COSTS OF EQUITY CAPITAL AND MODEL MISPRICING

Luboš Pástor
Robert F. Stambaugh

Working Paper 6490

NBER WORKING PAPER SERIES

COSTS OF EQUITY CAPITAL AND MODEL MISPRICING

Luboš Pástor
Robert F. Stambaugh

Working Paper 6490
http://www.nber.org/papers/w6490

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1998

The authors are grateful for helpful comments by Wayne Ferson, Craig MacKinlay, René Stulz, two anonymous referees, and seminar participants at Harvard University, the National Bureau of Economic Research, New York University, the University of Pennsylvania, the University of Toronto, and Virginia Polytechnic Institute. Stambaugh acknowledges financial support provided during his appointment as a Marvin Bower Fellow at Harvard Business School, where portions of this research were conducted. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Costs of Equity Capital and Model Mispricing
Luboš Pástor and Robert F. Stambaugh
NBER Working Paper No. 6490
April 1998
JEL Nos. G12, G31, C11

ABSTRACT

Costs of equity for individual firms are estimated in a Bayesian framework using several factor-based pricing models. Substantial prior uncertainty about mispricing often produces an estimated cost of equity close to that obtained with mispricing precluded, even for a stock whose average return departs significantly from the pricing model's prediction. Uncertainty about which pricing model to use is less important, on average, than within-model parameter uncertainty. In the absence of mispricing uncertainty, uncertainty about factor premiums is generally the largest source of overall uncertainty about a firm's cost of equity, although uncertainty about betas is nearly as important.

Luboš Pástor
Finance Department
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367
lubos23@wharton.upenn.edu

Robert F. Stambaugh
Finance Department
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367
and NBER
stambaugh@wharton.upenn.edu

Introduction

The expected rate of return on a firm's stock, the "cost of equity," is an important input for making decisions affecting the firm. Because it affects the discount rate at which expected future cash flows are valued, the cost of equity plays a key role in the firm's investment decisions. For a public utility, the estimated cost of equity capital has a direct impact on how the prices of the firm's services are regulated by its state public utility commission.

The cost of equity is equal to a riskless interest rate plus the expected excess return on the firm's stock. One approach to estimating the latter quantity uses a standard asset-pricing model, in which the expected excess return hinges on sensitivities of the firm's stock return to market-wide factors.[1] If r_t denotes the stock's excess return and f_t denotes a $K \times 1$ vector of the factors, all realized in period t, then the stock's sensitivities, or "betas," are the slope coefficients in the regression

$$r_t = \alpha + \beta' f_t + \epsilon_t, \tag{1}$$

where ϵ_t is the mean-zero regression disturbance. When the factors appropriate to the given model are constructed as excess portfolio returns or payoffs on zero-investment positions, as will be the case in the models analyzed below, then the pricing model implies $\alpha = 0$.[2] That is, the pricing model implies that the stock's expected excess return, μ, is given by $\beta'\lambda$, where λ is the vector of "factor premiums," the expected values of the factors.

The elements of β and λ must be estimated, so the true cost of equity is uncertain. Moreover, even if β and λ were known for certain, one might be skeptical about whether any pricing model could deliver the correct cost of equity for every stock. That is, a decision maker might be uncertain about whether the model misprices the stock in question, so that the expected excess return might actually be

$$\mu = \alpha + \beta'\lambda, \tag{2}$$

where α is some unknown non-zero amount by which the model misprices the stock. This "mispricing" uncertainty about α contributes further to the uncertainty about the cost of equity. Finally, if the decision maker has any doubts about which pricing model to use, then the uncertainty about μ also includes that "model" uncertainty. This study attempts to quantify these various sources of uncertainty and gauge the relative importance of each in estimating a firm's cost of equity.

In the presence of mispricing uncertainty, a decision maker might wish to combine a cost-of-equity estimate produced by a pricing model with an alternative estimate, such as the stock's historical average return. Suppose for example that, apart from any empirical evidence, a manager believes that there is a one-third probability that the difference between his stock's expected return and the value implied by the CAPM is at least 5% per annum in absolute value. In other words, the manager's "prior" standard deviation for his stock's α is about 5%. Given this degree of skepticism in the accuracy of the pricing model, how much attention should the manager pay to the historical average return on the firm's stock when estimating the firm's equity cost of capital? Specifically, suppose that, based on the last two decades, the firm's stock has a sample estimate of its CAPM α equal to 8% per annum with a t-statistic greater than two. To what extent should this skeptical manager make use of that historical information? With complete faith in the CAPM, such information would be ignored—only the market risk premium and the stock's beta would be used in estimating the cost of equity. With an extreme degree of uncertainty about the model's accuracy, one might simply ignore the CAPM and estimate the firm's cost of equity as the stock's historical average rate of return, which would be 8% higher than the CAPM value. We explore solutions to this problem with intermediate degrees of mispricing uncertainty, as in the example given here.

We develop and apply a method for estimating the cost of equity using a Bayesian approach. In this setting, the decision maker does not know the true expected excess return but instead uses the conditional expectation $\mathrm{E}(r_t|\Phi)$, where Φ denotes the information available at the time of the decision. We assume that excess returns have constant mean μ, so the decision maker's estimate for the expected excess return is then simply the posterior mean of μ given Φ, and the decision maker's uncertainty about the cost of equity is reflected in the posterior variance of μ. As Cornell, Hirshleifer, and James (1997) conclude, "judgment enters the process at numerous points," regardless of the method used to estimate the cost of capital. A basic feature of the Bayesian approach is that a decision maker's judgment, represented by prior beliefs, enters the estimation in a manner guided by a scientific principle (Bayes's theorem) as opposed to more ad hoc methods. As discussed above, one aspect of the decision maker's judgment that we can explore is uncertainty about whether a pricing model can deliver the precise expected excess return for a given stock. We find that, in many cases, mispricing uncertainty that seems important in economic terms does not impact greatly the estimated cost of equity. That is, the posterior mean of μ is close to the posterior mean obtained when mispricing is ruled out, even when the sample least-squares estimate of α departs significantly from zero. Suppose one's prior beliefs are that the stock of interest is typical in terms of its betas and its variance of ϵ_t. Then, when the prior standard deviation of α is 5% per annum, as in the above example, the estimated cost of equity is less than 1%

(per annum) above the CAPM value, even though the sample estimate of α is nearly 8% and its t-statistic exceeds two. In this sense, a pricing model that might be viewed by the decision maker as being only mediocre in its ability to price stocks accurately is still relied upon fairly heavily in estimating the cost of equity.

This study investigates factor-based models with a focus on the estimates they produce rather than on their asset-pricing abilities versus each other or versus non-factor-based approaches. Even though the latter issues continue to invite debate in the academic literature, we suggest that these factor-based models have received sufficient interest to merit investigating their potential use by decision makers. Three pricing models are used to illustrate our approach. The first is the CAPM, where the single factor is specified to be the excess return on a market index portfolio. The second model, proposed by Fama and French (1993), contains that market factor plus two additional factors: the difference in returns between small and large firms and the difference in returns between firms with high and low ratios of book value to market value. The third model also has three factors, but, instead of prespecifying them, we extract them from returns on a large cross-section of stocks using the asymptotic principal components method of Connor and Korajczyk (1986).

Uncertainty about which of the three factor-based models to use can contribute nontrivially to a decision maker's overall uncertainty about the cost of equity, but this source of uncertainty is typically less important than the parameter uncertainty within any given model. For example, when each model is assigned an equal probability of being the "correct" one, we obtain an overall posterior standard deviation for the cost of equity of 5% or more per year, depending on the prior uncertainty about α, but that value is typically no more than 0.75% above the posterior standard deviation of μ obtained within any single model.

Uncertainty about β contributes substantially to the overall uncertainty about the cost of equity for an individual firm, but somewhat more important is the uncertainty about λ, the vector of factor premiums. Fama and French (1997) estimate expected returns for industry portfolios using both the CAPM as well as the Fama-French (1993) three-factor model. Based on frequentist standard errors, they conclude that by far the largest source of imprecision in industry costs of equity arises from estimation of λ. Ferson and Locke (1997), also in a frequentist setting, examine sources of error in CAPM-based estimates of expected returns on portfolios of stocks grouped by industry or market capitalization. They similarly conclude that errors in β are likely to be less important than errors in estimating the market premium.[3] Although uncertainty about β, not surprisingly, is more important for individual firms than for portfolios, our conclusion regarding the importance of uncertainty about λ is otherwise similar to the conclusions of these studies. In all three of the models, the histories of the factors are available beginning in July 1963, but the factors are correlated with other series whose histories begin earlier. As a result, the longer-history series contain additional information about λ, as discussed by Stambaugh (1997). We find that, in the absence of uncertainty about mispricing, uncertainty about λ remains the most important source of uncertainty about a firm's cost of equity, even after incorporating information about λ that is contained in series whose histories begin in 1926.

In keeping with the spirit of a factor-based approach, much of our analysis assumes that the information set used by the decision maker consists of histories of factors and stock returns. That is, the decision maker does not make use of firm-specific characteristics, except perhaps in constructing the factors (as in, for example, the Fama-French (1993) model). Previous studies have recommended the use of firm-specific characteristics in estimating the cost of equity (e.g., Elton, Gruber, and Mei (1994) or Schink and Bower (1994)), and the usefulness of various firm-specific characteristics in explaining expected returns has been argued recently by Daniel and Titman (1997). Another feature of the Bayesian approach is that it allows the decision maker to introduce additional prior information about the firm whose cost of equity is to be estimated, and our methodology allows the decision maker to either ignore or incorporate such prior information. In specifying the prior, the firm can be regarded as a random draw either from the whole cross-section of stocks, when firm-specific characteristics are ignored, or from a group of firms with similar characteristics, when the firm's characteristics are incorporated. As a simple illustration of the latter case, we include a firm's industry classification as additional prior information and analyze estimates of expected excess returns on stocks of utilities, which constitute an industry in which estimated costs of equity have clear practical relevance.

The remainder of the paper is organized as follows. The methodology is developed in Section I, wherein we present the general form of the priors used in our Bayesian approach, explain how we obtain the resulting posterior distributions of μ and its components, and describe the empirical-Bayes procedure used to obtain parameters in the prior distributions. Sections II and III contain our empirical results. Section II reports and analyzes posterior moments of the expected excess return and its components for individual stocks. Those results include a detailed analysis for one stock as well as analyses for two cross-sections: a broad universe of 1,994 stocks and a smaller set of 135 utility stocks. Section III investigates the potential uncertainty about the cost of equity that arises from uncertainty about which pricing model to use. Section IV reviews the conclusions.

I. Methodology

A. Overview

The estimate of a stock's expected excess return is given by the posterior mean, $\mathrm{E}(\mu|\Phi)$, where μ is a function of the unknown parameters α, β, and λ (equation (2)) and Φ is the historical sample information available to the decision maker. The imprecision in the estimate of the expected excess return is characterized by the posterior variance, $\mathrm{Var}(\mu|\Phi)$. The posterior mean and variance are obtained by combining the sample information about the unknown parameters with the decision maker's prior beliefs about those parameters. A key feature of the prior beliefs is the mispricing uncertainty about α, represented by the prior standard deviation, σ_α. We let σ_α take different values on the interval $(0, \infty)$ in order to explore the role of mispricing uncertainty in estimating the cost of equity.

Prior beliefs about the elements of β and their correlations with α are constructed by viewing the firm as a random draw from a cross-section of firms. The prior mean of β, for example, is set equal to the average of the ordinary-least-squares (OLS) estimates of β for the firms in the cross-section. This cross-section can be selected either as a broad universe or as a subset of firms that share one or more characteristics with the firm whose cost of equity is to be estimated. That firm's posterior mean of β is then "shrunk" away from its own OLS estimate and toward the cross-sectional mean, in a manner similar to that discussed by Vasicek (1973). If, as in an example presented later, the firm is a public utility and the cross-section consists of other utilities, then the given firm's β is shrunk toward the average β for utilities. In estimating costs of equity for various industries, Fama and French (1997) follow a similar approach and shrink each industry's β toward the market-wide average β.

The prior mean of α is set equal to zero: in the absence of any observations of the firm's historical stock performance, the decision maker is viewed as unable to sign the potential mispricing. This assumption about the prior mean of α, although perhaps the most natural, is made for simplicity and is not required by the methodology. In particular, one could instead set the prior mean of α equal to the average OLS estimate of α for a cross-section of firms with characteristics similar to the given firm in question, as done with β. The latter approach would be one way to incorporate the type of characteristics-based pricing effects investigated by Daniel and Titman (1997).

The decision maker is assumed to have "diffuse" prior beliefs about λ, the vector of factor premiums. In other words, without observing any past realizations of the factors, the decision maker would have essentially no idea about their expected values. The histories of the factor realizations are often longer than the firm's return history used in the cost-of-capital estimation. Moreover, additional information about the factor premiums is obtained from variables whose histories are longer than those of the factors. In the one-factor CAPM and the two three-factor models used here, the histories of the factors begin in July 1963, but the returns on the factors are correlated with longer series that provide additional information about the factor premiums. For example, the Fama-French NYSE-AMEX-Nasdaq market index, which we use as the market factor in the CAPM and in the Fama-French three-factor model, has returns available starting in July 1963, but those returns are highly correlated with the returns on the value-weighted NYSE portfolio, which CRSP supplies beginning in 1926. As shown by Stambaugh (1997), that longer-history series contains additional information about the mean of the shorter-history market factor. For each of the pricing models used here, the cost-of-equity estimates incorporate the additional information about factor means that is contained in three series whose histories begin in January 1926: the value-weighted NYSE portfolio, the equally weighted NYSE portfolio, and the Ibbotson small-stock portfolio (all obtained from CRSP).

The remainder of this section provides the details of the methodology. The reader who wishes to proceed directly to the empirical results may skip to Section II.

B. Stochastic Setting

Let r denote the $T \times 1$ vector of returns on the stock of the firm whose cost of equity is to be estimated. In many cases, the stock's return history, or at least the portion of that history used in the estimation, may be shorter than the history of the factors. It is assumed that there are S observations of the factors, with $S \geq T$. Let $F^{(T)}$ denote the $T \times K$ matrix containing the T observations of the factors corresponding to the same periods as the returns in r. The regression disturbance ϵ_t in (1) is assumed to be, in each period t, an independent realization from a normal distribution with zero mean and variance σ^2, so the most recent T observations of the returns and the factors obey the regression relation

$$r = Xb + \epsilon, \qquad \epsilon \sim N(0, \sigma^2 I_T), \tag{3}$$

where $b' \equiv [\alpha \ \beta']$, $X = [\iota_T \ F^{(T)}]$, ϵ contains the T regression disturbances, ι_T is a T-vector of 1's, I_T is a $T \times T$ identity matrix, and the notation "$\sim$" is read "is distributed as."

In addition to the S observations of the K factors, there exist L observations of K_L variables that are correlated with the factors. If $L > S$, then, as shown by Stambaugh

(1997), the longer histories of these additional variables contain information about λ, the $K \times 1$ vector of factor means, beyond that contained in the factor histories alone. Let y_t denote the $K_L \times 1$ vector containing the observations of the additional variables in period t, and let $Y^{(L)}$ denote the $L \times K_L$ matrix containing all L observations of y_t. For each of the S periods over which both f_t and y_t are observed, define the "augmented" set of factors $f_t^{a\prime} = [f_t' \; y_t']$, and assume that

$$f_t^a \sim N(\theta,\ G), \tag{4}$$

where the realizations are independent across t, $\theta' = [\lambda' \; \theta_2']$, and G denotes the covariance matrix of f_t^a. For the $L - S$ periods in which only y_t is observed, it is also assumed that

$$y_t \sim N(\theta_2,\ G_{22}), \tag{5}$$

again with independent realizations across t, where G_{22} is the corresponding submatrix of G. That is, the marginal distribution of y_t is given by (5) for all L periods. Finally, it is assumed that f_t^a is independent of ϵ for all t.

Given the above assumptions, it follows that the likelihood function for the parameters $(b,\ \sigma,\ \theta,\ G)$ can be factored as

$$p(r, F^{(S)}, Y^{(L)}|b, \sigma, \theta, G) = p(r|F^{(T)}, b, \sigma)p(F^{(S)}, Y^{(L)}|\theta, G), \tag{6}$$

where the likelihood function for the regression parameters is

$$p(r|F^{(T)}, b, \sigma) \propto \frac{1}{\sigma^T} \exp\left\{-\frac{1}{2\sigma^2}(r - Xb)'(r - Xb)\right\}, \tag{7}$$

and the likelihood function for the moments of the factors and additional variables is

$$\begin{aligned} p(F^{(S)}, Y^{(L)}|\theta, G) \;\; \propto \;\; & |G_{22}|^{-\frac{L-S}{2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{L-S}(y_t - \theta_2)'(G_{22})^{-1}(y_t - \theta_2)\right\} \\ & \times |G|^{-\frac{S}{2}} \exp\left\{-\frac{1}{2}\sum_{t=L-S+1}^{L}(f_t^a - \theta)'(G)^{-1}(f_t^a - \theta)\right\}. \end{aligned} \tag{8}$$

(The notation "$\propto$" is read "is proportional to.")

C. Priors

C.1. General Specification

We propose a normal-inverted-gamma prior on the regression parameters b and σ:

$$b|\sigma \;\; \sim \;\; N(\bar{b}, \Psi(\sigma)) \tag{9}$$

$$\sigma^2 \;\; \sim \;\; \frac{\nu s_0^2}{\chi_\nu^2}, \tag{10}$$

where

$$\Psi(\sigma) = \begin{bmatrix} \left(\frac{\sigma^2}{E(\sigma^2)}\right)\sigma_\alpha^2 & \left(\frac{\sigma}{E(\sigma)}\right)\sigma_\alpha(\rho_{\alpha\beta}\sigma_\beta)' \\ \left(\frac{\sigma}{E(\sigma)}\right)\sigma_\alpha(\rho_{\alpha\beta}\sigma_\beta) & V_\beta \end{bmatrix}. \tag{11}$$

In the above, σ_β is a $K \times 1$ vector containing the square roots of the diagonal elements of V_β, the covariance matrix of β, and $\rho_{\alpha\beta}$ is a $K \times K$ diagonal matrix with the simple correlations between α and the elements of β on the main diagonal. Since $\bar{b}$ does not depend on σ, the marginal prior covariance matrix of b equals

$$\begin{aligned} V_b \equiv \operatorname{cov}(b, b') &= E(\Psi(\sigma)) \\ &= \begin{bmatrix} \sigma_\alpha^2 & \sigma_\alpha(\rho_{\alpha\beta}\sigma_\beta)' \\ \sigma_\alpha(\rho_{\alpha\beta}\sigma_\beta) & V_\beta \end{bmatrix}, \end{aligned} \tag{12}$$

and it is assumed that V_b is positive definite. In order to have $\Psi(\sigma)$ be positive definite, we also require

$$\sigma_\beta' \rho_{\alpha\beta}' V_\beta^{-1} \rho_{\alpha\beta}\sigma_\beta < \frac{[E(\sigma)]^2}{E(\sigma^2)} = \frac{\nu - 2}{2}\left(\frac{\Gamma[(\nu-1)/2]}{\Gamma(\nu/2)}\right)^2, \tag{13}$$

where the equality of the second and third expressions follows from the properties of the inverted gamma distribution for σ,[4]

$$E(\sigma^2) = \frac{\nu s_0^2}{\nu - 2} \tag{14}$$

and

$$E(\sigma) = \frac{\Gamma[(\nu-1)/2]}{\Gamma(\nu/2)}\left(\frac{\nu s_0^2}{2}\right)^{1/2}. \tag{15}$$

In specifying the parameters for the above priors, we use an empirical-Bayes procedure that relies on data for a cross-section of individual stocks. The effects of "mispricing" uncertainty are investigated by entertaining a wide range of values for σ_α. Details of that approach are provided in Subsection C.2.

Observe in (11) that the conditional prior variance of α is proportional to σ^2, the variance of ϵ_t. This feature of our prior recognizes that a high value of $|\alpha|$ accompanied by a low value of σ^2 implies a high Sharpe ratio for some combination of the asset, the factor-mimicking positions, and cash (earning the riskless rate).[5] In particular, $(\alpha/\sigma)^2$ is the difference between the maximum squared Sharpe ratio for such a combination and the maximum squared Sharpe ratio for combinations of only the factor-mimicking positions and cash.[6] Following MacKinlay (1995), a prior positive association between α and σ is imposed to reduce the probability of high Sharpe ratios as compared to priors that treat those parameters as independent. In contrast, we do assume independence between β and σ in the absence of a compelling a priori argument to the contrary.

The structure of the covariance matrix for b, $\Psi(\sigma)$ in (11), produces a prior that is essentially a hybrid of two more standard alternative priors for the regression model. In one alternative, the normal density for b and the inverted-gamma density for σ^2 are independent, so that no part of the covariance matrix for b involves σ^2 (e.g., Chib and Greenberg (1996)). As explained above, this prior would make α independent of σ^2 and hence assign greater probability to high Sharpe ratios. In the other alternative, the well-known natural-conjugate prior, the marginal prior for σ^2 is still inverted gamma, but the entire covariance matrix of b is proportional to σ^2 (e.g., Zellner (1971), Chapter 3). In the formula for the posterior mean of β for that case, the relative weights on the sample estimate and the prior mean do not depend on sample information about σ. That is, $\hat{\beta}$ is given no more weight when the sample residual variance is small than when it is large, and that property is unappealing. Vasicek (1973) argues that the natural-conjugate prior is inappropriate when the prior parameters are estimated from a cross-section of stocks.

We assume that the regression parameters are independent of the moments of f_t^a, the augmented set of factors:

$$p(b, \sigma, \theta, G) = p(b, \sigma)p(\theta, G). \tag{16}$$

The prior density for θ and G is specified as

$$p(\theta, G) \propto |G|^{-\frac{K+K_L+1}{2}}, \tag{17}$$

which is the standard diffuse prior used to represent "non-informative" beliefs about the parameters of a multivariate normal distribution (e.g., Box and Tiao (1973)).

C.2. Prior Parameters

In order to construct the prior distribution for the regression parameters in (9) and (10), we specify the elements in $\bar{b}$ and V_b and the scalar quantities s_0 and ν. (Note from (11) through (15) that V_b, s_0^2 and ν determine the conditional covariance matrix $\Psi(\sigma)$.) The prior values are chosen with the objective that the prior mean of b for any given stock be the mean of b in a given cross-section of stocks and that the prior unconditional covariance matrix of b for that stock, V_b, be the covariance matrix of b in the cross-section. Similarly, the prior mean and variance of σ^2 for the stock, determined by s_0 and ν, correspond to moments of σ^2 in the cross-section.

We construct prior distributions using two specifications for the cross-section of stocks. The first cross-section consists simply of all stocks on the NYSE and AMEX (subject to a data-availability requirement detailed below). In this first specification, which is used throughout much of our analysis, the stock to be analyzed is essentially viewed as a random draw from the universe of all stocks. Although this approach strikes us as a reasonable starting point, at least for our exploratory study, it is only one of many methods that might be used to specify the prior. In a statistical sense, the normal-inverted-gamma prior in (9) and (10) is generally characterized as "informative" as opposed to diffuse (non-informative), but this first specification of the prior does not rely on specific knowledge about the firm. In an economic sense, therefore, this prior is rather uninformative. In contrast, our second cross-section of stocks consists solely of utilities, so the prior thereby constructed can be viewed as economically informative. In other words, the prior incorporates knowledge of a characteristic—industry classification—of the firm whose cost of equity is to be estimated.

The cross-sectional moments of b and σ^2 are not directly observable. We take an empirical-Bayes approach and estimate those moments using values of $\hat{b}$ and $\hat{\sigma}^2$ computed for a large cross-section.[7] Fama and French (1997) apply a similar methodology, following Blattberg and George (1991), in computing shrinkage estimates of β for industry portfolios. The first prior, based on the broad cross-section, is constructed as follows. For each stock in the CRSP monthly NYSE-AMEX file with at least 24 months of data in the period from July 1963 through December 1995, we compute $\hat{b}$ and $\hat{\sigma}^2$ using all of that stock's available data during that period. The stock returns are in excess of the return on a one-month Treasury bill (from CRSP's SBBI file). For the CAPM and the Fama-French (1993) three-factor model (hereafter the FF model), the factor data begin in July 1963 and consist of monthly realizations of three factors: (i) the excess return on a market-index portfolio of NYSE, AMEX, and Nasdaq stocks, (ii) the difference in returns between a small-stock portfolio and a large-stock portfolio, and (iii) the difference in returns between a portfolio of high book-to-market (B/M) stocks and a portfolio of low B/M stocks.[8] Only the first of these factors is used in the CAPM. To construct the three factors for the Connor-Korajczyk (1986) model (hereafter the CK model), we take all stocks with at least one year of data on the NYSE-AMEX monthly CRSP file for the 7/63–12/95 period and then extract one set of factors for that entire period using the method in Connor and Korajczyk (1987) that allows for missing observations.[9]

The statistics $\hat{b}$ and $\hat{\sigma}^2$, computed for each stock, are used to construct the prior parameters b, V_b, s_0^2, and ν. The prior mean of b, $\bar{b}$, is set equal to the cross-sectional average of the $\hat{b}$'s, except that the first element, $\bar{\alpha}$, is set to zero. The prior covariance matrix of b, V_b, is constructed as follows. First, we compute the matrix

$$\hat{V}_b = \Xi(\hat{b}) - \overline{\hat{\sigma}_i^2 (X'X)_i^{-1}}, \tag{18}$$

where $\Xi(\hat{b})$ is the sample cross-sectional covariance matrix of the $\hat{b}$'s. The second term in (18) is the average across stocks of the usual estimate for the sampling variance of $\hat{b}$, where $\hat{\sigma}_i^2$ and $(X'X)_i$ are based on the observations available for stock i. (The bar denotes an average across stocks.) As noted by Fama and French (1997), under standard assumptions, $\hat{V}_b$ is an estimate of the cross-sectional covariance of the b's. For all three models, it happens that $\hat{V}_b$ is positive definite (not guaranteed in general). To construct the matrix V_b, as represented in (12), V_β is set equal to the corresponding submatrix of $\hat{V}_b$, and $\rho_{\alpha\beta}$ is taken from the correlation matrix associated with $\hat{V}_b$. Rather than set σ_α^2 equal to the (1,1) element of $\hat{V}_b$, however, we instead let it take a wide range of values, ranging from zero to infinity.[10] Each value of σ_α is then combined with the fixed values of V_β and $\rho_{\alpha\beta}$, using (12), to form the matrix V_b used in the prior.

The inverted gamma density for σ implies[11]

$$\nu = 4 + \frac{2(\mathrm{E}(\sigma^2))^2}{\mathrm{Var}(\sigma^2)}. \tag{19}$$

We substitute the cross-sectional mean of the $\hat{\sigma}^2$'s for $\mathrm{E}(\sigma^2)$ in (19), and for $\mathrm{Var}(\sigma^2)$ we substitute

$$\hat{v}_{\sigma^2} = \xi(\hat{\sigma}^2) - \overline{\frac{T_i - K - 1}{T_i^2} 2\hat{\sigma}_i^4}. \tag{20}$$

Equation (20) is the analog to equation (18). The first term on the right-hand side is the sample cross-sectional variance of the $\hat{\sigma}^2$'s, while the second term is the cross-sectional average of the estimates of the sampling variance of $\hat{\sigma}^2$. That is, $\hat{\sigma}_i^2$ denotes the estimated residual variance for stock i, based on T_i observations for that stock, and the estimated sampling variance of $\hat{\sigma}_i^2$ is the quantity under the bar in (20). The value of ν in the prior is set to the next largest integer of the resulting value on the right-hand side of (19). Given that value of ν, the value of s_0^2 used in the prior is obtained from (14), where the cross-sectional average of the $\hat{\sigma}^2$'s is substituted for $\mathrm{E}(\sigma^2)$.

Panel A of Table I reports the parameter values used in the prior constructed from the entire cross-section of stocks. Note that in the CAPM the prior correlation between α and β is positive. This occurs in spite of a negative cross-sectional correlation between the sample estimates $\hat{\alpha}$ and $\hat{\beta}$, as has been observed in previous studies (e.g., Black, Jensen, and Scholes (1972)). That is, the off-diagonal element of the first matrix on the right-hand side of (18) is negative. The positive correlation in the prior results from the fact that the average sampling covariance between $\hat{\alpha}$ and $\hat{\beta}$, appearing in the second term on the right-hand side of (18), is also negative, and the difference results in a positive estimate of the cross-sectional covariance between α and β. For the other two models, the prior correlations between α and the elements of β are generally negative. In particular, the prior correlation between α and the HML sensitivity (the last element of b in the FF model) is -0.55. This value, obtained here with individual stocks, is consistent with a similarly large negative correlation between α and HML sensitivities for industry portfolios observed by Fama and French (1997).

As noted earlier, the prior based on the entire cross-section can be viewed as economically non-informative compared to a prior that makes use of a cross-section selected according to one or more of the firm's characteristics. For example, if a public utility's cost of equity is to be estimated, the prior parameters can be obtained from a cross-section of utilities rather than the cross-section of all stocks. Our second prior uses the cross-section of 186 utility firms (SIC codes between 4900 and 4999) with at least 48 months of data in the period from July 1963 through December 1995. The same approach described earlier for the entire cross-section is applied here, except that the off-diagonal elements of $\hat{V}_b$ are set to zero. The latter simplification and the 48-month data requirement are imposed in order to obtain a positive-definite prior covariance matrix for b with this smaller cross-section. Panel B of Table I reports the parameter values in this utility-specific prior.

D. Posteriors

The posterior density for the parameters is proportional to the product of the prior density and the likelihood function. Given the factorizations of the likelihood function in (6) and the prior density in (16), the posterior density can also be factored as the posterior for b and σ multiplied by the posterior for θ and G. We analyze these two posteriors separately and then explain how we combine the posterior moments for b and λ to obtain posterior moments for the expected excess return.

D.1. Regression Parameters

The joint prior density $p(b, \sigma)$ is equal to the product $p(b|\sigma)p(\sigma)$, where the normal prior density for b given σ in (9) can be written as[12]

$$p(b|\sigma) \quad \propto \quad |\Psi(\sigma)|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(b-\bar{b})'\Psi(\sigma)^{-1}(b-\bar{b})\right\}$$

$$\propto \frac{1}{\sigma} \exp\left\{ -\frac{1}{2\sigma^2} (b - \bar{b})' (\frac{1}{\sigma^2} \Psi(\sigma))^{-1} (b - \bar{b}) \right\}, \tag{21}$$

and the marginal inverted-gamma prior density for σ in (10) can be written as

$$p(\sigma) \propto \frac{1}{\sigma^{\nu+1}} \exp\{-\frac{\nu s_0^2}{2\sigma^2}\}. \tag{22}$$

Multiplying the prior densities in (21) and (22) and the likelihood in (7) gives the joint posterior for b and σ, which can be written as

$$\begin{aligned} p(b, \sigma | r, F^{(T)}) \quad \propto \quad & \frac{1}{\sigma^{\nu+T+2}} \exp\left\{ -\frac{1}{2\sigma^2} \left[\nu s_0^2 + T\hat{\sigma}^2 + (b - \bar{b})' (\frac{1}{\sigma^2} \Psi(\sigma))^{-1} (b - \bar{b}) \right.\right. \\ & \left.\left. + (b - \hat{b})' X' X (b - \hat{b}) \right] \right\}, \end{aligned} \tag{23}$$

where $\hat{b} = (X'X)^{-1} X'r$ and $T\hat{\sigma}^2 = (r - X\hat{b})'(r - X\hat{b})$. We compute moments of this joint posterior using the Metropolis-Hastings (MH) algorithm, a Markov chain Monte Carlo procedure introduced by Metropolis et al. (1953) and generalized by Hastings (1970). (For an introduction to the MH algorithm, see Chib and Greenberg (1995) or Gilks, Richardson, and Spiegelhalter (1996).) Briefly, a sequence of draws of b and σ is constructed by making "candidate" draws from a "proposal" density and then accepting a new candidate or retaining the previous value based on a rule that assures the resulting sequence for (b, σ) forms a Markov chain whose invariant distribution is the "target" posterior density of interest. The posterior moments of the parameters are computed as the sample moments of a large number of draws. We use a "block-at-a-time" version of the MH algorithm, where b is drawn directly from the conditional density $p(b|\sigma, r, F^{(T)})$, but σ is drawn from a proposal density given by the conditional posterior density for σ that arises when σ and b are made independent in the normal-inverted-gamma prior.[13] The target density for σ is the conditional density $p(\sigma|b, r, F^{(T)})$, which is proportional to the right-hand side of (23) (since b is then viewed as a constant and the marginal density of b, by definition, does not involve σ). We simulate a MH chain of 50,500 draws, discard the first 500 draws, and estimate the posterior moments of b and σ over the remaining 50,000 draws. The number of draws is chosen such that, across repeated independent runs of the MH algorithm, differences in the computed first and second moments of b are small enough for us to report at least two decimal places in our results.

From (23) we see that the conditional posterior for b given σ can be written as

$$\begin{aligned} p(b|\sigma, r, F^{(T)}) \quad & \propto \quad \exp\left\{ -\frac{1}{2} \left[(b - \bar{b})' \Psi(\sigma)^{-1} (b - \bar{b}) + (b - \hat{b})' (\frac{1}{\sigma^2} X' X)(b - \hat{b}) \right] \right\} \\ & \propto \quad \exp\left\{ -\frac{1}{2} (b - \tilde{b}_\sigma)' M (b - \tilde{b}_\sigma) \right\}, \end{aligned} \tag{24}$$

where

$$M = \Psi(\sigma)^{-1} + \frac{1}{\sigma^2} X'X, \tag{25}$$

and

$$\tilde{b}_\sigma = M^{-1} \left[\Psi(\sigma)^{-1} \bar{b} + \frac{1}{\sigma^2} X'X\hat{b} \right]. \tag{26}$$

Hence, the conditional posterior distribution for b given σ is normal with mean $\tilde{b}_\sigma$ and covariance matrix M^{-1}, so it is easily sampled directly, as noted in the above description of the MH algorithm. Observe that $\tilde{b}_\sigma$ is a (matrix) weighted average of the prior mean $\bar{b}$ and the sample estimate $\hat{b}$, where the weights are the precisions of $\bar{b}$ and $\hat{b}$ conditional on σ. This weighting can be interpreted as shrinking the sample estimate $\hat{b}$ toward its prior mean $\bar{b}$, where the degree of shrinkage depends on the relative reliability of the sample estimate. This shrinkage effect is discussed further in Section II.

We find that the first and second posterior moments of b, computed using the MH algorithm, are approximated well by the moments of $p(b|\sigma, r, F^{(T)})$ evaluated at a reasonable estimate of σ (using (25) and (26)). An estimate of σ for this purpose is computed in two steps. Using (26), the posterior mean of b conditioned on $\sigma = \hat{\sigma}$ is computed, and its value is denoted as b^*. The final estimate of σ is computed as the posterior mean of σ conditioned on $b = b^*$, using the conditional posterior density for σ that arises when σ and b are made independent in the normal-inverted-gamma prior (i.e., the proposal density for σ in the MH algorithm). In the empirical analyses in Sections II and III, we present a one-stock example based on the MH algorithm, but we use the approximation to compute posterior moments for a large number of stocks, since performing the MH algorithm for each stock would be computationally prohibitive.

D.2. Factor Means

Define the first and second sample moments of y_t,

$$\hat{\theta}_2 = \frac{1}{L} Y^{(L)\prime} \iota_L, \tag{27}$$

and

$$\hat{G}_{22} = \frac{1}{L}(Y^{(L)} - \iota_L \hat{\theta}_2')'(Y^{(L)} - \iota_L \hat{\theta}_2'). \tag{28}$$

Let $Y^{(S)}$ denote the $S \times K_L$ matrix containing the S observations of y_t corresponding to the same S periods as those in $F^{(S)}$, and define $Z = [\iota_S \; Y^{(S)}]$. The least-squares coefficient matrix in a multivariate regression of $F^{(S)}$ on $Y^{(S)}$ is

$$\hat{H} = \begin{bmatrix} \hat{h}_1' \\ \hat{H}_2' \end{bmatrix} = (Z'Z)^{-1} Z' F^{(S)}, \tag{29}$$

where $\hat{h}_1$ is $K \times 1$ and $\hat{H}_2$ is $K \times K_L$, and the sample covariance matrix of the residuals is

$$\hat{\Sigma} = \frac{1}{S}(F^{(S)} - Z\hat{H})'(F^{(S)} - Z\hat{H}). \tag{30}$$

The sample statistics in (27) through (30) prove useful in computing the posterior first and second moments of λ, which are derived in Part A of the Appendix. The posterior mean of λ is

$$\tilde{\lambda} = \hat{h}_1 + \hat{H}_2\hat{\theta}_2, \tag{31}$$

and the posterior covariance matrix of λ is

$$\begin{aligned} \tilde{V}_\lambda \; = \; & \left(\frac{S}{S-K-2}\right) \operatorname{tr}\left((Z'Z)^{-1}\begin{bmatrix} 1 & \hat{\theta}_2' \\ \hat{\theta}_2 & \left(\frac{1}{L-K-K_L-2}\right)\hat{G}_{22} + \hat{\theta}_2\hat{\theta}_2' \end{bmatrix}\right) \cdot \hat{\Sigma} \\ & + \left(\frac{1}{L-K-K_L-2}\right)\hat{H}_2\hat{G}_{22}\hat{H}_2', \end{aligned} \tag{32}$$

where "tr" denotes the trace operator. When $S = L$, $\tilde{\lambda}$ in (31) simplifies to the vector of sample means of the factors over the S periods. That is, the more common estimate of the factor premia arises as a special case of our estimate when no longer-history asset returns are included in the estimation.

D.3. Cost of Equity

Recall that the stock's expected excess return is given by

$$\mu = \alpha + \lambda'\beta = [1 \;\; \lambda']b. \tag{33}$$

Once we have obtained the posterior first and second moments of λ and b, it is straightforward to compute the first and second moments of μ, since the posterior distributions of those parameters are independent. As noted at the outset, the decision maker's estimate of the expected excess return is the posterior mean of μ, which, given the independence of λ and β, is simply

$$\mathrm{E}(\mu|r, F^{(S)}, Y^{(L)}) = \tilde{\alpha} + \tilde{\lambda}'\tilde{\beta}, \tag{34}$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ denote posterior means of α and β. The posterior variance of μ is given by

$$\mathrm{Var}(\mu|r, F^{(S)}, Y^{(L)}) = \operatorname{tr}\left(\tilde{V}_b\begin{bmatrix} 1 & \tilde{\lambda}' \\ \tilde{\lambda} & \tilde{V}_\lambda + \tilde{\lambda}\tilde{\lambda}' \end{bmatrix}\right) + \tilde{\beta}'\tilde{V}_\lambda\tilde{\beta}, \tag{35}$$

where $\tilde{V}_b$ and $\tilde{V}_\beta$ denote the posterior covariance matrices of b and β.[14]

In the empirical results presented in the next section we compute the posterior variance of μ and its components, α and $\beta'\lambda$. For the latter quantity we report the unconditional variance as well as variances that condition on either β or λ set equal to their posterior means. The conditional variances provide additional insight into the sources of uncertainty about the cost of equity. These variances of $\beta'\lambda$ are computed as

$$\text{Var}(\beta'\lambda|r, F^{(S)}, Y^{(L)}) = \text{tr}\left(\tilde{V}_\beta\left[\tilde{V}_\lambda + \tilde{\lambda}'\tilde{\lambda}\right]\right) + \tilde{\beta}'\tilde{V}_\lambda\tilde{\beta}, \tag{36}$$

$$\text{Var}(\beta'\lambda|\lambda = \tilde{\lambda}, r, F^{(S)}, Y^{(L)}) = \tilde{\lambda}'\tilde{V}_\beta\tilde{\lambda}, \tag{37}$$

and

$$\text{Var}(\beta'\lambda|\beta = \tilde{\beta}, r, F^{(S)}, Y^{(L)}) = \tilde{\beta}'\tilde{V}_\lambda\tilde{\beta}. \tag{38}$$

II. Cost-of-Equity Estimates: Posterior Moments

A. Results for an Individual Stock

We first compute the moments of the posterior distribution for the expected excess return and its various components for the stock of a specific firm, Bay State Gas Company. One reason for this choice is that, across the three models, the OLS estimates of b for this company generally differ substantially from the cross-sectional averages, so the shrinkage effects discussed earlier can be illustrated. If we were instead to select a typical stock, $\hat{b}$ would be close to the cross-sectional average used to specify the prior mean, so any shrinkage effects would be minimal. In addition, selecting a utility allows us to compare results based on the first all-stock prior to those based on the second utility-specific prior.

As explained in the previous section, given the form of the likelihood and the assumed prior independence between the regression parameters (b and σ) and the factor means (λ), the posterior moments of the regression parameters depend only on the data used in the regression model. The monthly history of Bay State Gas begins in December 1974, so in this case, the regression-model data consist of monthly returns on the stock and the factors for the 253 months in the period from December 1974 through December 1995. For Bay State Gas, the Metropolis-Hastings algorithm is used to compute the posterior means and standard deviations of the regression parameters, as described in the previous section. Panel A of Table II reports the posterior means and standard deviations of the CAPM α and β. These posterior moments are reported for seven values of σ_α, the prior standard deviation of α that characterizes a decision maker's mispricing uncertainty about the given model.

A dogmatic belief in the ability of the model to deliver precisely the expected excess return is characterized by σ_α equal to zero. As σ_α moves from zero to infinity, the decision maker's confidence in the pricing model's ability declines, so greater weight is placed on the regression estimate $\hat{\alpha}$: the posterior mean of Bay State Gas's α moves from 0 to 7.66% (annualized). The latter value is close to Bay State's $\hat{\alpha}$ estimate of 7.92%—the small difference arises from correlation in the posterior between α and β.[15] Observe, however, that the posterior mean of α moves away from zero rather slowly. For example, the posterior mean of α is only 11 basis points (bp) above zero at $\sigma_\alpha = 3\%$ and only 73 bp above zero at $\sigma_\alpha = 5\%$. This slow movement away from the prior mean for α occurs in spite of the fact that the t-statistic associated with the Bay State's $\hat{\alpha}$ is equal to 2.07. (This case supplies the example discussed in the introduction.) In other words, even with substantial skepticism about the ability of the CAPM to capture precisely the expected excess return on any given stock, and even with a historical average return that departs substantially from the CAPM prediction, the posterior mean of the stock's excess return is still fairly close to the CAPM implied value.

In most cases, as with Bay State Gas, the posterior mean of α is shrunk away from the sample estimate $\hat{\alpha}$ and toward the prior mean $\bar{\alpha} = 0$. That is, the posterior mean for the expected excess return μ is shrunk away from the stock's sample average excess return and toward the value implied by the factor-based pricing model. The matrix expressions in (25) and (26) do not immediately reveal the weight given to $\hat{\alpha}$ in computing the posterior mean of α. An approximation that reveals the rough order of magnitude of the shrinkage effect for α is obtained by setting the prior correlations between α and β and the sample means of the factors equal to zero. In that case, (25) and (26) imply that the posterior mean of α is given by

$$\tilde{\alpha} = \bar{w}_\alpha \bar{\alpha} + (1 - \bar{w}_\alpha)\hat{\alpha}, \tag{39}$$

where

$$\bar{w}_\alpha = \frac{\mathrm{E}(\sigma^2)}{\mathrm{E}(\sigma^2) + T\sigma_\alpha^2}. \tag{40}$$

Recall that $\mathrm{E}(\sigma^2)$ is the prior mean of σ^2, which in this case is equal to 0.016 on a monthly basis, using (14) and the CAPM values in Panel A of Table I. If $\sigma_\alpha = 5\%$ on an annualized basis, then the corresponding monthly value of σ_α used in the calculations is 0.0042 (= .05/12). For $T = 253$, as with Bay State Gas, equation (40) implies $\bar{w}_\alpha = 0.78$. That is, even with mispricing uncertainty of 5% per annum, and a value of T that is fairly large compared to those often used in practice, the prior mean of α is still given heavy weight in computing the posterior mean of α. Of course, as σ_α becomes large, we see from (40) that the sample estimate $\hat{\alpha}$ is given increasingly greater weight, as illustrated by the results in Table II. Alternatively, $\hat{\alpha}$ would also be given greater weight if the prior mean $\mathrm{E}(\sigma^2)$ were lower. Such might be the case, for example, if one were to estimate the expected excess return on an asset known a priori to possess lower residual variance, such as a diversified portfolio of stocks. For the typical individual stock, however, as will be further demonstrated in the next subsection, $\hat{\alpha}$ is given heavy weight only when the prior mispricing uncertainty is very high.

The prior mean $\bar{\alpha}$ is set to zero in this study, but, as noted earlier, one could instead set $\bar{\alpha}$ to a non-zero value, possibly the average sample $\hat{\alpha}$ for a cross-section of stocks that share similar characteristics with the given stock. The degree of shrinkage toward that prior mean would, for a given σ_α, be otherwise similar to that demonstrated here.

Equations (25) and (26) essentially imply that the shrinkage weights applied to $\hat{\beta}$ and $\bar{\beta}$ in determining the posterior mean of β depend on the prior precision about β as compared to the sample precision of the regression estimates, and the latter increases in the sample length T. For Bay State Gas, with $T = 253$, the posterior mean of β is much closer to the least-squares estimate of 0.42 than the prior mean of 1.12, and the posterior mean moves only slightly, from 0.47 to 0.45, as σ_α goes from zero to infinity. For smaller values of T, the posterior mean of β is shrunk more toward the prior mean. The slight dependence on σ_α arises from correlation in the posterior between α and β, both through $\rho_{\alpha\beta}$ in the prior and through the off-diagonal elements in the first row of $X'X$.

The expected excess return has α as one of its components. Panel B of Table II reports posterior moments for the other component, $\beta'\lambda$, and the overall expected excess return, μ. Recall that information about λ is contained not only in the available histories of returns on the factors but also in the longer histories of other series that are correlated with the factors. The first part of Panel B reports posterior moments based on the longer period from January 1926 through December 1995, whereas the second part reports moments based on the shorter period beginning in July 1963. The posterior mean of λ, $\tilde{\lambda}$, is 8.05% based on the longer period but only 5.52% based on the shorter period. This difference reflects the fact that the average return on the NYSE portfolios is higher over the 1926–95 period than during the shorter 1963–95 period. Given the high positive correlations between the NYSE indexes and the Fama-French NYSE-AMEX-Nasdaq index, the posterior mean of the latter index is adjusted upward (see Stambaugh (1997)). This adjustment produces a cost-of-equity estimate for Bay State Gas that is above the shorter-period estimate by about 1.2%. For the overall period, the posterior mean of the expected excess return μ is about 3.8% based on a strict CAPM ($\sigma_\alpha = 0$) and, given the behavior of the posterior mean of α discussed above, the posterior mean of μ remains between 3% and 4% for values of σ_α smaller than 5%. That is, prior uncertainty about Bay State Gas's CAPM mispricing (α) that seems substantial in economic terms still results in a posterior mean fairly close to the CAPM value. As will be demonstrated below, this observation generalizes across stocks and across the three pricing models considered.

The results based on the shorter period, in the second part of Panel B of Table II, ignore the longer-history asset returns in the estimation of the factor premiums. As noted earlier, in such a case $\tilde{\lambda}$ is simply equal to $\bar{\lambda}$, the vector of sample averages of the factors. Also, due to the relatively large T for Bay State Gas, the posterior mean of β is very close to $\hat{\beta} = 0.42$. The posterior mean of α ranges from 0 to 7.66%, and the latter value is close to $\hat{\alpha}$, which equals 7.92%. As a result, in the extreme cases when σ_α equals zero and infinity, our estimates of the expected excess return in this shorter period are close to alternative textbook-recommended estimates (e.g., Benninga and Sarig (1997)). For $\sigma_\alpha = 0$, our estimate of 2.59% is close to the simpler CAPM-based estimate $\hat{\beta}\bar{\lambda}$, which equals 2.32%. For $\sigma_\alpha = \infty$, our estimate of 10.12% is close to the sample mean of 11.61% for the excess returns on Bay State Gas's stock, and the corresponding posterior standard deviation of 4.07% is close to the frequentist estimate of 4.01% for the standard error of the sample mean. (The difference between 10.12 and 11.61 arises primarily from the fact that the average market premium for the 1963–95 period, used in computing the posterior mean of μ, is slightly less than the average market premium over the somewhat shorter period for the returns on Bay State Gas.) The close correspondence between our estimates in the two extreme cases and the two common alternative estimates is also observed for both of the multifactor models.

Posterior standard deviations of μ, α, and $\beta\lambda$, reported in Table II, summarize the uncertainty about Bay State Gas's expected excess return and its components. The values reported for $\beta\lambda$ include both the unconditional standard deviation as well as standard deviations that condition on either β or λ set equal to their posterior means, $\tilde{\beta}$ and $\tilde{\lambda}$. (The calculations rely on equations (35) through (38) in the previous section.) Based on the 1926–95 period, the posterior standard deviation of Bay State Gas's (annualized) expected excess return ranges from 1.25%, in the case of a dogmatic belief in the CAPM ($\sigma_\alpha = 0$), to 4.02%, in the case of a diffuse prior about deviations from the model ($\sigma_\alpha = \infty$). The first value is essentially the posterior standard deviation of $\beta\lambda$, which is largely unaffected by σ_α. Further discussion of posterior standard deviations is deferred to the later analysis of cross-sectional averages.

Tables III and IV report posterior moments for the components of Bay State Gas's expected excess return under the Fama-French (FF) model and the Connor-Korajczyk (CK) model. In general, the observations made above for the CAPM apply to these three-factor models as well. In particular, Bay State Gas's $\hat{\alpha}$ is 5.04% in the FF model and 7.08% in the CK model, but, even with σ_α as large as 5%, the posterior means for α are only 0.85% and 1.90% in the two models. Also, the information about λ contained in the longer histories of the additional assets has a substantial effect on the estimated cost of equity. For both of the three-factor models, the expected excess return for Bay State Gas based on the longer 1926–95 period is higher than that based on the shorter 1963–95 period by about 1.5% for the FF model and 1.2% for the CK model. When mispricing uncertainty associated with each model is modest, the estimates of expected excess return for Bay State Gas differ substantially across the three models. The CAPM implies the lowest estimates, which often lie below the estimates from the three-factor models by 2% or more. The FF estimates exceed the CK estimates by more than 1% at the lowest values of σ_α, but, at $\sigma_\alpha = 5\%$ the differences between those models are less than 50 basis points. In Section III, we analyze the potential uncertainty about the cost of equity induced by such differenes across models, and we compare that component of uncertainty to the component that arises from uncertainty about the parameters within a given model.

B. Results for a Broad Cross-Section

For each stock on the NYSE and AMEX having at least 60 months of data continuing through December 1995, we compute the same posterior moments reported for Bay State Gas in Tables II–IV using the stock's available monthly history back through July 1963. Each value in Tables V–VII is the arithmetic average across those 1,994 stocks of the corresponding value reported in Tables II–IV. As explained earlier, computing the posterior moments for each of these stocks using the Metropolis-Hastings algorithm would be computationally prohibitive. Instead, in constructing Tables V–VII we use the approximations to the first and second posterior moments of b discussed in the previous section. The approximations appear to work well. For example, when the values in Tables II–IV are recomputed using the approximations, none of the posterior means and standard deviations change by more than 2 basis points.

Unless stated otherwise, our discussion in this subsection is confined to results obtained using the longer 1926–95 period. The FF and CK models yield posterior means of μ for the typical (average) stock in the range of 11% to 12%, roughly 3% higher than the corresponding mean under the CAPM. For the FF model, this difference relative to the CAPM is due largely to the second and third factors, since the average posterior means of the market betas are similar for the two models (1.01 versus 0.97). The average posterior means of the betas on

SMB_t and HML_t are 0.68 and 0.32, which indicates that the average firm in the cross-section is tilted toward smaller capitalization and higher book-to-market. When combined with the posterior means for SMB_t and HML_t of 3.6% and 5.3%, those betas account for the bulk of the difference between the CAPM and FF expected excess returns for the average stock. The difference between the CAPM and the CK model is more difficult to describe, given that the factors are less easily identified.

The average posterior standard deviations in Tables V–VII reveal various aspects of uncertainty about the cost of equity for a typical individual stock. An exact version of a pricing model, where $\alpha = 0$, implies an expected excess return equal to $\beta'\lambda$, and that quantity's average posterior standard deviation is largely unaffected by the prior uncertainty about α. The average posterior standard deviation of $\beta'\lambda$ is about 2.8% for the CAPM and 4.1% for the FF and CK models. These values reflect the uncertainty in both β and λ. For the typical stock, we see that uncertainty about β alone contributes substantially to the overall uncertainty about the cost of equity for an individual stock. Specifically, the average conditional standard deviation of $\beta'\lambda$ given $\lambda = \tilde{\lambda}$ is about 1.3% for the CAPM, 2.5% for the FF model, and 2.2% for the CK model. On average, uncertainty about β is less important than uncertainty about λ, but not dramatically so: the average conditional standard deviation of $\beta'\lambda$ given $\beta = \tilde{\beta}$ is about 2.4% for the CAPM and 3.1% for the FF and CK models. Note also from these conditional standard deviations that the higher unconditional posterior standard deviations of $\beta'\lambda$ in the three-factor models, as compared to the CAPM, reflect additional uncertainty about both β and λ.

In all three models, the posterior means of λ are affected substantially by augmenting the factor histories, which begin in July 1963, with the longer histories of additional series that begin in 1926. These effects on posterior means indicate an important reliance on the information in the longer histories of the additional variables, but the posterior standard deviations of $\beta'\lambda$ for the longer period are generally of about the same magnitude, or even slightly larger, than the posterior standard deviations for the shorter period. This outcome might seem puzzling, but the comparison of posterior standard deviations does not really provide a sensible measure of the additional information provided by the longer histories. The reason is that the longer histories can also provide additional information about uncertainty. In particular, since the sample volatility of the long-history series is higher prior to 1963 than after, the posterior beliefs about the factors' variances center on higher values when based on the overall period. This increase in posterior variance of the factors, ceteris paribus, raises the posterior variance of λ, the vector of factor means. In effect, more information can reveal greater volatility, and thus greater uncertainty, than otherwise perceived. That effect then works in opposition to the more obvious one (also present): longer histories provide more information about factor means and, ceteris paribus, lower their posterior variances.

When σ_α is very large, the posterior standard deviation of α is fairly close to the usual frequentist standard error for the estimated regression intercept. In that case, not surprisingly, the posterior uncertainty about α dominates the posterior uncertainty about the expected excess return. At lower values of σ_α, the posterior standard deviation of α is typically about 1/2 to 3/4 of σ_α. For example, when $\sigma_\alpha = 5\%$, the posterior standard deviation of α is just over 3% in all three models. The difference between the posterior standard deviation of μ and the posterior standard deviation of $\beta'\lambda$ arises due to uncertainty about α. In general, for values of σ_α between 3% and 5%, it seems that uncertainty about α is of roughly similar importance to uncertainty about β and λ in explaining the overall posterior uncertainty about a typical stock's expected excess return.

Recall that, for each of the three models, the estimate of the expected excess return for Bay State Gas is not very sensitive to the presence of economically plausible "pricing uncertainty," represented by σ_α. As the results in Tables II–IV demonstrate, for values of σ_α up to 5%, the posterior mean of Bay State Gas's α remains within 2% of its prior mean of zero, even though the least-squares estimate $\hat{\alpha}$, based on over 21 years of data, ranges between 5% and 8% for the three models. For the other 1,993 firms in our cross-section, the degree to which the cost of equity is sensitive to σ_α cannot be discerned from the averages reported in Tables V–VII. In order to explore this issue, we plot in Figures 1 through 3, for the three pricing models, each stock's posterior mean of μ obtained with $\sigma_\alpha = 0$ versus the stock's posterior mean of μ obtained with a non-zero value of σ_α. The latter value of σ_α is, in different plots, 3%, 5%, 10%, and infinity. A stock's vertical deviation from a 45-degree line is approximately $\tilde{\alpha}$, the posterior mean of α for that stock, since the values plotted are $\tilde{\beta}'\tilde{\lambda}$ (horizontal axis) versus $\tilde{\alpha} + \tilde{\beta}'\tilde{\lambda}$ (vertical axis), and $\tilde{\beta}$ is virtually unaffected by σ_α. In all three figures, the upper-left plot reveals that, across the 1,994 stocks in the cross-section, the estimate of the expected excess return obtained with $\sigma_\alpha = 3\%$ is generally quite close to that obtained with $\sigma_\alpha = 0$. The scatter of points becomes more disperse as the non-zero value of σ_α increases, but not very quickly. Even for $\sigma_\alpha = 10\%$, the estimates of expected excess return from all three models display a clear association with those obtained using an exact pricing relation.

Note that the elements of b are assumed to be constant during the T periods for which the stock's historical returns are used in (3) and (7). In the empirical analysis reported above, we take T to be the stock's entire history, at least back through July 1963. Thus, we essentially use "long-run" betas and ignore potential fluctuations in individual-stock betas over time. Several alternative approaches could be pursued. For example, T might be restricted to at most 60 months, as is consistent with common practice. We have redone the calculations for that case and find similar results, except that, not surprisingly, the estimate of the expected excess return is affected even less by $\hat{\alpha}$. In other words, for any economically reasonable prior uncertainty about mispricing, the estimate of the expected excess return is very close to the estimate produced by zero prior uncertainty. Also, the uncertainty associated with b rises somewhat for most stocks. Although we could have just as easily reported those results, we find the longer-period results, especially those involving α, to be more interesting. Another approach that might be a fruitful direction for research would be to reformulate the Bayesian model to allow changes in b. In a frequentist setting, for example, Shanken (1990) specifies b to be a linear function of observable state variables. Fama and French (1997) implement such a procedure by letting an industry's betas depend on its size and book-to-market ratio.[16]

C. An Industry-Specific Approach: Utilities

For the 135 utilities having at least 60 months of data continuing through December 1995, we compute the posterior moments in the same manner as above, except that the prior is constructed using the cross-section of utilities instead of the all-stock cross-section (as explained in Section I.C.2). These two priors result in different estimated expected excess returns for the 135 utilities. Figure 4 plots, for each model and for $\sigma_\alpha = 3\%$ and $\sigma_\alpha = 5\%$, the estimates of expected excess returns obtained with one prior versus those obtained with the other. Although the plots exhibit strong positive associations, and the ranges of estimates are similar for both priors, it is also clear that the differences between the two priors can produce non-trivial differences in estimated costs of equity.

Compared to the averages for the broad cross-section, the average posterior means of μ for the utilities are smaller, ranging roughly from 5% to 8%. (In the interest of space, we present only a brief summary of the results corresponding to those reported in Tables V through VII.) As before, the CAPM estimates are on average the smallest, and the FF estimates are the largest. The posterior standard deviations of μ are also smaller than their counterparts in the broad cross-section, by a factor of roughly two. This lower uncertainty about the expected excess return for utilities is due both to lower average betas and to lower posterior standard deviations of the betas. For example, the average posterior mean of the CAPM betas for utilities is only about 0.57, which is less than the average of 1.01 for the broad cross-section, and the average posterior standard deviation of the CAPM betas is only

0.07, which is less than half the corresponding value of 0.16 for the broad cross-section.

The uncertainty about λ is more important than the uncertainty about β and, not surprisingly, this difference is more pronounced for utilities than for a typical stock from the broad cross-section. Again with the CAPM as an example, the average conditional standard deviation of $\beta'\lambda$ given $\lambda = \tilde{\lambda}$ is about 0.57%, whereas the average conditional standard deviation of $\beta'\lambda$ given $\beta = \tilde{\beta}$ is about 1.35%. The lower beta-related posterior uncertainty for utilities also arises in small part from the utility-specific prior. Recall from Table I that the prior standard deviations for the betas are lower for the utility-specific prior than for the all-stock prior. As compared to the all-stock prior, those lower prior standard deviations produce lower posterior standard deviations as well as greater shrinkage of the posterior means of the betas toward their prior means. Both effects are modest, however. With the CAPM, for example, the beta-related uncertainty averaged across the 135 utilities is 0.63% based on the all-stock prior versus 0.57% based on the utility-specific prior. As reported earlier, the all-stock prior mean for the CAPM beta is 1.12, the OLS estimate of Bay State Gas's CAPM beta is 0.42, and the posterior mean of its beta lies between 0.45 and 0.47 based on the all-stock prior (depending on σ_α). The posterior means for Bay State's betas based on the utility-specific prior are nearly identical, 0.44 to 0.47, although these values represent a greater degree of shrinkage toward the prior mean of 0.64. Note that simply using the latter utility-average beta in estimating Bay State Gas's cost of equity places too little weight on that stock's sample beta. Of course, this result also depends on the relatively long 253-month sample period used here for Bay State Gas. For shorter sample periods, the shrinkage toward the industry average beta is stronger.

Figure 5 displays six plots corresponding to those displayed in Figures 1–3, where the non-zero values of σ_α are set equal to 3% and 5% (results for σ_α equal to 10% and infinity are not shown). That is, for all three models, each utility stock's expected excess return estimated with $\sigma_\alpha = 0$ is plotted against its expected excess return estimated with $\sigma_\alpha = 3\%$ or $\sigma_\alpha = 5\%$. As before, the plots exhibit clear positive associations. In Figure 5, the deviations from a 45-degree line for a given σ_α are of roughly the same magnitude as those in Figures 1–3. That is, the posterior means of α deviate from zero by similar amounts. This result combines two offsetting effects: the absolute values of $\hat{\alpha}$ tend to be somewhat lower for utilities, but those $\hat{\alpha}$ values receive relatively more weight in computing the posterior means. The latter effect arises from the utility-specific prior, wherein the prior mean for σ^2 is lower than that for the all-stock prior—roughly 0.0045 versus 0.015 (using the values for ν and s_0^2 in Table I and equation (14)). As implied by the approximation in (39) and (40), a lower value for $E(\sigma^2)$ results in greater weight placed on $\hat{\alpha}$ relative to $\bar{\alpha}$. The individual stock analyzed previously, Bay State Gas, belongs to the sample of utilities. Recall that its $\hat{\alpha}$ values across the three models are quite high—between 5% and 8% per annum. With the utility-specific prior, the posterior mean of Bay State Gas's α with $\sigma_\alpha = 5\%$ ranges between 2.4% and 3.8%, as compared to the range of 0.7% to 1.9% obtained using the all-stock prior (in Tables II–IV). Thus, when the decision maker's prior incorporates the belief that the stock of interest has a lower residual variance than the typical stock, due to the firm's industry classification or other characteristics, then the historical average return is given heavier weight in estimating that firm's cost of equity.

III. Model Uncertainty

Recall from Tables II through IV that estimates of the expected excess return on the stock of Bay State Gas differ by 2% or more across the three factor-based pricing models. In their analysis of industries, Fama and French (1997) find that the CAPM produces estimated industry costs of equity that can differ from those produced by the FF model by 2% or more. Such differences across models create additional uncertainty about the cost of equity for a decision maker who remains uncertain about which model to use. As a first step in exploring the potential importance of differences across models in costs of equity for individual firms, we simply plot the estimate of the expected excess return (posterior mean of μ) obtained using one model versus that obtained using another model. Figure 6 plots the estimated expected excess returns from the CAPM versus those from the FF model for the previously analyzed cross-section of 1,994 stocks and the all-stock prior. Figure 7 plots the CAPM estimates versus the CK estimates, and Figure 8 plots the FF estimates versus the CK estimates. Each figure contains four plots, produced with σ_α equal to 0, 5%, 10%, and infinity. In general, the plots reveal positive correlation between expected excess returns estimated using different models, although the degree of correlation depends on σ_α as well as the pair of models being compared. The plots in Figure 7, for the CAPM versus the CK model, exhibit the highest correlation, but even those plots exhibit more dispersion than any of the top two plots in Figures 1 through 3. That is, the disagreement across models in estimates of expected excess returns appears to be greater than the disagreement within a given model produced by changing the degree of prior mispricing uncertainty (σ_α) from 0 to 5%. Also recall that, as σ_α increases, the estimated expected excess returns from all three models generally move closer to the stock's historical average excess return. As a result, the closest agreement across models is observed for the plots in which $\sigma_\alpha = \infty$. The agreement in those plots is not perfect, however, due largely to the fact that the sample period used to estimate the factor premiums is longer than the period used to estimate the betas. Thus, the estimated expected excess return still differs from the historical average, and that difference varies across the pricing models.

The disagreements among models can be quantified further by associating a probability with each model and then computing the variance of a given stock's μ associated with model uncertainty. (Part B of the Appendix provides a more formal treatment.) For each model, the prior and posterior distributions of the parameters in the model are conditioned on that model's being the appropriate one. If there are Q models under consideration, $q = 1, \ldots, Q$, let $\tilde{\mu}_{[q]}$ denote the posterior mean of μ obtained under model q, and let π_q denote the decision maker's posterior probability associated with model q. Then, taking the unconditional posterior mean across models, the decision maker ultimately estimates the expected excess return to be

$$\mu^* = \sum_{q=1}^{Q} \pi_q \tilde{\mu}_{[q]}. \tag{41}$$

Combining estimates across models occurs in practice. For example, the New York State Public Service Commission has endorsed the use of equal weights across three different models to estimate the cost of equity for public utilities under its supervision. The three models used by the Commission are the CAPM (more precisely, an average of four CAPM-based estimates) and two non-factor-based models—the "Discounted Cash Flows" model and the "Comparable Earnings" model. The commission has also considered the inclusion of multifactor models in estimating costs of equity for public utilities (DiValentino (1994)).

Let $\tilde{v}_{\mu[q]}$ denote the posterior variance of μ obtained under model q. When estimating the expected excess return, the decision maker is left with overall uncertainty given by the unconditional posterior variance across models:

$$v^*_\mu = \sum_{q=1}^{Q} \pi_q \tilde{v}_{\mu[q]} + \sum_{q=1}^{Q} \pi_q (\tilde{\mu}_{[q]} - \mu^*)^2. \tag{42}$$

The first term on the right-hand side of (42), the expected value across models of the posterior variance of μ, is essentially the average within-model uncertainty about the expected excess return. This component of the overall uncertainty is analyzed in the previous section. The second term on the right-hand side of (42), the variance across models of the posterior mean of μ, can be termed "model" uncertainty, or the component of the overall posterior variance of μ attributable to uncertainty about which model to use.

Calculation of posterior model probabilities (π_q's) is beyond the intended scope of this study. As noted at the outset, we focus more on issues related to using various factor-based models for cost-of-equity estimation rather than on issues related to testing such models or evaluating their relative merits. In order to illustrate the calculation of model uncertainty, we consider various sets of candidate models and, for each set, the π_q's are made equal across models. When only two models are entertained, model uncertainty is bounded above by the value we report with equal π_q's. With three models, the greatest of those bounds for the three possible two-model combinations is the upper bound on model uncertainty for the three-model combination. Although assigning equal probabilities across the three models generally results in a value somewhat less than that upper bound, that simple specification still provides a fairly generous assessment of model uncertainty that is useful in revealing its potential importance relative to the within-model parameter uncertainty discussed previously.

We analyze the effects of model uncertainty using a range of values for σ_α, the prior within-model mispricing uncertainty, but each value of σ_α is held constant across models in order to limit the analysis to a manageable number of cases. Many more cases are possible, of course, since a decision maker's prior uncertainty about α can differ across models. When σ_α is (nearly) zero, so that the decision maker essentially believes a priori that a given model prices stocks without error, it seems unreasonable that the same decision maker would still assign non-zero probabilities to other models. Although a decision maker might know that *one* of the models is exactly correct—just not *which* one—such a scenario seems unlikely. In general, uncertainty about which model to use would be accompanied by uncertainty about whether any one model prices all stocks accurately. Since estimates of expected excess return tend to differ less across models as σ_α increases, as can be seen in Figures 6–8, the values of model uncertainty obtained with equal model probabilities but $\sigma_\alpha = 0$ in each model will, for most stocks, tend to overstate the model uncertainty that would be encountered in practice.

Table VIII reports the model uncertainty about μ as well as the amount of overall uncertainty, which includes the within-model parameter uncertainty. Calculations are reported for the various two-model subsets as well as for the set of all three models. The results are based on the longer 1926–95 period and are computed for the same alternative values of σ_α used in Section II. All values are reported as annualized percentage standard deviations. Also shown, for comparison, are (square roots of) the expected values across the three models of the posterior variances of μ, α, and $\beta'\lambda$. Panel A of Table VIII displays results for Bay State Gas, the individual stock examined previously. Recall from Tables II through IV that, when $\sigma_\alpha = 0$, the estimate for the expected excess return on Bay State's equity is lowest for the CAPM (3.77%) and highest for the FF model (6.94%). The model uncertainty for that pairing of models is 1.58% (annualized standard deviation), which is the highest value among those for the two-model sets in Panel A of Table VIII. The model uncertainty associated with the three-model set is 1.29%, which is less than the average within-model uncertainty of 1.51%. As σ_α grows large, within-model uncertainty increases, since it includes uncertainty about α, and model uncertainty typically declines (although the latter effect is somewhat non-monotonic for Bay State). Thus, in terms of contributions to the overall uncertainty about Bay State's cost of equity, uncertainty about the values of the parameters within a given model is greater than uncertainty about which model to use.

Panel B of Table VIII reports the averages across the 1,994 stocks of each value in Panel A. For the typical stock, both model uncertainty and overall uncertainty are higher than for Bay State Gas, a utility. Otherwise, the conclusions are similar. In particular, even with $\sigma_\alpha = 0$, the average model uncertainty is less than the average within-model parameter uncertainty: 2.26% versus 3.71%. The average overall uncertainty about μ in that case is 4.40%, only 0.69% higher than the average within-model uncertainty. As σ_α grows large, the average model uncertainty decreases and the average within-model uncertainty increases. In general, although model uncertainty is substantial, it appears to be less than the within-model parameter uncertainty in estimating costs of equity for individual firms using the factor-based models entertained here.

We conduct a similar analysis for the utilities industry. Figure 9 displays the plots corresponding to those in Figures 6–8 for σ_α set to 3% and 5%, the plots corresponding to those in Figures 6–8. That is, each utility's expected excess returns estimated using two different models are plotted against each other, where the utility-specific prior is used instead of the all-stock prior. The associations between the estimates obtained from different models appear to be stronger than those observed in Figures 6–8 for the whole cross-section of stocks. All three models typically produce rather similar estimates, and the fit between the estimates from the CAPM and the three-factor CK model is especially close. Note that, contrary to the observation for the whole cross-section, the cross-model plots in Figure 9 are less disperse than the within-model plots in Figure 5. In other words, the disagreements across models in utilities' estimated expected excess returns appear to be smaller than the disagreements within a given model produced by changing the degree of prior mispricing uncertainty (σ_α) from 0 to 5%.

Table IX is the equivalent of Table VIII, except that it is constructed for the utilities industry and based on the utility-specific prior. The results for Bay State Gas in Panel A are quite similar to those obtained with the all-stock prior. In Panel B, which reports averages across the 135 utility stocks, both model uncertainty and overall uncertainty about the expected excess return are smaller for utilities than for the whole cross-section. In particular, even with $\sigma_\alpha = 0$, the model uncertainty for the pairing of CAPM and CK is only 0.74%, which is consistent with the close correspondence between the estimates of expected excess returns from those models displayed in Figure 9. Despite some differences in magnitude, the relative proportions of model uncertainty and overall uncertainty are similar to those observed in Table VIII. Thus, in the utilities industry, uncertainty about which model to use again appears to be less important than uncertainty about the parameters within a given model.

IV. Conclusions

Costs of equity capital implied by factor-based pricing models can be estimated in a Bayesian setting. After using the available data, a decision maker possesses uncertainty about a firm's cost of equity that is characterized by the posterior standard deviation of μ, the expected excess return on the firm's stock. The posterior standard deviation of μ is typically at least 3% per year in a one-factor model and 4% per year in a three-factor model, even if the possibility that the model might misprice the stock is ruled out a priori. For utilities, this standard deviation is smaller but generally at least 2% per year. Uncertainty about a pricing model's potential mispricing of the stock (α) increases the uncertainty about μ, but the posterior mean of μ—the decision maker's estimate of the expected excess return—is generally not affected greatly by uncertainty about α. When a decision maker is uncertain about which factor-based model to use, the estimate of the stock's expected excess return is then a weighted average of estimates from different models. The model uncertainty associated with that estimate is nontrivial, typically adding another 0.7% to the overall posterior standard deviation of μ, but the model uncertainty on average is less than the within-model parameter uncertainty.

The framework introduced here allows a decision maker to adjust a stock's estimated expected excess return away from the value implied by a pricing model and toward the historical average excess return on the firm's stock (since the posterior mean of α is adjusted away from zero and toward the OLS intercept $\hat{\alpha}$). That is, instead of either taking the strict implication of a pricing model or completely abandoning the model in favor of the simpler historical average return, the decision maker can combine those estimates. The weight on the historical average essentially depends, for a given sample length, on the decision maker's prior uncertainty about the model's mispricing (σ_α) and his prior expectation of the stock's residual variance ($E(\sigma^2)$). If the stock is a priori judged likely to possess average residual variance, and if the prior mispricing uncertainty is, say, less than 5% per annum, then the weight on the stock's historical average return is low. In such a case, even if the mispricing uncertainty seems substantial in economic terms, the traditional use of the pricing model—taking its exact implication as the cost-of-equity estimate—generally yields a reasonably close approximation to the posterior mean. Of course, that simpler estimate does differ somewhat from the posterior mean, and the latter can be computed using our methodology.

There are scenarios in which even those who favor simpler methods might be advised to estimate the cost of equity using our Bayesian approach. Adjusting the cost-of-equity estimate toward the stock's historical average return becomes more important when one believes a priori that the stock's residual variance is lower than that of the typical stock. The weight on the historical average return is generally decreasing in $E(\sigma^2)$. Therefore, when the stock's characteristics lead one to assign a lower prior mean for σ^2, the cost-of-equity estimate should place less weight on the traditional pricing-model estimate and more weight on the average return. Such a scenario is illustrated in this study for the case of utility stocks. In such a scenario, if it happens that the historical average return for the stock of interest is far from the pricing model's prediction, and if the sample generating that estimate is fairly long, then the information contained in the stock's historical average excess return should probably be incorporated, even for modest values of σ_α. Our framework provides a method for doing so.

As noted at the outset, the Bayesian approach provides a coherent framework for permitting a decision maker's judgment, expressed as prior beliefs, to enter the cost-of-equity estimation. A key feature of those prior beliefs explored in this study is the degree of mispricing uncertainty (σ_α). We set the prior mean of the pricing error ($\bar{\alpha}$) equal to zero, but that specification could be relaxed, as discussed previously (Section I). In particular, the prior mean for the pricing error could depend on one or more characteristics of the firm. The posterior mean for α is then adjusted away from that non-zero prior mean and toward $\hat{\alpha}$, and the degree of that adjustment would depend on the prior parameters σ_α and $E(\sigma^2)$ in essentially the same manner as discussed previously.

Given the imprecision associated with estimates of factor premiums, found here and in previous studies, it seems essential that those quantities be estimated using as much information as possible. Our methodology allows that information to include series whose histories are longer than those of the factors—over twice as long in this study. We find that the additional information in those series produces posterior means for the factors, and thus for

μ, that differ substantially from those based on the factor histories alone. We also find that, even after incorporating the additional information in series beginning in 1926, uncertainty about factor premiums still makes the largest contribution to overall uncertainty about the expected excess return (in the absence of uncertainty about α), although uncertainty about betas is nearly as important for the typical individual stock. The priors for the factor premiums are specified in this study as diffuse (non-informative). One might instead be able to construct reasonable informative prior beliefs about one or more of the factor premiums, and the posterior uncertainty about a stock's expected excess return would then no doubt be less than we report. Alternatively, introducing additional historical data, possibly within a different stochastic setting, might also prove helpful. In general, the uncertainty about factor premiums present in cost-of-equity estimation offers payoffs to future research.

Appendix

A. Posterior Moments of Factor Premiums

This section extends results in Stambaugh (1997) and derives the posterior mean and variance-covariance matrix of λ in (31) and (32) when the likelihood function is given by (8) and the prior is given by (17). Recall that λ contains the first K elements of θ. Let Φ denote the data set consisting of $F^{(S)}$ and $Y^{(L)}$, the sample information about the moments of f_t^a. Define the population counterparts to the quantities in (29) and (30),

$$H_2 = G_{12}G_{22}^{-1}, \qquad \text{(A.1)}$$

$$h_1 = \lambda - H_2\theta_2, \qquad \text{(A.2)}$$

and

$$\Sigma = G_{11} - H_2 G_{22} H_2', \qquad \text{(A.3)}$$

where G_{11}, G_{12}, and G_{22} are the submatrices of G in (4) that correspond to the partitioning of $f_t^{a'} = [f_t' \; y_t']$, and let

$$H = \begin{bmatrix} h_1' \\ H_2' \end{bmatrix}. \qquad \text{(A.4)}$$

It is shown in Stambaugh (1997) that

$$p(H, \Sigma, \theta_2, G_{22}|\Phi) = p(H, \Sigma|\Phi)p(\theta_2, G_{22}|\Phi), \qquad \text{(A.5)}$$

where

$$p(H, \Sigma|\Phi) \propto |\Sigma|^{-\frac{S+K+K_L+1}{2}} \exp\left\{-\frac{1}{2}\text{tr}\left(S\hat{\Sigma} + (H-\hat{H})'Z'Z(H-\hat{H})\right)\Sigma^{-1}\right\}, \qquad \text{(A.6)}$$

and

$$p(\theta_2, G_{22}|\Phi) \propto |G_{22}|^{-\frac{L-K+K_L+1}{2}} \exp\left\{-\frac{1}{2}L \cdot \text{tr}\left(\hat{G}_{22} + (\theta_2-\hat{\theta}_2)(\theta_2-\hat{\theta}_2)'\right)G_{22}^{-1}\right\}. \qquad \text{(A.7)}$$

From (A.7), the conditional posterior of θ_2 given G_{22} is

$$p(\theta_2|G_{22}, \Phi) \propto |G_{22}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}L(\theta_2-\hat{\theta}_2)'G_{22}^{-1}(\theta_2-\hat{\theta}_2)\right\}, \qquad \text{(A.8)}$$

which is a multivariate normal density with

$$\text{E}(\theta_2|G_{22}, \Phi) = \text{E}(\theta_2|\Phi) = \hat{\theta}_2 \qquad \text{(A.9)}$$

and

$$\mathrm{Cov}(\theta_2, \theta_2' | G_{22}, \Phi) = \frac{1}{L} G_{22}. \tag{A.10}$$

From (A.7) and (A.8), the marginal posterior density of G_{22} is

$$p(G_{22}|\Phi) \propto |G_{22}|^{-\frac{L-K+K_L}{2}} \exp\left\{-\frac{1}{2} L \cdot \mathrm{tr}\,(\hat{G}_{22} G_{22}^{-1})\right\}, \tag{A.11}$$

which is an inverted Wishart density with

$$\mathrm{E}(G_{22}|\Phi) = \frac{L}{L-K-K_L-2} \hat{G}_{22}, \tag{A.12}$$

where (A.12) follows from properties of the inverted Wishart distribution. (See, for example, Anderson (1984), pp. 268–270.) Therefore, since the conditional mean in (A.9) does not involve G_{22}, the unconditional posterior covariance matrix of θ_2 is the expectation of (A.10), which, using (A.12), is

$$\mathrm{Cov}(\theta_2, \theta_2' | \Phi) = \frac{1}{L-K-K_L-2} \hat{G}_{22}. \tag{A.13}$$

Next rewrite equation (A.2) as

$$\lambda = Dc, \tag{A.14}$$

where

$$D = I_K \otimes [1\ \theta_2'], \tag{A.15}$$

$$c = \mathrm{vec}\{H\}, \tag{A.16}$$

and "vec$\{H\}$" denotes the $K \times (K_L + 1)$ column vector formed by stacking the successive columns of H. Similarly, define

$$\hat{c} = \mathrm{vec}\{\hat{H}\}. \tag{A.17}$$

From (A.6) and the analysis of the multivariate regression model in Zellner (1971, p. 227), the conditional posterior density of c given Σ can be written as

$$p(c|\Sigma, \Phi) \propto |\Sigma|^{-\frac{K_L+1}{2}} \exp\left\{-\frac{1}{2}(c-\hat{c})'(\Sigma^{-1} \otimes Z'Z)(c-\hat{c})\right\}, \tag{A.18}$$

which is a multivariate normal density with

$$\mathrm{E}(c|\Sigma, \Phi) = \hat{c} \tag{A.19}$$

and

$$\mathrm{Cov}(c, c'|\Sigma, \Phi) = \mathrm{Cov}(c, c'|\Phi) = \Sigma \otimes (Z'Z)^{-1}. \tag{A.20}$$

Because c and θ_2 are independent (cf. (A.5)), it follows immediately from (A.9) and (A.14) through (A.17) that

$$\mathrm{E}(\lambda|\theta_2, \Phi) = \hat{h}_1 + \hat{H}_2\theta_2, \tag{A.21}$$

and the unconditional posterior mean of λ, $\tilde{\lambda}$, is given by (31).

From (A.6) and (A.18), and again relying on the analysis in Zellner (1971), p. 227, the marginal posterior density of Σ is given by

$$p(\Sigma|\Phi) \propto |\Sigma|^{-\frac{S+K}{2}} \exp\left\{-\frac{1}{2}S \cdot \mathrm{tr}\,(\hat{\Sigma}\Sigma^{-1})\right\}, \tag{A.22}$$

and, using the same property of the inverted Wishart distribution as in (A.12), the unconditional posterior mean of Σ is

$$\mathrm{E}(\Sigma|\Phi) = \frac{S}{S-K-2}\hat{\Sigma}. \tag{A.23}$$

Given (A.19), the unconditional posterior covariance matrix of c is the expectation of the conditional covariance matrix in (A.20), which, using (A.23), is equal to

$$\mathrm{Cov}(c, c'|\Phi) = \frac{S}{S-K-2}\hat{\Sigma} \otimes (Z'Z)^{-1}. \tag{A.24}$$

Combining (A.14) and (A.24) gives

$$\begin{aligned} \mathrm{Cov}(\lambda, \lambda'|\theta_2, \Phi) &= \frac{S}{S-K-2} D(\hat{\Sigma} \otimes (Z'Z)^{-1})D' \\ &= \frac{S}{S-K-2}\left([1\ \theta_2']\,(Z'Z)^{-1}\begin{bmatrix} 1 \\ \theta_2 \end{bmatrix}\right)\hat{\Sigma}, \end{aligned} \tag{A.25}$$

and taking the unconditional expectation of (A.25), using (A.9) and (A.12), gives

$$\begin{aligned} &\mathrm{E}\left(\mathrm{Cov}(\lambda, \lambda'|\theta_2, \Phi)|\Phi\right) = \\ &\left(\frac{S}{S-K-2}\right) \mathrm{tr}\left((Z'Z)^{-1}\begin{bmatrix} 1 & \hat{\theta}_2' \\ \hat{\theta}_2 & \left(\frac{1}{L-K-K_L-2}\right)\hat{G}_{22} + \hat{\theta}_2\hat{\theta}_2' \end{bmatrix}\right) \cdot \hat{\Sigma}. \end{aligned} \tag{A.26}$$

Also, from (A.21) and (A.13),

$$\begin{aligned} \mathrm{Cov}\left(\mathrm{E}(\lambda|\theta_2, \Phi), \mathrm{E}(\lambda'|\theta_2, \Phi)|\Phi\right) &= \mathrm{Cov}\left(\hat{H}_2\theta_2, \theta_2'\hat{H}_2'|\Phi\right) \\ &= \hat{H}_2\mathrm{Cov}(\theta_2, \theta_2'|\Phi)\hat{H}_2' \\ &= \frac{1}{L-K-K_L-2}\hat{H}_2\hat{G}_{22}\hat{H}_2'. \end{aligned} \tag{A.27}$$

By the variance decomposition rule, the sum of the matrices in (A.26) and (A.27) gives $\tilde{V}_\lambda$, the unconditional variance-covariance matrix of λ, and that result is displayed in (32).

B. Model Uncertainty: Details

We briefly summarize here the framework underlying equations (41) and (42). Interested readers may also consult Kass and Raftery (1995) and Poirier (1995), Chapter 10, for related discussions. Let m denote a random discrete quantity that takes values $q = 1, \ldots, Q$ and serves as an index for the Q models under consideration. Let $\underline{\pi}_q \equiv p(m = q)$ denote the prior probability for model q, and let δ_q denote the vector of parameters in model q. Let D denote the observed data, which in our case consist of the stock's return history, the histories of the factors (six in total, across the three models) and the three additional longer-history series used to augment the factor histories in each model.

The posterior model probability $\pi_q \equiv p(m = q|D)$ is, from Bayes's theorem, given by

$$\pi_q = \frac{\underline{\pi}_q p(D|m = q)}{\sum_{j=1}^{Q} \underline{\pi}_j p(D|m = j)}, \tag{A.28}$$

and the marginal density of the data under model j ($j = 1, \ldots, Q$) is given by

$$p(D|m = j) = \int p(D|\delta_j, m = j) p(\delta_j|m = j) d\delta_j, \tag{A.29}$$

where $p(\delta_j|m = j)$ and $p(D|\delta_j, m = j)$ are the prior parameter density and the likelihood function for model j, respectively.

If the prior parameter density for one or more of the models is improper, then obtaining π_q can be problematic, since undefined constants appear in the numerator and/or denominator of (A.28). In our setting, even though the prior density for θ and G in (17) is improper, model probabilities can still be defined, because that improper prior density can be made identical across the three models. In our setting, the parameter vector for model q can be partitioned as $\delta_q = [\delta_{q(1)}\ \delta_{(2)}]$, where $\delta_{q(1)}$ contains the elements of b and σ and $\delta_{(2)}$ contains the elements of θ and G. The elements of $\delta_{q(1)}$ differ in number and identity across models, but $\delta_{(2)}$ can be made identical across models (as discussed below). In that case, from (16),

$$p(\delta_q|m = q) = p(\delta_{q(1)}|m = q) p(\delta_{(2)}), \quad q = 1, \ldots, Q. \tag{A.30}$$

Therefore, even though $p(\delta_{(2)})$ is defined only up to an undetermined constant, that constant appears in both the numerator and denominator of (A.28), and thus the ratio can be defined. As discussed by Kass and Raftery (1995), this treatment of ratios of improper priors for parameters that are common across models is due to Jeffreys (1961) and has been widely adopted.

The above statement that $\delta_{(2)}$ can be made identical across models requires some clarification. As θ and G are defined in Section I, they contain different elements across models. Recall that they are the mean and covariance matrix of the vector of augmented factors, $f_t^{a'} = [f_t' \ y_t']$, where f_t is an observation of the K factors for the given model and y_t is an observation of the three long-history series. Although y_t is the same across all three models, the factors differ across models. If, however, $f_t^{a'}$ is instead defined as a larger vector containing y_t and the union of all six factors across the three models, and θ and G is defined as the mean and covariance matrix of that larger vector, all of the posterior moments reported are essentially unchanged. Specifically, $\hat{h}_1$ and $\hat{H}_2$ in (31) and (32) are redefined as submatrices of larger arrays, but their values are unchanged. As a result, equation (31), which gives the posterior mean of λ, is unaffected. In equation (32), which gives the posterior covariance matrix of λ, the value of K changes to 6 (from either 1 or 3), but both S and L are large enough (390 and 840) such that any resulting changes in the reported standard deviations are trivial. Initially defining θ and G in the above fashion would complicate the presentation of the methodology, so, given that the choice of definitions is essentially irrelevant to the empirical results, we adopt the simpler definition in Section I.

In Section II, we report and analyze the first and second moments of $p(\mu|D, m = q)$, the marginal posterior density of μ for model q, where μ is a function of δ_q. With well-defined posterior model probabilities, the conditioning on model q is removed by computing the overall (unconditional) density

$$p(\mu|D) = \sum_{q=1}^{Q} \pi_q p(\mu|D, m = q), \tag{A.31}$$

and the first and second moments of this density are given by (41) and (42). (See Leamer (1978), pp. 117–118.)

In the model uncertainty calculations presented in Tables VIII and IX we simply set the posterior model probabilities to be equal across models. Computing those probabilities using (A.28) is beyond the scope of this study. Moreover, if the set of prior model probabilities ($\underline{\pi}_q$'s) is the same for each stock, then the posterior probabilities would differ across stocks. Rather than take that course, we instead specify equal posterior probabilities in order to simplify the analysis and, as discussed previously, obtain what is likely to be a generous assessment of model uncertainty. (Of course, assuming the same posterior model probabilities across stocks implies that the prior probabilities would differ across stocks.) Explicit posterior model probabilities would probably be more interesting in a multi-asset setting, and such an extension is a possible direction for future research.

Table I

Parameters Used in the Priors

In Panel A, for each stock with at least 24 months of data in the period 7/1963–12/1995, $\hat{b}$ is the ordinary-least-squares estimate of b defined by the regression

$$r_t = [1\ f_t']b + \epsilon_t,$$

where r_t is the excess return on the stock and f_t is a vector of factors. The sample variance of the residuals from that regression is $\hat{\sigma}^2$, an estimate of σ^2, the variance of ϵ_t. In Panel B, $\hat{b}$ and $\hat{\sigma}^2$ are obtained for every utility stock with at least 48 months of data in the above period. The prior mean of b, $\bar{b}$, is computed as the cross-sectional average of the $\hat{b}$'s, except that its first element, the mean of α, is set to zero. The prior standard deviations and correlations are obtained from $\hat{V}_b$, which is computed as the cross-sectional covariance matrix of the $\hat{b}$'s minus the cross-sectional average of the time-series sampling variances of the $\hat{b}$'s. The prior covariance matrix of b, V_b, is computed from $\hat{V}_b$ by varying the prior standard deviation of α (σ_α) between zero and infinity while preserving the correlation structure of $\hat{V}_b$. In Panel B, the off-diagonal elements of V_b are set equal to zero in order to have that matrix be positive definite. The prior mean of σ^2 is computed as the cross-sectional average of the $\hat{\sigma}^2$'s, and the prior variance of σ^2 is computed as the difference between the cross-sectional variance of the $\hat{\sigma}^2$'s and the cross-sectional average of the time-series sampling variances of the $\hat{\sigma}^2$'s. Given the prior mean and variance of σ^2, properties of the inverted gamma density imply the values of ν and s_0^2, which are the two parameters used to define the prior density of σ.

	Prior moments of $b(= [\alpha\ \beta']')$						Prior parameters for σ	
Model	Mean	Std. dev.	Correlations				ν	s_o^2
			Panel A. All Stocks					
CAPM	0	0–∞	1	0.26			5	0.0097
	1.122	0.424		1				
3-factor FF	0	0–∞	1	-0.15	-0.28	-0.55	5	0.0086
	1.006	0.384		1	0.26	0.20		
	0.967	0.860			1	0.45		
	0.382	0.654				1		
3-factor CK	0	0–∞	1	-0.19	-0.24	0.08	5	0.0084
	1.051	0.563		1	0.00	-0.42		
	0.017	0.750			1	-0.12		
	0.056	0.424				1		
			Panel B. Utilities					
CAPM	0	0–∞	1	0			6	0.0030
	0.641	0.228		1				
3-factor FF	0	0–∞	1	0	0	0	6	0.0028
	0.735	0.176		1	0	0		
	-0.001	0.326			1	0		
	0.420	0.213				1		
3-factor CK	0	0–∞	1	0	0	0	6	0.0029
	0.493	0.293		1	0	0		
	-0.091	0.019			1	0		
	0.420	0.213				1		

Table II

Posterior Means and Standard Deviations for the Components of Bay State Gas's Expected Excess Return from the CAPM

The expected excess return on the stock, μ, is given by $\mu = \alpha + \beta\lambda$, where λ is the expected excess market return, and α and β are parameters in the regression of the stock's monthly excess return on the excess market return:

$$r_t = \alpha + \beta r_{M,t} + \epsilon_t.$$

The moments for the parameters of the regression model, reported in Panel A, are based on monthly excess returns for the period 12/1974–12/1995 (253 months). The ordinary least-squares estimates are $\hat{\alpha} = 7.92\%$ (annualized) and $\hat{\beta} = 0.42$. The moments for the quantities involving λ, reported in Panel B, are based on monthly excess returns for the periods indicated and, for the longer period, use the additional information in the history of returns on the value-weighted and equally weighted NYSE portfolios and the Ibbotson small-stock portfolio. Also reported for each period is $\hat{\mu} = \hat{\alpha} + \hat{\beta}\tilde{\lambda}$, which is the posterior mean of μ obtained with diffuse priors on all parameters, where $\tilde{\lambda}$ denotes the posterior mean of λ. Except for the moments of β, all posterior means and standard deviations are reported as annualized percentage values.

	Prior Standard Deviation of α (σ_α)						
	0	1%	3%	5%	10%	30%	∞
Panel A. Regression parameters							
				Means			
α	0.00	-0.10	0.11	0.73	2.77	6.26	7.66
β	0.47	0.47	0.47	0.47	0.46	0.45	0.45
				Standard deviations			
α	0.00	0.38	1.10	1.72	2.73	3.72	3.94
β	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Panel B. Components involving the expected market return							
			1/1926–12/1995; $\hat{\mu} = 11.30$, $\tilde{\lambda} = 8.05$				
				Means			
μ	3.77	3.68	3.89	4.51	6.51	9.90	11.25
$\beta\lambda$	3.77	3.78	3.79	3.78	3.74	3.65	3.60
				Standard Deviations			
μ	1.25	1.32	1.67	2.11	2.96	3.83	4.02
$\beta\lambda$	1.25	1.26	1.26	1.25	1.24	1.22	1.21
$\beta\lambda \mid \lambda = \tilde{\lambda}$	0.58	0.58	0.58	0.58	0.58	0.59	0.59
$\beta\lambda \mid \beta = \tilde{\beta}$	1.10	1.10	1.10	1.10	1.09	1.06	1.04
			7/1963–12/1995; $\hat{\mu} = 10.24$, $\tilde{\lambda} = 5.52$				
				Means			
μ	2.59	2.49	2.70	3.33	5.34	8.76	10.12
$\beta\lambda$	2.59	2.59	2.60	2.59	2.56	2.50	2.47
				Standard Deviations			
μ	1.31	1.37	1.71	2.15	3.00	3.87	4.07
$\beta\lambda$	1.31	1.31	1.32	1.31	1.30	1.27	1.26
$\beta\lambda \mid \lambda = \tilde{\lambda}$	0.40	0.40	0.40	0.40	0.40	0.40	0.40
$\beta\lambda \mid \beta = \tilde{\beta}$	1.24	1.24	1.24	1.24	1.22	1.19	1.18

Table III

Posterior Means and Standard Deviations for the Components of Bay State Gas's Expected Excess Return from the Three-Factor Fama-French Model

The expected excess return on the stock, μ, is given by $\mu = \alpha + \beta'\lambda$, where λ is the vector of expected values of the three Fama-French factors ($r_{M,t}$, SMB_t, and HML_t), and α and $\beta = [\beta_1\ \beta_2\ \beta_3]'$ are parameters in the regression of the stock's monthly excess return on the factors:

$$r_t = \alpha + \beta_1 r_{M,t} + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_t.$$

The moments for the parameters of the regression model, reported in Panel A, are based on monthly data for the period 12/1974–12/1995 (253 months). The ordinary least-squares estimates are $\hat{\alpha} = 5.04\%$ (annualized) and $\hat{\beta} = [0.50\ 0.08\ 0.40]'$. The moments for the quantities involving λ, reported in Panel B, are based on monthly data for the periods indicated and, for the longer period, use the additional information in the history of returns on the value-weighted and equally weighted NYSE portfolios and the Ibbotson small-stock portfolio. Also reported for each period is $\hat{\mu} = \hat{\alpha} + \hat{\beta}'\tilde{\lambda}$, which is the posterior mean of μ obtained with diffuse priors on all parameters, where $\tilde{\lambda}$ denotes the posterior mean of λ. Except for the moments of β, all posterior means and standard deviations are reported as annualized percentage values.

	Prior Standard Deviation of α (σ_α)						
	0	1%	3%	5%	10%	30%	∞
Panel A. Regression parameters							
				Means			
α	0.00	0.05	0.35	0.85	2.25	4.47	4.97
β_1	0.54	0.53	0.53	0.53	0.52	0.52	0.52
β_2	0.10	0.10	0.10	0.10	0.09	0.09	0.09
β_3	0.43	0.42	0.42	0.41	0.40	0.39	0.39
				Standard deviations			
α	0.00	0.34	0.99	1.57	2.62	3.74	3.98
β_1	0.08	0.08	0.08	0.08	0.08	0.08	0.08
β_2	0.12	0.12	0.12	0.12	0.12	0.12	0.12
β_3	0.13	0.13	0.13	0.13	0.13	0.13	0.13
Panel B. Components involving the expected factors							
	1/1926–12/1995; $\hat{\mu} = 11.48$, $\tilde{\lambda} = [8.05\ 3.63\ 5.32]'$						
				Means			
μ	6.94	6.97	7.23	7.66	8.94	11.02	11.54
$\beta'\lambda$	6.94	6.92	6.87	6.82	6.69	6.55	6.56
				Standard Deviations			
μ	1.74	1.75	1.90	2.18	2.88	3.80	4.04
$\beta'\lambda$	1.74	1.74	1.74	1.74	1.74	1.72	1.72
$\beta'\lambda\ \lvert\lambda = \tilde{\lambda}$	1.09	1.09	1.10	1.11	1.12	1.13	1.12
$\beta'\lambda\ \lvert\beta = \tilde{\beta}$	1.33	1.32	1.32	1.31	1.29	1.27	1.27
	7/1963–12/1995; $\hat{\mu} = 10.06$, $\tilde{\lambda} = [5.52\ 3.01\ 5.05]'$						
				Means			
μ	5.41	5.44	5.71	6.15	7.44	9.55	10.07
$\beta'\lambda$	5.41	5.39	5.35	5.30	5.20	5.08	5.10
				Standard Deviations			
μ	1.69	1.69	1.86	2.16	2.88	3.83	4.07
$\beta'\lambda$	1.69	1.69	1.69	1.68	1.68	1.66	1.66
$\beta'\lambda\ \lvert\lambda = \tilde{\lambda}$	0.91	0.91	0.92	0.93	0.94	0.94	0.94
$\beta'\lambda\ \lvert\beta = \tilde{\beta}$	1.38	1.38	1.37	1.37	1.35	1.32	1.32

Table IV

Posterior Means and Standard Deviations for the Components of Bay State Gas's Expected Excess Return from the Three-Factor Connor-Korajczyk Model

The expected excess return on the stock, μ, is given by $\mu = \alpha + \beta'\lambda$, where λ is the vector of expected values of the three Connor-Korajczyk factors ($F_{1,t}$, $F_{2,t}$, and $F_{3,t}$), and α and $\beta = [\beta_1\ \beta_2\ \beta_3]'$ are parameters in the regression of the stock's monthly excess return on the factors:

$$r_t = \alpha + \beta_1 F_{1,t} + \beta_2 F_{2,t} + \beta_3 F_{3,t} + \epsilon_t.$$

The moments for the parameters of the regression model, reported in Panel A, are based on monthly data for the period 12/1974–12/1995 (253 months). The ordinary least-squares estimates are $\hat{\alpha} = 7.08\%$ (annualized) and $\hat{\beta} = [0.35\ \ -0.04\ 0.18]'$. The moments for the quantities involving λ, reported in Panel B, are based on monthly data for the periods indicated and, for the longer period, use the additional information in the history of returns on the value-weighted and equally weighted NYSE portfolios and the Ibbotson small-stock portfolio. Also reported for each period is $\hat{\mu} = \hat{\alpha} + \hat{\beta}'\tilde{\lambda}$, which is the posterior mean of μ obtained with diffuse priors on all parameters, where $\tilde{\lambda}$ denotes the posterior mean of λ. Except for the moments of β, all posterior means and standard deviations are reported as annualized percentage values.

	Prior Standard Deviation of α (σ_α)						
	0	1%	3%	5%	10%	30%	∞
Panel A. Regression parameters							
	Means						
α	0.00	0.19	0.92	1.90	4.12	6.61	6.93
β_1	0.39	0.39	0.38	0.38	0.37	0.36	0.36
β_2	-0.06	-0.06	-0.06	-0.05	-0.05	-0.05	-0.04
β_3	0.18	0.18	0.18	0.18	0.18	0.18	0.18
	Standard deviations						
α	0.00	0.40	1.16	1.80	2.83	3.77	3.96
β_1	0.07	0.07	0.07	0.07	0.07	0.07	0.07
β_2	0.06	0.06	0.06	0.06	0.06	0.06	0.06
β_3	0.06	0.06	0.06	0.06	0.06	0.06	0.06
Panel B. Components involving the expected factors							
	1/1926–12/1995; $\hat{\mu} = 12.04$, $\tilde{\lambda} = [10.85\ \ -2.22\ 5.94]'$						
	Means						
μ	5.42	5.60	6.29	7.22	9.34	11.73	12.04
$\beta'\lambda$	5.42	5.41	5.37	5.32	5.22	5.12	5.11
	Standard Deviations						
μ	1.51	1.56	1.86	2.27	3.07	3.88	4.05
$\beta'\lambda$	1.51	1.51	1.51	1.50	1.48	1.46	1.46
$\beta'\lambda\ \|\lambda = \tilde{\lambda}$	0.83	0.83	0.83	0.83	0.84	0.84	0.85
$\beta'\lambda\ \|\beta = \tilde{\beta}$	1.23	1.22	1.21	1.20	1.18	1.15	1.15
	7/1963–12/1995; $\hat{\mu} = 10.90$, $\tilde{\lambda} = [7.65\ \ -3.17\ 5.66]'$						
	Means						
μ	4.18	4.36	5.06	6.01	8.15	10.56	10.87
$\beta'\lambda$	4.18	4.17	4.14	4.11	4.03	3.95	3.94
	Standard Deviations						
μ	1.51	1.55	1.86	2.28	3.09	3.91	4.08
$\beta'\lambda$	1.51	1.50	1.50	1.49	1.47	1.45	1.45
$\beta'\lambda\ \|\lambda = \tilde{\lambda}$	0.66	0.66	0.66	0.66	0.66	0.67	0.67
$\beta'\lambda\ \|\beta = \tilde{\beta}$	1.31	1.31	1.30	1.29	1.27	1.24	1.24

Table V

Averages Across Stocks of Posterior Means and Standard Deviations for Components of Expected Excess Returns from the CAPM

Posterior means and standard deviations for components of expected excess return are computed for each of the 1,994 stocks that have data through December 1995 and have at least 60 months of historical returns, and the cross-sectional averages of those values across all stocks are reported below. The expected excess return on a stock, μ, is given by $\mu = \alpha + \beta\lambda$, where λ is the expected excess market return, and α and β are parameters in the regression of the stock's monthly excess return on the excess market return:

$$r_t = \alpha + \beta r_{M,t} + \epsilon_t.$$

The moments for the parameters of the regression model, reported in Panel A, are based on each stock's available history of monthly returns, back through July 1963. The moments for the quantities involving λ, reported in Panel B, are based on monthly excess returns for the periods indicated and, for the longer period, use the additional information in the history of returns on the value-weighted and equally weighted NYSE portfolios and the Ibbotson small-stock portfolio. Also reported for each period is $\tilde{\lambda}$, the posterior mean of λ. Except for the moments of β, all posterior means and standard deviations are reported as annualized percentage values.

	Prior Standard Deviation of α (σ_α)						
	0	1%	3%	5%	10%	30%	∞
Panel A. Regression parameters							
				Means			
α	0.00	0.00	0.11	0.29	0.71	1.24	1.51
β	1.01	1.01	1.01	1.01	1.01	1.01	1.01
				Standard deviations			
α	0.00	0.72	2.09	3.30	5.48	8.32	9.30
β	0.16	0.16	0.16	0.16	0.16	0.16	0.16
Panel B. Components involving the expected market return							
				1/1926–12/1995; $\tilde{\lambda} = 8.05$			
				Means			
μ	8.11	8.12	8.22	8.40	8.83	9.35	9.60
$\beta\lambda$	8.11	8.12	8.12	8.12	8.11	8.10	8.09
				Standard Deviations			
μ	2.78	2.93	3.60	4.45	6.23	8.76	9.61
$\beta\lambda$	2.78	2.78	2.78	2.78	2.77	2.77	2.77
$\beta\lambda \mid \lambda = \tilde{\lambda}$	1.28	1.28	1.28	1.27	1.27	1.27	1.28
$\beta\lambda \mid \beta = \tilde{\beta}$	2.36	2.36	2.36	2.36	2.36	2.35	2.35
				7/1963–12/1995; $\tilde{\lambda} = 5.52$			
				Means			
μ	5.57	5.57	5.67	5.86	6.28	6.80	7.06
$\beta\lambda$	5.57	5.57	5.57	5.57	5.57	5.56	5.55
				Standard Deviations			
μ	2.87	3.01	3.67	4.50	6.29	8.84	9.72
$\beta\lambda$	2.87	2.87	2.87	2.87	2.87	2.86	2.87
$\beta\lambda \mid \lambda = \tilde{\lambda}$	0.88	0.88	0.88	0.87	0.87	0.87	0.88
$\beta\lambda \mid \beta = \tilde{\beta}$	2.66	2.66	2.66	2.66	2.66	2.65	2.65

Table VI

Averages Across Stocks of Posterior Means and Standard Deviations for Components of Expected Excess Returns from the Three-Factor Fama-French Model

Posterior means and standard deviations for components of expected excess return are computed for each of the 1,994 stocks that have data through December 1995 and have at least 60 months of historical returns, and the cross-sectional averages of those values across all stocks are reported below. The expected excess return on a stock, μ, is given by $\mu = \alpha + \beta'\lambda$, where λ is the vector of expected values of the three Fama-French factors ($r_{M,t}$, SMB_t, and HML_t), and α and $\beta = [\beta_1\ \beta_2\ \beta_3]'$ are parameters in the regression of the stock's monthly excess return on the factors:

$$r_t = \alpha + \beta_1 r_{M,t} + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_t.$$

The moments for the parameters of the regression model, reported in Panel A, are based on each stock's available history of monthly returns, back through July 1963. The moments for the quantities involving λ, reported in Panel B, are based on monthly data for the periods indicated and, for the longer period, use the additional information in the history of returns on the value-weighted and equally weighted NYSE portfolios and the Ibbotson small-stock portfolio. Also reported for each period is $\tilde{\lambda}$, the posterior mean of λ. Except for the moments of β, all posterior means and standard deviations are reported as annualized percentage values.

	Prior Standard Deviation of α (σ_α)						
	0	1%	3%	5%	10%	30%	∞
Panel A. Regression parameters							
				Means			
α	0.00	0.03	0.09	0.16	0.29	0.42	0.35
β_1	0.97	0.97	0.97	0.97	0.97	0.97	0.97
β_2	0.68	0.68	0.68	0.68	0.68	0.68	0.68
β_3	0.32	0.32	0.31	0.31	0.31	0.31	0.32
				Standard deviations			
α	0.00	0.66	1.93	3.09	5.27	8.20	9.05
β_1	0.16	0.16	0.16	0.16	0.16	0.16	0.16
β_2	0.26	0.26	0.26	0.26	0.26	0.26	0.26
β_3	0.26	0.27	0.27	0.27	0.27	0.26	0.25
Panel B. Components involving the expected factors							
		1/1926–12/1995; $\tilde{\lambda} = [8.05\ 3.63\ 5.32]'$					
				Means			
μ	11.97	11.99	12.05	12.11	12.23	12.36	12.31
$\beta'\lambda$	11.97	11.96	11.96	11.95	11.94	11.94	11.96
				Standard Deviations			
μ	4.08	4.02	4.19	4.65	6.04	8.59	9.62
$\beta'\lambda$	4.08	4.08	4.09	4.09	4.09	4.06	4.01
$\beta'\lambda\ \|\lambda = \tilde{\lambda}$	2.45	2.46	2.47	2.48	2.48	2.43	2.37
$\beta'\lambda\ \|\beta = \tilde{\beta}$	3.05	3.05	3.05	3.05	3.05	3.05	3.05
		7/1963–12/1995; $\tilde{\lambda} = [5.52\ 3.01\ 5.05]'$					
				Means			
μ	9.00	9.03	9.09	9.15	9.27	9.40	9.35
$\beta'\lambda$	9.00	9.00	8.99	8.99	8.98	8.98	9.00
				Standard Deviations			
μ	3.94	3.89	4.10	4.59	6.05	8.64	9.65
$\beta'\lambda$	3.94	3.94	3.95	3.95	3.95	3.91	3.88
$\beta'\lambda\ \|\lambda = \tilde{\lambda}$	2.05	2.06	2.07	2.07	2.07	2.03	1.97
$\beta'\lambda\ \|\beta = \tilde{\beta}$	3.17	3.17	3.17	3.17	3.17	3.17	3.16

Table VII

Averages Across Stocks of Posterior Means and Standard Deviations for Components of Expected Excess Returns from the Three-Factor Connor-Korajczyk Model

Posterior means and standard deviations for components of expected excess return are computed for each of the 1,994 stocks that have data through December 1995 and have at least 60 months of historical returns, and the cross-sectional averages of those values across all stocks are reported below. The expected excess return on a stock, μ, is given by $\mu = \alpha + \beta'\lambda$, where λ is the vector of expected values of the three Connor-Korajczyk factors ($F_{1,t}$, $F_{2,t}$, and $F_{3,t}$), and α and $\beta = [\beta_1\ \beta_2\ \beta_3]'$ are parameters in the regression of the stock's monthly excess return on the factors:

$$r_t = \alpha + \beta_1 F_{1,t} + \beta_2 F_{2,t} + \beta_3 F_{3,t} + \epsilon_t.$$

The moments for the parameters of the regression model, reported in Panel A, are based on each stock's available history of monthly returns, back through July 1963. The moments for the quantities involving λ, reported in Panel B, are based on monthly data for the periods indicated and, for the longer period, use the additional information in the history of returns on the value-weighted and equally weighted NYSE portfolios and the Ibbotson small-stock portfolio. Also reported for each period is $\tilde{\lambda}$, the posterior mean of λ. Except for the moments of β, all posterior means and standard deviations are reported as annualized percentage values.

	Prior Standard Deviation of α (σ_α)						
	0	1%	3%	5%	10%	30%	∞
Panel A. Regression parameters							
				Means			
α	0.00	-0.01	0.01	0.09	0.35	0.88	1.13
β_1	0.97	0.97	0.97	0.97	0.97	0.97	0.97
β_2	0.02	0.02	0.02	0.02	0.02	0.02	0.02
β_3	0.14	0.14	0.14	0.14	0.14	0.14	0.14
				Standard deviations			
α	0.00	0.73	2.13	3.35	5.51	8.17	8.95
β_1	0.22	0.22	0.22	0.22	0.22	0.22	0.22
β_2	0.12	0.12	0.12	0.12	0.12	0.12	0.12
β_3	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Panel B. Components involving the expected factors							
				1/1926–12/1995; $\tilde{\lambda} = [10.85\ -2.22\ 5.94]'$			
				Means			
μ	11.36	11.34	11.36	11.44	11.68	12.18	12.43
$\beta'\lambda$	11.36	11.35	11.35	11.34	11.33	11.31	11.31
				Standard Deviations			
μ	4.09	4.12	4.53	5.15	6.64	8.88	9.63
$\beta'\lambda$	4.09	4.09	4.09	4.09	4.09	4.09	4.08
$\beta'\lambda\ \|\lambda = \tilde{\lambda}$	2.23	2.23	2.23	2.24	2.24	2.24	2.24
$\beta'\lambda\ \|\beta = \tilde{\beta}$	3.10	3.10	3.10	3.09	3.09	3.09	3.09
				7/1963–12/1995; $\tilde{\lambda} = [7.65\ -3.17\ 5.66]'$			
				Means			
μ	8.19	8.17	8.20	8.27	8.52	9.03	9.28
$\beta'\lambda$	8.19	8.19	8.18	8.18	8.17	8.15	8.15
				Standard Deviations			
μ	3.83	3.87	4.32	4.99	6.56	8.86	9.61
$\beta'\lambda$	3.83	3.83	3.83	3.83	3.83	3.82	3.82
$\beta'\lambda\ \|\lambda = \tilde{\lambda}$	1.65	1.65	1.66	1.66	1.66	1.67	1.66
$\beta'\lambda\ \|\beta = \tilde{\beta}$	3.18	3.18	3.18	3.18	3.18	3.17	3.17

Table VIII

Model Uncertainty and Overall Uncertainty About the Cost of Equity

The table reports the uncertainty about a stock's expected excess return (μ) that arises from entertaining multiple pricing models, the overall uncertainty about μ that incorporates both model uncertainty and within-model parameter uncertainty, and the average within-model parameter uncertainty. The three pricing models are the CAPM, the three-factor Fama-French model (FF), and the three-factor Connor-Korajczyk model (CK). For any given subset of models entertained, each model is assigned equal probability. Each stock's α and β are defined by the regression

$$r_t = \alpha + \beta' f_t + \epsilon_t,$$

where r_t is the stock's excess return, f_t is a vector of factors, and $\lambda = \mathrm{E}(f_t)$. All values are reported as annualized percentage standard deviations. In the third part of Panel A, the posterior variances are averaged across models before taking the square root to obtain the values reported, and those latter values are then averaged across stocks to obtain the values reported in the third part of Panel B.

	Prior Standard Deviation of α (σ_α)						
	0	1%	3%	5%	10%	30%	∞
Panel A. Results for Bay State Gas							
Model uncertainty about μ when the set of models entertained is							
CAPM and FF	1.58	1.65	1.67	1.57	1.21	0.55	0.14
CAPM and CK	0.82	0.96	1.19	1.35	1.41	0.91	0.39
FF and CK	0.76	0.69	0.47	0.22	0.20	0.36	0.25
CAPM, FF, and CK	1.29	1.35	1.40	1.39	1.24	0.75	0.32
Overall uncertainty about μ when the set of models entertained is							
CAPM and FF	2.19	2.26	2.44	2.66	3.15	3.84	4.02
CAPM and CK	1.61	1.73	2.13	2.57	3.32	3.94	4.04
FF and CK	1.80	1.79	1.93	2.23	2.96	3.83	4.02
CAPM, FF, and CK	1.99	2.05	2.29	2.59	3.21	3.89	4.03
Average across the three models of the within-model uncertainty of							
μ	1.51	1.55	1.81	2.18	2.96	3.81	4.02
α	0.00	0.37	1.08	1.69	2.71	3.72	3.93
$\beta'\lambda$	1.51	1.51	1.51	1.51	1.50	1.48	1.48
Panel B. Average Across 1,994 Stocks of the Values in Panel A							
Model uncertainty about μ when the set of models entertained is							
CAPM and FF	2.10	2.08	2.00	1.91	1.72	1.53	1.42
CAPM and CK	1.68	1.67	1.63	1.58	1.50	1.47	1.47
FF and CK	1.50	1.41	1.22	1.04	0.77	0.61	0.79
CAPM, FF, and CK	2.26	2.20	2.07	1.93	1.71	1.55	1.57
Overall uncertainty about μ when the set of models entertained is							
CAPM and FF	4.24	4.23	4.51	5.03	6.46	8.87	9.78
CAPM and CK	3.94	4.00	4.46	5.12	6.68	9.00	9.78
FF and CK	4.48	4.43	4.63	5.09	6.43	8.78	9.69
CAPM, FF, and CK	4.40	4.40	4.67	5.19	6.59	8.93	9.79
Average across the three models of the within-model uncertainty of							
μ	3.71	3.74	4.13	4.76	6.31	8.75	9.62
α	0.00	0.71	2.06	3.25	5.42	8.23	9.10
$\beta'\lambda$	3.71	3.71	3.72	3.72	3.72	3.70	3.68

Table IX

Model Uncertainty and Overall Uncertainty About the Cost of Equity for Utilities

The table reports the uncertainty about a stock's expected excess return (μ) that arises from entertaining multiple pricing models, the overall uncertainty about μ that incorporates both model uncertainty and within-model parameter uncertainty, and the average within-model parameter uncertainty. The three pricing models are the CAPM, the three-factor Fama-French model (FF), and the three-factor Connor-Korajczyk model (CK). For any given subset of models entertained, each model is assigned equal probability. Each stock's α and β are defined by the regression

$$r_t = \alpha + \beta' f_t + \epsilon_t,$$

where r_t is the stock's excess return, f_t is a vector of factors, and $\lambda = \mathrm{E}(f_t)$. All values are reported as annualized percentage standard deviations. In the third part of Panel A, the posterior variances are averaged across models before taking the square root to obtain the values reported, and those latter values are then averaged across stocks to obtain the values reported in the third part of Panel B.

	Prior Standard Deviation of α (σ_α)						
	0	1%	3%	5%	10%	30%	∞
Panel A. Results for Bay State Gas							
Model uncertainty about μ when the set of models entertained is							
CAPM and FF	1.68	1.62	1.29	0.93	0.46	0.18	0.14
CAPM and CK	0.88	0.86	0.75	0.63	0.46	0.36	0.34
FF and CK	0.79	0.76	0.54	0.31	0.00	0.18	0.21
CAPM, FF, and CK	1.37	1.33	1.06	0.78	0.43	0.29	0.28
Overall uncertainty about μ when the set of models entertained is							
CAPM and FF	2.23	2.30	2.70	3.10	3.60	3.88	3.93
CAPM and CK	1.62	1.76	2.44	2.99	3.59	3.90	3.95
FF and CK	1.77	1.89	2.49	3.01	3.59	3.90	3.95
CAPM, FF, and CK	2.01	2.10	2.60	3.06	3.60	3.90	3.94
Average across the three models of the within-model uncertainty of							
μ	1.47	1.63	2.37	2.96	3.57	3.89	3.93
α	0.00	0.74	1.95	2.69	3.41	3.77	3.82
$\beta'\lambda$	1.47	1.47	1.46	1.45	1.44	1.44	1.43
Panel B. Average Across 135 Utilities of the Values in Panel A							
Model uncertainty about μ when the set of models entertained is							
CAPM and FF	1.59	1.52	1.16	0.83	0.48	0.30	0.27
CAPM and CK	0.74	0.71	0.56	0.43	0.32	0.32	0.33
FF and CK	1.01	0.97	0.74	0.54	0.39	0.36	0.36
CAPM, FF, and CK	1.41	1.35	1.04	0.77	0.51	0.42	0.41
Overall uncertainty about μ when the set of models entertained is							
CAPM and FF	2.38	2.46	2.91	3.34	3.89	4.26	4.32
CAPM and CK	1.78	1.95	2.68	3.25	3.90	4.29	4.35
FF and CK	2.14	2.26	2.83	3.33	3.92	4.29	4.35
CAPM, FF, and CK	2.24	2.35	2.86	3.33	3.91	4.28	4.35
Average across the three models of the within-model uncertainty of							
μ	1.71	1.88	2.65	3.23	3.87	4.26	4.32
α	0.00	0.81	2.08	2.82	3.57	4.00	4.07
$\beta'\lambda$	1.71	1.71	1.71	1.71	1.71	1.71	1.71

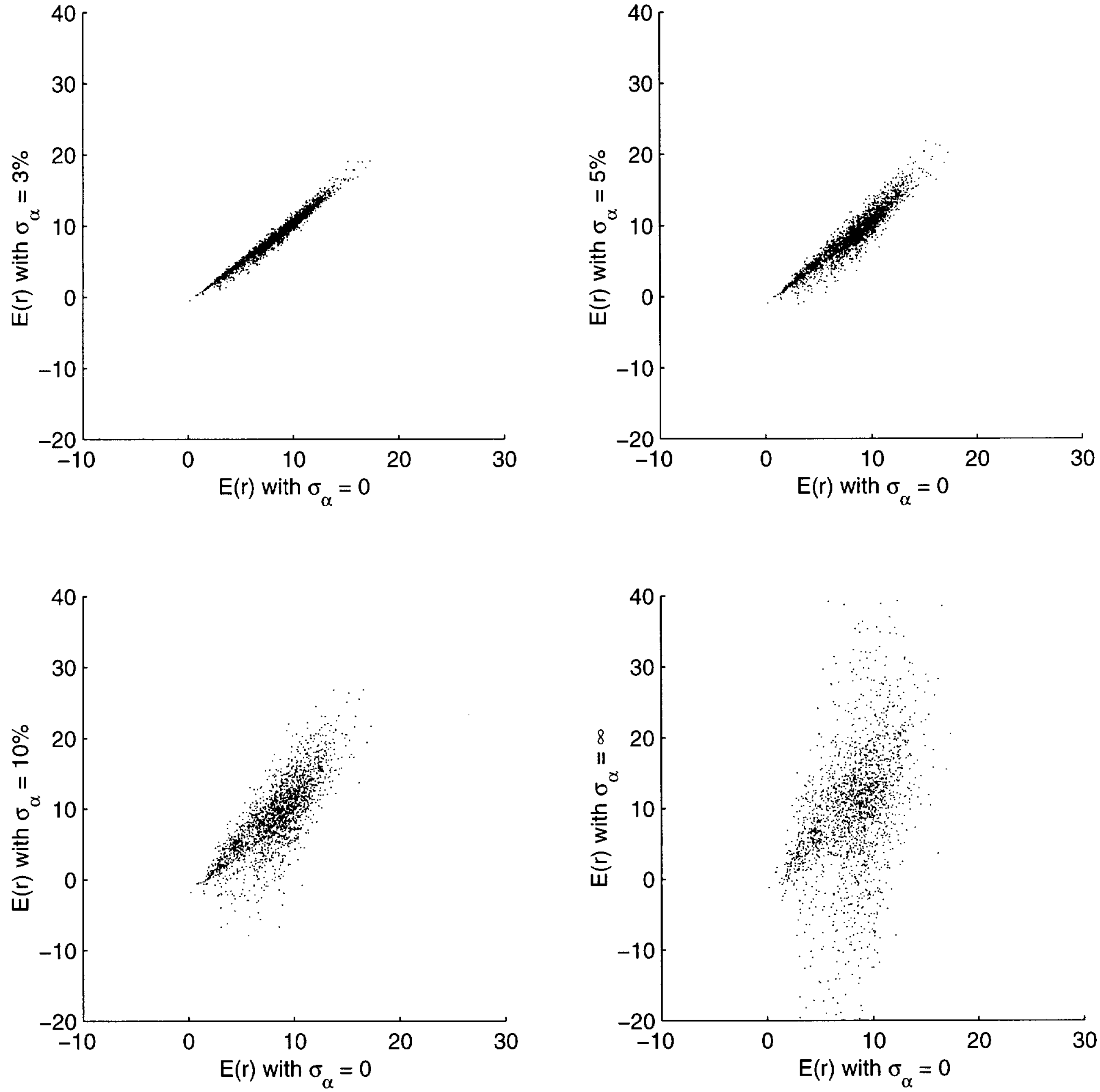

Figure 1. Effects of mispricing uncertainty in the CAPM. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α. Each graph plots, for 1,994 individual stocks, the estimate of the expected excess return ($\mathrm{E}(r)$) with $\sigma_\alpha = 0$ versus $\mathrm{E}(r)$ with a non-zero value of σ_α.

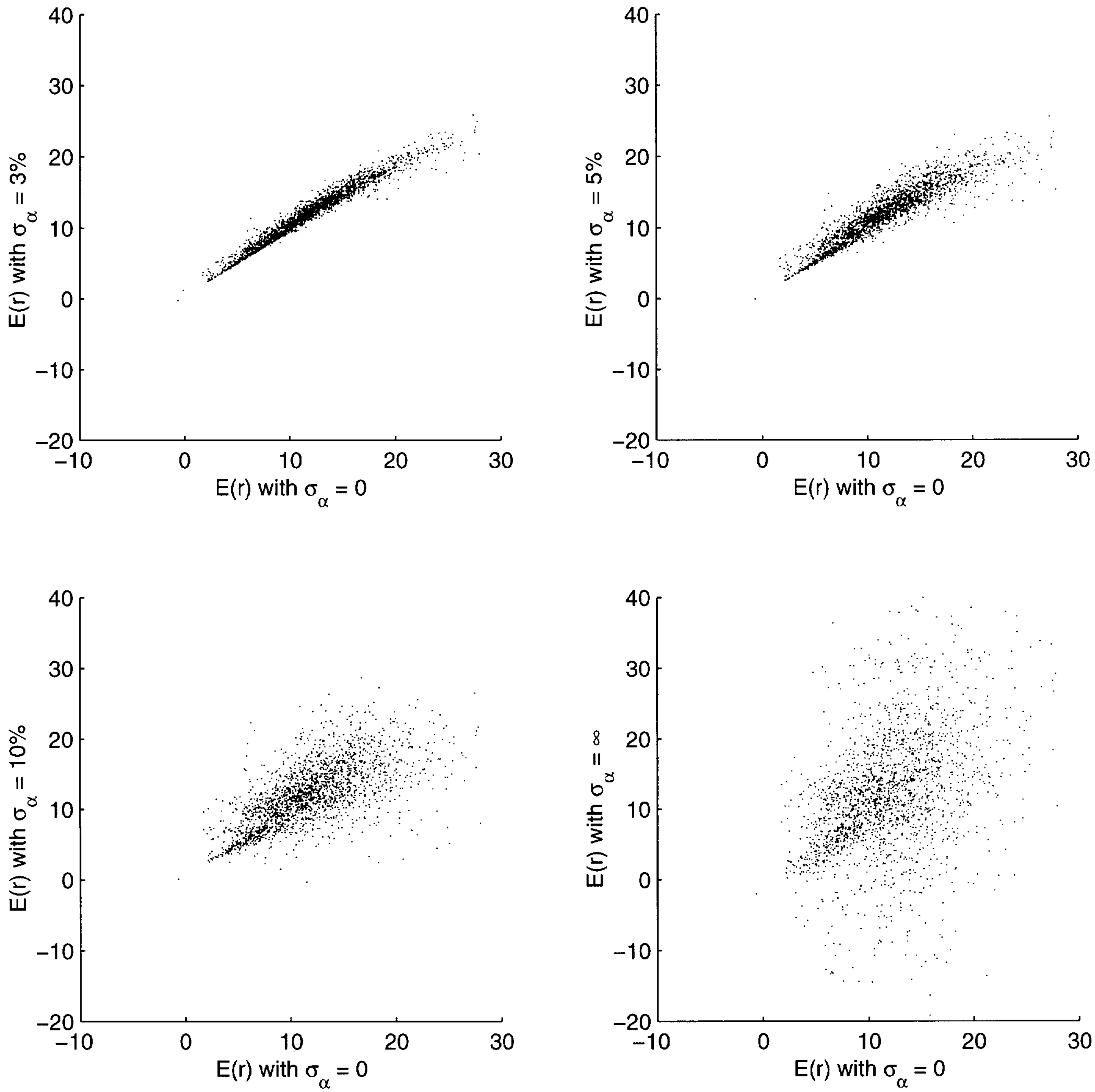

Figure 2. Effects of mispricing uncertainty in the three-factor Fama-French (FF) model. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α. Each graph plots, for 1,994 individual stocks, the estimate of the expected excess return (E(r)) with $\sigma_\alpha = 0$ versus E(r) with a non-zero value of σ_α.

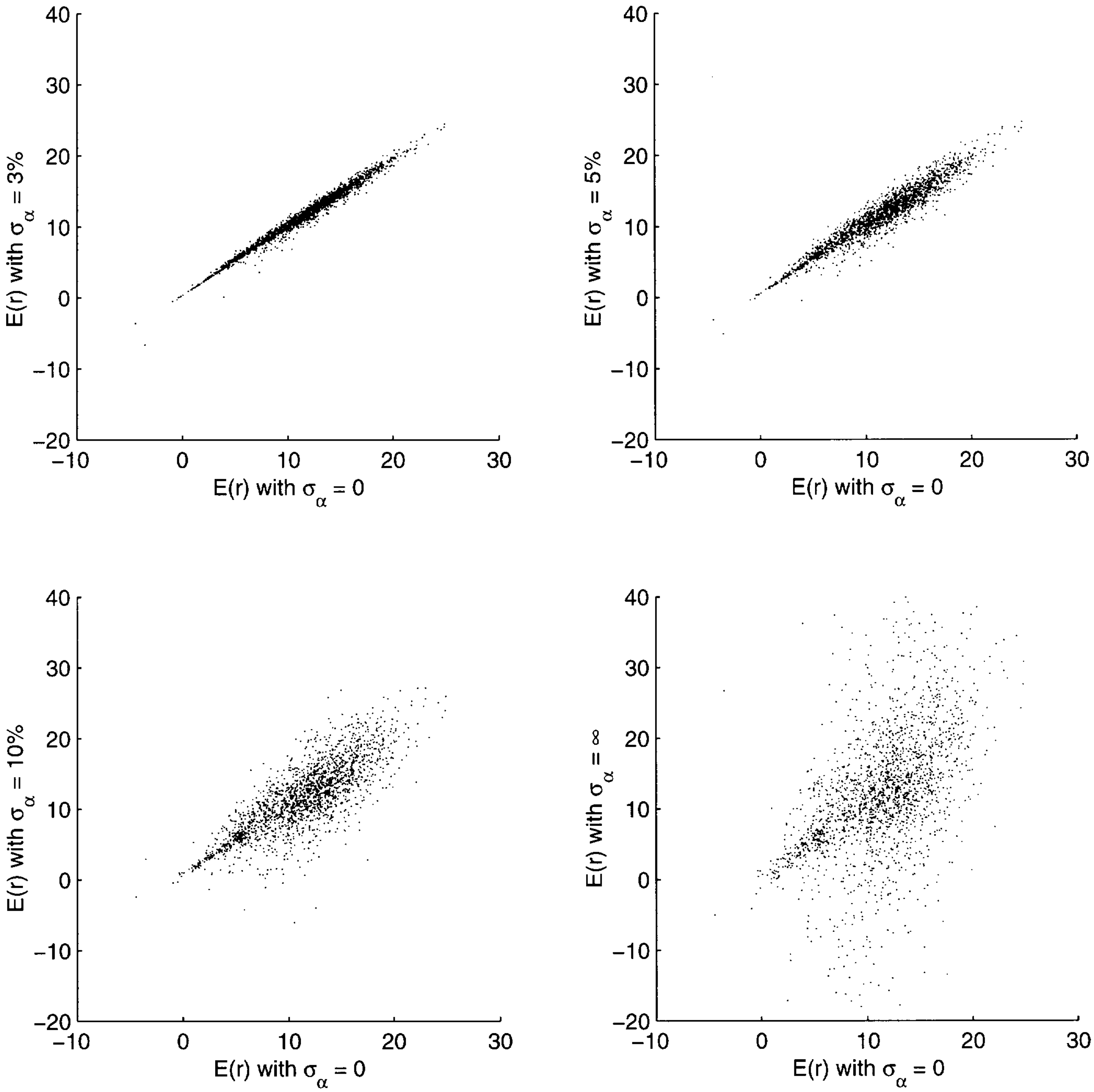

Figure 3. Effects of mispricing uncertainty in the three-factor Connor-Korajczyk (CK) model. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α. Each graph plots, for 1,994 individual stocks, the estimate of the expected excess return (E(r)) with $\sigma_\alpha = 0$ versus E(r) with a non-zero value of σ_α.

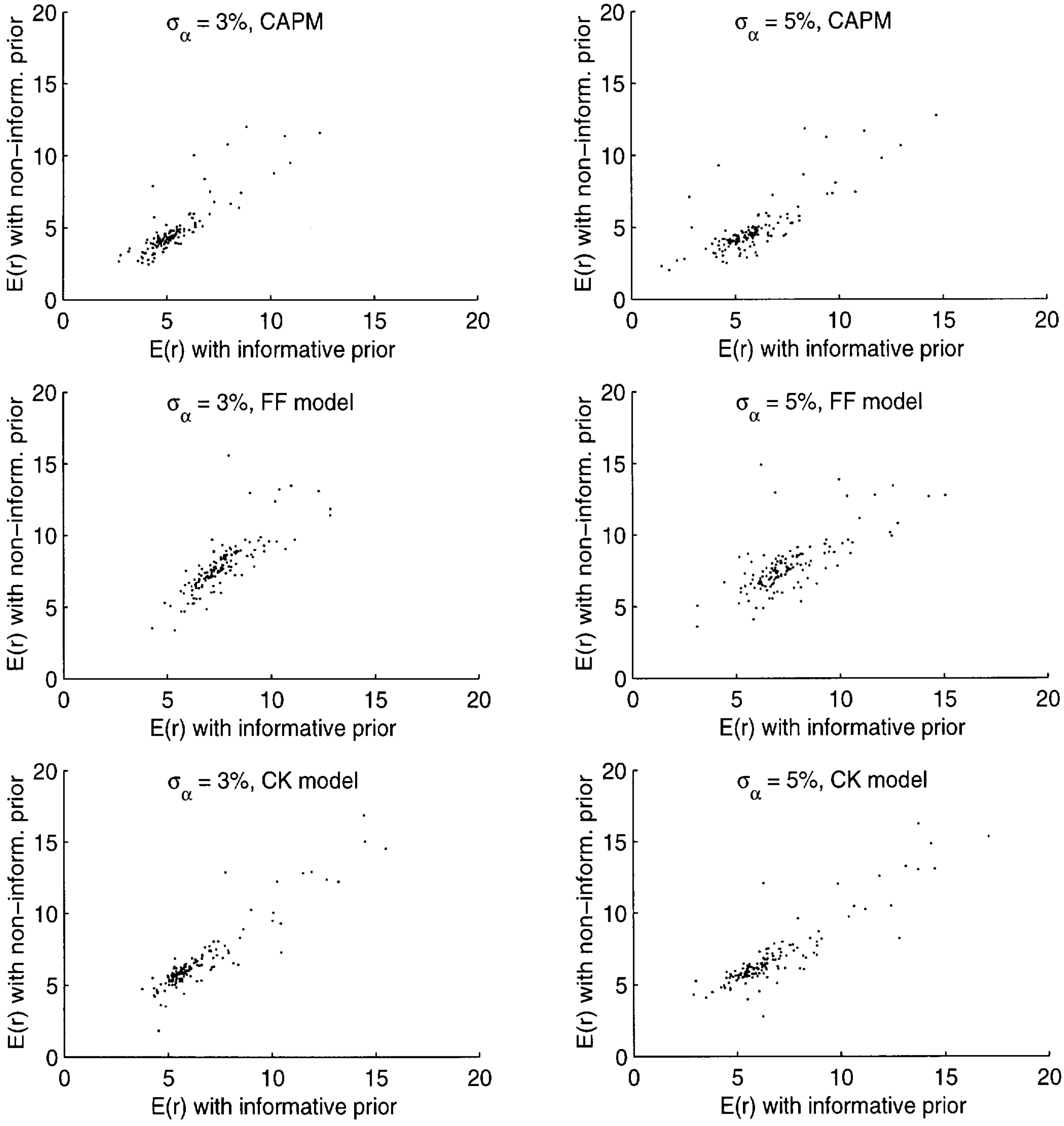

Figure 4. Effects of economically informative versus non-informative priors on estimates of expected excess returns for utility stocks. Each graph plots, for 135 utility stocks, the estimate of the expected excess return (E(r)) with an economically informative prior versus E(r) with an economically non-informative prior. Results are displayed for three pricing models: the CAPM, the three-factor Fama-French (FF) model, and the three-factor Connor-Korajczyk (CK) model. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α.

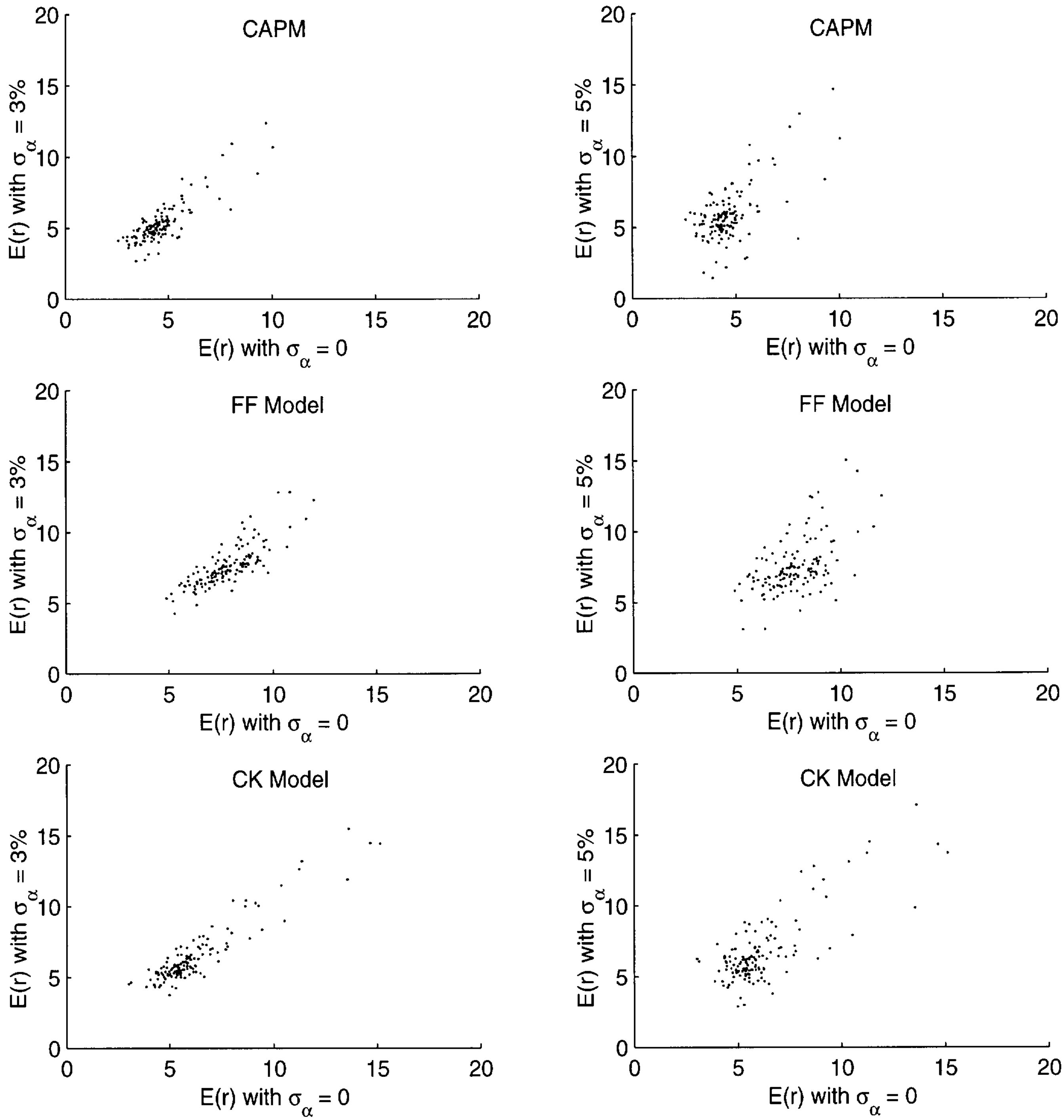

Figure 5. Effects of mispricing uncertainty on estimates of expected excess returns for utility stocks. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α. Each graph plots, for 135 utility stocks, the estimate of the expected excess return (E(r)) with $\sigma_\alpha = 0$ versus E(r) with a non-zero value of σ_α. Results are displayed for three pricing models: the CAPM, the three-factor Fama-French (FF) model, and the three-factor Connor-Korajczyk (CK) model.

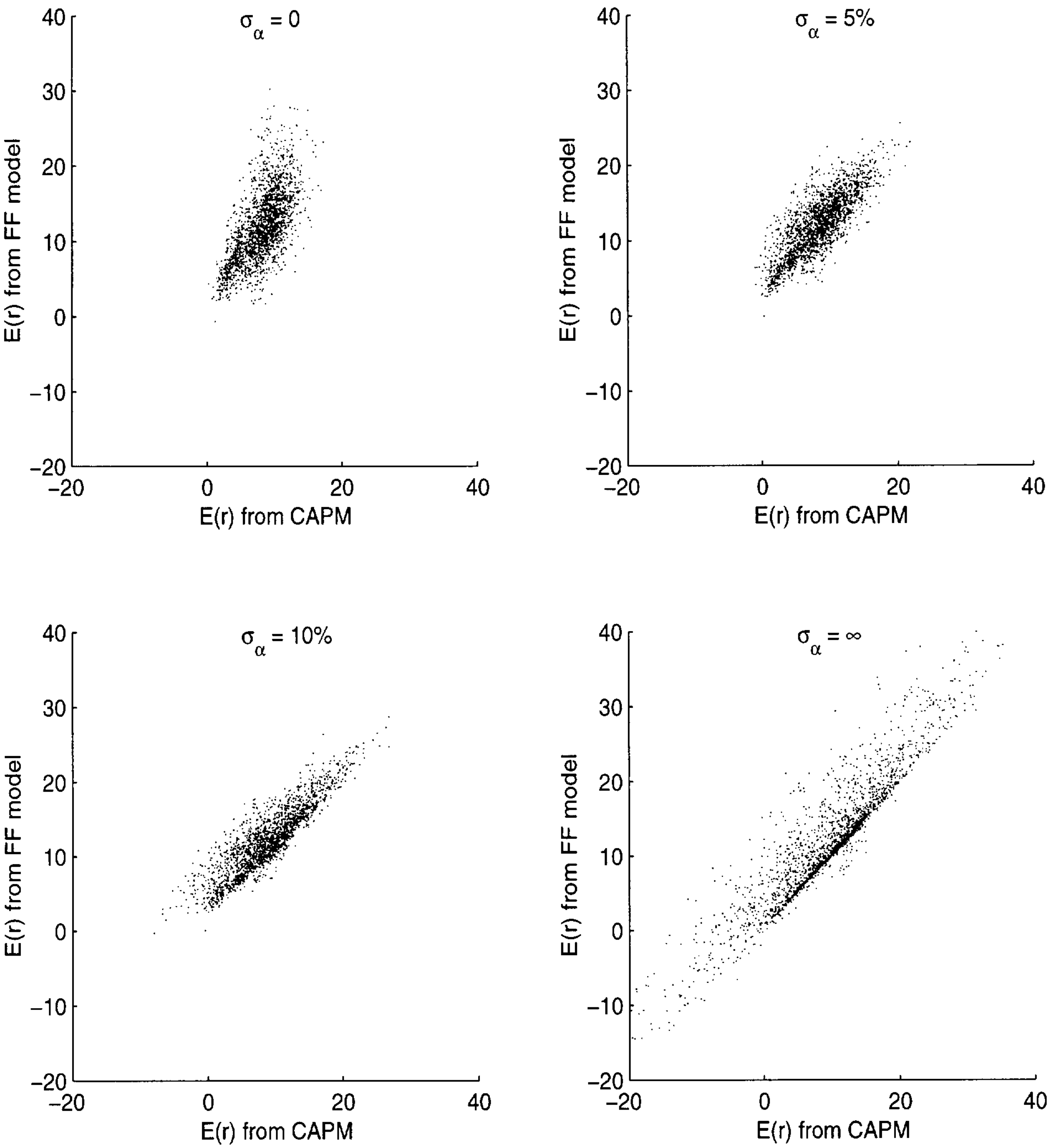

Figure 6. Comparison of expected excess returns from the CAPM and the Fama-French model. Each graph plots, for 1,994 individual stocks, the estimate of the expected excess return (E(r)) from the CAPM versus E(r) from the three-factor Fama-French (FF) model. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α.

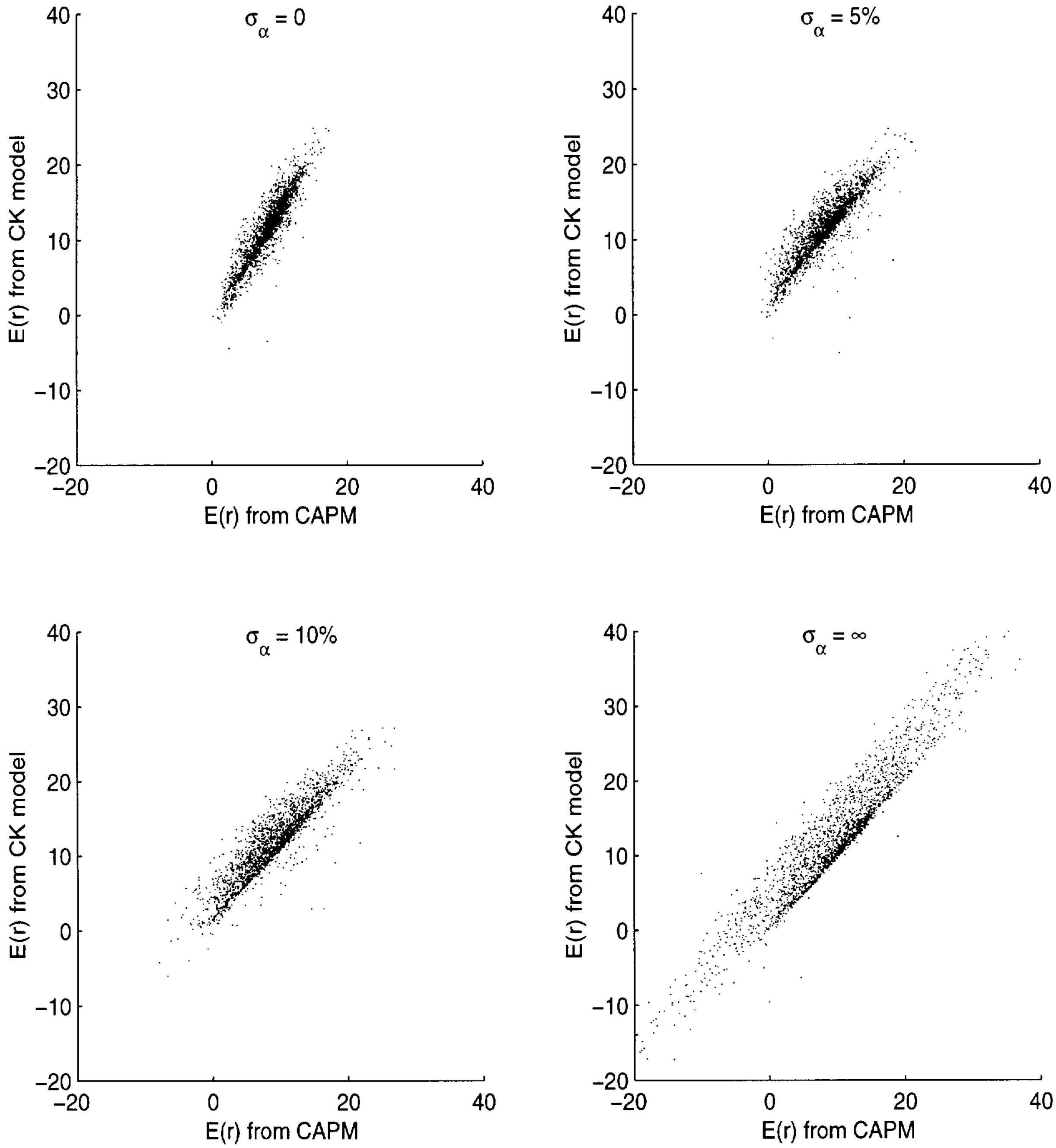

Figure 7. Comparison of expected excess returns from the CAPM and the Connor-Korajczyk model. Each graph plots, for 1,994 individual stocks, the estimate of the expected excess return (E(r)) from the CAPM versus E(r) from the three-factor Connor-Korajczyk (CK) model. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α.

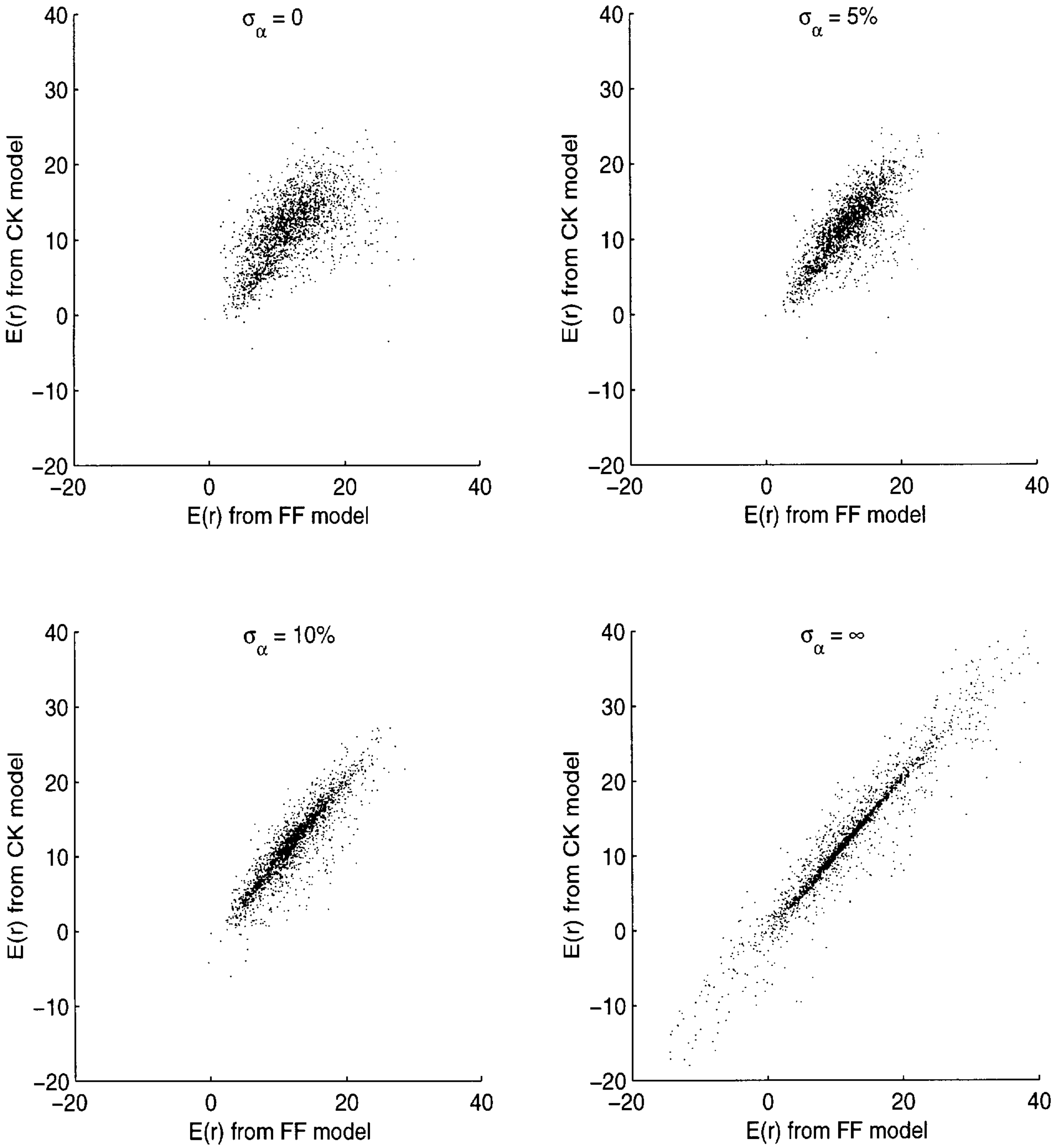

Figure 8. Comparison of expected excess returns from the Fama-French and Connor-Korajczyk models. Each graph plots, for 1,994 individual stocks, the estimate of the expected excess return (E(r)) from the three-factor Fama-French (FF) model versus E(r) from the three-factor Connor-Korajczyk (CK) model. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α.

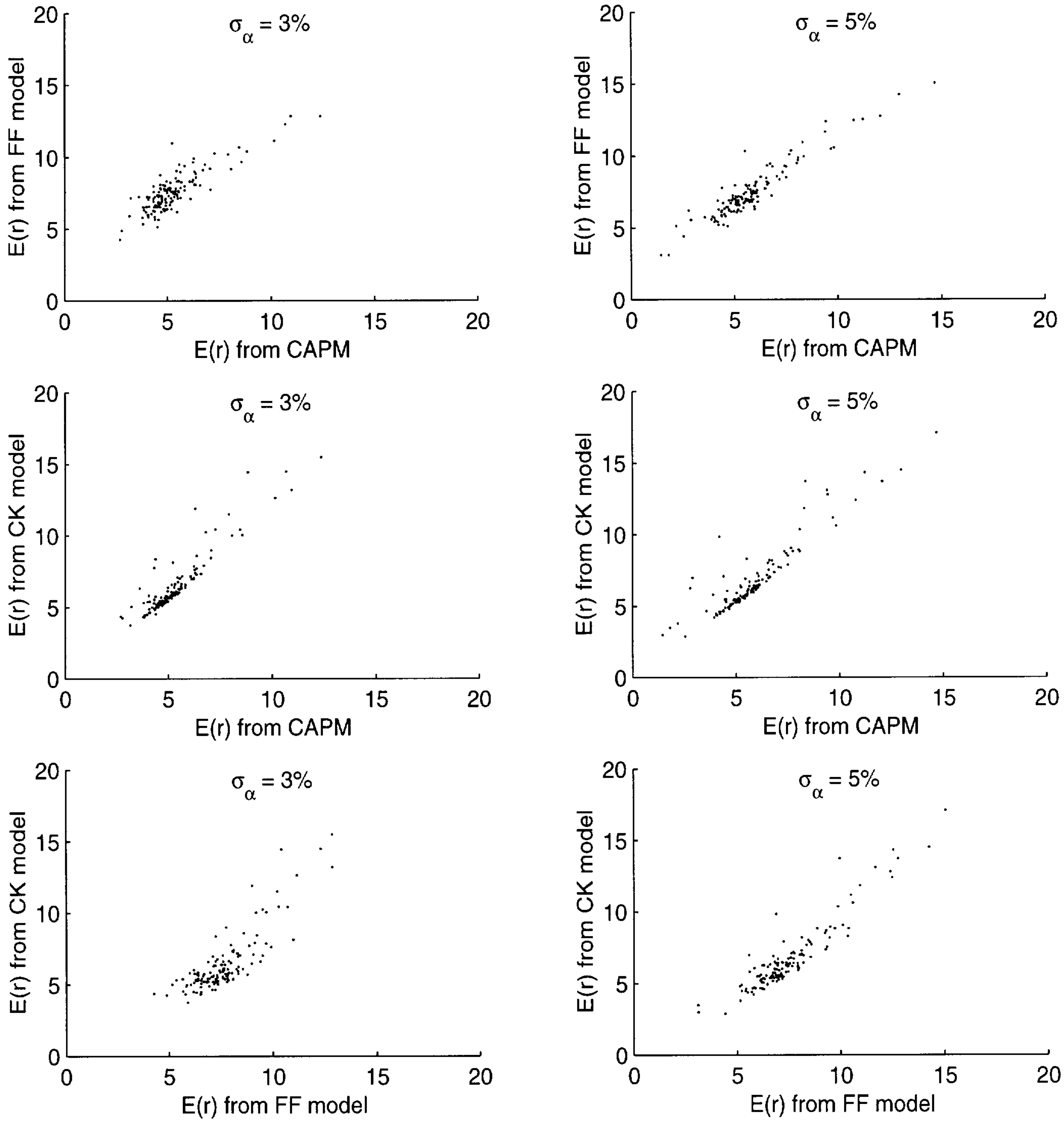

Figure 9. Comparison of expected excess returns on utility stocks from different pricing models. Each graph plots, for 135 utility stocks, the estimate of the expected excess return (E(r)) from one model versus E(r) from another model. Three pricing models are entertained: the CAPM, the three-factor Fama-French (FF) model, and the three-factor Connor-Korajczyk (CK) model. The prior mispricing uncertainty, σ_α, is the annualized prior standard deviation of α.

Notes

[1]Such models include the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965), the intertemporal CAPM of Merton (1973), and the Arbitrage Pricing Theory of Ross (1976). An issue beyond the intended scope of this study is the selection of the appropriate riskless rate for computing the total cost of equity. See, for example, Cornell, Hirshleifer, and James (1997).

[2]See Huberman, Kandel, and Stambaugh (1987) for a deeper discussion of this point.

[3]Ferson and Locke (1997) reach this conclusion even after allowing the errors in betas to encompass errors in constructing the market index.

[4]See Zellner (1971), p.372.

[5]A portfolio's Sharpe ratio is its expected excess return divided by its standard deviation of return.

[6]See, for example, Gibbons, Ross, and Shanken (1989).

[7]Vasicek (1973) proposes using a cross-section of stocks to obtain the parameters of the prior distribution for the market beta. See Berger (1985) for a general discussion of empirical-Bayes methods.

[8]We thank Ken French for providing these data.

[9]The factors are the first three eigenvectors of the $T \times T$ matrix ($T = 390$) whose (s, t) element is $(1/N_{s,t}) \sum_{i=1}^{N_{s,t}} r_{i,s} r_{i,t}$, where $r_{i,t}$ is the excess return on stock i in month t and $N_{s,t}$ denotes the number of stocks that have returns in months s and t.

[10]Technically, the priors and posteriors given in our formulas are defined only for finite positive values of σ_α, so the results reported for "zero" and "infinity" are actually computed by setting σ_α to very small and very large values.

[11]This follows directly from the moments given by Zellner (1971), p. 372.

[12]The second line in (21) follows from

$$\begin{aligned} |\Psi(\sigma)| &= \sigma_\alpha^2 \left| \left(\frac{\sigma^2}{E(\sigma^2)} \right) - \left(\frac{\sigma}{E(\sigma)} \right)^2 \sigma_\beta' \rho_{\alpha\beta}' V_\beta^{-1} \rho_{\alpha\beta} \sigma_\beta \right| \cdot |V_\beta| \\ &\propto \sigma^2. \end{aligned}$$

[13]The prior for σ is the same as in (10), and the prior for b is specified as normal with mean $\bar{b}$ and covariance matrix V_b. The conditional posterior for σ given b is inverted gamma in that case, and the acceptance rate for candidates drawn from this proposal density typically ranges between 75% and 79% across the three pricing models. The procedure described above, where the proposal density for some of the parameters (b in this case) is the same as the target "full conditional" density, is sometimes referred to as "Metropolis within Gibbs," although Chib and Greenberg (1995) suggest that such terminology is inappropriate. As they point out, Gibbs sampling, introduced by Geman and Geman (1984), is a special case of MH in which *all* parameters are drawn from their full conditional densities. Casella and George (1992) provide an introduction to the Gibbs sampler, and an early finance application appears in Kandel, McCulloch, and Stambaugh (1995).

[14]From (33) and the independence of λ and β, the conditional variance of μ given λ (where all moments are posterior) is

$$\mathrm{Var}(\mu|\lambda) = [1\ \lambda']\tilde{V}_b \begin{bmatrix} 1 \\ \lambda \end{bmatrix} = \mathrm{tr}\left(\tilde{V}_b \begin{bmatrix} 1 & \lambda' \\ \lambda & \lambda\lambda' \end{bmatrix}\right),$$

and the expectation of this quantity, taken with respect to λ, is the first term on the right-hand side of (35). The conditional mean of μ given λ is $\tilde{\alpha} + \tilde{\beta}'\lambda$, and the variance of this quantity, taken with respect to λ, is the second term on the right-hand side of (35).

[15]Recall that α is the first element of b in equations (25) and (26). With non-zero off-diagonal elements in the first row of M^{-1}, the posterior covariance matrix of b (conditional on σ), the posterior mean of α generally depends, for a finite T, on all of the elements of $\bar{b}$ and $\hat{b}$. The slight non-monotonicity in $\tilde{\alpha}$ observed in the first row of Table II is most likely due to the influence of the β-related terms on α.

[16]Fama and French (1997) find support for such a specification, although they do not find its merits over the simpler procedure to be clear cut. Moreover, they also suggest (p. 170) that, because variables such as size and book-to-market may be somewhat under management's control, "firms might be better off using full-period constant-slope [costs of equity] for capital budgeting." Schink and Bower (1994), for example, use full-period betas in estimating the costs of equity for individual public utilities.

References

Anderson, Theodore W., 1984, *An Introduction to Multivariate Statistical Analysis* (John Wiley and Sons, New York, NY).

Benninga, Simon Z., and Oded H. Sarig, 1997, *Corporate Finance, A Valuation Approach* (McGraw-Hill, New York, NY).

Berger, James O., 1985, *Statistical Decision Theory and Bayesian Analysis* (Springer-Verlag, New York, NY).

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The Capital Asset Pricing Model: Some empirical tests, in Michael C. Jensen, ed.: *Studies in the Theory of Capital Markets* (Praeger, New York, NY).

Box, George E.P., and George C. Tiao, 1973, *Bayesian Inference in Statistical Analysis* (Addison-Wesley, Reading, MA).

Blattberg, Robert C., and Edward I. George, 1991, Shrinkage estimation of price and promotional elasticities: Seemingly unrelated equations, *Journal of the American Statistical Association* 86, 304–315.

Casella, G., and E.I. George, 1992, Explaining the Gibbs sampler, *The American Statistician* 46, 167–174.

Chib, Siddhartha, and Edward Greenberg, 1995, Understanding the Metropolis-Hastings algorithm, *American Statistician* 49, 327–335.

Chib, Siddhartha, and Edward Greenberg, 1996, Markov chain Monte Carlo simulation methods in econometrics, *Econometric Theory* 12, 409–431.

Connor, Gregory, and Robert A. Korajczyk, 1986, Performance measurement within the arbitrage pricing theory: A new framework for analysis, *Journal of Financial Economics* 15, 373-394.

Connor, Gregory, and Robert A. Korajczyk, 1987, Estimating pervasive economic factors with missing observations, Working paper, Northwestern University.

Cornell, Bradford, John I. Hirshleifer, and Elizabeth P. James, 1997, Estimating the cost of equity capital, *Contemporary Finance Digest* 1, 5–26.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

DiValentino, L. Mario, 1994, Preface, *Financial Markets, Institutions, and Instruments* 3, 46-73

Elton, Edwin J., Martin J. Gruber, and Jianping Mei, 1994, Cost of capital using arbitrage pricing theory: A case study of nine New York utilities, *Financial Markets, Institutions, and Instruments* 3, 46-73

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Ferson, Wayne E., and Dennis H. Locke, 1997, Estimating the cost of capital through time: an analysis of the sources of error, *Management Science*, forthcoming.

Geman, S., and D. Geman, 1984, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, 609-628.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, 1996, *Markov Chain Monte Carlo in Practice* (Chapman and Hall, London, UK).

Hastings, W. K., 1970, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* 57, 97–109.

Huberman, Gur, Shmuel Kandel, and Robert F. Stambaugh, 1987, Mimicking portfolios and exact arbitrage pricing, *Journal of Finance* 42, 1–9.

Jeffreys, Harold B., 1961, *Theory of Probability* (Oxford University Press, Oxford, U.K.).

Kandel, Shmuel, Robert McCulloch, and Robert F. Stambaugh, 1995, Bayesian inference and portfolio efficiency, *Review of Financial Studies* 8, 1–53.

Kass, Robert E., and Adrian E. Raftery, 1995, Bayes factors, *Journal of the American*

Statistical Association 90, 773–795.

Leamer, Edward E., 1978, *Specification Searches: Ad Hoc Inference with Non-Experimental Data* (John Wiley, New York, NY).

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

MacKinlay, A. Craig, 1995, Multifactor models do not explain the CAPM, *Journal of Financial Economics* 38, 3–28.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953, Equations of state calculations by fast computing machines, *Journal of Chemical Physics* 21, 1087–1092.

Poirier, Dale J., 1995, *Intermediate Statistics and Econometrics* (MIT Press, Cambridge, MA).

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Schink, George R., and Richard S. Bower, 1994, Application of the Fama-French model to utility stocks, *Financial Markets, Institutions, and Instruments* 3, 74-96.

Shanken, Jay, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99–120.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Stambaugh, Robert F., 1997, Analyzing investments whose histories differ in length, *Journal of Financial Economics*, 45, 285–331.

Tanner, Martin A., 1993, *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions* (Springer-Verlag, New York, NY).

Vasicek, Oldrich A., 1973, A note on using cross-sectional information in Bayesian estimation of security betas, *Journal of Finance*, 28, 1233-1239.

Zellner, Arnold, 1971, *An Introduction to Bayesian Inference in Econometrics* (John Wiley and Sons, New York, NY).