JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS VOL. 40, NO. 4, DECEMBER 2005
COPYRIGHT 2005, SCHOOL OF BUSINESS ADMINISTRATION, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

Bayesian Analysis of Stochastic Betas

Gergana Jostova and Alexander Philipov*

Abstract

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

Beta estimates are essential for many areas of modern finance, including asset pricing, cash flow valuation, risk management, and performance evaluation. The modeling and measurement of this unobservable parameter can have far-reaching implications. For example, Avramov and Chordia (2005) suggest that previously documented anomalies such as the book-to-market (BM), size, and firm characteristics effects in the cross section of returns may in fact arise from the dynamic behavior of betas. Petkova and Zhang (2005) and Zhang (2004) link the BM effect to the different behavior of value and growth betas across the business cycle. Furthermore, Ang and Liu (2004) show that beta variation is also essential for cash flow valuation, especially at longer horizons. From a portfolio management perspective, betas are crucial for effective hedging and they underlie important performance measures. Therefore, the precise measurement of latent betas is an issue of primary importance.

This study improves the precision of beta estimates by introducing a new beta model that describes systematic risk as a general mean-reverting stochastic process. This specification offers several advantages. First, it includes existing beta models as special cases, which allows one to test their imposed restrictions. Indeed, our model combines both time variation and a stochastic component in

*Jostova, jostova@gwu.edu, George Washington University, Department of Finance, 2023 G St, Washington, DC 20052, and Philipov, philipov@american.edu, American University, Kogod School of Business, Department of Finance, 4400 Massachusetts Ave, Washington, DC 20016. We thank Doron Avramov, Pierluigi Balduzzi, Stephen Brown (the editor), Eric Jacquier, Pedro Santa-Clara (associate editor and referee), Robert Savickas, and Lu Zhang for insightful comments.

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

the beta process, whereas early beta models assume that beta has a constant mean and a random component, while later models assume a time-varying mean but no separate random component for beta (i.e., multivariate GARCH models and models conditioning beta on observed instruments).[1] Second, our model accounts for the documented excess kurtosis in security returns. Specifically, in previous time-varying beta models randomness in beta is solely driven by return or instrument innovations, thereby constraining beta persistence and return kurtosis by a set of common parameters and restricting kurtosis below empirically observed levels. We relax this restriction by allowing for both time variation and randomness in beta. Third, our model captures the theoretically motivated and empirically observed high beta persistence. According to Berk, Green, and Naik (1999), beta persistence accounts for the momentum in returns, while Ang and Chen (2003) show that it also generates the BM effect in the cross section of stock returns.

Imposing mean reversion in the beta process is important from both a theoretical and empirical standpoint. Theoretically, within a single-factor framework, since returns are linear functions of betas, mean reversion in betas is essential to ensure return stationarity (see Gomes, Kogan, and Zhang (2003)). At the aggregate level, stationarity in returns is in turn necessary to ensure that the growth rate of output does not explode (in line with the standard growth literature; e.g., see Kaldor (1961)). Empirically, mean reversion in betas is confirmed by Ang and Chen (2003), Campbell and Vuolteenaho (2005), and Petkova and Zhang (2005).

The stochastic beta (SBETA) model is estimated using Bayesian methods, which offer a number of desirable features.[2] For one, Bayesian prior-posterior relations reflect the investors' learning process. A Bayesian approach also incorporates the inherent uncertainty about model parameters. Finally, our proposed Markov chain Monte Carlo (MCMC) algorithm provides exact finite sample inference, unlike maximum likelihood estimation (MLE) or generalized method of moments (GMM), which run into convergence problems and rely on asymptotic approximations. Reliance on asymptotics can be misleading, as Ang and Chen (2003) document.

Our SBETA estimates, which we examine in a simulation study, are highly precise relative to competing GARCH betas, betas scaled by aggregate and firm-level variables, and rolling regression betas, even when the true betas are generated based on these competing models. We also find that while ordinary least squares (OLS) betas approximate well the average level of systematic risk, they fail to capture time variation in beta, especially when betas are persistent. This may explain the implausibly large time variation in beta, which Lewellen and Nagel (2003) compute by rolling regressions, that is necessary to justify departures from the capital asset pricing model (CAPM).

We apply our framework to U.S. industry portfolios and show that betas exhibit both mean reversion and stochastic behavior. Financial and utilities firms

[1] Early models include Blume (1971), Levy (1971), Vasicek (1973), and Chen and Lee (1982). Bollerslev, Engle, and Wooldridge (1988), Turtle, Buse, and Korkie (1994), and Ledoit, Santa-Clara, and Wolf (2003) offer multivariate GARCH models for variances, while Braun, Nelson, and Sunier (1995) introduce a GARCH model for beta. Shanken (1990) and Ferson and Harvey (1993) condition betas on aggregate instruments, while Avramov and Chordia (2005) use firm-level instruments.

[2] Bayesian estimation of beta is also used by Tamayo (2002) and Ang and Chen (2003).

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

exhibit strikingly high beta persistence, manufacturing betas exhibit the strongest mean reversion, and retail and wholesale firm betas possess the largest stochastic component.

In addition, the SBETA model generates superior beta forecasts. Out-of-sample hedging results show that the SBETA model provides an economically and statistically significant improvement in hedging effectiveness, reducing the volatility of hedging errors relative to OLS by up to 30%.

In asset pricing tests, SBETA estimates provide strong support for the conditional CAPM. The asset pricing tests are based on individual stock SBETA betas and demonstrate a strongly significant market risk premium without any risk-related firm-specific returns. Specifically, SBETA betas explain 23% of the cross-sectional variation in individual stock returns, while OLS, GARCH, and conditional betas explain between 1% and 6%. Furthermore, SBETA estimates help resolve previously documented asset pricing anomalies, such as the size, BM, and idiosyncratic risk effects in the cross section of stock returns. These findings lend empirical support to the theoretical prediction of Gomes, Kogan, and Zhang (2003) that firm-level size and BM are associated with expected returns because they are correlated with the true beta.

The remainder of the paper is organized as follows. Section II introduces the model. Section III develops the Bayesian setup. Section IV describes the simulation study. Section V illustrates the behavior of industry betas. Section VI discusses the out-of-sample hedging application, and Section VII presents tests of the CAPM. Section VIII concludes. The Appendix provides derivation details.

II. The Model

A long history of evidence suggests that betas respond to macroeconomic and company-specific news.[3] The news arrives randomly, implying a stochastic process for beta. The proposed mean-reverting SBETA is an integral part of the single-factor asset pricing model,

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

where r_{pt} and r_{mt} are the portfolio and market excess returns, β_{pt} is portfolio p's sensitivity to market movements, α_p, δ_p, and $\sigma^2_{\beta p}$ are the unconditional mean, persistence, and conditional volatility of portfolio p's beta, σ^2_p is portfolio p's idiosyncratic return volatility, and $p(\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p)$ is the joint distribution of the model parameters, and ε_{pt} and ν_{pt} are the stochastic components of excess return and beta, respectively.[4]

[3] Factors such as changes in GNP, inflation, as well as changes in the firm's leverage and assets, are found to affect beta in studies by Baxter (1967), Beaver, Kettler, and Scholes (1970), Ben-Zion and Shalit (1975), and Turnbull (1977).

[4] Negative correlation between the two terms indicates leverage effects in beta. While leverage effects are present in time-varying stochastic volatility (SVOL), Braun, Nelson, and Sunier (1995) do not find such effects in the beta process. Similarly, in the estimation, we do not find any leverage effects for either single-security or portfolio betas, so we do not impose it explicitly.

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

Mean reversion in beta is important. In the single-factor framework of model (1), since returns are a linear combination of beta and a stationary random component, stationarity in beta is a necessary condition for stationarity in returns. Stationarity in returns is in turn necessary in equilibrium to ensure that the aggregate growth rate of output does not explode (see Kaldor (1961)). Mean reversion in beta is also justified by the theoretical models of Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003). In these models, beta is determined by the number of available positive net present value (NPV) projects (growth opportunities), which themselves are mean reverting and vary with the business cycle. These models imply that beta reverts to its long-term mean over a period of time consistent with the business cycle, a prediction that is confirmed empirically by Petkova and Zhang (2005) and Zhang (2004). Ang and Chen (2003) also find evidence of time-varying betas with a slow mean-reverting component. Similarly, Campbell and Vuolteenaho (2005) show that discount rate betas exhibit economically large and statistically significant reversal patterns linked to growth and value characteristics of stocks.

Our beta process is sufficiently general to capture a wide variety of economic scenarios. For example, discrete and rare one-time shocks to beta—such as those caused by changes in leverage due to issuance or retirement of debt or changes in cash flows due to mergers or acquisitions—can be captured by our model through a high level of persistence, namely one that is so close to unity that the value of beta does not revert to a long-term mean for a long period of time. Even in these scenarios, mean reversion is necessary to ensure that the value of beta cannot increase or decrease indefinitely, i.e., to ensure stationarity. The dynamics of beta are also affected by more frequent and dynamic events, such as changing economic conditions over the business cycle or changing firm characteristics, which cause reversion to the long-term mean over shorter periods of time and can be captured by a lower level of beta persistence.

The mean-reverting beta process is described by the parameters, α_p, δ_p, and $\sigma^2_{\beta_p}$. Based on the model, the unconditional mean and variance of the beta are, respectively,

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

$$\text{var}(\beta_p) = \frac{\sigma^2_{\beta_p}}{1-\delta_p^2}. \tag{3}$$

The parameter α_p represents the unconditional mean of beta, i.e., the long-run level of beta (equation (2)). The parameter $\sigma^2_{\beta_p}$ represents the conditional variance of beta. Persistence (or clustering) in the β_p process is measured by the parameter δ_p. This parameter reveals the strength of mean reversion toward the unconditional mean after a shock in beta and is less than one, $|\delta_{\beta_p}| < 1$, to assure stationarity in beta and hence in returns. In equation (3), note that the stronger the mean reversion (lower persistence, δ_p), the lower the unconditional beta volatility.

Our model formulation is sufficiently general to include a variety of existing specifications for the beta-generating process as special cases. With perfect mean reversion (i.e., no persistence, $\delta_p = 0$), betas have a constant mean and variance,

$E(\beta_p) = E(\beta_{pt}|\beta_{p,t-1}, \text{rest}) = \alpha_p$ and $\text{var}(\beta_{pt}) = \text{var}(\beta_{pt}|\beta_{p,t-1}, \text{rest}) = \sigma^2_{\beta p}$ (see equation (3)). In this case, the model reduces to that of Chen and Lee (1982), wherein beta represents a constant with "noise" and exhibits no persistence of shocks. With zero beta volatility, $\sigma^2_\beta = 0$, the model can be transformed into a GARCH type filter similar to the model of Braun, Nelson, and Sunier (1995), into the conditional beta model of Shanken (1990), or into the characteristic-based beta model of Avramov and Chordia (2005) (as long as the instruments are stationary themselves). When both the persistence and conditional volatility in beta are zero ($\delta = 0$ and $\sigma^2_\beta = 0$), the model reduces to a constant beta model. Thus, the significance of the parameters of the beta process allows us to test the restrictions imposed by these embedded models.

III. Estimation Methodology

We estimate the parameters of the model with Bayesian methods. The Bayesian estimation approach involves specifying a joint prior distribution, identifying the likelihood function, and computing the joint posterior distribution of the parameters and the data. We construct a MCMC algorithm to sample the parameters from their joint posterior distribution. The estimates of the parameters are summarized by the sample statistics of the MCMC draws.

A. Choice of Prior Distributions

We assume the joint prior distribution is the product of the independent priors of each parameter. It is necessary to specify a prior distribution for the parameters $\alpha_p, \delta_p, \sigma^2_{\beta_p}$, and σ^2_p. We choose to work with diffuse priors.

A diffuse but proper prior reflects some belief about the distributions of the parameters, while also reflecting the lack of information about the mean and variance of these distributions. The lack of information is reflected in the large variances of the priors. For the unconditional mean parameter, α_p, we suggest a normal prior with mean 1 and standard deviation 10. The unity mean is motivated by the fact that α_p is the unconditional mean of beta. Values for δ_p that ensure stationarity lie in the interval $(-1, 1)$, so we select a prior of δ_p that has support within this interval. Hence, for δ_p, we consider a truncated normal prior with mean 0.5 and standard deviation 10. For the variance of the beta, $\sigma^2_{\beta_p}$, we suggest an inverse gamma (*IG*) prior (typically used in the literature to model the distribution of unknown variances (see Box and Tiao (1992)) with shape and scale parameters equal to 0.001. The scale parameter implies a variance of 1/0.001. For the variance of returns parameter, σ^2_p, we select an *IG* prior with shape and scale parameters equal to 0.001. The joint prior distribution for the vector of parameters $\boldsymbol{\omega}_p = (\alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p)$ therefore becomes

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

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

B. Likelihood Function

Based on the SBETA model (equation (1)), at time t beta follows a normal distribution,

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

The model also assumes that each portfolio return follows a normal distribution,

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

Based on the specified conditional distributions of the betas and portfolio returns, the SBETA model has the likelihood function,

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

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

C. Joint Posterior Distribution

By Bayes theorem, the joint posterior distribution of the augmented parameter set, $\boldsymbol{\theta}_p = (\boldsymbol{\beta}_p, \boldsymbol{\omega}_p) = (\boldsymbol{\beta}_p, \alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p)$, is proportional to the product of the joint prior distribution and the likelihood function,

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

Using the specified prior distribution (equation (5)) and the derived likelihood (equation (7)), we obtain the following joint posterior distribution of all parameters,

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

Ideally, we would sample all parameters directly from their joint posterior distributions. However, this may be infeasible because the posterior is not a standard distribution. Therefore, we use the Gibbs sampler, which successively samples from the conditional posterior distribution of each parameter given the remaining parameters and the data (see Gelfand and Smith (1990)). After a large number of iterations, under broad regularity conditions, this process provides samples from the joint posterior distribution. Summary statistics from these

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

samples are used to draw inferences about the true parameter values. The first n draws (burn-in stage) are discarded to eliminate dependence on initial values, with summary statistics used for inference computed based on the remaining iterations (monitoring stage). The Gibbs sampler thus simplifies the problem of drawing from a joint posterior to the problem of drawing from sections of this posterior that are smaller and that can be represented by standard distributions. The conditional posterior distributions of each parameter are derived in Section A of the Appendix. Note that β_{pt} and α_p have normal conditional posteriors, whereas due to its truncated prior, δ_p's conditional posterior is truncated normal. The conditional posteriors for the two volatility parameters are inverse χ^2 distributions. Since all parameters have standard conditional posterior distributions, they can be drawn directly, significantly simplifying the Gibbs sampling procedure. Section B of the Appendix provides details on the implementation of the Gibbs sampler.

IV. Simulation Study

The simulation study allows us to know the unobservable true beta process and to measure the precision of the SBETA estimates relative to alternative beta estimates under a variety of specifications for the true beta-generating process. We compare the SBETA estimates to rolling OLS, GARCH, and conditional beta (i.e., betas scaled by instruments) estimates, using root mean square error (RMSE) measures. We assess the performance of the SBETA estimates for situations in which true betas are generated according to different specifications of the SBETA model, as well as according to the GARCH model of Ledoit, Santa-Clara, and Wolf (2003) and the conditional beta model of Shanken (1990).

A. SBETA Data-Generating Process

The simulation involves two stages: a generation stage and an estimation stage. In the first stage, portfolio returns and true beta series are generated according to the SBETA model using a set of pre-specified true parameters, $(\alpha, \delta, \sigma_\beta, \sigma)$, and actual realized market excess returns for a period of 10 years. In the second stage, betas are estimated with the SBETA Gibbs sampler and OLS based on the generated return series. These beta estimates are then compared to the true beta values.

In the first stage, we analyze several scenarios for the true parameters. The unconditional expectation of beta, α_p, is set to 0.7 and 1.3. The beta persistence parameter, δ_p, takes values of 0.4, 0.6, 0.8, and 0.95 to reflect degrees of relatively low, medium, and high persistence. For the volatility parameters, we set σ_{β_p}=0.25 and $\sigma_p = 0.01$ to match beta and return volatilities from actual monthly industry portfolios. For each parameter combination, $\alpha, \delta, \sigma_\beta$, and σ, we generate 500 sets of true beta series, each of length $T = 120$. We then generate 500 corresponding samples of portfolio return series based on the true betas, idiosyncratic volatility, σ, and monthly excess market returns from January 1991 to December 2000.

In the second stage, for each of the 500 samples, we use the Gibbs sampler to estimate the model parameters. The algorithm iterates through 3,000 draws. The first 1,000 draws (burn-in stage) are discarded to allow for convergence, which

usually occurs within 200 iterations. The last 2,000 draws (monitoring stage) are used to calculate the posterior means of the model parameters. We then compare the posterior means to the true parameters we use to generate the model. Estimation precision is assessed by comparing each of the estimates to the true parameters over the 500 different samples using the RMSE statistic,

$$\text{RMSE}_{\widehat{\theta}} = \sqrt{\frac{1}{500}\sum_{i=1}^{500}\left(\widehat{\theta}_i - \theta\right)^2}, \tag{10}$$

where $\widehat{\theta}_i$ is the posterior mean of each parameter, $\alpha, \delta, \sigma_\beta$, and σ, in sample i and θ is the corresponding true parameter value.

Table 1 summarizes the sampling performance of the Bayesian parameter estimates and reports for each parameter the average of the 500 posterior means with their corresponding RMSEs in parentheses. The beta and return volatility parameters $\widehat{\sigma}_\beta$ and $\widehat{\sigma}$ are estimated with high precision and with virtually no bias under all true parameter sets. The estimate for the unconditional mean of beta, $\widehat{\alpha}$, is also unbiased. The persistence parameter, $\widehat{\delta}$, has a slight negative bias due to the negative skewness in its sampling distribution (see column 2 in Figure 1).[5] However, this bias decreases sharply and the precision increases significantly with higher values of the true beta persistence. Moreover, we notice the negative correlation between the precisions of $\widehat{\alpha}$ and $\widehat{\delta}$: the higher the true beta persistence, δ, the lower the RMSE of $\widehat{\delta}$ and the higher the RMSE of $\widehat{\alpha}$. This result is intuitive—the higher the persistence in beta, the longer betas are further away from the unconditional mean after a shock thus making it more difficult to determine the long-term mean. Indeed, statistically, equation (3) shows that the higher the persistence, δ, the higher the unconditional beta variance, and thus the larger the variance of its unconditional mean, α. Figure 1 illustrates the sampling distribution of each parameter estimate under the various scenarios. We notice that the sampling distribution of all parameters is normal, except for the δs, which exhibit slight negative skewness.

Next we compare the precision of the SBETA and rolling OLS estimates of beta. We estimate the average RMSE of the SBETA and OLS beta estimates, $\widehat{\boldsymbol{\beta}}_{\mathcal{M},i}$ ($\widehat{\boldsymbol{\beta}}_{\text{SBETA},i}$ or $\widehat{\boldsymbol{\beta}}_{\text{OLS},i}$), relative to each of the 500 true beta series, $\boldsymbol{\beta}_i$, as

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

Table 2 summarizes the performance of the SBETA and OLS beta estimates. The first column presents the mean of the true simulated betas across the 500 samples of 120 periods each. Columns 4 and 5 show that the SBETA and OLS betas estimate this mean beta with almost negligible small sample bias (ranging between -0.0005 and 0.0016 for SBETA and between -0.0036 and 0.0034 for OLS). The SBETA estimates, however, significantly outperform the OLS estimates in capturing the time variation in the true beta. The RMSE of the SBETA

[5]Similar skewness is observed in the sampling distribution of the persistence parameter in the SVOL model of Jacquier, Polson, and Rossi (1994).

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

TABLE 1

Sampling Properties of the SBETA Estimates

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

We simulate 500 data samples of 120 observations each according to the SBETA model (equation (1)), using the following true values of the parameters: $\alpha = 0.7, 1.3$; $\delta_p = 0.40, 0.60, 0.80, 0.95$; $\sigma_{\beta p} = 0.25$; and $\sigma_p = 0.01$. Monthly returns for the CRSP value-weighted index from January 1991 to December 2000 are used as the market returns. For each data sample, the model parameters are estimated based on 2,000 draws of the sampler, after discarding 1,000 draws. The first row of each cell reports the average of the 500 posterior means for each parameter. The number in parentheses is the RMSE of the posterior mean of each parameter, $\hat{\theta}_i$, with respect to the true parameter, θ_i,

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

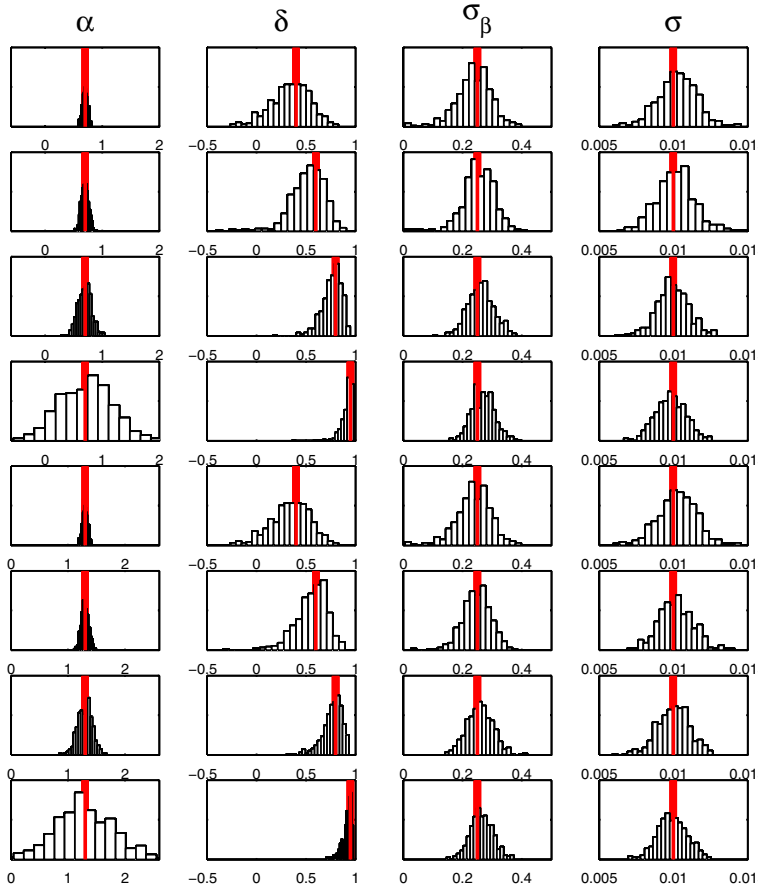
estimates with respect to the true simulated beta series (see last two columns) is much smaller than that of the OLS estimate. Moreover, as the true beta persistence increases, the precision of the SBETA estimates increases while that of the OLS estimates drastically deteriorates. For example, when the true persistence is 0.95, OLS betas produce a RMSE that is almost three times as large as the SBETA (RMSE of 0.5630 for OLS vs. 0.2031 for SBETA).

Figure 2 illustrates that SBETA estimates provide an extremely close approximation to the true beta process. In contrast, the figure shows that although OLS betas capture the average beta risk, they fail to reflect changes in the underlying systematic risk process, and their performance worsens with increases in beta persistence.

Table 3 summarizes the percentage of times the true beta lies outside the 90% SBETA and OLS confidence bounds. If the estimated bounds are correct, then 5% of the true betas should be below the 5th percentile bound and 5% should be above the 95th percentile bound for each estimator. We notice from the middle two columns that the SBETA provides reliable statistical bounds: 6.85% and 6.83% of the true beta lie outside the 5th and 95th percentile bounds, respectively, when the true persistence is 0.4, and 5.01% and 4.91% when the true persistence is 0.95. In contrast, the true betas lie outside the OLS confidence bounds too often. For example, when persistence is low, 37.55% of the true betas lie below the 5th percentile bound and 37.76% of them are above the 95th bound. The numbers become 40.60% and 40.37% when beta persistence is 0.95. These results

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

FIGURE 1
Sampling Distribution of the SBETA Estimates



For each parameter set ($\alpha = 0.7, 1.3$; $\delta = 0.4, 0.6, 0.8, 0.95$; $\sigma_\beta = 0.25$; and $\sigma = 0.01$), we simulate 500 data series according to the SBETA model (equation (1)). For each simulated series, we make 3,000 draws from the conditional posterior of each parameter. The last 2,000 iterations are used to estimate the posterior means. The histograms illustrate the sampling distribution of the 500 posterior means for each parameter. The thick vertical line shows the true value of the parameter used to simulate the model. Each row represents a different combination of true parameters.

suggest that the OLS beta estimates significantly misspecify the true underlying risk, especially for persistent betas. This finding is consistent with Ang and Chen (2003) who show that such results can be attributed to the reliance on asymptotic standard errors, which are especially misleading when betas are persistent.

In summary, our MCMC algorithm is computationally efficient, converges quickly, and achieves high precision. Estimating the SBETA model for 120 time periods with 3,000 Gibbs iterations takes 30 seconds on a 2GHz PC. Moreover, the estimator is not sensitive to the choice of initial values and converges in less than 200 iterations.

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

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

We simulate 500 return and beta series of $T = 120$ observations from the SBETA model (model (1)) for each of the following true values of the parameters: $\alpha = 0.7, 1.3$; $\delta_p = 0.40, 0.60, 0.80, 0.95$; $\sigma_{\beta p} = 0.25$; and $\sigma_p = 0.01$. Monthly returns for the CRSP value-weighted index from January 1991 to December 2000 are used as the market return. For each sample, the posterior means of beta are based on 2,000 draws of the sampler, after discarding 1,000 draws. Each OLS beta, $\widehat{\beta}_t$, is obtained by linear regression of the simulated return series on the market excess returns over the 60-month window extending from $t - 29$ to $t + 30$. The average value is reported for the 500 simulated true series, the bias in SBETA and OLS estimates of beta, and the estimated average RMSEs of the 500 beta series with respect to the true simulated beta series,

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

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%

We simulate 500 return and beta series of $T = 120$ observations from the SBETA model (model (1)). Monthly returns for the CRSP value-weighted index from January 1991 to December 2000 are used as the market return. The first two columns provide the true α and δ values used to generate the true beta series. The third and fifth columns show the percentage of times in which the true beta is smaller than the 5th percentile confidence bound of the estimated SBETA and OLS betas, respectively. Similarly, the fourth and sixth columns present the percentage of times in which the true beta is larger than the 95th percentile confidence bound of the estimated SBETA and OLS betas, respectively. The confidence bounds for the SBETA are constructed based on 2,000 draws from the posterior distribution of beta. The OLS beta confidence intervals are based on the $t - 29$ to $t + 30$ 60-month window regression standard errors of the estimated betas.

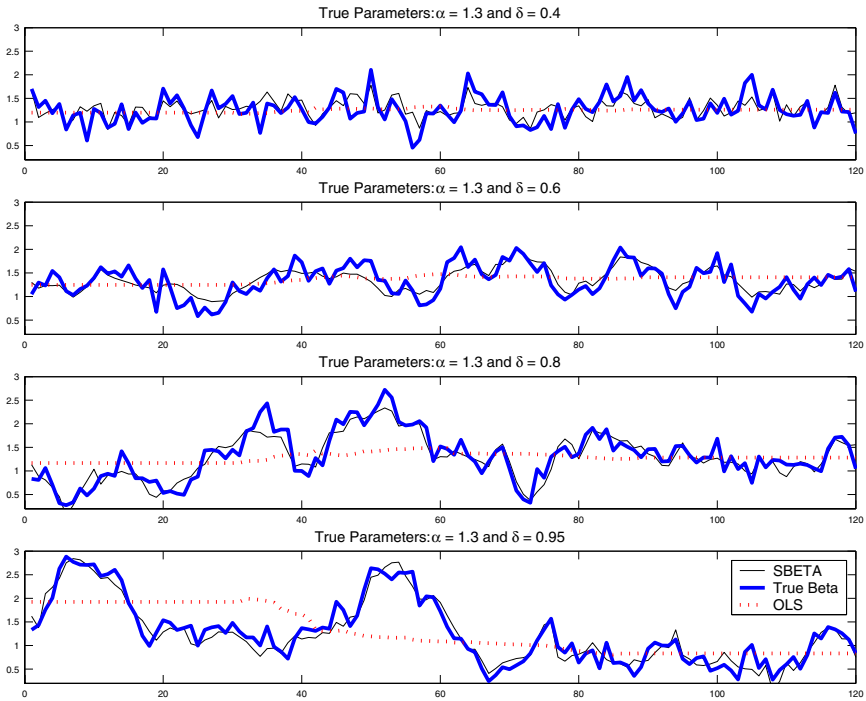
B. GARCH Data-Generating Process

In the previous simulation exercise, the true time-varying betas are generated using the SBETA model specification. Although this specification is very general and can describe a wide variety of data-generating processes, we further investigate whether our proposed methodology can provide precise beta estimates when the data are generated according to a different model. We use data generated by the flexible multivariate GARCH model of Ledoit, Santa-Clara, and Wolf

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

FIGURE 2

OLS and SBETA Estimates vs. Generated True Betas



We simulate 500 data samples of return and beta series of 120 observations each under different parameter sets. Figure 2 illustrates the simulated and estimated beta series in a randomly selected sample. The thick line represents the simulated true beta series according to model (1) assuming $\sigma_\beta = 0.25$, $\sigma = 0.01$, and α and δ as given on the top of each plot. The thin continuous line gives the estimated posterior mean of the SBETA β_t series using 2,000 draws of the Gibbs sampler. The dotted line illustrates OLS betas, $\widehat{\beta}_t$, estimated in a 60-month rolling regression using the simulated return series and the market returns from periods $t - 29$ to $t + 30$.

(2003).[6] These authors propose a new decentralized estimation approach of a general multivariate GARCH,

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

where $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are matrix parameters, $\mathbf{H}_t$ is the conditional covariance matrix, and $\boldsymbol{\Sigma}_t$ is a cross-product of variables observed at time t, i.e., $[x_{i,t}x_{j,t}]$, $i, j = 1, \ldots, k$. The model parameters are estimated using a two-step procedure for the diagonal and off-diagonal covariance matrix elements. While the time-varying betas are not directly represented by the GARCH model, they are implied by the covariance matrices $\mathbf{H}_t$. A direct representation of the beta process is offered by the EGARCH model of Braun, Nelson, and Sunier (1995). However, this heavily parameterized model is often infeasible, exhibiting an almost flat likeli-

[6] We are grateful to the authors for providing the program code for estimating the parameters of their model.

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

hood function and unreliable standard errors. The EGARCH MLE algorithm fails to converge with our data sample.

In the generation stage, a set of true model parameters, $(\mathbf{A}, \mathbf{B}, \mathbf{C})$, is determined following the compatibility constraints of the model (see Ledoit, Santa-Clara, and Wolf (2003)). Using the true parameters, a data sample of 300 observations of returns and covariance matrices is generated according to the GARCH model, equation (12). Table 4 provides the true model parameters. In the estimation stage, the GARCH model parameters, covariance matrices, and implied betas are estimated. In addition, we estimate SBETA and rolling OLS betas. The three sets of beta estimates are compared using the RMSE (equation (11)) from the true GARCH-implied beta series. The procedure is repeated for 100 simulated samples. Table 4 indicates that the SBETA model outperforms the rest of the models based on RMSE. The SBETA RMSE (0.1051) is smaller than the GARCH RMSE (0.1180) and is half of the OLS RMSE (0.2104). The SBETA RMSE also has the smallest standard error. Figure 3 shows these results.

TABLE 4

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

Panel A. Predetermined True Parameters of the GARCH Model

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

Panel B. Parameter Estimates of the GARCH Model

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

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	0.1051	0.2104
	(0.0183)	(0.0029)	(0.0086)

Using predetermined true GARCH parameters, we simulate 100 data samples of 300 observations each from the GARCH model of Ledoit, Santa-Clara, and Wolf (2003), $\mathbf{x}_t \sim N_k(0, \mathbf{H}_t)$ and $\mathbf{H}_t = \mathbf{C} + \mathbf{A}\boldsymbol{\Sigma}_{t-1} + \mathbf{B}\mathbf{H}_{t-1}$. The true beta series are implied by the 2×2 covariance matrices generated by the GARCH model. The parameters of the GARCH model are estimated using maximum likelihood techniques. The SBETA estimates are based on 2,000 draws of the Gibbs sampler, after discarding the initial 1,000 draws. The OLS betas are estimated based on the 60-month rolling regression window.

C. Beta Process with Conditioning Information

We also test the SBETA model on another popular data-generating process with time-varying betas: we condition the betas on a set of instruments (either aggregate or firm specific). Such beta models are reviewed by Shanken (1990), Ferson and Harvey (1993), and Avramov and Chordia (2005). We use a one-factor model in which the beta is a deterministic function of two instruments, i.e.,

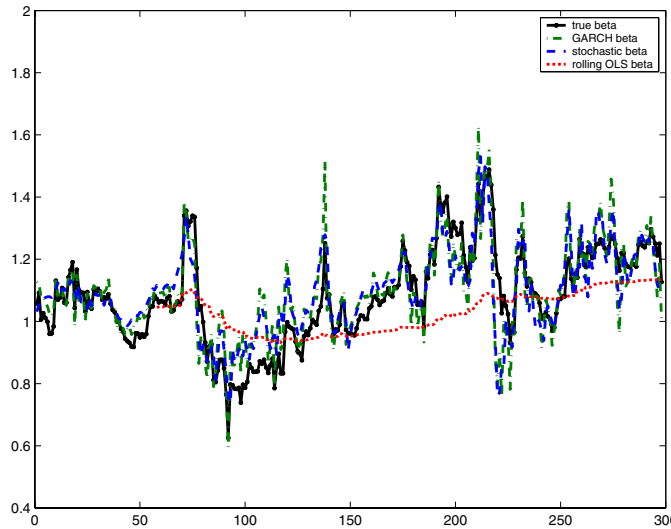
$$y_t = \alpha + \beta_t x_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2) \tag{13}$$

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

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

FIGURE 3

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



The betas of the three models are estimated using simulated returns over 300 periods. The GARCH parameters are estimated using maximum likelihood techniques, while the SBETA parameters are estimated using the Gibbs sampler. The OLS estimates are based on a 60-month rolling window.

In the first stage of the simulation exercise, we generate data samples using prespecified true parameters. We use the value-weighted index from Center for Research in Security Prices (CRSP) as the market return series, $\mathbf{x}$. As instruments, we use monthly series for the yield on the three-month T-bill and the dividend yield of the Dow Jones Index.

The true parameters are presented in Table 5. We select values for the true parameters, α_β, δ_1, δ_2, and σ^2, which produce reasonable (not dramatically different from a unit beta) time-series for the conditional beta. The true conditional beta series are deterministic functions of the instruments (equation (14)). Using the true beta series, we draw 100 data samples of 300 observations each (equation (13)). We select values for the volatility parameter $\sigma^{2^{\text{true}}}$ that will result in a roughly equal proportion of systematic and idiosyncratic risk in the data series $\mathbf{y}$.

We estimate the conditional betas with OLS using the cross-product of the instruments with the market factor, following Shanken (1990). We estimate the conditional beta series based on all true instruments, based on a subsample of the true instruments, and based on proxy instruments (i.e., the true instruments are unknown). We also estimate the SBETA, rolling OLS, and GARCH betas. These beta estimates are compared to the true betas using the RMSE from the 100 simulations (see Table 5 and Figure 4).

We draw the following conclusions from this simulation exercise. First, not surprisingly, if all true instruments are known, the conditional beta model is the preferred model (RMSE $= 0.101$). Second, in the case of partial misspecification (only some of the true instruments are known), the conditional model may still be

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

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

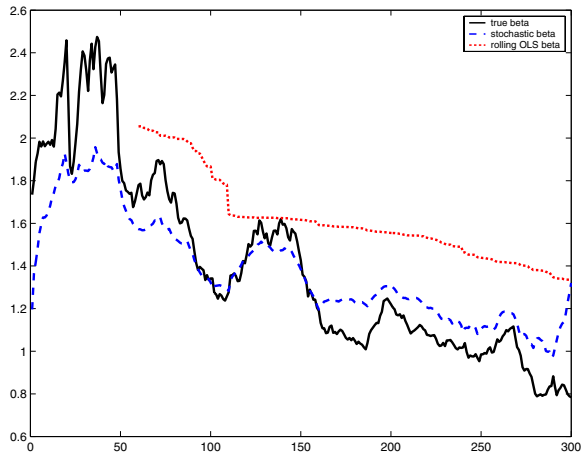
Panel B. RMSE Estimates

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)

Using predetermined true parameters, we simulate 100 data samples of 300 observations from the conditional beta model of the type proposed by Shanken (1990), $y_t = \alpha + \beta_t x_t + \varepsilon_t$, and $\beta_t = \alpha_\beta + \delta_1 z_{1,t-1} + \delta_2 z_{2,t-1}$. The yield of the three-month T-bill and the dividend yield of the Dow Jones index are used as the two instruments. The true betas are deterministic functions of the instruments. The SBETA estimates are based on 2,000 draws of the Gibbs sampler, after discarding the initial 1,000 draws. The second column presents the average RMSE measures with their standard errors in parentheses.

FIGURE 4

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



The betas of the SBETA and conditional models are estimated using simulated returns over 300 periods. The OLS estimates are based on a 60-month rolling window.

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

preferred (RMSE = 0.154,0.161). However, if the true instruments are not known, as is the case in practice, the SBETA model is preferred (RMSE = 0.296). The SBETA estimates are much more precise than OLS (RMSE = 0.414), GARCH (RMSE = 0.527), or conditional betas based on proxy instruments (RMSE = 0.423 and 0.447).

V. Industry Betas over the Past 40 Years

Section V presents an empirical investigation of the behavior of equity betas. Specifically, we evaluate the degree of time variation, persistence, and volatility exhibited by industry portfolio betas over the January 1962–December 2001 period.

The SBETA model is fitted to the monthly excess return series of five U.S. industry portfolios.[7] The returns are value-weighted across all NYSE, AMEX, and NASDAQ stocks assigned to a portfolio based on their four-digit SIC code as follows.

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

The market excess return is proxied by the value-weighted index return from CRSP in excess of the one-month T-bill rate (from Ibbotson Associates).

The statistical significance of the SBETA estimates is evaluated using Geweke's (1992) numerical standard errors (NSE). Geweke's method addresses the bias due to autocorrelation in sample draws, which is in turn caused by the Markov nature of the Gibbs algorithm. Geweke's diagnostic uncovers this bias through spectral analysis. Geweke shows that, when estimating the mean of the parameter θ, the estimate for $E(\theta)$ based on n sample draws is $\overline{\theta}_n = (\sum_{i=1}^{n} \theta^{(i)})/n$, with an asymptotic variance of $S_\theta(0)/n$. Taking the square root, we obtain Geweke's NSE of the estimate, $\overline{\theta}_n$,

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

If the Gibbs sampler draws were i.i.d., Geweke's NSE reduces to the familiar standard error of the mean used in classical statistics, i.e.,

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

Table 6 summarizes the estimation results for the SBETA parameters, $\widehat{\alpha}$, $\widehat{\delta}$, $\widehat{\sigma}_\beta$, and $\widehat{\sigma}$, of the beta process for each industry. We report the posterior mean, Geweke's NSE, and the 90% Bayesian credibility interval (BCI) for each estimate based on 45,000 draws (i.e., 90% of the draws lie within this BCI).

[7]The industry return series are obtained from Kenneth French's Web site data library at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

TABLE 6

Analysis of Industry Portfolio Betas

	Manufacturing	Utilities	Shops	Finance	Other
$\widehat{\alpha}$					
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]
$\widehat{\delta}$					
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]
$\widehat{\sigma}_\beta$					
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]
$\widehat{\sigma}$					
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]

The SBETA model (equation (1)) is fitted to monthly value-weighted industry portfolio returns from January 1962 to December 2001. The first line reports the posterior mean of the parameter. The NSE are computed as in equation (15). The brackets in the third row denote the 90% BCI, which is constructed from the 5th and 95th percentiles of the simulated posterior draws. Each estimation is based on 45,000 draws (after discarding the first 5,000). Shops represents the wholesale and retail sector.

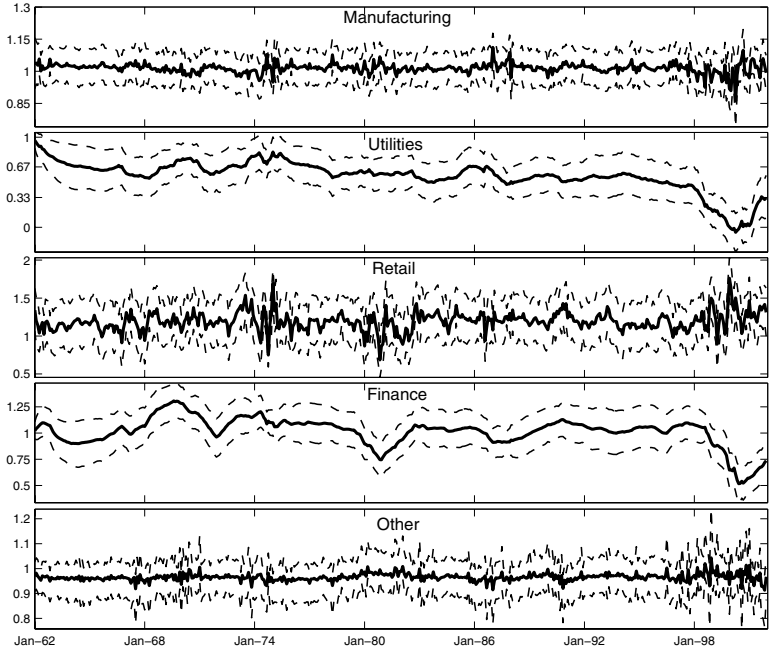
Manufacturing firms have average systematic risk ($\widehat{\alpha} = 1.0138$) and the smallest idiosyncratic volatility ($\widehat{\sigma} = 0.0075$). Beta persistence is low ($\widehat{\delta} = 0.1975$) and so is the beta volatility ($\widehat{\sigma}_\beta = 0.0017$). The results suggest that most of the variation in manufacturing returns is explained by variation in market returns, rather than variations in beta or idiosyncratic noise. Figure 5 illustrates that manufacturing betas are strongly mean reverting and remain largely unchanged over the period under analysis.

Utilities tend to have the lowest systematic risk ($\widehat{\alpha} = 0.5207$) and are among the highest in beta persistence ($\widehat{\delta} = 0.9299$), suggesting that, while market shocks tend to have a relatively small impact on utility returns, shocks to utility betas tend to have a long life. Utility betas tend to have the second largest stochastic component ($\widehat{\sigma}_\beta = 0.0963$) and utility returns tend to have the highest idiosyncratic risk ($\widehat{\sigma} = 0.0296$). Figure 5 illustrates that utility betas declined over the past 40 years, with a sudden increase in the last two years. Financial firms have average systematic risk ($\widehat{\alpha} = 0.9983$) and beta volatility ($\widehat{\sigma}_\beta = 0.0742$), but financial betas are the most persistent ($\widehat{\delta} = 0.9348$). The persistence of shocks is prominent in the large cycles in the beta process (see Figure 5). The high persistence of shocks to finance and utility betas could result from their tightly regulated nature. Indeed, if regulatory changes cause shocks in betas, these shocks would be of a more permanent nature. Industry-specific regulation could also be driving the larger idiosyncratic volatility in these two industries.

Betas in the retail and wholesale sector are the highest ($\widehat{\alpha} = 1.1916$) and most volatile ($\widehat{\sigma}_\beta = 0.2362$), implying that returns in this sector are strongly affected by variation in both market returns and betas. Figure 5 shows that the highest beta volatility occurs both around the beginning of 1974 and at the turn of the millennium. Firms in this industry have average beta persistence ($\widehat{\delta} = 0.3821$) and average idiosyncratic return volatility ($\widehat{\sigma} = 0.0209$).

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

FIGURE 5
Estimated Industry Portfolio SBETA Betas, $\widehat{\beta}_t$, with 90% Bayesian Posterior Bounds



The SBETA model (equation (1)) is fitted to industry portfolio returns over a 480-month period from January 1962 to December 2001, and 45,000 draws are used to calculate the posterior means (after discarding the first 5,000 draws). The thick line in each plot illustrates the posterior mean of $\widehat{\beta}_t$, while the two thin dashed lines represent the 90% Bayesian posterior interval.

Since a diverse set of industries is classified as Other, it is not surprising that this portfolio has average systematic risk ($\widehat{\alpha} = 0.9647$), one of the lowest levels of idiosyncratic risk ($\widehat{\sigma} = 0.0170$), and the lowest time variation in beta ($\widehat{\delta} = 0.0136$ and $\widehat{\sigma}_\beta = 0.0430$). Figure 5 illustrates that despite the noise in beta, the average systematic risk remains virtually unchanged over our sample period.

The behavior of industry betas over the past 40 years is consistent with the SBETA model: industry betas exhibit mean reversion, persistence, and stochastic behavior. A major benefit of the SBETA model is its ability to separate and measure these beta characteristics. Further, it offers the advantage of separating return variability into three distinct components: market shocks, beta shocks, and a third idiosyncratic component. This feature has valuable implications for risk management and asset pricing, as we demonstrate next.

VI. Out-of-Sample Hedging Application

The out-of-sample hedging application compares the predictive power of the SBETA and rolling OLS beta estimates. The comparison is based on the relative effectiveness of the two beta estimates in hedging the market risk of the five

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

industry portfolios. Effectiveness is measured by the volatility of the hedged positions.

For each industry, the market hedge at time T is implemented as follows. For each dollar invested in portfolio i, we sell short $\widehat{\beta}_{i,T+1}$ dollars of the market portfolio. The return of the hedged (long-short) position over month $T + 1$ is the hedging error,

$$\Delta^{\mathcal{M}}_{i,T+1} = r_{i,T+1} - \widehat{\beta}^{\mathcal{M}}_{i,T+1} r_{m,T+1}, \tag{16}$$

where $r_{i,T+1}$ and $r_{m,T+1}$ are industry i and the market returns over period $T + 1$, and $\widehat{\beta}^{\mathcal{M}}_{i,T+1}$ is the predicted beta based on either the SBETA or OLS models. A perfect hedge should yield zero return. For each out-of-sample period $T + 1$, the SBETA betas are estimated using a rolling 120-month window ending in month T. The OLS betas are based on a 60-month rolling regression window ending in month T. The shorter OLS window is intended to induce greater variation in beta forecasts.[8]

Table 7 presents the results of the market hedge for the 300-month out-of-sample window between January 1977 and December 2001. The first column reports the volatility of the unhedged returns for each industry. Their monthly standard deviations range between 3.82% (utilities) and 5.91% (retail). The volatility of the hedged returns is reported in the second and third columns. For each industry, the volatility of the hedged returns is always smaller than that of the unhedged returns. The market hedge is especially effective in manufacturing, retail, and finance, where the reduction in monthly volatility ranges between 87% (manufacturing: from 4.67% to 0.63%) and 61% (finance: from 4.88% to 1.89%). The lower hedging effectiveness in utilities can be explained by their highest idiosyncratic volatility (see Table 6).

The hedging errors are always smaller when using the SBETA estimates. There is an additional 17%–29% reduction in monthly volatility when the SBETA estimates are used. For example, the volatility of a hedged position in manufacturing is 0.63% under SBETA vs. 0.89% under OLS. For utilities, the use of the SBETA estimator reduces the volatility of the hedged position from 3.10% to 2.58%; for retail, from 2.32% to 1.78%; and for finance, from 2.49% to 1.89%. Note that even for Other industries, which contain a diverse set of stocks and is a close proxy for the market portfolio, the use of SBETA estimates improves hedging effectiveness from 1.85% to 1.31%.

The SBETA improvement in hedging effectiveness over OLS is strongly statistically significant. The last column of Table 7 reports the sample t-statistics of the differences between the absolute SBETA and OLS hedging errors. The negative sign of the t-statistics implies that the SBETA-based hedging errors are smaller than the OLS-based hedging errors in absolute values, i.e., the hedge is more effective. The t-statistics range between -11.66 (manufacturing) and -7.62 (utilities), indicating that the SBETA hedging errors are significantly smaller than their OLS analogs. In summary, the hedging exercise shows that the use of

[8]We also estimate SBETA and OLS betas using an extended (rather than rolling) window. The SBETA results are virtually identical, while the OLS results are slightly worse.

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%	2.32%	−9.97
Finance	4.88%	1.89%	2.49%	−9.22
Other	4.83%	1.31%	1.85%	−11.83

Table 7 compares the results from an out-of-sample market hedge of the five industry portfolios based on SBETAs to those based on rolling regression OLS betas. The SBETAs are estimated using a rolling 120-month window ending in month T. The OLS betas are based on a 60-month rolling regression window ending in month T. The market hedge is constructed at the end of month T and evaluated out of sample at the end of month $T + 1$. The results are based on a 300-month out-of-sample window, extending over the January 1977–December 2001 period. The market hedge at time T is implemented as follows: for each dollar invested in portfolio i, we sell short $\widehat{\beta}_{i,T+1}$ dollars of the market portfolio. The return of the hedged (long-short) position over $T + 1$ is the hedging error,

$$\Delta^{\mathcal{M}}_{i,T+1} \quad = \quad r_{i,T+1} - \widehat{\beta}^{\mathcal{M}}_{i,T+1} r_{m,T+1},$$

where $r_{i,T+1}$ and $r_{m,T+1}$ are the industry i and the market returns, respectively, over period $T + 1$, and $\widehat{\beta}^{\mathcal{M}}_{i,T+1}$ is the predicted beta based on either the SBETA or OLS models. The first column reports the volatility (standard deviation) of the unhedged monthly industry returns. The second and third columns report the volatility of the hedging errors based on the two models. The last column reports the sample t-statistic of the differences between the absolute SBETA and OLS hedging errors. A negative sign of this t-statistic indicates that the SBETA-based hedging errors are smaller than the OLS-based hedging errors.

SBETA estimates in hedging brings economically and statistically significant improvement in hedging effectiveness relative to the use of OLS betas.

VII. Asset Pricing Application: Test of the CAPM

The results so far are encouraging: the SBETA model appears to produce more precise beta estimates both in- and out-of-sample. The higher precision in estimating latent betas may lead to more powerful asset pricing tests. We investigate the SBETA performance in a Fama and MacBeth (1973) type two-step test of the CAPM relative to four alternative beta estimates.

Our first goal is to assess whether the SBETA estimator produces a more significant market risk premium and reduces asset pricing errors. Second, we investigate whether the SBETA estimates can help explain the size, BM, and idiosyncratic volatility effects in the cross section of stock returns that are documented in previous work and considered to be anomalies relative to the canonical CAPM.

The asset pricing tests are based on individual stock, rather than portfolio returns as in Brennan, Chordia, and Subrahmanyam (1998) and Avramov and Chordia (2005). Using individual stocks in asset pricing tests has several advantages. First, it guards against data-snooping biases inherent in portfolio-based asset pricing tests (see Lo and MacKinlay (1990)). Second, it eliminates the error in the variables problem, which arises because extreme beta errors are aggregated into the same portfolios, thereby introducing correlation between the errors and the independent variable (beta) in the cross-sectional regressions. Third, it avoids the loss of information that results when stocks are sorted into groups (Litzenberger

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

and Ramaswamy (1979)) and circumvents the problem raised by Berk (2000) that sorting data into portfolios introduces a bias in favor of rejecting the model.

The individual stock data for the 40-year period between January 1964 and December 2003 are extracted from CRSP and Compustat and are subject to filtering rules similar to those used in previous asset pricing tests based on individual data. First, each stock needs to have at least 10 years of monthly data. Second, sufficient data has to be available to calculate the market capitalization (or size) of each firm. The size variable is calculated as the product of the stock's price times the number of shares outstanding, standardized to sum to one in each month by dividing each firm's size by the total market capitalization of all sample firms in the given month. Third, sufficient data has to be available in Compustat to calculate the BM ratio as of December of the previous year. Following Fama and French (1992), we control for Compustat survival bias by dropping the first two years of Compustat data for every firm. Finally, outliers outside the 0.005 and 0.995 fractile values for each variable are removed. This filtering yields a sample of 809 stocks.

We start the asset pricing tests by first estimating individual stock betas for the 809 return series with respect to the value-weighted return index from CRSP. Apart from the SBETA model, four alternative beta estimators are considered: i) the 60-month rolling regression OLS betas, ii) the characteristic-scaled betas of the type proposed by Avramov and Chordia (2005), where factor loadings vary with firm-specific size and the BM ratio, iii) the conditional betas proposed by Shanken (1990), where individual betas are conditioned on aggregate variables such as the level and volatility of interest rates, and iv) the GARCH betas implied by the GARCH model of Ledoit, Santa-Clara, and Wolf (2003).

Table 8 reports descriptive statistics for the 809 stock betas estimated by the different models. Note that except for the linear regression model, by construction no other model guarantees that the capitalization-weighted average stock beta (across all stocks) equals one. Still, except for the characteristics-based betas, the weighted average beta of the 809 sample stocks is very close to one (see first row in Table 8). In fact, the SBETA average is the closest to one (0.9852), apart from the OLS average (1.0016). The second row in Table 8 shows that the SBETA estimates are the most time varying (the average beta standard deviation is 0.3913 vs. 0.2406 for OLS and less for the alternatives). Yet, the standard deviation of the SBETA weighted average or market beta has the second lowest time variation (0.0466). The lower the market beta time variation the better because such variation may reflect bias in the estimation given the market beta should equal one in theory. These results are encouraging as they show that the SBETA model produces reasonable values for the betas of individual stocks.

The SBETA estimates provide strong support for the CAPM. Panel A in Table 9 reports results from the second pass cross-sectional regressions, in which for each month t excess stock returns are regressed on their estimated betas,

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

where r_{it} are excess returns, $\widehat{\beta}_{it}$ are the estimated betas, α_t is the return unexplained by beta, and λ_t is the market risk premium. The estimate of the risk premium when the SBETA model is used is strongly significant (t-statistic $= 2.92$)

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

The first row of Table 8 reports the capitalization-weighted average beta across the 809 sample stocks over the January 1964 to December 2003 period. The second row reports the average monthly standard deviation of individual stock betas across the sample period. The third row presents the monthly standard deviation of the capitalization-weighted average beta. The columns present results from five different beta estimators that include the following.

SBETA: SBETA estimated using model (1).

OLS: Regression betas estimated using a 60-month rolling regression window.

Char. Scaled: Characteristics-scaled betas estimated, as in Avramov and Chordia (2005), by conditioning betas on firm-level size and BM ratio.

Conditional: Conditional betas estimated, as in Shanken (1990), by conditioning betas on aggregate variables such as the level and volatility of interest rates.

GARCH: Betas estimated by the GARCH model of Ledoit, Santa-Clara, and Wolf (2003). All betas, except the OLS rolling regression betas, are estimated using the entire sample period between January 1964 and December 2003.

and the intercept is insignificant, as implied by the CAPM. In contrast, the alternative beta measures produce a strongly significant intercept in the cross-sectional regressions (t-statistics between 2.50 and 3.79) and a market risk premium that is insignificant at the 5% level (the risk premium is significant at the 10% level only when betas are estimated with OLS (t-statistic $= 1.68$) or the conditional beta model (t-statistic $= 1.88$). The significant intercepts imply that there are priced risks not captured by these alternative betas.

The SBETA estimates also explain a much larger proportion of the cross-sectional variation in individual security returns than the alternatives studied. Panel A shows that the SBETA estimates explain on average 22.90% of the cross-sectional variation in individual stock excess returns. Alternative estimates capture between 1.03% (for characteristic-scaled betas) and 6.43% (for OLS betas) of this cross-sectional variation. This implies that the SBETA model provides a significant reduction in pricing errors even when working with individual stocks.

Next we examine whether the SBETA estimator can also help resolve some asset pricing anomalies previously documented in the literature. Panels B–E in Table 9 report results from the second pass cross-sectional regressions when firm-level characteristics are included along with the betas as regressors,

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

where $\mathbf{C}_{it}$ is a vector of firm-level characteristics and $\boldsymbol{\gamma}_t$ is a vector of premiums for the firm-level characteristics. If the predictive power of such firm-level attributes is explained by asset pricing models, then these attributes should be statistically insignificant in the second pass regressions.

The SBETA estimator helps explain the size and BM effects in the cross section of stock returns. Note that in Table 9, Panel B the intercept, size, and BM are indeed significant (at least at the 10% level) when the four alternative models are considered. In contrast, when the SBETA estimates are used, only the market risk premium is significant (t-statistic $= 2.93$), while the intercept, size, and BM effects are not at any conventional statistical level. The adjusted R^2 for the SBETA model is still strikingly higher at 24.25% than the alternatives (2.88% to 8.09%).

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

The SBETA model also explains the previously documented pricing of idiosyncratic volatility in the cross section of stock returns. Table 9, Panel C presents results when idiosyncratic volatility, σ_i^2, an output of the SBETA model, is included among the regressors. When the alternative beta models are considered, idiosyncratic volatility is positively priced (t-statistics are between 2.48 and 3.95) as documented in previous work. Note that idiosyncratic volatility helps partially explain the size effect in the OLS and conditional beta models. Yet in the SBETA model, beta still remains the single significantly priced source of risk (t-statistic $= 2.79$), while the intercept, size, BM, and idiosyncratic volatility remain insignificant, consistent with the CAPM.

The CAPM states that beta should be the only source of excess returns in equilibrium. Empirically, it is interesting to know whether characteristics of the beta process such as persistence and volatility are themselves priced separately from the level of beta. We find that beta persistence is positively priced in the cross section of individual stocks returns. Table 9, Panel D shows that when the persistence of beta shocks, δ_i, estimated in the SBETA model, is included among the explanatory variables, it is strongly significant in all models. The t-statistic for beta persistence is 2.27 in the SBETA case and between 1.75 and 2.43 in the

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	1.07**	0.39	0.07	0.44	0.29
	(2.92)	(1.68)	(0.74)	(1.88)	(1.38)
Adj. R^2	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%

(continued on next page)

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%

Individual stock betas for the 809 stocks in the sample are first estimated using the following five alternative methods.

SBETA: SBETA estimated using model (1).

OLS: Regression betas estimated using a 60-month rolling regression window.

Char. Scaled: Characteristics-scaled betas estimated, as in Avramov and Chordia (2005), by conditioning betas on firm-level size and the BM ratio.

Conditional: Conditional betas estimated, as in Shanken (1990), by conditioning betas on aggregate variables such as the level and volatility of interest rates.

GARCH: Betas estimated by the GARCH model of Ledoit, Santa-Clara, and Wolf (2003). All betas, except for the OLS rolling regression betas, are estimated using the entire sample period between January 1964 and December 2003. Next, for each month, we run cross-sectional regressions of stock excess returns on the estimated betas and firm-level characteristics. The different panels report results from different subsets of independent variables. Each panel reports the average coefficients from the cross-sectional regressions, $r_{it} = \alpha_t + \lambda_t \hat{\beta}_{it} + \boldsymbol{\gamma}_t \mathbf{C}_{it} + e_{it}$, where r_{it} are excess returns, $\hat{\beta}_{it}$ are the estimated betas, α_t is the return unexplained by beta, λ_t is the market risk premium, $\mathbf{C}_{it}$ is a vector of firm characteristics, and $\boldsymbol{\gamma}_t$ is a vector of premiums for the firm characteristics. The time-series t-statistics of the coefficients are reported in parentheses.

*, ** indicate significance at the 5% and 1% levels, respectively.

alternatives. We find that persistent beta stocks are typically from the finance and utilities industry. The remaining results on size, BM, and idiosyncratic volatility remain unchanged. These results show that, in practice, investors require a positive premium for holding stocks for which shocks to beta are highly persistent.

We note that yet there is no evidence that the noise (or uncertainty) in beta is priced. Panel E of Table 9 shows that the conditional volatility in beta, $\sigma^2_{\beta_i}$, estimated in the SBETA framework, is insignificant across all models, while the beta persistence variable remains strongly significant. This implies that investors

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

require a premium for persistent, rather than temporary, shocks to beta. This result, consistent with the predictions of the CAPM that beta risk is priced, is sensible because shocks that affect beta over a longer period have a lasting impact on the portfolio's systematic risk and may thus require adjustments in portfolio evaluation, risk management, and portfolio management.

The asset pricing tests suggest that it may be premature to discard rational asset pricing models in explaining the cross section of equity excess returns. The SBETA model provides beta estimates that explain the size, BM, and idiosyncratic volatility effects, and point at beta as the only significantly priced source of risk. The results support Ghysels (1998) who argues that if the beta dynamics are appropriately captured, then a time-varying beta model outperforms constant beta models. However, if beta is misspecified, the pricing errors potentially could be larger than with a constant beta model. Indeed, we find that in many cases, OLS betas produce more support for the CAPM than certain time-varying alternatives.

VIII. Conclusion

This paper builds on the literature on time-varying systematic risk by modeling the dynamics of beta as a general mean-reverting stochastic process. This relaxed beta structure allows for randomness in both the systematic and idiosyncratic components of asset returns and thus introduces more kurtosis in the return distribution. The excess kurtosis has long been a stylized fact in equity markets. In addition, as the proposed general beta model contains existing beta models as special cases, it allows for testing the restriction on the beta's mean or volatility that these models impose. An empirical investigation of industry returns over the past 40 years provides evidence that industry betas are indeed both time varying and random, and shocks to systematic risk, especially in the utilities and finance industries, are highly persistent.

A simulation study documents that SBETA estimates outperform competing beta estimates even when the true betas are generated according to these competing models. We document that the popular rolling regression betas often severely misspecify true systematic risk, especially when beta shocks are truly persistent. The SBETA model also produces superior beta forecasts. An out-of-sample hedging application over a 25-year period shows SBETA estimates reduce hedging errors relative to OLS by up to 30%. In addition, our asset pricing tests show that the SBETA model provides much stronger support for the conditional CAPM than popular alternatives and helps resolve asset pricing anomalies such as the size, book-to-market, and idiosyncratic volatility effects in the cross section of stock returns.

Appendix

A. Derivation of Posterior Densities

Following Bayes' rule, the posterior density of the augmented parameter vector is

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

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

where $\mathbf{r}_p = (r_{p1}, \ldots, r_{pT})'$, $\mathbf{r}_m = (r_{m1}, \ldots, r_{mT})'$, $\boldsymbol{\theta}_p = (\boldsymbol{\beta}_p, \boldsymbol{\omega}_p) = (\boldsymbol{\beta}_p, \alpha_p, \delta_p, \sigma^2_{\beta_p}, \sigma^2_p)$, and $p(\boldsymbol{\omega}_p)$ equals the product of independent priors, $p(\sigma^2_p) \times p(\alpha_p) \times p(\delta_p) \times p(\sigma^2_{\beta p})$. The density of $\boldsymbol{\beta}_p$ is part of the model and is therefore included in the likelihood.

We assume the following proper priors for every portfolio p,

$$\text{(A-2)} \quad p(\sigma^2_p) = IG(a_1, b_1) \propto \left(\sigma^2_p\right)^{-(a1+1)} \exp\left(-\frac{b_1}{\sigma^2_p}\right)$$

$$\text{(A-3)} \quad p(\alpha_p) = N(\mu_\alpha, \sigma_\alpha) \propto \frac{1}{\sigma_\alpha} \exp\left(-\frac{(\alpha_p - \mu_\alpha)^2}{2\sigma^2_\alpha}\right)$$

$$\text{(A-4)} \quad p(\delta_p) = \text{truncated}_{(-1,1)} - N(\mu_\delta, \sigma_\delta) \propto \frac{1}{\sigma_\delta} \exp\left(-\frac{(\delta_p - \mu_\delta)^2}{2\sigma^2_\delta}\right)$$

$$\text{(A-5)} \quad p(\sigma^2_{\beta p}) = IG(a_2, b_2) \propto \left(\sigma^2_{\beta p}\right)^{-(a_2+1)} \exp\left(-\frac{b_2}{\sigma^2_{\beta p}}\right).$$

Based on the SBETA model, the likelihood is represented by

$$\text{(A-6)} \quad L(\boldsymbol{\theta}_p|\mathbf{r}_p, \mathbf{r}_m) \propto \prod_{t=1}^{T} N(\alpha_p + \delta_p(\beta_{p,t-1} - \alpha_p), \sigma^2_{\beta_p}) \times \prod_{t=1}^{T} N(\beta_{pt} r_{mt}, \sigma^2_p)$$

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

$$\times \frac{1}{\sigma^T_p} \exp\left(-\frac{\sum_{t=1}^{T} (r_{pt} - \beta_{pt} r_{mt})^2}{2\sigma^2_p}\right).$$

Following equation (A-1), the joint posterior becomes

$$\text{(A-7)} \quad p(\boldsymbol{\theta}_p|\mathbf{r}_p, \mathbf{r}_m) \propto \frac{1}{\sigma_\alpha} \exp\left(-\frac{(\alpha_p - \mu_\alpha)^2}{2\sigma^2_\alpha}\right) \times \frac{1}{\sigma_\delta} \exp\left(-\frac{(\delta_p - \mu_\delta)^2}{2\sigma^2_\delta}\right)$$

$$\times \left(\sigma^2_{\beta p}\right)^{-(a_2+1)} \exp\left(-\frac{b_2}{\sigma^2_{\beta p}}\right) \times \left(\sigma^2_p\right)^{-(a1+1)} \exp\left(-\frac{b_1}{\sigma^2_p}\right)$$

$$\times \exp\left(-\frac{\sum_{t=1}^{T} (r_{pt} - \beta_{pt} r_{mt})^2}{2\sigma^2_p}\right) \frac{1}{\sigma^T_{\beta p}}$$

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

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

The conditional posteriors used in the Gibbs sampler are as follows. For each time $t = 1, \ldots, T-1$,

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

For time $t=1$, the conditional posterior density of β_{p1} depends on $\beta_{p,t-1}=\beta_{p0}$. We therefore need to provide initial values β_{p0}. In addition, to draw from β_{pt} we need $\beta_{p,t+1}$, which also depends on β_{pt}. This draw is made possible by conditioning $\beta^{(i)}_{pt}$ on $\beta^{(i)}_{p,t-1}$ and $\beta^{(i-1)}_{p,t+1}$ from the previous iteration.

For $t=T$, on the other hand, $\beta_{p,t+1}=\beta_{p,T+1}$ does not exist, so the conditional posterior simplifies to

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

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

The conditional posteriors of the remaining SBETA model parameters are

$$
\begin{aligned}
\text{(A-10)} \qquad p(\alpha_p|\text{rest}) \;&\propto\; \exp\left(-\frac{\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2}{2\sigma_{\beta p}^2}\right)\\
&\quad\times\exp\left(-\frac{(\alpha_p-\mu_\alpha)^2}{2\sigma_\alpha^2}\right)\\
&\propto\; \exp\left(-\frac{\alpha_p^2T(1-\delta_p)^2-2\alpha_p(1-\delta_p)\sum_{t=1}^{T}\left(\beta_{pt}-\delta_p\beta_{p,t-1}\right)}{2\sigma_{\beta p}^2}\right.\\
&\quad\left.-\frac{\alpha_p^2-2\alpha_p\mu_\alpha}{2\sigma_\alpha^2}\right)\\
&\propto\; \exp\left(-\frac{\left(\alpha_p-\frac{(1-\delta_p)\sigma_\alpha^2\sum_{t=1}^{T}\left(\beta_{pt}-\delta_p\beta_{p,t-1}\right)+\mu_\alpha\sigma_{\beta p}^2}{T(1-\delta_p)^2\sigma_\alpha^2+\sigma_{\beta p}^2}\right)^2}{2\left(\frac{\sigma_{\beta p}\sigma_\alpha}{\sqrt{T(1-\delta_p)^2\sigma_\alpha^2+\sigma_{\beta p}^2}}\right)^2}\right)\\
&\propto\; N\left(\frac{(1-\delta_p)\sigma_\alpha^2\sum_{t=1}^{T}\left(\beta_{pt}-\delta_p\beta_{p,t-1}\right)+\mu_\alpha\sigma_{\beta p}^2}{T(1-\delta_p)^2\sigma_\alpha^2+\sigma_{\beta p}^2},\right.\\
&\quad\left.\left(\frac{\sigma_{\beta p}\sigma_\alpha}{\sqrt{T(1-\delta_p)^2\sigma_\alpha^2+\sigma_{\beta p}^2}}\right)^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\text{(A-11)} \qquad p(\delta_p|\text{rest}) \;&\propto\; \exp\left(-\frac{\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2}{2\sigma_{\beta p}^2}\right)\\
&\quad\times\exp\left(-\frac{(\delta_p-\mu_\delta)^2}{2\sigma_\delta^2}\right)\\
&\propto\; \exp\left(-\frac{\delta_p^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2-2\delta_p\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)(\beta_{pt}-\alpha_p)}{2\sigma_{\beta p}^2}\right.\\
&\quad\left.-\frac{\delta_p^2-2\delta_p\mu_\delta}{2\sigma_\delta^2}\right)\\
&\propto\; \exp\left(-\frac{\left(\delta_p-\frac{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)(\beta_{pt}-\alpha_p)+\sigma_{\beta p}^2\mu_\delta}{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2+\sigma_{\beta p}^2}\right)^2}{2\left(\frac{\sigma_{\beta p}\sigma_\delta}{\sqrt{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2+\sigma_{\beta p}^2}}\right)^2}\right)\\
&\propto\; \text{truncated}_{(-1,1)}N\left(\frac{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)(\beta_{pt}-\alpha p)+\sigma_{\beta p}^2\mu_\delta}{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2+\sigma_{\beta p}^2},\right.\\
&\quad\left.\left(\frac{\sigma_{\beta p}\sigma_\delta}{\sqrt{\sigma_\delta^2\sum_{t=1}^{T}(\beta_{p,t-1}-\alpha_p)^2+\sigma_{\beta p}^2}}\right)^2\right)
\end{aligned}
$$

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

$$
\begin{aligned}
\text{(A-12)}\quad p(\sigma^2_{\beta p}|\text{rest}) &\propto \frac{1}{\sigma^T_{\beta p}} \exp\left(-\frac{\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2}{2\sigma^2_{\beta p}}\right) \\
&\quad \times \left(\sigma^2_{\beta_p}\right)^{-(a_2+1)} \exp\left(-\frac{b_2}{\sigma^2_{\beta_p}}\right) \\
&\propto \left(\sigma^2_{\beta p}\right)^{-\left(\frac{T+2a_2}{2}+1\right)} \\
&\quad \times \exp\left(-\frac{(T+2a_2)\,\frac{\left(\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2+2b_2\right)}{T+2a_2}}{2\sigma^2_{\beta p}}\right) \\
&\propto \text{Inv}-\chi^2\left(T+2a_2, \frac{\left(\sum_{t=1}^{T}\left(\beta_{pt}-\alpha_p-\delta_p(\beta_{p,t-1}-\alpha_p)\right)^2+2b_2\right)}{T+2a_2}\right).
\end{aligned}
$$

$$
\begin{aligned}
\text{(A-13)}\quad p(\sigma^2_p|\text{rest}) &\propto \frac{1}{\sigma^T_p} \exp\left(-\frac{\sum_{t=1}^{T}\left(r_{pt}-\beta_{pt}r_{mt}\right)^2}{2\sigma^2_p}\right)\left(\sigma^2_p\right)^{-(a1+1)} \exp\left(-\frac{b_1}{\sigma^2_p}\right) \\
&\propto \left(\sigma^2_p\right)^{-\left(\frac{T+2a_1}{2}+1\right)} \exp\left(-\frac{(T+2a_1)\,\frac{\sum_{t=1}^{T}\left(r_{pt}-\beta_{pt}r_{mt}+2b_1\right)^2}{(T+2a_1)}}{2\sigma^2_p}\right) \\
&\propto \text{Inv}-\chi^2\left(T+2a_1, \frac{\sum_{t=1}^{T}\left(r_{pt}-\beta_{pt}r_{mt}\right)^2+2b_1}{T+2a_1}\right).
\end{aligned}
$$

B. The Gibbs Sampler

The description provided here does not go beyond issues related to sampling from the joint posterior distribution (see Gelfand and Smith (1990) for additional details). Suppose that we need to sample a parameter set $\boldsymbol{\theta}$ with K elements from a joint posterior distribution, $p(\boldsymbol{\theta} \mid y)$, where $\mathbf{y}$ is a data set. However, it could be inconvenient to sample directly from $p(\boldsymbol{\theta} \mid y)$. The Gibbs sampler iteratively draws, in sequence, each of the K parameters over a total of I iterations from the conditional posterior distribution for each parameter given the remaining parameters,

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

The algorithm has the following steps: i) pick an arbitrary value $\theta^{(0)}$; ii) at each iteration, i, sample each $\theta_k^{(i)}$ from (B-1); and iii) continue the iterations until I becomes sufficiently large.

For large values of I, this procedure is equivalent to drawing the parameter set θ from $p(\theta \mid \mathbf{y})$. The Gibbs sampler simplifies the problem of drawing from the joint posterior to a problem of drawing from sections of it that are smaller and easier to draw from.

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

C. Multivariate Extension

Extending the methodology to a multivariate factor asset pricing structure implies i) a multivariate distribution of betas and ii) a vector specification for the factor sensitivities. Based on the SBETA model, the likelihood becomes

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

where $\boldsymbol{\beta}_p=(\beta_{p1},\ldots,\beta_{pT})'$, $\boldsymbol{\mu}_{\beta_t}=\boldsymbol{\alpha}_p-\boldsymbol{\delta}_p(\beta_{p,t-1}-\boldsymbol{\alpha}_p)$, and $\boldsymbol{\delta}$ is a diagonal matrix that gives the persistence parameters.

We specify the joint uninformative prior distribution as

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

The joint posterior distribution is

$$
\begin{aligned}
\text{(C-3)}\quad p(\boldsymbol{\theta}_p|\mathbf{r}_p,\mathbf{f}) &\propto \frac{1}{\sigma_p^{T+2}}\exp\left(-\frac{\sum_{t=1}^{T}\left(r_{pt}-\mathbf{f}_t\boldsymbol{\beta}_{pt}\right)^2}{2\sigma_p^2}\right) \\
&\quad\times\frac{1}{|\Sigma_{\beta_p}|^{T+1}}\exp\left(-\frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{\beta}_{pt}-\boldsymbol{\mu}_{\beta t}\right)'\Sigma_{\beta_p}^{-1}\left(\boldsymbol{\beta}_{pt}-\boldsymbol{\mu}_{\beta t}\right)\right).
\end{aligned}
$$

The conditional posterior distributions need to be calculated as in Subsection A, and a Gibbs sampler is specified according to the steps in Subsection B.

Conditional Posterior Distributions

The conditional posterior distribution for the inverse beta covariance matrix, $\Sigma_{\beta_p}^{-1}$, is a Wishart distribution,[9]

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

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

The conditional posterior distribution of the beta vector at time t is

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

[9]For properties of the Wishart distribution, see Gelman, Carlin, Stern, and Rubin (1995).

where

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

The conditional posterior distribution of the variance of returns is

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

The conditional posterior distribution of $\boldsymbol{\alpha}$ is

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

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

References

Ang, A., and J. Chen. "CAPM over the Long-Run: 1926–2001." Working Paper, Columbia Univ. and Univ. of Southern California (2003).

Ang, A., and J. Liu. "How to Discount Cashflows with Time-Varying Expected Returns." *Journal of Finance*, 59 (2004), 2745–2783.

Avramov, D., and T. Chordia. "Asset Pricing Models and Financial Market Anomalies." *Review of Financial Studies*, forthcoming (2005).

Baxter, N. D. "Leverage, Risk of Ruin and the Cost of Capital." *Journal of Finance*, 22 (1967), 395–403.

Beaver, W.; P. Kettler; and M. Scholes. "The Association between Market Determined and Accounting Determined Risk Measures." *Accounting Review*, 45 (1970), 654–682.

Ben-Zion, U., and S. S. Shalit. "Size, Leverage, and Dividend Record As Determinants of Equity Risk." *Journal of Finance*, 30 (1975), 1015–1026.

Berk, J. "Sorting Out Sorts." *Journal of Finance*, 55 (2000), 407–427.

Berk, J.; R. Green; and V. Naik. "Optimal Investment, Growth Options, and Security Returns." *Journal of Finance*, 54 (1999), 1553–1607.

Blume, M. E. "On The Assessment of Risk." *Journal of Finance*, 26 (1971), 177–181.

Bollerslev, T.; R. Engle; and J. Wooldridge. "A Capital Asset Pricing Model with Time Varying Covariances." *Journal of Political Economy*, 96 (1988), 116–131.

Box, G. E. P., and G. C. Tiao. *Bayesian Inference in Statistical Analysis.* Wiley Classics Library ed. New York: John Wiley & Sons (1992).

Braun, P.; D. Nelson; and A. Sunier. "Good News, Bad News, Volatility, and Betas." *Journal of Finance*, 50 (1995), 1575–1603.

Brennan, M.; T. Chordia; and A. Subrahmanyam. "Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns." *Journal of Financial Economics*, 49 (1998), 345–373.

Campbell, J., and T. Vuolteenaho. "Bad Beta, Good Beta." *American Economic Review*, forthcoming (2005).

Chen, S., and C. F. Lee. "Bayesian and Mixed Estimators of Time-Varying Betas." *Journal of Economics and Business*, 34 (1982), 291–301.

Fama, E., and K. French. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, 47 (1992), 427–465.

Fama, E., and J. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy*, 81 (1973), 607–636.

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press

Ferson, W., and C. Harvey. "The Risk and Predictability of International Equity Returns." *Review of Financial Studies*, 6 (1993), 527–566.
Gelfand, A. E., and A. F. M. Smith. "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association*, 85 (1990), 398–409.
Gelman, A.; J. B. Carlin; H. S. Stern; and D. B. Rubin. *Bayesian Data Analysis*, 1st ed. London: Chapman & Hall (1995).
Geweke, J. "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." In *Bayesian Statistics 4*, J. O. Berger, J. M. Bernardo, A. P. David, and A. F. M. Smith, eds. Oxford: Oxford Univ. Press (1992).
Ghysels, E. "On Stable Factor Structures in the Pricing: Do Time-Varying Betas Help or Hurt?" *Journal of Finance*, 53 (1998), 549–573.
Gomes, J.; L. Kogan; and L. Zhang. "Equilibrium Cross-Section of Returns." *Journal of Political Economy*, 111 (2003), 693–732.
Jacquier, E.; N. G. Polson; and P. E. Rossi. "Bayesian Analysis of Stochastic Volatility Models (with discussion)." *Journal of Business and Economic Statistics*, 12 (1994), 281–300.
Kaldor, N. "Capital Accumulation and Economic Growth." In Proceedings of a Conference Held by the International Economics Association. F. A. Lutz and D. C. Hague, eds. London: Macmillan (1961).
Ledoit, O.; P. Santa-Clara; and M. Wolf. "Flexible Multivariate GARCH Modeling with an Application to International Stock Markets." *Review of Economics and Statistics*, 85 (2003), 735–747.
Levy, R. A. "Stationarity of Beta Coefficients." *Financial Analysts Journal*, 27 (1971), 55–62.
Lewellen, J., and S. Nagel. "The Conditional CAPM Does Not Explain Asset-Pricing Anomalies." Working Paper, MIT and Harvard Univ. (2003).
Litzenberger, R., and K. Ramaswamy. "The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence." *Journal of Financial Economics*, 7 (1979), 163–196.
Lo, A., and C. MacKinlay. "Data-Snooping Biases in Tests of Financial Asset Pricing Models." *Review of Financial Studies*, 3 (1990), 431–468.
Petkova, R., and L. Zhang. "Is Value Riskier than Growth?" *Journal of Financial Economics*, forthcoming (2005).
Shanken, J. "Intertemporal Asset Pricing: An Empirical Investigation." *Journal of Econometrics*, 45 (1990), 99–120.
Tamayo, A. "Stock Return Predictability, Conditional Asset Pricing Models and Portfolio Selection." Working Paper, London Business School (2002).
Turnbull, S. M. "Market Value and Systematic Risk." *Journal of Finance*, 32 (1977), 1125–1142.
Turtle, H.; A. Buse; and B. Korkie. "Test of Conditional Asset Pricing with Time-Varying Moments and Risk Prices." *Journal of Financial and Quantitative Analysis*, 29 (1994), 15–29.
Vasicek, O. "A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas." *Journal of Finance*, 28 (1973), 1233–1239.
Zhang, L. "The Value Premium." *Journal of Finance*, 60 (2004), 67–103.

https://doi.org/10.1017/S0022109000001964 Published online by Cambridge University Press