.092]	[0.029 0.147]	[0.167 0.301]	[0.038 0.115]	[0.001 0.1712]
$\hat{\sigma}$					
Mean	0.0075	0.0296	0.0209	0.0229	0.0170
NSE	0.0000	0.0000	0.0000	0.0000	0.0001
BCI	[0.007 0.008]	[0.028 0.031]	[0.020 0.022]	[0.022 0.024]	[0.016 0.018]

Implies that manufacturing returns are largely determined by just the variation in market returns r_{mt}

Highly persistent: implies tight regulation

Low persistence due to diversification

Utilities exhibit highest idiosyncratic risk

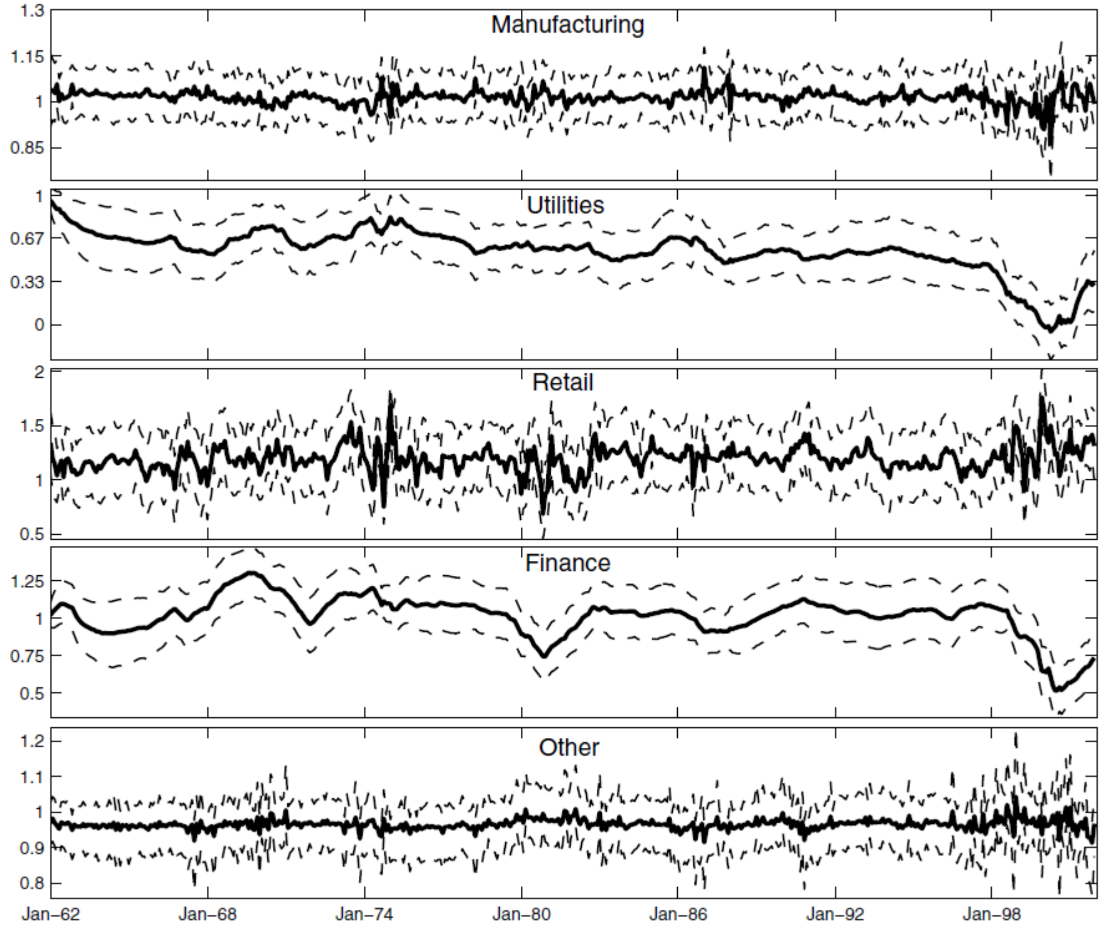
$$r_p = \beta_{pt} r_{mt} + \sigma_p \varepsilon_{pt}$$

$$\beta_{pt} = \alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p) + \sigma_{\beta_p} \nu_{pt}$$

Figure 5: βs Vary Differently by Industry

FIGURE 5

Estimated Industry Portfolio SBETA Betas, $\widehat{\beta}_t$, with 90% Bayesian Posterior Bounds



Manufacturing
→low δ

Utilities
→high δ, low α

Retail
→high δ_β

Finance
→high δ

Other
→very low δ

VI. Out-of-Sample Hedging Application

Table 7: SBETA Facilitates Effective Hedging

(16) This should be close to 0 $\boxed{\Delta^{\mathcal{M}}_{i,T+1}} = \boxed{r_{i,T+1} - \widehat{\beta}^{\mathcal{M}}_{i,T+1} r_{m,T+1},}$ According to CAPM (normative) (if β is correctly estimated)

TABLE 7

Out-of-Sample Hedging Results

	Volatility of Returns	Volatility of Hedging Errors			Significance of Hedging Differences
	$\sigma(\mathbf{r}_i)$	$\sigma(\Delta_{\text{SBETA}})$		$\sigma(\Delta_{\text{OLS}})$	$t\text{-stat}_{\lvert \Delta_{\text{SBETA}} \rvert - \lvert \Delta_{\text{OLS}} \rvert}$
Manufacturing	4.67%	0.63%		0.89%	−11.77
Utilities	3.82%	2.58%		3.10%	−8.65
Retail	5.91%	1.78%	$\ll$	2.32%	−9.97
Finance	4.88%	1.89%		2.49%	−9.22
Other	4.83%	1.31%		1.85%	−11.83

Better hedging performance by SBETA

Better than just OLS (statistically significant)

VII. Asset Pricing Application: Test of the CAPM

- Fama and MacBeth (1973, JPE): 2-step Test of the CAPM
- Whether the SBETA Estimates
  - Produce a More Significant Market Risk Premium
  - Reduce Asset Pricing Errors
  - Can Help Explain the Size, B/M and Volatility Effects (i.e. Anomalies)
- The Asset Pricing Tests Are Based on Individual Stock
  - Brennan, Chordia and Subrahmanyam (1998, JFE)
  - Avramov and Chordia (2006, RFS)
  - Guards against Data-snooping Biases
  - Eliminates the Error in the Variables Problem
  - Avoids the Loss of Information Caused by Sorting Stocks into Groups
    - Litzenberger and Ramaswamy (1979, JFE)
    - Circumvents Berk's (2000, JF) Problem: a Bias in Favor of Rejecting the Model

VII. Asset Pricing Application (Cont'd)

- Monthly Stock Data from CRSP & Compustat: Jan 1964–Dec 2003
  - At Least 10 Years of Monthly Data (Each Stock)
  - Sufficient Data for Calculating Size (CRSP), B/M (Compustat)
  - Winsorizing at .005 and .995 (Each Variable)→809 Stocks as a Result
- Estimate βs by Using CRSP VW Return Index
  1. 60-month Rolling Regression OLS βs
  2. Characteristic-scaled βs Suggested by Avramov and Chordia (2005, RFS)
  3. Conditional βs Suggested by Shanken (1990, Journal of Econometrics)
  4. GARCH βs Implied by Ledoit, Santa-Clara and Wolf (2003, RES)

$$r_{it} = \alpha_t + \lambda_t \widehat{\beta}_{it} + e_{it}, \tag{17}$$

$$r_{it} = \alpha_t + \lambda_t \widehat{\beta}_{it} + \gamma_t \mathbf{C}_{it} + e_{it}, \tag{18}$$

  - Where r_{it}=Excess Returns, $\widehat{\beta}_{it}$=Estimated βs, α_t=Returns Unexplained by βs, λ_t=Market Risk Premium, $\mathbf{C}_{it}$=Firm-level Characteristics & $\boldsymbol{\gamma}_t$=Their Premiums

Table 8: Candidates for the Best β

TABLE 8

Descriptive Statistics of Estimated Stock Betas

	SBETA	OLS	Char. Scaled	Conditional	GARCH
Weighted average beta	0.9852	1.0016	0.4637	0.9818	0.9169
Std. dev. of individual stock betas	0.3913	0.2406	0.0551	0.1629	0.2014
Std. dev. of weighted average beta	0.0466	0.0884	0.2327	0.0366	0.1169

$$(1) \quad r_{pt} = \beta_{pt} r_{mt} + \sigma_p \varepsilon_{pt}, \qquad \varepsilon_{pt} \sim N(0,1)$$

$$\boxed{\beta_{pt}} = \alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p) + \sigma_{\beta_p} \nu_{pt}, \qquad \nu_{pt} \sim N(0,1)$$

$$\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right) \sim p\left(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p\right),$$

$$(13) \quad y_t = \alpha + \beta_t x_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

$$(14) \quad \boxed{\beta_t} = \alpha_\beta + \delta_1 z_{1,t-1} + \delta_2 z_{2,t-1}.$$

$$(12) \quad \mathbf{x}_t \sim N_k(0, \mathbf{H}_t)$$

$$\boxed{\mathbf{H}_t} = \mathbf{C} + \mathbf{A}\boldsymbol{\Sigma}_{t-1} + \mathbf{B}\mathbf{H}_{t-1},$$

Table 9: SBETA Better Explains r_{it}s

TABLE 9

Test of CAPM: Cross-Sectional Regression Results

	SBETA	OLS	Char. Scaled	Conditional	GARCH
Panel A.					
α	−0.19	0.36**	0.66**	0.36*	0.48**
	(−1.14)	(2.78)	(3.79)	(2.50)	(3.31)
λ	$\boxed{1.07^{**}}$	0.39	0.07	0.44	0.29
	$\boxed{(2.92)}$	(1.68)	(0.74)	(1.88)	(1.38)
Adj. R^2	$\boxed{22.90\%}$	6.43%	1.03%	5.57%	4.38%
Panel B.					
α	−0.16	0.40**	0.62**	0.39**	0.51**
	(−0.99)	(3.06)	(3.43)	(2.71)	(3.59)
λ	1.08**	0.31	−0.06	0.35	0.16
	(2.93)	(1.31)	(−0.74)	(1.46)	(0.77)
γ(size)	−38.40	58.66	100.99**	61.19	83.46*
	(−1.35)	(1.85)	(2.91)	(1.84)	(2.48)
γ(book-to-market)	−0.06	−0.18*	−0.15	−0.16*	−0.17*
	(−0.88)	(−2.26)	(−1.68)	(−1.96)	(−2.11)
Adj. R^2	24.25%	8.09%	2.88%	7.28%	6.07%
Panel C.					
α	−0.08	0.30*	0.42**	0.32*	0.41**
	(−0.53)	(2.12)	(2.70)	(2.09)	(2.71)
λ	1.11**	0.18	−0.10	0.20	−0.03
	(2.79)	(0.79)	(−1.20)	(0.89)	(−0.14)
γ(size)	−0.04	0.04	0.07*	0.04	0.06*
	(−1.62)	(1.60)	(2.23)	(1.40)	(2.21)
γ(book-to-market)	−0.07	−0.18*	−0.15	−0.16*	−0.17*
	(−1.01)	(−2.31)	(−1.77)	(−1.96)	(−2.17)
γ(idiosyncratic volatility)	−0.00	0.00**	0.00*	0.00**	0.00**
	(−1.17)	(3.47)	(2.48)	(3.48)	(3.95)
Adj. R^2	25.70%	8.78%	5.00%	7.73%	6.75%

β of SBETA explains r_{it}s

High R^2 for SBETA

No size, B/M anomalies for SBETA case

No volatility anomalies For SBETA case

Significant pricing errors for other βs

Other βs cannot explain the returns

Size anomalies and B/M anomalies for the other models

Volatility anomalies for the other models

But... Fama and French (1992, JF)

Table III
(p. 439)

Coefficients on the size usually have negative signs

Those of the B/M usually have positive signs

β	ln(ME)	ln(BE/ME)	ln(A/ME)	ln(A/BE)	E/P Dummy	E(+)/P
0.15						
(0.46)						
	−0.15					
	(−2 58)					
−0.37	−0.17					
(−1.21)	(−3.41)					
		0.50				
		(5.71)				
			0.50	−0.57		
			(5 69)	(−5.34)		
					0.57	4.72
					(2.28)	(4.57)
	−0.11	0.35				
	(−1.99)	(4.44)				
	−0 11		0.35	−0 50		
	(−2 06)		(4.32)	(−4.56)		
	−0.16				0.06	2.99
	(−3.06)				(0 38)	(3.04)
	−0.13	0.33			−0.14	0.87
	(−2.47)	(4.46)			(−0.90)	(1.23)
	−0.13		0.32	−0.46	−0.08	1.15
	(−2.47)		(4.28)	(−4.45)	(−0.56)	(1.57)

But… Ang, Hodrick, Xing and Zhang (2009, JFE)

Table 3 (p. 9)

– for volatility

– for size
+ for B/M

	Geographic areas		G7 countries		All countries	
	Europe	Asia	G7	G7 Ex. U.S.	All	All Ex. U.S.
Constant	0.823	1.402	1.382	0.871	1.320	0.861
	[2.11]	[2.27]	[3.64]	[2.11]	[3.58]	[2.15]
W-FF idiosyncratic volatility	−0.668	−1.177	−1.747	−1.069	−1.536	−0.604
	[−2.33]	[−3.17]	[−6.40]	[−4.14]	[−5.82]	[−2.32]
$\beta(MKT^W)$	0.145	0.209	0.367	0.331	0.314	0.238
	[1.31]	[2.18]	[4.52]	[3.73]	[3.94]	[2.78]
$\beta(SMB^W)$	0.026	−0.020	−0.055	−0.031	−0.048	−0.039
	[0.39]	[−0.26]	[−1.38]	[−0.59]	[−1.15]	[−0.71]
$\beta(HML^W)$	−0.071	−0.039	−0.057	−0.067	−0.048	−0.051
	[−1.48]	[−0.59]	[−1.77]	[−1.22]	[−1.57]	[−1.02]
Size	−0.087	−0.190	−0.111	−0.099	−0.107	−0.107
	[−2.45]	[−3.19]	[−2.89]	[−2.73]	[−2.95]	[−3.16]
Book-to-market	0.189	0.517	0.293	0.275	0.268	0.241
	[5.51]	[3.52]	[6.01]	[5.15]	[6.79]	[5.85]
Lagged return	0.010	−0.006	0.000	0.003	0.001	0.004
	[3.57]	[−1.45]	[0.12]	[1.31]	[0.58]	[1.78]
Dummy Canada			−0.054	0.240	−0.055	0.190
			[−0.26]	[0.81]	[−0.26]	[0.64]
Dummy France	0.254		−0.060	0.275	−0.024	0.278
	[0.79]		[−0.15]	[0.84]	[−0.06]	[0.84]
Dummy Germany	−0.190		−0.552	−0.195	−0.527	−0.190
	[−0.59]		[−1.49]	[−0.58]	[−1.41]	[−0.58]
Dummy Italy	0.517		0.291	0.636	0.324	0.630
	[1.01]		[0.52]	[1.22]	[0.58]	[1.22]
Dummy Japan			−0.128	−0.043	−0.133	−0.040
			[−0.25]	[−0.10]	[−0.26]	[−0.08]
Dummy U.K.			−0.311		−0.280	
			[−0.94]		[−0.84]	
Dummy other country	0.081				−0.104	0.176
	[0.34]				[−0.33]	[0.79]
Adjusted R^2	0.114	0.115	0.105	0.168	0.099	0.144

Table 9: SBETA Better Explains r_{it}s (Cont'd)

TABLE 9 (continued)

Test of CAPM: Cross-Sectional Regression Results

	SBETA	OLS	Char.Scaled	Conditional	GARCH
Panel D.					
α	−0.15	0.24	0.35*	0.23	0.36*
	(−1.06)	(1.62)	(2.18)	(1.41)	(2.23)
λ	1.12**	0.18	−0.10	0.21	−0.03
	(2.78)	(0.77)	(−1.24)	(0.91)	(−0.18)
γ(size)	−0.04	0.05	0.07*	0.04	0.06*
	(−1.59)	(1.62)	(2.26)	(1.42)	(2.24)
γ(book-to-market)	−0.05	−0.17*	−0.14	−0.15	−0.17*
	(−0.82)	(−2.21)	(−1.66)	(−1.82)	(−2.11)
γ(idiosyncratic volatility)	−0.00	0.00**	0.00*	0.00**	0.00**
	(−1.29)	(3.35)	(2.38)	(3.23)	(3.83)
γ(beta persistence)	0.17*	0.17	0.19*	0.23*	0.16
	(2.27)	(1.91)	(2.03)	(2.43)	(1.75)
Adj. R^2	25.83%	8.93%	5.16%	7.93%	6.92%
Panel E.					
α	−0.16	0.24	0.35*	0.23	0.35*
	(−1.09)	(1.59)	(2.21)	(1.40)	(2.21)
λ	1.15**	0.19	−0.10	0.22	−0.03
	(2.80)	(0.82)	(−1.24)	(0.96)	(−0.16)
γ(size)	−0.04	0.04	0.07*	0.04	0.06*
	(−1.69)	(1.58)	(2.24)	(1.39)	(2.22)
γ(book-to-market)	−0.05	−0.17*	−0.14	−0.14	−0.17*
	(−0.77)	(−2.16)	(−1.64)	(−1.79)	(−2.09)
γ(idiosyncratic volatility)	−0.00	0.00**	0.00*	0.00**	0.00**
	(−1.36)	(3.18)	(2.30)	(3.02)	(3.70)
γ(beta persistence)	0.16*	0.17	0.19*	0.23*	0.17
	(2.07)	(1.90)	(2.01)	(2.40)	(1.79)
γ(beta volatility)	−0.02	−0.00	−0.01	−0.00	−0.00
	(−1.22)	(−0.15)	(−0.50)	(−0.29)	(−0.02)
Adj. R^2	26.05%	9.21%	5.50%	8.23%	7.23%

β persistence (δ) explains r_{it}s (i.e. it is priced)

β volatility (σ_β^2) cannot explain (i.e. not priced)

VIII. Conclusion

- Dynamics of β as a Mean-Reverting Stochastic Process
  - Randomness in the Systematic/Idiosyncratic r_{pt} Components
  - More Kurtosis in the r_{pt} Distribution
  - Testing the Restriction on β's Mean and Volatility
- SBETA Estimates Outperform Competing β Estimates
  - Even under the Competing Environments: GARCH and Conditional β
  - Rolling βs Are Poor: Especially When the True βs Are Persistent
  - Superior Predictability: 30% Smaller Hedging Errors than Rolling OLS
- SBETA Provides Stronger Support for the Conditional CAPM
  - Helps Resolve Asset Pricing Anomalies: Size, B/M and Volatility Effects
  - Better Explains r_{pt}: High R^2s with Insignificant Pricing Errors

Q&A Session

Thanks for Listening

Appendix A. Derivation of Posteriors

$$\text{(A-1)} \quad p(\boldsymbol{\theta}_p \mid \mathbf{r}_p, \mathbf{r}_m) \propto p(\boldsymbol{\omega}_p) L(\boldsymbol{\theta}_p \mid \mathbf{r}_p, \mathbf{r}_m),$$

Priors

$$\begin{aligned}
&\text{(A-2)} & p(\sigma_p^2) &= IG(a_1, b_1) &&\propto \left(\sigma_p^2\right)^{-(a1+1)} \exp\left(-\frac{b_1}{\sigma_p^2}\right) \\
&\text{(A-3)} & p(\alpha_p) &= N(\mu_\alpha, \sigma_\alpha) &&\propto \frac{1}{\sigma_\alpha} \exp\left(-\frac{(\alpha_p - \mu_\alpha)^2}{2\sigma_\alpha^2}\right) \\
&\text{(A-4)} & p(\delta_p) &= \text{truncated}_{(-1,1)} - N(\mu_\delta, \sigma_\delta) &&\propto \frac{1}{\sigma_\delta} \exp\left(-\frac{(\delta_p - \mu_\delta)^2}{2\sigma_\delta^2}\right) \\
&\text{(A-5)} & p(\sigma_{\beta p}^2) &= IG(a_2, b_2) &&\propto \left(\sigma_{\beta p}^2\right)^{-(a_2+1)} \exp\left(-\frac{b_2}{\sigma_{\beta p}^2}\right).
\end{aligned}$$

Likelihood

$$\begin{aligned}
\text{(A-6)} \quad L(\boldsymbol{\theta}_p | \mathbf{r}_p, \mathbf{r}_m) &\propto \prod_{t=1}^{T} N(\alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p), \sigma_{\beta_p}^2) \times \prod_{t=1}^{T} N(\beta_{pt} r_{mt}, \sigma_p^2) \\
&\propto \frac{1}{\sigma_{\beta p}^T} \exp\left(-\frac{\sum_{t=1}^{T} (\beta_{pt} - \alpha_p - \delta_p(\beta_{p,t-1} - \alpha_p))^2}{2\sigma_{\beta p}^2}\right) \\
&\quad \times \frac{1}{\sigma_p^T} \exp\left(-\frac{\sum_{t=1}^{T} (r_{pt} - \beta_{pt} r_{mt})^2}{2\sigma_p^2}\right).
\end{aligned}$$

Appendix A. Derivation of Posteriors (Cont'd)

(A-7)

Joint posterior

$$
\begin{aligned}
p(\boldsymbol{\theta}_p|\mathbf{r}_p,\mathbf{r}_m) \quad \propto \quad & \frac{1}{\sigma_\alpha}\exp\left(-\frac{(\alpha_p-\mu_\alpha)^2}{2\sigma_\alpha^2}\right)\times\frac{1}{\sigma_\delta}\exp\left(-\frac{(\delta_p-\mu_\delta)^2}{2\sigma_\delta^2}\right) \\
& \times\left(\sigma_{\beta p}^2\right)^{-(a_2+1)}\exp\left(-\frac{b_2}{\sigma_{\beta p}^2}\right)\times\left(\sigma_p^2\right)^{-(a1+1)}\exp\left(-\frac{b_1}{\sigma_p^2}\right) \\
& \times\exp\left(-\frac{\sum_{t=1}^T(r_{pt}-\beta_{pt}r_{mt})^2}{2\sigma_p^2}\right)\frac{1}{\sigma_{\beta p}^T} \\
& \times\exp\left(-\frac{\sum_{t=1}^T(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p))^2}{2\sigma_{\beta p}^2}\right).
\end{aligned}
$$

Appendix A. Derivation of Posteriors (Cont'd)

$$
\begin{aligned}
\text{(A-8)}\quad \boxed{p(\beta_{pt}|\text{rest})} &\propto p(\beta_{pt}|\beta_{p,t-1})p(\beta_{p,t+1}|\beta_{pt})p(r_{pt}|\beta_{pt}) \\
&\propto \exp\left(-\frac{(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p))^2}{2\sigma^2_{\beta p}}\right) \\
&\quad \times \exp\left(-\frac{(\beta_{p,t+1}-\alpha_p-\delta_p(\beta_{pt}-\alpha_p))^2}{2\sigma^2_{\beta p}}\right)\exp\left(-\frac{(r_{pt}-\beta_{pt}r_{mt})^2}{2\sigma^2_p}\right) \\
&\propto \exp\left(-\frac{\beta^2_{pt}-2\beta_{pt}\left(\alpha_p+\delta_p(\beta_{p,t-1}-\alpha_p)\right)}{2\sigma^2_{\beta p}}\right) \\
&\quad \times \exp\left(-\frac{\beta^2_{pt}\delta^2_p-2\beta_{pt}\delta_p\left(\beta_{p,t+1}-\alpha_p+\delta_p\alpha_p\right)}{2\sigma^2_{\beta p}}-\frac{\beta^2_{pt}r^2_{mt}-2r_{pt}\beta_{pt}r_{mt}}{2\sigma^2_p}\right) \\
&\propto \exp\left(-\frac{\left(\beta_{pt}-\frac{\left(\sigma^2_{\beta p}r_{mt}r_{pt}+\sigma^2_p\left(\alpha_p(1-\delta_p)^2+\delta_p(\beta_{p,t-1}+\beta_{p,t+1})\right)\right)}{\sigma^2_p\left(1+\delta^2_p\right)+\sigma^2_{\beta p}r^2_{mt}}\right)^2}{2\left(\frac{\sigma_{\beta p}\sigma_p}{\sqrt{\sigma^2_p\left(1+\delta^2_p\right)+\sigma^2_{\beta p}r^2_{mt}}}\right)^2}\right) \\
&\boxed{\propto N\left(\frac{\left(\sigma^2_{\beta p}r_{mt}r_{pt}+\sigma^2_p\left(\alpha_p(1-\delta_p)^2+\delta_p(\beta_{p,t-1}+\beta_{p,t+1})\right)\right)}{\sigma^2_p\left(1+\delta^2_p\right)+\sigma^2_{\beta p}r^2_{mt}},\left(\frac{\sigma_{\beta p}\sigma_p}{\sqrt{\sigma^2_p\left(1+\delta^2_p\right)+\sigma^2_{\beta p}r^2_{mt}}}\right)^2\right).}
\end{aligned}
$$

β_{pt}'s posteriors for t=1,…,T−1

Appendix A. Derivation of Posteriors (Cont'd)

$$
\begin{aligned}
\text{(A-9)} \quad \boxed{p(\beta_{pT}|\text{rest})} &\propto p(\beta_{pT}|\beta_{p,T-1})p(r_{pT}|\beta_{pT}) \\
&\propto \exp\left(-\frac{(\beta_{pT}-\alpha_p-\delta_p(\beta_{p,T-1}-\alpha_p))^2}{2\sigma_{\beta p}^2}\right) \\
&\quad \times \exp\left(-\frac{(r_{pT}-\beta_{pT}r_{mT})^2}{2\sigma_p^2}\right) \\
&\propto \exp\left(-\frac{\left(\beta_{pT}-\frac{\sigma_{\beta p}^2 r_{mT}r_{pT}+\sigma_p^2(\alpha_p+\delta_p(\beta_{p,T-1}-\alpha_p))}{\sigma_p^2+\sigma_{\beta p}^2 r_{mT}^2}\right)^2}{2\left(\frac{\sigma_{\beta p}\sigma_p}{\sqrt{\sigma_p^2+\sigma_{\beta p}^2 r_{mT}^2}}\right)^2}\right) \\
&\boxed{\propto N\left(\frac{\sigma_{\beta p}^2 r_{mT}r_{pT}+\sigma_p^2(\alpha_p+\delta_p(\beta_{p,T-1}-\alpha_p))}{\sigma_p^2+\sigma_{\beta p}^2 r_{mT}^2}, \left(\frac{\sigma_{\beta p}\sigma_p}{\sqrt{\sigma_p^2+\sigma_{\beta p}^2 r_{mT}^2}}\right)^2\right).}
\end{aligned}
$$

β_{pT}'s posterior

Appendix A. Derivation of Posteriors (Cont'd)

$$
\text{(A-10)} \qquad \boxed{p(\alpha_p|\text{rest})} \propto \exp\left(-\frac{\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2}{2\sigma_{\beta p}^2}\right)
$$

$$
\times \exp\left(-\frac{(\alpha_p-\mu_\alpha)^2}{2\sigma_\alpha^2}\right)
$$

$$
\propto \exp\left(-\frac{\alpha_p^2 T(1-\delta_p)^2 - 2\alpha_p(1-\delta_p)\sum_{t=1}^{T}\left(\beta_{pt}-\delta_p\beta_{p,t-1}\right)}{2\sigma_{\beta p}^2}\right.
$$

$$
\left. -\frac{\alpha_p^2 - 2\alpha_p\mu_\alpha}{2\sigma_\alpha^2}\right)
$$

$$
\propto \exp\left(-\frac{\left(\alpha_p - \frac{(1-\delta_p)\sigma_\alpha^2\sum_{t=1}^{T}\left(\beta_{pt}-\delta_p\beta_{p,t-1}\right)+\mu_\alpha\sigma_{\beta p}^2}{T(1-\delta_p)^2\sigma_\alpha^2+\sigma_{\beta p}^2}\right)^2}{2\left(\frac{\sigma_{\beta p}\sigma_\alpha}{\sqrt{T(1-\delta_p)^2\sigma_\alpha^2+\sigma_{\beta p}^2}}\right)^2}\right)
$$

$$
\boxed{\propto N\left(\frac{(1-\delta_p)\sigma_\alpha^2\sum_{t=1}^{T}\left(\beta_{pt}-\delta_p\beta_{p,t-1}\right)+\mu_\alpha\sigma_{\beta p}^2}{T(1-\delta_p)^2\sigma_\alpha^2+\sigma_{\beta p}^2},\ \left(\frac{\sigma_{\beta p}\sigma_\alpha}{\sqrt{T(1-\delta_p)^2\sigma_\alpha^2+\sigma_{\beta p}^2}}\right)^2\right),}
$$

α_p's posterior

Appendix A. Derivation of Posteriors (Cont'd)

$$\text{(A-11)} \quad \boxed{p(\delta_p|\text{rest})} \propto \exp\left(-\frac{\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2}{2\sigma_{\beta p}^2}\right)$$

$$\times \exp\left(-\frac{(\delta_p-\mu_\delta)^2}{2\sigma_\delta^2}\right)$$

$$\propto \exp\left(-\frac{\delta_p^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2-2\delta_p\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)(\beta_{pt}-\alpha_p)}{2\sigma_{\beta p}^2}\right.$$

$$\left.-\frac{\delta_p^2-2\delta_p\mu_\delta}{2\sigma_\delta^2}\right)$$

$$\propto \exp\left(-\frac{\left(\delta_p-\frac{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)(\beta_{pt}-\alpha_p)+\sigma_{\beta p}^2\mu_\delta}{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2+\sigma_{\beta p}^2}\right)^2}{2\left(\frac{\sigma_{\beta p}\sigma_\delta}{\sqrt{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2+\sigma_{\beta p}^2}}\right)^2}\right)$$

$$\boxed{\propto \text{truncated}_{(-1,1)}N\left(\frac{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)(\beta_{pt}-\alpha p)+\sigma_{\beta p}^2\mu_\delta}{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2+\sigma_{\beta p}^2}, \left(\frac{\sigma_{\beta p}\sigma_\delta}{\sqrt{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2+\sigma_{\beta p}^2}}\right)^2\right)}$$

δ_p's posterior

Appendix A. Derivation of Posteriors (Cont'd)

$$\text{(A-12)} \quad \boxed{p(\sigma^2_{\beta p}|\text{rest})} \propto \frac{1}{\sigma^T_{\beta p}} \exp\left(-\frac{\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2}{2\sigma^2_{\beta p}}\right)$$

$$\times \left(\sigma^2_{\beta_p}\right)^{-(a_2+1)} \exp\left(-\frac{b_2}{\sigma^2_{\beta_p}}\right)$$

$$\propto \left(\sigma^2_{\beta p}\right)^{-\left(\frac{T+2a_2}{2}+1\right)}$$

$$\times \exp\left(-\frac{(T+2a_2)\frac{\left(\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2+2b_2\right)}{T+2a_2}}{2\sigma^2_{\beta p}}\right)$$

$$\boxed{\propto \operatorname{Inv}-\chi^2\left(T+2a_2, \frac{\left(\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2+2b_2\right)}{T+2a_2}\right).}$$

$\sigma_{\beta p}{}^2$'s posterior

Appendix A. Derivation of Posteriors (Cont'd)

$$
\text{(A-13)}\quad \boxed{p(\sigma_p^2|\text{rest})} \propto \frac{1}{\sigma_p^T}\exp\left(-\frac{\sum_{t=1}^{T}(r_{pt}-\beta_{pt}r_{mt})^2}{2\sigma_p^2}\right)\left(\sigma_p^2\right)^{-(a1+1)}\exp\left(-\frac{b_1}{\sigma_p^2}\right)
$$

$$
\propto \left(\sigma_p^2\right)^{-\left(\frac{T+2a_1}{2}+1\right)}\exp\left(-\frac{(T+2a_1)\frac{\sum_{t=1}^{T}\left(r_{pt}-\beta_{pt}r_{mt}+2b_1\right)^2}{(T+2a_1)}}{2\sigma_p^2}\right)
$$

$$
\boxed{\propto \operatorname{Inv}-\chi^2\left(T+2a_1,\frac{\sum_{t=1}^{T}(r_{pt}-\beta_{pt}r_{mt})^2+2b_1}{T+2a_1}\right).}
$$

$\sigma_p{}^2$'s Posterior

Appendix C. Multivariate Extension

i. Multivariate Distribution of βs

ii. Vector Specification for the Factor Sensitivities

$$
\begin{aligned}
\text{(C-1)} \quad L(\boldsymbol{\theta}_p|\mathbf{r}_p,\mathbf{r}_m) &\propto p(\mathbf{r}_p,\boldsymbol{\beta}_p|\boldsymbol{\theta}_p^{\text{rest}},\mathbf{f}) \\
&\propto \prod_{t=1}^{T} p(r_{pt}|\boldsymbol{\theta}_p,\mathbf{f}_t)\times p(\boldsymbol{\beta}_{pt}|\boldsymbol{\theta}_p^{\text{rest}},r_{pt},\mathbf{f}_t) \\
&= \prod_{t=1}^{T} N\left(\mathbf{f}_t\boldsymbol{\beta}_{pt},\sigma_p^2\right)\times N_k\left(\boldsymbol{\alpha}_p+\boldsymbol{\delta}_p(\boldsymbol{\beta}_{p,t-1}-\boldsymbol{\alpha}_p),\Sigma_{\beta_p}\right) \\
&\propto \frac{1}{\sigma_p^T}\exp\left(-\frac{\sum_{t=1}^{T}\left(r_{pt}-\mathbf{f}_t\boldsymbol{\beta}_{pt}\right)^2}{2\sigma_p^2}\right) \\
&\quad \times\frac{1}{|\Sigma_{\beta_p}|^T}\exp\left(-\frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{\beta}_{pt}-\boldsymbol{\mu}_{\beta t}\right)'\Sigma_{\beta_p}^{-1}\left(\boldsymbol{\beta}_{pt}-\boldsymbol{\mu}_{\beta t}\right)\right),
\end{aligned}
$$

- Where $\boldsymbol{\beta}_p=(\beta_{p1},\ldots,\beta_{pT})^T$, $\boldsymbol{\mu}_{\beta t}=\boldsymbol{\alpha}_p-\boldsymbol{\delta}_p(\boldsymbol{\beta}_{p,t-1}-\boldsymbol{\alpha}_p)$, $\boldsymbol{\delta}$ Is a Diagnoal Matrix That Gives the Persistence Parameters

Appendix C. Multivariate Extension

- Joint Uninformative Prior Distribution

$$\text{(C-2)} \qquad p(\sigma_p^2, \alpha_p, \delta_p, \sigma_{\beta p}^2) \quad \propto \quad \sigma_p^{-2} \left| \Sigma_{\beta_p} \right|^{-1}.$$

- Joint Posterior Distribution

$$\text{(C-3)} \quad p(\boldsymbol{\theta}_p | \mathbf{r}_p, \mathrm{f}) \quad \propto \quad \frac{1}{\sigma_p^{T+2}} \exp\left(-\frac{\sum_{t=1}^{T} \left(r_{pt} - \mathrm{f}_t \boldsymbol{\beta}_{pt} \right)^2}{2\sigma_p^2} \right)$$

$$\times \frac{1}{\left| \Sigma_{\beta_p} \right|^{T+1}} \exp\left(-\frac{1}{2} \sum_{t=1}^{T} \left(\boldsymbol{\beta}_{pt} - \boldsymbol{\mu}_{\beta t} \right)' \Sigma_{\beta_p}^{-1} \left(\boldsymbol{\beta}_{pt} - \boldsymbol{\mu}_{\beta t} \right) \right).$$

- Posterior Distribution of $\boldsymbol{\Sigma}_{\beta p}{}^{-1} \approx \sigma_{\beta p}{}^{-2} \sim$ Wishart Distribution

$$\text{(C-4)} \qquad \Sigma_{\beta_p}^{-1} \sim \mathrm{Wish}_k(\nu, \mathbf{S}),$$

where $\nu = 2T - k + 1$ and $\mathbf{S} = \sum_{t=1}^{T} (\boldsymbol{\beta}_{pt} - \boldsymbol{\mu}_{\beta t})(\boldsymbol{\beta}_{pt} - \boldsymbol{\mu}_{\beta t})'$.

Appendix C. Multivariate Extension (Cont'd)

- Posterior Distribution of $\boldsymbol{\beta}_{pt} \approx \beta_{pt} \sim$ Multivariate Normal Distribution

(C-5) $$p(\boldsymbol{\beta}_t \mid \text{rest}) \quad \sim \quad N_k(\boldsymbol{\mu}_t^*, \Sigma_\beta^*),$$

where

$$\boldsymbol{\mu}_t^* = \left(\frac{1}{\sigma_p^2}\mathbf{f}_t\mathbf{f}_t' + \Sigma_\beta^{-1} + \boldsymbol{\delta}'\Sigma_\beta^{-1}\boldsymbol{\delta}\right)^{-1}\left(\frac{r_p}{\sigma_p^2}\mathbf{f}_t + \Sigma_\beta^{-1}\left(\boldsymbol{\alpha} + \boldsymbol{\delta}(\boldsymbol{\beta}_{t-1} - \boldsymbol{\alpha})\right)\right.$$
$$\left. + \boldsymbol{\delta}'\Sigma_\beta^{-1}\boldsymbol{\delta}\left(\boldsymbol{\delta\alpha} - \boldsymbol{\alpha} + \boldsymbol{\beta}_{t+1}\boldsymbol{\delta}^{-1}\right)\right)$$
$$\Sigma^* = \left(\frac{1}{\sigma_p^2}\mathbf{f}_t\mathbf{f}_t' + \Sigma_\beta^{-1} + \boldsymbol{\delta}'\Sigma_\beta^{-1}\boldsymbol{\delta}\right)^{-1}.$$

- Posterior Distribution of $\sigma_p^2 \sim$ Chi-square Distribution

(C-6) $$p(\sigma_p^2 \mid \text{rest}) \quad \propto \quad \text{scaled} - \chi^{-2}\left(T, \frac{1}{T}\sum_{t=1}^{T}\left(r_{pt} - \mathbf{f}_t'\boldsymbol{\beta}_{pt}\right)^2\right).$$

- Posterior Distribution of $\boldsymbol{\alpha}_p \approx \alpha_p \sim$ Multivariate Normal Distribution

(C-7) $$p(\alpha_p|\text{rest}) \propto N_k\left(\left(T(\mathbf{I} - \boldsymbol{\delta}_p)\right)^{-1}\sum_{t=1}^{T}\left(\boldsymbol{\beta}_{pt} - \boldsymbol{\delta}_p\boldsymbol{\beta}_{p,t-1}\right), \Sigma_\alpha\right),$$

where $\Sigma_\alpha = \left(\Sigma_{\beta p}\left(\sqrt{T}(\mathbf{I} - \boldsymbol{\delta}_p)\right)^{-1}\right)$.