ELSEVIER

Journal of Financial Economics 56 (2000) 335–381

JOURNAL OF Financial ECONOMICS

www.elsevier.com/locate/econbase

Comparing asset pricing models: an investment perspective☆

Ľuboš Pástor[a], Robert F. Stambaugh[b,c,*]

[a]*Graduate School of Business, University of Chicago, Chicago, IL 60637, USA*
[b]*Finance Department, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA*
[c]*The National Bureau of Economic Research, USA*

Abstract

We investigate the portfolio choices of mean-variance-optimizing investors who use sample evidence to update prior beliefs centered on either risk-based or characteristic-based pricing models. With dogmatic beliefs in such models and an unconstrained ratio of position size to capital, optimal portfolios can differ across models to economically significant degrees. The differences are substantially reduced by modest uncertainty about the models' pricing abilities. When the ratio of position size to capital is subject to realistic constraints, the differences in portfolios across models become even less important and are nonexistent in some cases.

JEL classification: G11; G12; C11

Keywords: Portfolio selection; Asset pricing models; Investment constraints; Bayesian analysis

☆We are grateful for comments by Nick Barberis, Gene Fama, Ravi Jagannathan, Don Keim, David Modest, Toby Moskowitz, William Tyson, an anonymous referee, and workshop participants at London Business School, Stanford University, University of British Columbia, University of Chicago, University of Maryland, University of Pennsylvania, the Spring 1999 NBER Asset Pricing meeting, and the 2000 AFA meetings.

* Correspondence address. Finance Department, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6367, USA. Tel.: 215-8985734; fax: 215-8986200.

E-mail address: stambaugh@wharton.upenn.edu (R.F. Stambaugh).

PII: S0304-405X(00)00044-1

1. Introduction

To an investor seeking a mean-variance efficient portfolio, a risk-based pricing model offers a powerful insight. If expected excess returns are linear combinations of exposures (or 'betas') to k sources of risk (or 'factors'), then the risky portion of any mean-variance efficient portfolio is a combination of k benchmark portfolios that mimic those factors.[1] In attempting to apply this insight, an investor confronts a variety of issues. Alternative pricing models vie for consideration by the investor, who realizes that no model is likely to be completely accurate. In one class of contenders to risk-based models, asset characteristics unrelated to risk exposures enter expected returns due to behavioral phenomena such as overreaction, and these 'characteristic-based' models do not identify a set of benchmark portfolios for investment. Moreover, the investor faces constraints on borrowing and short sales, to at least some degree, whereas a theory's investment implications are often derived in the absence of such constraints.

This study conducts an empirical comparison of asset pricing models. In order to provide an economic metric for judging differences between models, we analyze the portfolio-choice problem from the perspective of investors who face the issues described above. Our investigation reveals the extent to which alternative pricing models imply different investment choices. An investor views a pricing model as providing a way to center prior beliefs, specified with varying degrees of confidence in the model. These prior beliefs about the distribution of returns are then updated by the data and used to compute an optimal portfolio, subject to margin requirements ranging from 50% to none at all. Our objective is not to choose one pricing model over another but instead to shed some light on the economic importance of deliberating such a choice. We find that, in the presence of mispricing uncertainty and margin requirements, models with fundamentally different views about the economic determinants of expected returns often imply similar portfolio choices.

Fama and French (1993) and Daniel and Titman (1997) differ in their explanations of the apparent empirical relation between a firm's expected equity return and its market capitalization and book-to-market ratio. In the Fama–French model, the latter characteristics are associated with risk exposures, whereas they reflect mispricing in the Daniel–Titman model. From an investment perspective, we compare these models to each other and to the Sharpe–Lintner CAPM. The investment universes are similar to those analyzed

[1] Examples of risk-based models that contain this implication are the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965), the intertemporal model of Merton (1973), and the Arbitrage Pricing Theory of Ross (1976). For further discussion see Jobson and Korkie (1985), Grinblatt and Titman (1987), and Huberman et al. (1987).

by Daniel and Titman (1997) and Davis, Fama, and French (2000). In particular, portfolios are formed by sorting stocks on total equity capitalization ('size'), the ratio of book value to market value of common equity ('book-to-market'), and betas with respect to the 'HML' book-to-market factor of Fama and French (1993). One investment universe consists of the three benchmark positions from the Fama–French specification of a factor-based model plus nine spread positions that are long stocks with low HML betas and short stocks with high HML betas, holding size and book-to-market constant. These spread positions are designed to exploit differences between the Fama–French model and the Daniel–Titman model: the expected payoffs are negative under the first model but zero under the second. Optimal portfolios are computed for various hypothetical investors who face this investment universe but believe in different pricing models, with perfect confidence or with some uncertainty about a model's pricing ability.

The optimal portfolio for an investor with perfect confidence in a model's pricing ability can exhibit economically significant differences from that of an investor with equally strong beliefs in an alternative pricing model. Suppose, for example, that risk aversion is specified such that all of an investor's wealth is allocated to the value-weighted stock market portfolio if that is the only risky asset available. Then, if an investor with perfect confidence in the Daniel–Titman characteristic-based model is forced to accept the portfolio of an investor with equally strong beliefs in the Fama–French factor-based model, the first investor perceives a certainty-equivalent loss of about 8% per year. A similar loss is perceived by the second investor if forced to accept the portfolio of the first. When each investor has some uncertainty about a model's pricing ability, the losses are substantially reduced. For example, if this 'mispricing' uncertainty is such that a spread with long and short positions of one dollar each has an expected payoff whose deviation from the model has a standard deviation of two cents per year, then the annual certainty-equivalent losses described above drop below 2%.

Expected returns are associated with characteristics in both the Daniel–Titman and Fama–French models. Given the latter model's specification of the factors SMB and HML, betas on those factors are correlated with size and book-to-market (e.g., high book-to-market firms tend to have high HML betas). Indeed, based on comparisons of investment implications under strict beliefs in both models, each model is closer to the other than to the CAPM, in which characteristics play a weaker role. An investor who believes in the complete accuracy of the Daniel–Titman model perceives an annual certainty-equivalent loss of 20% if forced to hold a 100% allocation in the market portfolio, the choice of an investor with complete confidence in the CAPM. The same loss for the investor with complete confidence in the Fama–French model is about 12%. These losses are larger than the corresponding 8% value in the earlier comparison, but mispricing uncertainty again reduces these differences substantially. For example, with the same two-cent mispricing uncertainty described previously, the loss for the Daniel–Titman investor drops below 5%, and the loss for the Fama–French investor drops to about 2%.

The above results all describe cases in which an investor is permitted to establish long and short equity positions of any size. Most investors, however, are likely to face some limit on the aggregate value of risky positions that can be established per dollar of invested capital. For example, Regulation T, which applies to customers of U.S. broker/dealers, requires 50% margin, or a ratio of total position size to invested capital of no more than two. There are practices by which some investors exceed this limit, such as dealing with non-U.S. brokers or engaging in joint-back-office arrangements, but our understanding is that Regulation T governs much of the U.S. investment industry at both the individual and institutional levels. Nevertheless, in addition to a 50% margin requirement, we also consider margins of only 20% and 10%.

Margin requirements can dramatically reduce, and even eliminate, the cross-model differences in investment implications. With margin requirements of 50% and 20%, the Fama–French and Daniel–Titman models yield identical portfolios from the asset universe described above, even for investors whose prior beliefs preclude any mispricing. Imposing even a 10% margin still has large effects. For example, an investor with complete confidence in the Daniel–Titman model who must hold the portfolio of an equally confident Fama–French investor perceives an annual certainty-equivalent loss of about 2%, as compared to 8% when positions are unrestricted. With the two-cent mispricing uncertainty described earlier, the loss drops to 65 basis points, compared to about 2% in the unrestricted case. (Essentially the same statements apply for the Fama–French investor who must hold the portfolio of a Daniel–Titman believer.) These results seem especially noteworthy, given that the asset universe is constructed to exploit differences between these two models.

We also examine another asset universe that is similar to the first, except that the three-way-sorted portfolios enter individually instead of being paired in long-short spreads. Margin requirements give rise to a striking result. Under the typical 50% margin requirement, an investor with complete confidence in the accuracy of the Fama–French model bears substantial risk but allocates nothing to that model's three benchmark positions. Of course, without margin requirements, those benchmarks constitute the entire risky portion of that investor's portfolio. This example illustrates a general point that, for an investor facing constraints, a set of benchmarks can be correct for pricing but not for investing.

With the alternative asset universe, it is less straightforward to represent a prior belief in a characteristic-based model, since the simple zero-expectation implication for the spread positions no longer applies. We develop an alternative approach, applicable more generally, that requires the size and book-to-market characteristics of each portfolio. A comparison of the optimal portfolios chosen by investors with beliefs in different pricing models gives results similar to those of the preceding analysis, except that the differences between this alternative characteristic-based representation and the Fama–French model are no longer eliminated completely by imposing margin requirements.

The remainder of the study is organized as follows. Section 2 discusses mean-variance portfolio optimization in the presence of margin requirements and presents examples using the second asset universe discussed above. The moments of the return distribution for that illustration are taken as known and constructed assuming the Fama–French model holds exactly, with sample moments replacing unknown quantities. Section 3 incorporates parameter uncertainty, including uncertainty about a model's pricing ability. The Bayesian approaches we develop for this purpose are described, and the comparison of portfolio choices across pricing models is conducted using the first asset universe, containing the characteristic-paired spreads. Section 4 reports similar results for the second universe, after describing the alternative characteristic-based approach developed to accomodate a more general set of non-paired portfolios. Section 5 discusses briefly a potential extension to a model-uncertainty setting in which an investor assigns probabilities to multiple models in making a portfolio choice. Section 6 reviews the study's conclusions.

2. Portfolio choice under investment constraints

Define spread position i, established at the end of period $t-1$, as a purchase of one asset coupled with a short sale of an equal amount of another. The two assets are denoted as L_i and S_i, and their rates of return in period t are denoted as $R_{L_i,t}$ and $R_{S_i,t}$. The spread position involves at least one risky asset, which, without loss of generality, is designated as asset L_i. Asset S_i can be either risky or riskless. The investment universe consists of a riskless asset plus n such spread positions, and we assume that some amount of margin capital is required to establish each position. Consider a spread position of size X_i with a dollar payoff equal to $X_i(R_{L_i,t} - R_{S_i,t})$, where X_i can be positive or negative. For a specified $c > 0$, the margin requirements are as follows. If asset S_i is risky, then establishing a spread position of size X_i requires $(2/c)|X_i|$ dollars of capital. If asset S_i is riskless, then establishing a spread position of size X_i requires $(1/c)|X_i|$ dollars of capital.

The total capital required to establish the spread positions is less than the investor's wealth, W_{t-1}. That is,

$$\sum_{i\in\Lambda} (2/c)|X_i| + \sum_{i\notin\Lambda} (1/c)|X_i| \leqslant W_{t-1}, \tag{1}$$

where Λ denotes the set of positions in which S_i is risky, or

$$\sum_{i\in\Lambda} 2|w_i| + \sum_{i\notin\Lambda} |w_i| \leqslant c, \tag{2}$$

where $w_i \equiv X_i/W_{t-1}$. In other words, c is the maximum permitted total value of risky long and short positions per dollar of the investor's wealth. A value of $c = 2$, for example, corresponds to the 50% margin requirement for common stocks specified by Regulation T.

The amount of total wealth in excess of the margin capital required to establish the n spread positions is invested in the riskless asset, earning rate $R_{f,t}$, and we assume that the margin capital also earns that rate. The rate of return on the total portfolio is then given by

$$R_{p,t} = \frac{\sum_{i=1}^{n} X_i(R_{L_i,t} - R_{S_i,t}) + W_{t-1}R_{f,t}}{W_{t-1}}, \tag{3}$$

so the excess portfolio return is simply

$$R_{p,t} - R_{f,t} = \sum_{i=1}^{n} w_i(R_{L_i,t} - R_{S_i,t}). \tag{4}$$

We also assume a common borrowing and lending rate, so that $R_{S_i,t} = R_{f,t}$ for all of the spread positions in which S_i is riskless ($i \notin \Lambda$). In essence, the proceeds from \$1 of short sales can be invested in \$1 of long positions, but an interest-bearing margin deposit of $2/c$ dollars is required. Equivalently, each dollar of a long risky position can be financed by borrowing $[1 - (1/c)]$ dollars, each dollar of a short position requires a margin deposit of $(1/c)$ dollars, and interest is earned on margin deposits as well as the proceeds of short sales.

Note that assuming margin capital earns the riskless rate, as stated above, does not imply that the investor's portfolio contains cash (the riskless asset). The cash position is given by

$$X_f = W_{t-1} - \sum_{i \notin \Lambda} X_i, \tag{5}$$

which can be zero or negative (representing borrowing) if there are some positions in which S_i is riskless, i.e., if the investment universe does not consist solely of spreads in which both legs are risky. For example, suppose $n = 2$ and the first position has common stocks constituting assets L_1 and S_1 but the second position has common stocks in only L_2, so asset S_2 is riskless. Let $W_{t-1} = 100$, $X_1 = 50$, and $X_2 = 100$. The portfolio contains no cash and has stock positions of 150 long and 50 short, or 200 in total, thereby meeting exactly a 50% margin requirement ($c = 2$). In essence, 100 in interest-bearing margin capital supporting positions X_1 (requiring 50) and X_2 (requiring 50) is offset by a 100 short position in cash implied by X_2.

Let w denote the n-vector with ith element w_i. The investor is assumed to choose w so as to maximize the mean-variance objective function

$$U = \mathrm{E}\{R_{p,t}\} - \tfrac{1}{2}A \,\mathrm{Var}\{R_{p,t}\}, \tag{6}$$

subject to the constraint in (2), where A is interpreted as the coefficient of relative risk aversion. Let r_t denote an n-vector with ith element $r_{i,t} \equiv R_{L_i,t} - R_{S_i,t}$, and denote the mean vector and variance–covariance matrix of r_t as E and V. Then the optimal portfolio choice w can be rewritten as the solution to

$$\max_{w} \; (w'E - \tfrac{1}{2}Aw'Vw)$$

$$\text{s.t.} \sum_{i \in \Lambda} 2|w_i| + \sum_{i \notin \Lambda} |w_i| \leqslant c. \tag{7}$$

When c is infinite, so that there is no margin requirement, it is well known that the solution to (7) is given by

$$w = \frac{1}{A} V^{-1} E, \tag{8}$$

which gives the usual tangent portfolio of risky assets. That is, if each risky 'asset' i consists of the zero-investment spread position i plus an investment in the riskless asset, then w in (8) is proportional to the weights in the risky portfolio having the maximum Sharpe ratio (expected excess return divided by standard deviation). When c is finite, the solution to (7) need not produce a portfolio with the maximum Sharpe ratio.[2]

In the next section, E and V are replaced by moments of Bayesian predictive distributions corresponding to varying degrees of prior confidence in alternative pricing models. Before proceeding to that analysis, we use a simpler specification of E and V to illustrate the potential effects of investment constraints on portfolio choice. Suppose expected asset returns are known to be given exactly by, say, the three-factor model of Fama and French. That is, the expected payoff on spread position i is given by

$$\mathrm{E}\{r_{it}\} = \beta_{1,i}\mathrm{E}\{\mathrm{MKT}_t\} + \beta_{2,i}\mathrm{E}\{\mathrm{SMB}_t\} + \beta_{3,i}\mathrm{E}\{\mathrm{HML}_t\}, \tag{9}$$

where MKT_t is the excess return on a value-weighted market index, SMB_t is the difference in returns between small and large firms, and HML_t is the difference in returns between firms with high and low book-to-market ratios. Assume that, in the universe of n spread positions, those benchmark positions are included as

---

[2] Lintner (1965) observes that paying interest on margin deposits and short-sale proceeds gives rise to a case in which (4) holds and allocations obey the restriction $\sum_{i=1}^{n}|w_i| \leqslant 1$ (he assumes 100% margin and all $i \notin \Lambda$). Lintner allows unlimited borrowing and lending at the riskless rate, however, and therefore maximizes the Sharpe ratio. As he observes, the solution in that case simply amounts to rescaling w in (8) to satisfy the constraint.

the last three, corresponding to the second subset in the following partition:

$$E = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}, \qquad V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}. \tag{10}$$

The vector of betas for each of the first $(n-3)$ non-benchmark positions, $(\beta_{1,i}\ \beta_{2,i}\ \beta_{3,i})$, is a row of the $(n-3)\times 3$ matrix

$$B = V_{12}V_{22}^{-1}, \tag{11}$$

and (9) implies

$$E_1 = BE_2. \tag{12}$$

When $c = \infty$, Eqs. (8), (11), and (12) give

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{A}\begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}^{-1}\begin{bmatrix} V_{12}V_{22}^{-1}E_2 \\ E_2 \end{bmatrix} = \frac{1}{A}\begin{bmatrix} 0 \\ V_{22}^{-1}E_2 \end{bmatrix}. \tag{13}$$

That is, the optimal combination involves only the three benchmark positions, and w_2 is proportional to the weights in the tangent portfolio corresponding to those positions. When c is finite, however, the solution to (7) can yield $w_1 \neq 0$ and, as demonstrated below, it can even yield $w_2 = 0$.

We present here an example with a universe containing $n = 11$ risky positions, the last three of which are the Fama–French (FF) benchmark positions, SMB, HML, and MKT. The eight non-benchmark positions are selected from a larger universe of 27 equity portfolios, constructed in essentially the same manner as those in Davis et al. (2000). At the end of June of year t, all NYSE, AMEX, and NASDAQ stocks in the intersection of the CRSP and Compustat files are sorted on market capitalization ('size') and assigned to three categories. The same stocks are also assigned to three categories in an independent sort on the ratio of book value of equity to market capitalization ('book-to-market'). There are equal numbers of NYSE stocks in each of the three size categories as well as the three book-to-market categories. The intersection of these categories produces nine groups of stocks. The stocks within a group are then sorted by HML beta and assigned to one of three subgroups containing equal numbers of stocks. Using up to 60 months of data through December of year $t-1$, the 'pre-formation' HML betas are computed in a regression of the stock's excess returns on 'fixed-weight' versions of the FF factors, which hold the weights on the constituent stocks constant at their June-end values of year t. (The latter procedure, suggested by Daniel and Titman, is designed to increase the dispersion in the 'post-formation' betas of the resulting portfolios.) This three-way grouping procedure produces 27 value-weighted portfolios, which we identify by a combination of three letters, designating increasing values of size (S, M, B), book-to-market (L, M, H), and HML beta (l, m, h). For example, portfolio SHh contains stocks with the smallest size (S), highest book-to-market

(H), and highest HML beta (h). The eight non-benchmark positions form the subset of the 27 portfolios with only the high and low (i.e., no medium) values of size, book-to-market, and HML beta. Each portfolio is combined with a short position in the riskless asset to construct the spread positions consistent with the framework presented earlier.

Values of E and V are constructed to satisfy exact Fama–French three-factor pricing. The values of V and E_2 are set equal to sample estimates based on monthly returns from July 1963 through December 1997, and the 8×1 vector E_1 is then specified using (11) and (12).[3] The value of A in (7) is set to 2.83, which is the value that results in an unconstrained allocation of all wealth to MKT when that is the only risky position available, i.e., the investor chooses neither to borrow nor lend. For the sample estimates used here, the optimal market allocation is actually 101%; the corresponding allocation is exactly 100% in the next section, which accounts for estimation risk.

Table 1 reports, for different values of c, the optimal position sizes (X_i's) per \$100 of invested wealth ($W_{t-1} = 100$). The 'cash' row reports the overall (net) allocation in (5). In the unconstrained case ($c = \infty$), the optimal allocation calls for borrowing 71 and for position sizes of about 64, 358, and 171 in SMB, HML, and MKT. The fact that no allocations are made to any of the non-benchmark positions ($w_1 = 0$) is consistent with the unconstrained solution in (5). Recall that a position size of 358 in HML implies that, for each \$100 of invested wealth, the investor establishes a \$358 long position in the FF 'high' book-to-market stocks along with a \$358 short position in the 'low' book-to-market stocks, while a position of 171 in MKT implies only a one-way (long) position in a risky asset. The sum of the unconstrained risky positions, long and short, is equal to $2(64) + 2(358) + 171 = 1{,}015$, a bit more than ten times wealth. Thus, the constraint on w binds slightly at $c = 10$, as evidenced by the small nonzero allocations to the non-benchmark positions. The investment constraint in (7) treats the n assets as indivisible, which ignores the fact that they are constructed as portfolios of stocks, and a given stock can appear in a non-benchmark portfolio as well as in each of the FF three benchmarks. A different (more complicated) constraint applies if the latter fact is incorporated. If, for example, all of the underlying stock positions are held with a single broker, then short and long positions in the same stock cancel each other, reducing the margin required.

For smaller values of c, properties of the optimal portfolio are affected significantly by the constraints on w. The standard deviations of the return on

[3] To obtain the size and book-to-market characteristics of the Fama–French factors required in the later analysis in Section 4, the factor portfolios are reconstructed following the procedure in Fama and French (1993). The returns on those portfolios are used in Section 4 to ensure the compatibility of factor returns and characteristics, and they are used throughout the rest of the paper for uniformity. The reconstructed factors are virtually identical to the original Fama–French series, supplied generously by Ken French.

Table 1
Optimal allocations under the Fama–French model

The table reports optimal allocations (position sizes) per \$100 of wealth for a mean-variance-optimizing investor with relative risk aversion equal to 2.83. The maximum value of risky positions that can be established per dollar of wealth is denoted by c. Sample estimates based on monthly returns from July 1963 through December 1997 are used to specify the expected payoffs on the Fama–French benchmark positions (SMB, HML, and MKT) as well as all betas, variances, and covariances. The expected payoffs on the first eight (non-benchmark) positions obey exact Fama–French pricing. The risky components of those positions are a subset of value-weighted portfolios constructed by a three-way sort and identified by a combination of three letters, designating increasing values of size (S, M, B), book-to-market (L, M, H), and HML beta (l, m, h). The 'cash' row reports the overall amount invested in the riskless asset, including the amounts implied by the spread positions in which the second asset is riskless, which are the first eight positions and MKT. (Negative amounts represent borrowing.) Also reported are the annualized mean and standard deviation (std) of the portfolio's return as well as the correlation ($\rho_{c,\infty}$) between the returns on the overall portfolio with and without the constraint.

	c			
	2	5	10	∞
SLl	− 8.1	− 0.6	0.0	0.0
SLh	0.0	0.0	0.2	0.0
SHl	0.0	0.0	1.4	0.0
SHh	76.0	64.3	1.9	0.0
BLl	− 37.1	0.0	− 0.3	0.0
BLh	0.0	0.0	0.0	0.0
BHl	6.6	33.5	0.5	0.0
BHh	72.1	41.8	0.3	0.0
SMB	0.0	0.0	59.9	64.1
HML	0.0	179.8	354.6	357.7
MKT	0.0	0.0	166.4	170.7
Cash	− 9.6	− 39.1	− 70.4	− 70.7
Mean	17.4	25.7	30.8	30.8
Std	21.8	28.1	33.0	33.0
$\rho_{c,\infty}$	0.86	0.98	1.00	1.00

the optimal portfolio are lower under the constraints: 21.8% per annum with $c = 2$ versus 33.0% with $c = \infty$. The correlation between the return on the unconstrained portfolio and the return on the constrained portfolio (denoted $\rho_{c,\infty}$) is equal to 0.86 for $c = 2$. For $c = 5$, HML is the only FF benchmark receiving a nonzero allocation (of 180), and SHh, BHl, and BHh receive allocations of 64, 34, and 42. For $c = 2$, none of the three FF benchmarks enter the optimal portfolio, while SHh, BHl, and BHh receive allocations of 76, − 37, and 72. To reiterate, with a 50% margin requirement, the FF benchmark positions receive *zero weight* in this investment universe, even though E and V conform *exactly* to FF pricing.

When $c = 2$, substantial long positions are taken in SHh and BHh, whose HML betas are large (0.80 and 0.84, respectively), since they both contain high book-to-market stocks with high HML betas. The portfolios of low book-to-market stocks with low HML betas, SLl and BLl, both receive short positions, and their HML betas are -0.42 and -0.65, respectively. In the unconstrained case, recall that the optimal portfolio includes a positive exposure of 358 to HML. With $c = 2$, that exposure can be no more than 100, even with all other exposures set to zero. For the constrained investor who believes in Fama–French pricing, going long and short the non-benchmark portfolios essentially provides an alternative path to high HML exposure that makes better use of the permitted overall risky-asset position. That investor largely avoids the long positions in the stocks with medium and low HML betas present in the long (H) leg of HML as well as the short positions in the stocks with the high and medium HML betas present in the short (L) leg of HML. As a result, the HML beta of the constrained optimal portfolio is 1.51, which exceeds the value of 1.00 obtainable by allocating only to HML. This example illustrates the point that, for a constrained investor, a set of benchmarks that is correct for pricing need not be correct for investing.

The example presented here simply uses sample estimates of E_2 and V. Of course, an investor who relies on a finite sample of data remains uncertain about the true values of those parameters. Moreover, E_1 obeys exact three-factor FF pricing in this example, whereas an investor might be uncertain about whether any given pricing model holds exactly. In the next section, we incorporate parameter uncertainty, or 'estimation risk', which includes this potential mispricing uncertainty.

3. Comparing investments under parameter uncertainty

Our objective is to compare pricing models in terms of their implications for portfolio choice. Specifically, for a given investment universe containing cash plus n risky positions, we compare the portfolios selected by investors who base their prior beliefs on three different pricing models. Two are risk-based, the FF model and the CAPM, and relate expected returns to betas on one or more risk factors. In the Daniel–Titman (DT) characteristic-based model, expected returns depend on size and book-to-market, not betas.

To incorporate uncertainty about parameter values, including mispricing uncertainty, we apply Bayesian methods.[4] Recall that r_t is the n-vector with ith

[4] Early applications of Bayesian methods to portfolio choice include Zellner and Chetty (1965), Klein and Bawa (1976), and Brown (1979). Bayesian posterior distributions of measures of portfolio inefficiency are analyzed by Shanken (1987), Kandel et al. (1995), and Wang (1998), who incorporates short sale restrictions.

element $R_{L_i,t} - R_{S_i,t}$. We assume that r_t is drawn independently across t from a multivariate normal distribution with unknown parameters E and V. An investor has prior beliefs $p(E, V)$, shaped in part by a prior belief about the accuracy of a given pricing model. The investor forms posterior beliefs $p(E, V|R)$, based on the data $\{R: r_t, t = 1, \ldots, T\}$, and forms the predictive distribution for r_{T+1},

$$p(r_{T+1}|R) = \int_E \int_V p(r_{T+1}|R, E, V)\, p(E, V|R)\, \mathrm{d}E\, \mathrm{d}V. \tag{14}$$

The investor then solves (7) with E and V replaced by E^* and V^*, the moments of the predictive distribution in Eq. (14). As detailed in the Appendix, E^* and V^* are obtained analytically for the risk-based models and through Gibbs sampling for the characteristic-based model. The prior beliefs about E and V are discussed below. The prior beliefs about V are noninformative, and since the monthly data are fairly informative about second moments, the predictive covariance matrix V^* is quite similar across different pricing models and degrees of mispricing uncertainty.

3.1. Framework

It is useful to cast the problem in a regression setting. Let $r_t = (r'_{1,t}\ r'_{2,t})'$, following the same partitioning applied to E and V in Eq. (10). That is, $r_{2,t}$ contains the payoffs on k benchmark positions from a factor-based model, and $r_{1,t}$ contains the payoffs on m $(= n - k)$ non-benchmark positions. Consider the multivariate regression,

$$r_{1,t} = \alpha + Br_{2,t} + u_t, \tag{15}$$

where u_t obeys a multivariate normal distribution with mean zero and variance-covariance matrix equal to Σ. In this regression framework, the set of parameters (E, V) is replaced by $(\alpha, B, \Sigma, E_2, V_{22})$, where B is defined in Eq. (11),

$$\alpha = E_1 - BE_2, \tag{16}$$

and

$$\Sigma = V_{11} - BV_{22}B'. \tag{17}$$

The factor-based and characteristic-based models imply different restrictions on α, and prior beliefs about α are centered on these restrictions, as will be explained below. The models impose no restrictions on B, Σ, E_2, and V_{22}, so the prior distributions for these parameters are noninformative. The prior distribution for Σ is specified as inverted Wishart,

$$\Sigma^{-1} \sim W(H^{-1}, \nu), \tag{18}$$

with degrees of freedom $\nu = 15$, so that the prior contains only about as much information as a sample of 15 observations (' $\sim$ ' is read 'is distributed as'). From the properties of the inverted Wishart distribution (e.g., Anderson, 1984), the prior expectation of Σ equals $H/(\nu - m - 1)$. We specify $H = s^2(\nu - m - 1)I_m$, so that $\mathrm{E}(\Sigma) = s^2 I_m$. Following an 'empirical Bayes' approach, the value of s^2 is set equal to the average of the diagonal elements of the sample estimate of Σ. The joint prior distribution for the remaining parameters (B, E_2, V_{22}) is assumed to be diffuse and independent of α and Σ.

The CAPM, in which $k = 1$, and the three-factor FF model, in which $k = 3$, are treated in the same manner. Under the CAPM, the payoffs on the two nonmarket positions SMB and HML are simply included in $r_{1,t}$, which then has $n - 1$ elements, whereas those payoffs are included in $r_{2,t}$ under the FF model, so $r_{1,t}$ then has only $n - 3$ elements. The factor-based pricing restriction in Eq. (12) is equivalent to $\alpha = 0$. To allow for mispricing uncertainty, the prior distribution for α is specified as a normal distribution,

$$\alpha|\Sigma \sim \mathrm{N}\left(0,\ \sigma_\alpha^2\left(\frac{1}{s^2}\Sigma\right)\right). \tag{19}$$

The unconditional prior variance of each element of α, σ_α^2, reflects the investor's prior degree of mispricing uncertainty. When $\sigma_\alpha = 0$, the investor believes dogmatically in the model, and mispricing is ruled out completely. When $\sigma_\alpha = \infty$, the investor regards the model as useless, since mispricing is completely unrestricted. Pástor and Stambaugh (1999) introduce this measure of mispricing uncertainty.

Observe that the conditional prior covariance matrix of α is proportional to Σ. This specification is motivated by the recognition that there exist portfolios with high Sharpe ratios if the elements of α are large when the elements of Σ are small. When w is unrestricted (no margin requirements), the maximum squared Sharpe ratio from the universe containing the n positions is given by

$$S_n^2 = S_k^2 + \alpha'\Sigma^{-1}\alpha, \tag{20}$$

where S_k^2 is the maximum squared Sharpe ratio from the k benchmark positions. The importance of bounding $\alpha'\Sigma^{-1}\alpha$ as n grows large arises in the arbitrage-pricing literature (e.g., Ingersoll, 1984). For finite values of n, MacKinlay (1995) demonstrates empirically the importance of a positive association between α and Σ in reducing the value of S_n. With no mispricing in the factor-based model, $S_n^2 = S_k^2$. The prior in (19) reflects a belief that, even with some mispricing, the risk-based nature of the model makes large values of $S_n^2 - S_k^2$ less likely than under non-risk-based alternatives. For a given σ_α, large values of $S_n^2 - S_k^2$ receive lower prior probabilities under (19) than when each element of α has standard deviation σ_α but is distributed independently of all other parameters. Pástor and Stambaugh (1999) introduce the same type of prior for a single element of α, and Pástor (2000) applies the multivariate version in (19) to portfolio-choice problems. MacKinlay (1995) develops a functional relation between α and Σ when the mispricing is entirely risk based, i.e., when α is proportional to sensitivities to a risk factor not included in $r_{2,t}$. MacKinlay and Pástor (2000) use maximum likelihood procedures to investigate the implications of this relation for estimation of expected returns and for portfolio selection.

The investment universe contains n risky positions, and in this section the last three are again the FF benchmarks (as in Section 2). Recall that exact FF pricing implies the restriction on E_1 in Eq. (12). The DT model yields a simple alternative restriction on E_1 if the non-benchmark positions are constructed in a particular fashion. For this purpose, we construct a set of nine non-benchmark positions as beta spreads within categories of size and book-to-market. For a given joint classification of size and book-to-market, the payoff on each spread, per \$1 of position size, is produced by going long \$1 of the portfolio with the lowest HML beta and short \$1 of the portfolio with the highest HML beta. For example, SH(l-h) denotes the spread position that is long portfolio SHl and short portfolio SHh. (These portfolios are described in the previous section.) The key to such spreads, proposed by Daniel and Titman, is that the characteristics of the long and short positions are (approximately) the same. Thus, we represent the DT model by the restriction $E_1 = 0$, which is equivalent to

$$\alpha = -BE_2. \tag{21}$$

To allow for mispricing uncertainty, we assume

$$\alpha|\Sigma \sim \mathrm{N}(-BE_2, \sigma_\alpha^2 I_m). \tag{22}$$

In contrast to the risk-based specification in (19), Σ does not appear in (22). The difference $S_n^2 - S_k^2$ is positive even when (21) holds exactly (using (20) and assuming $BE_2 \neq 0$), and this non-risk-based model provides no reason to limit this difference if (21) is violated ($\sigma_\alpha > 0$). Thus, α and Σ are made independent in the prior for the DT model. The three elements of E_2 are left unrestricted, with the rationale that at least three unknown parameters would relate those values to size and book-to-market (e.g., an intercept and two slopes).

The nine non-benchmark positions are formed as spreads between portfolios selected from the set produced by a three-way sort on size, book-to-market, and HML beta. As demonstrated earlier, margin requirements can induce an investor to select positions that are extreme in terms of their betas or characteristics. Grouping stocks into 27 portfolios, following Davis et al. (2000), essentially maximizes the differences among the extreme portfolios, given the limitations of the data (avoiding portfolios containing very few or no stocks during the early years of the sample). The construction of the investment universe is likely to be important to our analysis, especially in the presence of margin requirements, and in Section 4 we explore the robustness of our results to the choice of investment universe.

In each of the nine non-benchmark spreads, the long and short legs are assumed to be matched in terms of their size and book-to-market characteristics. Davis et al. (2000) observe that, despite this objective of the sorting procedure, the short leg of each spread exhibits a tendency to contain stocks with slightly higher book-to-market ratios than the stocks in the long leg. The reason is that there remains some positive correlation between book-to-market ratios and HML betas within the nine portfolios formed by a two-way sort on size and book-to-market. (Recall that the spreads are long low HML betas and short high HML betas.) Under a DT model in which expected return is increasing in the book-to-market characteristic, that mismatch implies that the elements of E_1 are slightly less than zero, as opposed to the equality implying (21). The elements of BE_2 are likely to be negative (the HML betas are negative by construction), so under the DT model a slight book-to-market mismatch in the spreads would require that the prior mean of α ($= E_1 - BE_2$) be somewhat closer to the zero vector than is $-BE_2$, the prior mean used in (22) corresponding to $E_1 = 0$. In other words, accounting for the mismatches would move the prior for the DT model closer to that of the FF model. As detailed below, we find that the differences in portfolio choices implied by the DT and FF models are often made small or nonexistent by incorporating mispricing uncertainty and margin requirements. Accounting for slight characteristic mismatches in the spreads would tend to further reduce those differences.

Black and Litterman (1991, 1992) present a portfolio-selection framework in which CAPM-implied expected returns are combined with subjective beliefs about violations of that model. In that sense, their approach also centers beliefs about expected returns on a pricing model, and the divergence of the investor's portfolio from the market benchmark depends on the strengths of the investor's beliefs about model violations. Black and Litterman define no role for sample information about expected returns; beliefs about model violations are assigned means and variances a priori. In contrast, our approach relies on the strength of the sample's information about violations of the pricing model, as well as the investor's prior confidence in the model, to determine how the investor's portfolio departs from the strict implication of the model.

3.2. Results

Table 2 reports optimal allocations per \$100 of wealth when prior beliefs are centered on each of the three pricing models, with varying degrees of mispricing uncertainty (σ_α). As mispricing uncertainty increases, optimal allocations must

Table 2
Optimal allocations under parameter uncertainty

The table reports optimal allocations (position sizes) per $100 of wealth for a mean-variance-optimizing investor with relative risk aversion equal to 2.83. The maximum value of risky positions that can be established per dollar of wealth is denoted by c. Mispricing uncertainty, denoted by σ_α, is the prior standard deviation of the difference between each position's annualized expected payoff and the pricing model's exact implication, expressed as a percentage of initial position size. Allocations are reported for prior beliefs centered on three different pricing models: the three-factor Fama–French model (FF), the Daniel–Titman characteristic-based model (DT), and the CAPM (CM). Optimization is based on the predictive distribution, obtained by updating the prior beliefs using monthly returns from July 1963 through December 1997. The Fama–French benchmark positions are denoted by SMB, HML, and MKT. The other nine positions are spreads between value-weighted portfolios constructed by a three-way sort and identified by a combination of three letters, designating increasing values of size (S, M, B), book-to-market (L, M, H), and HML beta (l, m, h); the portfolios on each side of a spread differ in HML beta (high vs. low) but are matched in size and book-to-market. The 'cash' row reports the overall amount invested in the riskless asset, which includes the amounts implied by the position MKT (which shorts the riskless asset). Negative cash amounts represent borrowing. Also reported are the annualized mean and standard deviation (std) of the portfolio's return with respect to the given predictive distribution.

	$\sigma_\alpha = 0$			$\sigma_\alpha = 1\%$			$\sigma_\alpha = 2\%$			$\sigma_\alpha = \infty$ (all)
	DT	FF	CM	DT	FF	CM	DT	FF	CM	
					$c = 2$					
SL(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SM(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SH(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ML(l-h)	0.0	0.0	0.0	0.0	0.0	− 12.7	0.0	0.0	− 12.3	− 11.9
MM(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MH(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BL(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BM(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BH(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SMB	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HML	62.1	62.2	0.0	62.1	62.2	35.1	62.3	62.2	41.8	51.0
MKT	75.8	75.6	100.0	75.8	75.6	104.3	75.4	75.6	91.8	74.1
Cash	24.2	24.4	0.0	24.2	24.4	−4.3	24.6	24.4	8.2	25.9
Mean	8.0	8.0	6.3	8.0	8.0	6.8	8.0	8.0	7.1	7.9
Std	10.5	10.5	15.0	10.5	10.5	14.6	10.5	10.5	12.7	10.3

					$c = 5$					
SL(l-h)	0.0	0.0	0.0	0.0	0.0	10.0	0.0	0.0	3.0	0.0
SM(l-h)	0.0	0.0	0.0	0.0	0.0	3.4	0.0	0.0	0.0	0.0
SH(l-h)	0.0	0.0	0.0	0.0	0.0	28.6	0.0	0.0	13.3	0.0
ML(l-h)	0.0	0.0	0.0	0.0	0.0	−33.8	0.0	−8.2	−46.1	−50.0
MM(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MH(l-h)	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0
BL(l-h)	0.0	0.0	0.0	0.0	0.0	17.0	0.0	0.0	0.0	0.0
BM(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BH(l-h)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SMB	0.0	0.0	0.0	0.0	0.0	15.1	0.0	0.0	4.4	0.0
HML	189.2	189.3	0.0	189.2	189.3	82.3	189.5	181.3	119.4	141.2
MKT	121.6	121.4	100.0	121.6	121.4	119.0	121.1	120.9	127.7	117.6
Cash	−21.6	−21.4	0.0	−21.6	−21.4	−19.0	−21.1	−20.9	−27.7	−17.6
Mean	− 17.4	17.4	6.3	17.4	17.4	8.1	17.4	17.3	12.1	17.2
Std	19.4	19.4	15.0	19.4	19.4	16.4	19.4	19.2	18.5	18.6
					$c = 10$					
SL(l-h)	0.0	0.0	0.0	0.0	3.9	15.2	4.2	8.6	22.3	15.0
SM(l-h)	0.0	0.0	0.0	0.0	0.0	18.4	0.0	0.0	5.8	0.0
SH(l-h)	41.3	0.0	0.0	50.5	13.1	43.0	59.1	29.1	64.4	50.4
ML(l-h)	0.0	0.0	0.0	0.0	−24.3	−45.2	−4.5	−54.3	−76.7	−91.9
MM(l-h)	0.0	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0	0.0
MH(l-h)	0.0	0.0	0.0	0.0	0.0	7.4	0.0	0.0	0.3	0.0
BL(l-h)	57.6	0.0	0.0	47.8	2.0	34.6	31.5	4.3	37.2	8.3
BM(l-h)	0.0	0.0	0.0	0.0	0.0	−8.3	0.0	0.0	0.0	0.0
BH(l-h)	0.0	0.0	0.0	0.0	0.0	−5.0	0.0	0.0	0.0	0.0
SMB	15.6	63.2	0.0	16.1	51.7	22.6	16.5	37.6	33.7	20.3
HML	308.2	352.5	0.0	307.6	322.3	108.2	305.8	285.2	187.3	236.2
MKT	154.9	168.2	100.0	156.0	165.3	124.9	156.9	161.9	144.6	155.5
Cash	−54.9	−68.2	0.0	−56.0	−65.3	−24.9	−56.9	−61.9	−44.6	−55.5
Mean	26.0	30.4	6.3	26.0	28.7	8.7	26.0	27.5	15.7	26.9
Std	25.2	32.8	15.0	25.3	30.8	17.6	25.6	28.6	21.9	26.3

Table 2 (*continued*)

	$\sigma_\alpha = 0$			$\sigma_\alpha = 1\%$			$\sigma_\alpha = 2\%$			$\sigma_\alpha = \infty$ (all)
	DT	FF	CM	DT	FF	CM	DT	FF	CM	
					$c = \infty$					
SL(l-h)	17.0	0.0	0.0	37.2	16.5	15.2	55.6	37.5	35.8	63.8
SM(l-h)	60.4	0.0	0.0	76.8	20.0	18.4	81.0	45.4	43.2	78.5
SH(l-h)	178.6	0.0	0.0	177.9	46.5	43.0	181.2	105.4	101.0	181.6
ML(l-h)	62.3	0.0	0.0	− 26.8	− 48.0	− 45.2	− 107.0	− 109.0	− 106.1	− 187.2
MM(l-h)	− 29.1	0.0	0.0	− 1.3	4.1	3.7	14.6	9.3	8.8	15.6
MH(l-h)	35.3	0.0	0.0	32.4	8.0	7.4	33.4	18.2	17.3	30.4
BL(l-h)	179.9	0.0	0.0	179.5	37.8	34.6	168.6	85.7	81.2	146.9
BM(l-h)	3.5	0.0	0.0	− 3.4	− 8.9	− 8.3	− 12.6	− 20.2	− 19.5	− 35.1
BH(l-h)	− 0.8	0.0	0.0	− 2.5	− 5.3	− 5.0	− 7.4	− 12.0	− 11.8	− 20.6
SMB	81.6	63.2	0.0	89.7	71.1	22.6	94.8	81.7	53.1	94.4
HML	613.0	352.5	0.0	579.6	378.5	108.2	537.9	413.8	253.8	455.6
MKT	209.4	168.2	100.0	209.3	178.3	124.9	210.1	192.3	160.1	208.9
Cash	− 109.4	− 68.2	0.0	− 109.3	− 78.3	− 24.9	− 110.1	− 92.3	− 60.1	− 108.9
Mean	46.8	30.4	6.3	45.4	31.5	8.7	45.8	35.9	19.1	46.0
Std	40.7	32.8	15.0	40.1	33.3	17.6	40.2	35.6	26.0	40.3

approach those based on the sample moments of returns, whatever the pricing model. Our aim is to explore the extent to which this behavior occurs at interesting levels of σ_α. Results in Table 2 are reported for $\sigma_\alpha = 1\%$ and $\sigma_\alpha = 2\%$ (per annum) as well as the limiting values $\sigma_\alpha = 0$ (exact pricing) and $\sigma_\alpha = \infty$ (no use of a pricing model).

Given the true values of B and E_2, σ_α represents the prior volatility of the errors in E_1 obtained from the pricing model. In the absence of a pricing model, one might construct E_1 by computing sample means. The volatilities of the errors in the elements of E_1 obtained in that manner provide one benchmark for assessing the magnitude of σ_α. Of course, such volatilities depend on the investment universe. For the universe of spread positions analyzed in this section, the elements of $r_{1,t}$ have annualized standard deviations that average about 10%. Thus, the errors in E_1 specified as sample means on average have a 2% volatility in samples of about 25 years and a 1% volatility in samples of about 100 years. Sections 2 and 4 analyze an investment universe composed largely of non-spread positions, whose annualized standard deviations average about 20%. The sample sizes corresponding to error volatilities of 2% and 1% are then about 100 and 400 years. When viewed in this context, values of σ_α equal to 1%, or even 2%, seem to represent modest degrees of prior mispricing uncertainty for these investment universes.

The first three columns of Table 2, with $\sigma_\alpha = 0$, display the allocations corresponding to dogmatic beliefs in each of the three pricing models. As in Table 1, the row labeled 'cash' includes the overall (net) amount in (5). Recall that a value for risk aversion of $A = 2.83$ implies that all of the investor's wealth is allocated to the market portfolio when that is the only risky position available. With a dogmatic belief in the CAPM (shortened to 'CM' in the table), the market portfolio has the maximum Sharpe ratio, so MKT is again the only risky position with a nonzero allocation for such an investor, and that allocation equals 100 for each value of c considered.

Table 2 presents portfolio allocations for the different levels of investment constraints analyzed previously in Table 1 ($c = 2, 5, 10, \infty$). The most striking result is that, when $c = 2$, the optimal allocations under the FF model are virtually identical to those under the DT model. This is true even for dogmatic beliefs in each model ($\sigma_\alpha = 0$) and remains true for the nonzero values of σ_α. In other words, it makes no difference whether a mean-variance investor who must allocate funds across the 12 risky positions considered here has strong beliefs in the three-factor model or strong beliefs in the characteristic-based model. When constrained by a 50% margin requirement, the optimal portfolio is the same under either model. This result seems noteworthy, since the nine beta-spread positions included here are constructed to exploit differences between the FF and DT models.

For the three finite values of σ_α, the only risky positions receiving non-zero allocations under the DT and FF models when $c = 2$ are HML and MKT, with position sizes of 62 and 76, and 24 is placed in cash. The fact that the HML and MKT allocations are virtually identical under the DT and FF models but different from the allocations under the CAPM can be explained by the predictive means of the factors. With $\sigma_\alpha = 0$, the predictive mean of MKT is around 6.3% per annum across all three models. The predictive mean of HML, however, is 5.1% under the DT and FF models as compared to -1.4% under the CAPM, which prices HML by its MKT beta. With $c = 2$, no funds are allocated to SMB or any of the nine beta-spread positions; the intuition for this outcome is deferred to the discussion of risk aversion in the next subsection. The return on the optimal portfolio has a predictive standard deviation of 10.5% (annualized), substantially lower than those of the optimal portfolios under dogmatic beliefs in each model in the unconstrained case (40.7% for the DT model and 32.8% for the FF model). Note, however, that the investment constraints do not preclude higher standard deviations. For example, a simple two-to-one leveraging of MKT, permitted under $c = 2$, produces a standard deviation of about 30%.

To a large extent, the above observations for $c = 2$ also apply when $c = 5$ (corresponding to a 20% margin requirement). The allocations are again virtually identical under the FF and DT models when $\sigma_\alpha = 0$ and $\sigma_\alpha = 1\%$. As with $c = 2$, HML and MKT are the only risky positions receiving nonzero allocations, although the HML position is now larger than that of MKT (189 versus 121). The optimal portfolios under the DT and FF models diverge slightly at $\sigma_\alpha = 2\%$ only in that ML(l-h) receives a short position of -8 under the FF model. When $\sigma_\alpha = \infty$, the optimal allocation again includes a short position in ML(l-h).

For margin requirements in the range of 20–50%, the above results reveal little or no role for mispricing uncertainty (σ_α) in determining optimal allocations among the opportunities considered. As investment constraints are relaxed, the degree of mispricing uncertainty exerts more influence. When $c = \infty$, the unconstrained case, the nine beta spreads receive zero allocations under dogmatic beliefs ($\sigma_\alpha = 0$) in the CAPM or FF model, as they must. In contrast, some large nonzero positions in those spreads arise under dogmatic beliefs in the DT model, such as the positions of roughly 180 in both SH(l-h) and BL(l-h). As σ_α increases, the allocations change, and the beta-spread positions receive substantial nonzero allocations under all three models. With $\sigma_\alpha = \infty$, where the pricing models are not used at all, the optimal unconstrained portfolio has a number of large long and short positions, such as 182 in SH(l-h), -187 in ML(l-h), 456 in HML, and 209 in MKT. The optimal portfolio also calls for borrowing 109.

Because the payoffs on the 12 risky positions are correlated, large differences in position-by-position allocations need not produce economically significant differences in the overall portfolio characteristics. To gauge the economic importance of such differences across pricing models, we compare certainty-equivalent returns on the portfolios as follows. Let E^* and V^* denote the predictive moments of r_t formed under a given σ_α and a given pricing model. For a given c, we first compute the certainty-equivalent excess return of the allocation w_O that is optimal under that predictive distribution,

$$CE_O = w_O' E^* - \tfrac{1}{2} A w_O' V^* w_O. \tag{23}$$

Then we compute the certainty-equivalent excess return of a suboptimal allocation w_S,

$$CE_S = w_S' E^* - \tfrac{1}{2} A w_S' V^* w_S, \tag{24}$$

where w_S is an allocation that is optimal for the same c and σ_α under the predictive distribution from a different pricing model. The difference $CE_O - CE_S$ provides an economic measure of the difference between the two portfolios. It is the perceived certainty-equivalent loss to an investor with a given degree of belief in one pricing model who is forced to accept the portfolio selection of another investor with the same degree of belief in a different pricing model, where both investors face the same constraints. Kandel and Stambaugh (1996) propose the approach wherein a single predictive distribution is used to compute the certainty equivalents of both portfolios.

Fig. 1 displays the certainty-equivalent losses for an investor who believes in the characteristic-based model with varying degrees of mispricing uncertainty. For each of four values of c, the figure displays a plot of the certainty-equivalent loss versus σ_α. Losses are computed for portfolios from the FF model (solid lines) and from the CAPM (dashed lines). When $c = \infty$ and $\sigma_\alpha = 0$, the certainty-equivalent loss for the FF allocation is about 8% per annum, an economically large magnitude. With mispricing uncertainty, the FF loss drops to about 5% at $\sigma_\alpha = 1\%$ and to less than 2% at $\sigma_\alpha = 2\%$. Note that a loss of 2% for the FF model also occurs with no mispricing uncertainty but with a 10% margin requirement ($c = 10$). Although a 2% loss is still economically significant, it is only one-fourth the size of the loss under the same sets of beliefs but no investment constraints. When $c = 2$ and $c = 5$, the certainty-equivalent losses for the FF portfolios plot as essentially a flat line at zero, since as observed previously, the optimal allocations in the DT and FF models are virtually identical in those cases.

In all cases considered, larger certainty-equivalent losses are associated with the CAPM portfolios. When $c = \infty$ and $\sigma_\alpha = 0$, the loss for the CAPM portfolio is about 20%. Mispricing uncertainty again reduces that loss to about 12% at $\sigma_\alpha = 1\%$ and 4.5% at $\sigma_\alpha = 2\%$, but those values are considerably larger than the corresponding losses for the FF model. Unlike the FF losses, the CAPM losses do not completely disappear at the lower values of c. For example, with $c = 5$ and dogmatic beliefs in the models, the CAPM portfolio produces a loss of almost 9% for an investor who believes in the DT model, whereas the FF portfolio produces no loss for such an investor. Nevertheless, a combination

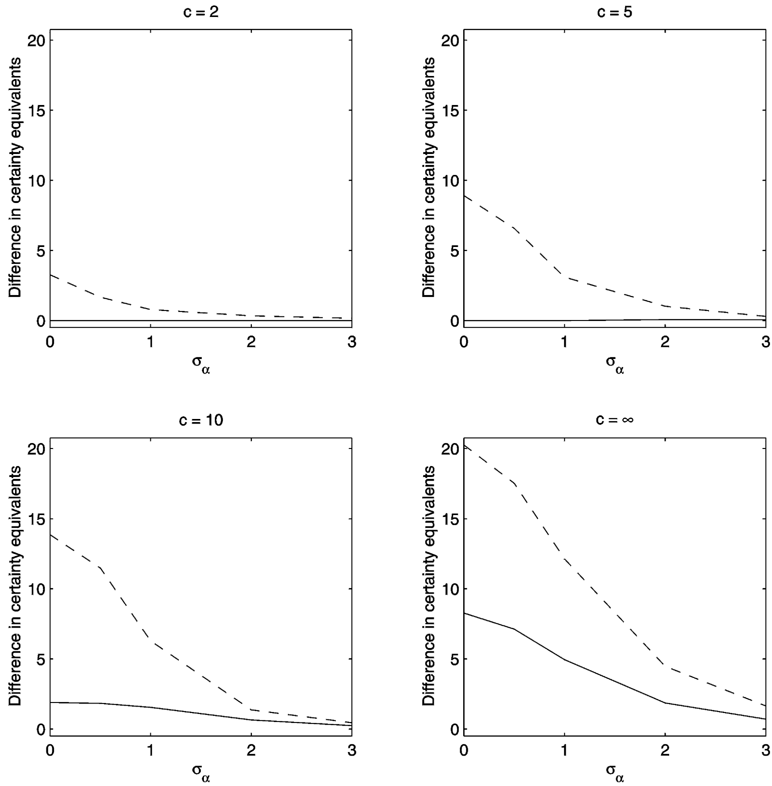

Fig. 1. Certainty-equivalent losses for other models' portfolios from the perspective of a Daniel–Titman investor. The figure displays the certainty-equivalent loss (in % per year) for a mean-variance-optimizing investor whose prior beliefs are centered on the Daniel–Titman characteristic-based model, with mispricing uncertainty σ_α, but who is forced to hold portfolios chosen by investors with the same degree of belief in either the Fama–French model (solid line) or the CAPM (dashed line). Investor risk aversion is set to $A = 2.83$. The maximum value of risky positions that can be established per dollar of wealth is denoted by c. Mispricing uncertainty, denoted by σ_α, is the prior standard deviation of the difference between each position's annualized expected payoff and the pricing model's exact implication, expressed as a percentage of initial position size.

of realistic investment constraints and modest mispricing uncertainty is still sufficient to reduce the losses for the CAPM portfolio to rather low levels. With $c = 2$ and $\sigma_\alpha = 2\%$, the annualized loss for the CAPM allocation is only 33 basis points.

Fig. 2 displays precisely the same analysis except that the certainty-equivalent losses are computed for an investor who believes in the FF model instead of the DT model as in Fig. 1. The dashed lines still represent the losses for the CAPM portfolios, but the solid line now represents the losses associated with the DT

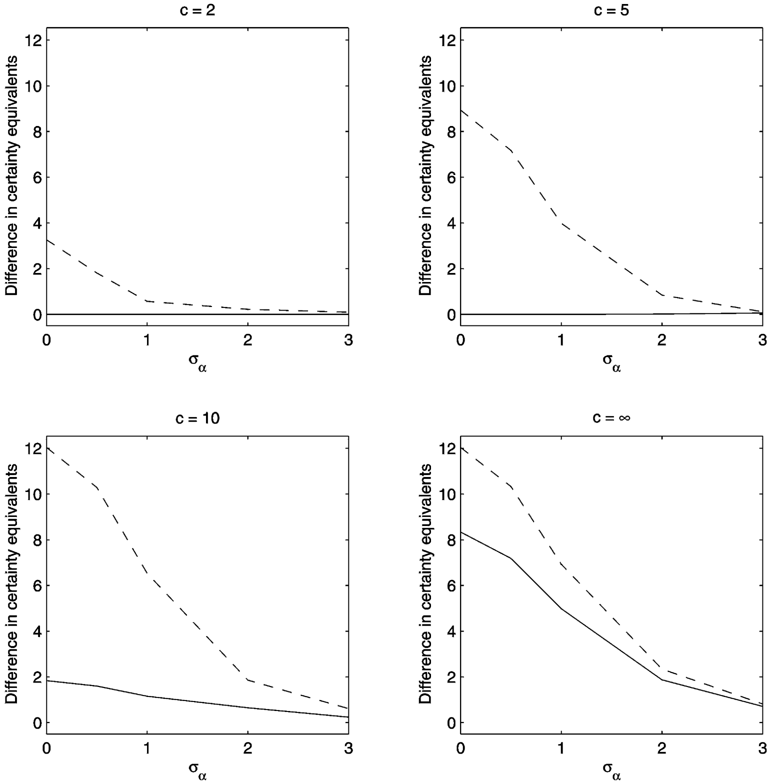

Fig. 2. Certainty-equivalent losses for other models' portfolios from the perspective of a Fama–French investor. The figure displays the certainty-equivalent loss (in % per year) for a mean-variance-optimizing investor whose prior beliefs are centered on the Fama–French factor-based model, with mispricing uncertainty σ_α, but who is forced to hold portfolios chosen by investors with the same degree of belief in either the Daniel–Titman model (solid line) or the CAPM (dashed line). Investor risk aversion is set to $A = 2.83$. The maximum value of risky positions that can be established per dollar of wealth is denoted by c. Mispricing uncertainty, denoted by σ_α, is the prior standard deviation of the difference between each position's annualized expected payoff and the pricing model's exact implication, expressed as a percentage of initial position size.

portfolios. The losses for the DT portfolios perceived by the investor with FF beliefs (Fig. 2) are, in all cases, very close to the losses for the FF portfolios perceived by the investor with DT beliefs (Fig. 1). The magnitudes for some of the CAPM losses are somewhat different, most notably at $c = \infty$ and $\sigma_\alpha = 0$, where the loss in Fig. 2 is 12% versus 20% in Fig. 1. This ordering is perhaps not surprising given that the CAPM and the FF model are both factor-based and have the MKT factor in common. In general, however, the observations made for Fig. 1 are unchanged when based on Fig. 2. Again, in terms of investment implications, the FF and DT models are significantly closer to each other than is either model to the CAPM.

We also compare portfolios by computing the correlation between their returns. As before, let V^* denote the predictive covariance matrix of r_t formed under a given σ_α and a given pricing model. The predictive correlation between the return on the optimal allocation w_O and suboptimal allocation w_S is given by

$$\rho_\mathrm{OS} = \frac{w'_\mathrm{O} V^* w_\mathrm{S}}{\sqrt{w'_\mathrm{O} V^*}} w_\mathrm{O} \sqrt{w'_\mathrm{S} V^* w_\mathrm{S}}, \tag{25}$$

where w_S is again an allocation that is optimal for the same c and σ_α but a different pricing model. These correlations for the DT predictive distribution are displayed in Fig. 3, which follows precisely the same format used to display the certainty-equivalent losses in Fig. 1. (As before, the results based on the FF predictive distribution are close to those based on the DT predictive distribution and are omitted in the interest of space.) The correlations between the FF and DT portfolios plot as flat lines at 1.0 for $c = 2$ and $c = 5$. For $c = 10$, the correlations between the FF and DT allocations are 0.95 or higher. Thus, in the presence of even weak investment constraints, the FF and DT models imply highly correlated optimal portfolios. The correlations between either of those portfolios and the CAPM portfolio are substantially lower, especially in the absence of mispricing uncertainty. For $c = 5$ and $c = 10$, the correlations between the CAPM portfolio and the FF or DT portfolios are 0.6 or less. When $c = \infty$ and $\sigma_\alpha = 0$, the correlation between the FF and DT portfolios is about 0.8, while the correlation of either of those portfolios with the CAPM portfolio is in the vicinity of 0.4. With mispricing uncertainty of $\sigma_\alpha = 2\%$, those same correlations are both around 0.95.

3.3. Risk aversion

The optimal portfolios discussed so far are computed for an investor with a risk aversion coefficient of $A = 2.83$. Recall that this risk aversion implies that all of the investor's wealth is allocated to the market portfolio when that is the only risky position available. Different values of A imply different optimal portfolios and hence different certainty-equivalent losses for allocations from alternative models. For different levels of risk aversion, Fig. 4 displays certainty-equivalent losses analogous to those plotted as solid lines in Fig. 1, where $A = 2.83$. That is, for an investor with a given degree of belief in the DT model, we plot that investor's certainty-equivalent loss when forced to hold the portfolio of an investor with the same degree of belief in the FF model. The four graphs in Fig. 4 display results for risk aversion levels of 1, 5, 10, and 15. Each graph plots the certainty-equivalent loss versus mispricing uncertainty (σ_α) for four levels of investment constraints, $c = 2$, 5, 10, and ∞.

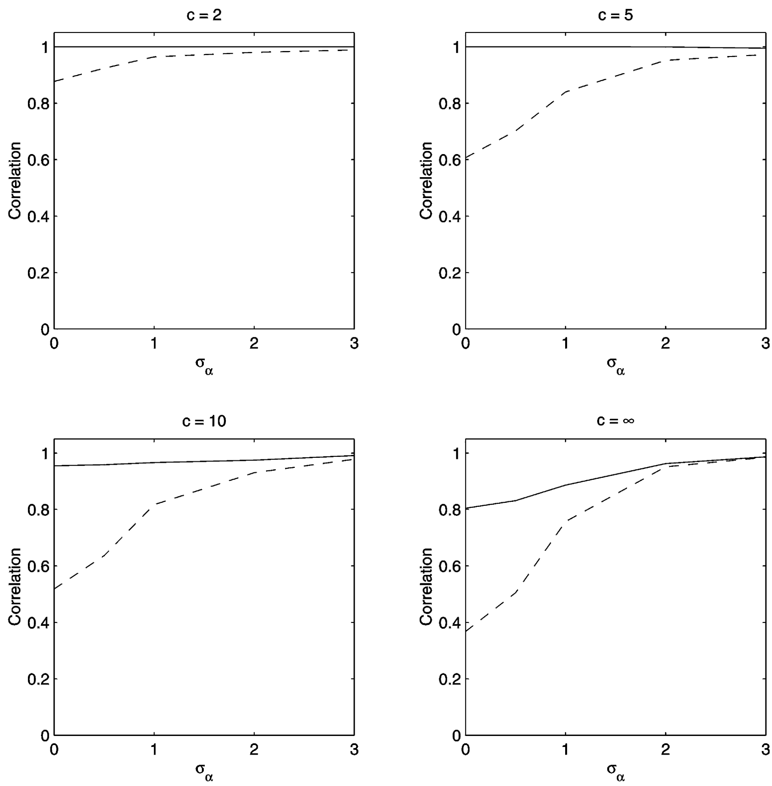

Fig. 3. Correlations with other models' portfolios from the perspective of a Daniel–Titman investor. The figure displays the correlation between the portfolio of a mean-variance-optimizing investor whose prior beliefs are centered on the Daniel–Titman characteristic-based model, with mispricing uncertainty σ_α, and the portfolios chosen by investors with the same degree of belief in either the Fama–French model (solid line) or the CAPM (dashed line). Investor risk aversion is set to $A = 2.83$. The maximum value of risky positions that can be established per dollar of wealth is denoted by c. Mispricing uncertainty, denoted by σ_α, is the prior standard deviation of the difference between each position's annualized expected payoff and the pricing model's exact implication, expressed as a percentage of initial position size.

With no investment constraints ($c = \infty$), a large certainty-equivalent loss is perceived by an investor with dogmatic DT beliefs ($\sigma_\alpha = 0$) who is forced to accept the portfolio of a dogmatic FF investor. When risk aversion $A = 2.83$, the annualized loss is about 8%, as observed previously in Fig. 1. Fig. 4 reveals that the loss is decreasing in A: the loss is roughly 24% for $A = 1$ but less than 2% for $A = 15$. (Note that the vertical scale is different in the $A = 1$ graph.) As A increases, the optimal portfolios from both models involve larger cash positions,

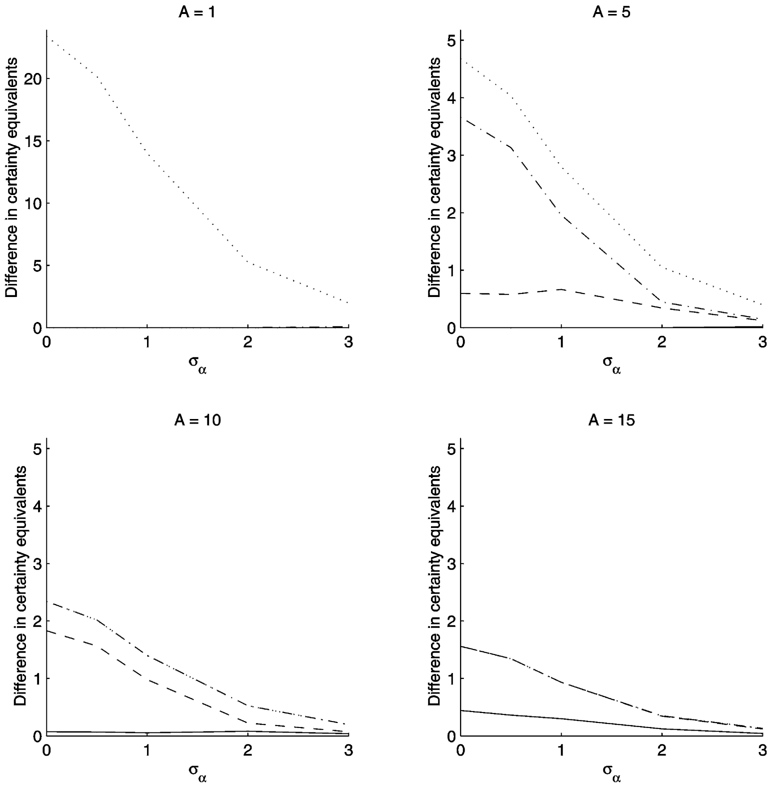

Fig. 4. Effects of risk aversion on the certainty-equivalent losses for the Fama–French model's portfolios from the perspective of a Daniel–Titman investor. The figure displays the certainty-equivalent loss (in % per year) for a mean-variance-optimizing investor whose prior beliefs are centered on the Daniel–Titman characteristic-based model, with mispricing uncertainty σ_α, but who is forced to hold portfolios chosen by investors with the same degree of belief in the Fama–French model. Investor risk aversion is denoted by A. Mispricing uncertainty, denoted by σ_α, is the prior standard deviation of the difference between each position's annualized expected payoff and the pricing model's exact implication, expressed as a percentage of initial position size. The maximum value of risky positions that can be established per dollar of wealth is denoted by c, and each plot displays results for $c = 2$ (solid), $c = 5$ (dash), $c = 10$ (dash-dot), and $c = \infty$ (dots).

and this effect reduces the certainty-equivalent loss from holding the alternative portfolio.

In Fig. 4, a 50% margin requirement ($c = 2$) eliminates the loss for the first three levels of risk aversion ($A = 1$, 5, and 10), duplicating the result obtained earlier with $A = 2.83$ in Fig. 1. As before, the portfolio choices under the constraint are virtually identical for investors with dogmatic beliefs in the FF and DT models. As in Table 2, those optimal portfolios call for zero allocations to SMB and the nine beta-spread portfolios. (In the absence of constraints, a DT investor chooses substantial nonzero allocations to those positions, and a FF investor chooses a positive allocation to SMB.) When $A = 1$, Fig. 4 reveals a zero certainty-equivalent loss with only a 10% margin, whereas the corresponding loss with $A = 2.83$ is about 2%. Dogmatic DT and FF investors with $A = 1$ continue to allocate nothing to SMB and the nine beta spreads. With $A = 2.83$, both investors choose positive (but different) allocations to SMB, and the DT investor allocates funds to some of the beta spreads as well (Table 2). The fact that these positions receive zero allocations under the lower A can be understood by considering their expected payoffs. Under dogmatic DT beliefs, the nine beta spreads have zero expected payoffs ($E_1 = 0$). The posterior mean of SMB is positive but lower than that of HML and the highest-mean payoff, MKT. The positions with low means offer diversification, but capturing that diversification requires allocating capital to those positions. With margin requirements and sufficiently low risk aversion, that required capital is better used in positions with higher expected payoffs.

With investment constraints, the certainty-equivalent loss can increase with risk aversion. For example, with $c = 2$ and $\sigma_\alpha = 0$, the loss for the FF portfolio is zero for $A = 1$ but about 45 basis points for $A = 15$. Although the optimal portfolios under both models indeed involve larger cash positions for $A = 15$, the constrained optimal portfolios also include some positions that are not included for $A = 1$. In particular, the optimal portfolios from both models include only HML and MKT when $A = 1$, as observed above. In contrast, for $A = 15$, the optimal portfolio from the FF model also includes a \$12 position in SMB (per \$100 invested), while the optimal portfolio from the DT model includes nonzero allocations (between 10 cents and \$12) to all 11 risky positions. With this higher level of risk aversion, the DT investor finds the zero-mean HML-beta spreads attractive for diversification reasons, whereas the dogmatic FF investor allocates nothing to those positions, whether constrained or not.

Fig. 5 reports losses for an investor who believes in the DT model but is forced to hold the portfolio of an investor with CAPM beliefs. This figure confirms the observations made for Fig. 4. Certainty-equivalent losses decrease as risk aversion increases as well as when investment constraints and mispricing uncertainty are introduced. Note that the scales for the vertical axes in Fig. 5 are different from those in Fig. 4 since, to an investor with DT beliefs, the losses from holding a CAPM investor's portfolio are larger than those from holding the portfolio of an FF investor. This result is observed for all levels of risk aversion and is consistent with the evidence for $A = 2.83$ presented in the earlier figures. With $A = 1$, however, the DT portfolio is quite close to the CAPM portfolio, with the latter producing less than a 50 basis point loss for all values of σ_α. For risk aversion so low, the highest-mean position MKT constitutes the bulk of each portfolio's allocation. In general, though, the FF and DT models are substantially closer to each other than to the CAPM in terms of their implications for portfolio choice. With either realistic investment constraints or

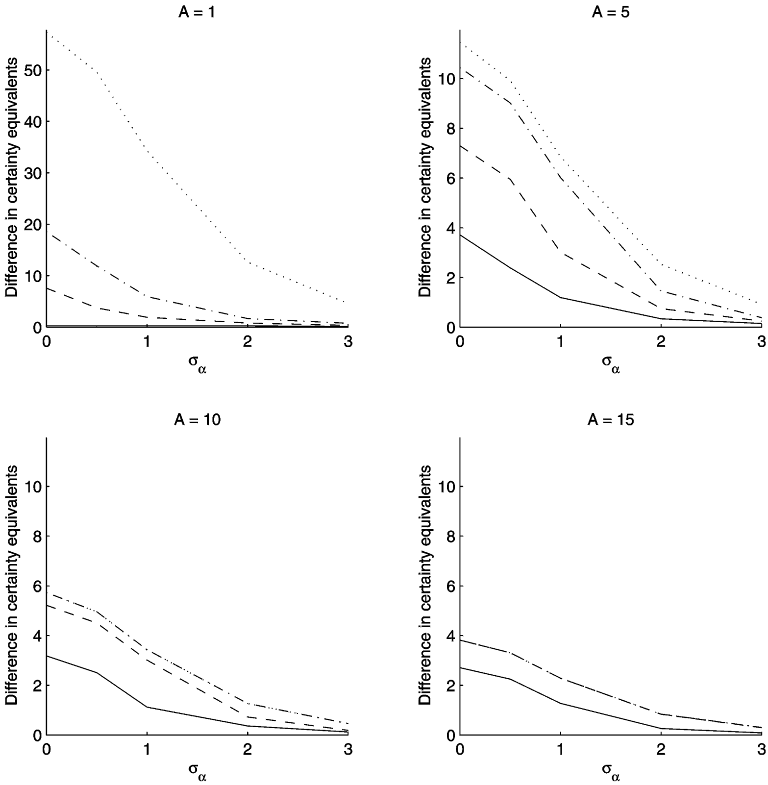

Fig. 5. Effects of risk aversion on the certainty-equivalent losses for the CAPM's portfolios from the perspective of a Daniel–Titman investor. The figure displays the certainty-equivalent loss (in % per year) for a mean-variance-optimizing investor whose prior beliefs are centered on the Daniel–Titman characteristic-based model, with mispricing uncertainty σ_α, but who is forced to hold portfolios chosen by investors with the same degree of belief in the CAPM. Investor risk aversion is denoted by A. Mispricing uncertainty, denoted by σ_α, is the prior standard deviation of the difference between each position's annualized expected payoff and the pricing model's exact implication, expressed as a percentage of initial position size. The maximum value of risky positions that can be established per dollar of wealth is denoted by c, and each plot displays results for $c = 2$ (solid), $c = 5$ (dash), $c = 10$ (dash-dot), and $c = \infty$ (dots).

modest mispricing uncertainty, the portfolio implications from the FF and DT models are very similar.

4. An alternative characteristic-based model

The representation of the characteristic-based model in the previous section requires positions to be constructed as spreads between assets with matched characteristics. This section proposes an alternative representation that can, in principle, be applied to any set of equity positions. At the same time, this alternative model makes stronger assumptions about the relation between expected returns and characteristics. Our main goal in pursuing this second approach is to explore the robustness of the empirical results in the previous section to changes in the specification of the investment universe.

In this second characteristic-based model ('CB2'), the expected excess return on a (positive-cost) equity portfolio is assumed to be a linear function of (known) characteristics. Let C denote an $n \times L$ matrix in which each of columns 2 through L contains the values of a characteristic, such as book-to-market. In the empirical application presented here, the two characteristics are size and book-to-market ($L = 3$). Then CB2 can be represented by the restriction

$$E = C\gamma, \tag{26}$$

where γ is an $L \times 1$ vector. For a position in which asset S_i is risky, such as SMB and HML, the corresponding element in the first column of C is equal to zero. For the other positions, the first column of C contains the value one, to include an intercept in the linear relation between expected excess returns and characteristics. Note that this model restricts all n elements of E, whereas the models examined in the previous section imply restrictions only on the $n - k$ elements of E_1, the expected payoffs on the non-benchmark positions.

To allow for mispricing uncertainty in this characteristic-based model, the prior distribution for E is specified as

$$E \mid \gamma \sim \mathrm{N}(C\gamma, \sigma_\alpha^2 I_n). \tag{27}$$

As in the previous models, σ_α represents the degree of mispricing uncertainty. The lower the value of σ_α, the higher the prior confidence in the model's pricing restriction in (26). As before, r_t is assumed to be normally distributed with mean E and covariance matrix V. The prior distributions for V and γ are specified as noninformative. Optimal portfolios are computed as in (7) using the predictive mean and covariance matrix, which are obtained through Gibbs sampling (as explained in the Appendix).

The investment universe analyzed here consists of cash plus the same set of $n = 11$ positions used in the example in Section 2. The second and third columns of C contain values of each position's size and book-to-market, which are the same two characteristics used in the DT model analyzed in Section 3 (i.e., the same as used to construct the characteristic-matched spreads). The characteristics in C are computed as follows. For each stock, size is the natural logarithm of total equity capitalization, and book-to-market is the ratio of book value to market value of common equity. In each month, the values of size and book-to-market for each three-way-sorted portfolio (described in Section 2) are computed as the value-weighted averages of those characteristics for the stocks in the portfolio. This computation is also performed for the portfolios 'S' and 'B' used to construct the FF factor SMB, and similar calculations are performed for the portfolios used to construct the FF factors HML and MKT. In the positions constructed as spreads between equity portfolios (SMB and HML), the characteristic of one portfolio is subtracted from that of another. The formulation in (26) treats the characteristics as known constants, which are obtained from the monthly time series as follows. In each month, the characteristic of each position is divided by the cross-sectional average of that characteristic across the n positions. The time-series averages of these standardized series are then used as the values in C. Plots of the standardized characteristics exhibit no apparent trend and exhibit much less time-series variation than do plots of the raw characteristics.

Table 3 reports optimal allocations per \$100 of wealth when prior beliefs are centered on each of three pricing models, CB2, FF, and CAPM (CM). (The format is otherwise identical to Table 2.) Consider first the optimal allocations obtained with a dogmatic belief in the FF model ($\sigma_\alpha = 0$). Recall that these allocations are reported in Table 1 in a simplified setting that does not account for estimation risk. Not surprisingly, the optimal portfolio in the unconstrained case ($c = \infty$) includes only the three FF positions as before. With a 10% margin requirement ($c = 10$), the optimal portfolio in Table 1 includes some small non-benchmark positions, whereas the portfolio in Table 3 is still unaffected by the constraint. With a 50% margin requirement ($c = 2$), the striking result seen earlier in Table 1 remains after accounting for estimation risk: an investor who believes dogmatically in the FF model invests only in non-benchmark positions.

In Table 3, the optimal portfolios from the characteristic-based and FF models for $\sigma_\alpha = 0$ and $c = 2$ are no longer identical, as in Table 2. For example, SHl receives a zero allocation under the FF model but the largest allocation (of 114) under CB2. The latter allocation is not surprising. The posterior distributions for the elements of γ in (26) are such that expected returns are decreasing in size and increasing in book-to-market. In particular, nearly 95% of the posterior mass of the size coefficient lies below zero, while virtually all of the posterior mass of the book-to-market coefficient lies above zero. Thus, SHl and SHh have the highest expected returns. Although SHl has a lower expected return than SHh (14.2% versus 14.8% per year), it also has a lower standard deviation (19.9% versus 21.4% per year) and somewhat smaller covariances with the other positions, so it is more attractive to a CB2 investor. Nevertheless, SHh receives the second largest allocation (of 28) under CB2. Although the compositions of the optimal portfolios under CB2 and FF differ, neither contains allocations in the FF benchmark positions. Also, the annual means and standard deviations of the two portfolios are quite similar: 18.1% and 21.0% for the CB2 portfolio, and 17.1% and 21.6% for the FF portfolio. The two portfolios become even more similar when dogmatic beliefs in both models are relaxed. As in Table 2, relaxing the investment constraints leads to more borrowing and hence to optimal portfolios with higher means and standard deviations. For $c = \infty$ and $\sigma_\alpha = \infty$,

Table 3
Optimal allocations under parameter uncertainty (with the alternative characteristic-based model)

The table reports optimal allocations (position sizes) per $100 of wealth for a mean-variance-optimizing investor with relative risk aversion equal to 2.83. The maximum value of risky positions that can be established per dollar of wealth is denoted by c. Mispricing uncertainty, denoted by σ_α, is the prior standard deviation of the difference between each position's annualized expected payoff and the pricing model's exact implication, expressed as a percentage of initial position size. Allocations are reported for prior beliefs centered on three different pricing models: the three-factor Fama–French model (FF), the CAPM (CM), and the alternative characteristic-based model (CB2). Optimization is based on the predictive distribution, obtained by updating the prior beliefs using monthly returns from July 1963 through December 1997. The Fama–French benchmark positions are denoted by SMB, HML, and MKT. The risky components of the first eight positions are a subset of value-weighted portfolios constructed by a three-way sort and identified by a combination of three letters, designating increasing values of size (S, M, B), book-to-market (L, M, H), and HML beta (l, m, h). The 'cash' row reports the overall amount invested in the riskless asset and includes the amounts implied by the positions that short the riskless asset (all positions except SMB and HML). Negative cash amounts represent borrowing. Also reported are the annualized mean and standard deviation (std) of the portfolio's return with respect to the given predictive distribution.

	$\sigma_\alpha = 0$			$\sigma_\alpha = 1\%$			$\sigma_\alpha = 2\%$			$\sigma_\alpha = \infty$ (all)
	CB2	FF	CM	CB2	FF	CM	CB2	FF	CM	
					$c = 2$					
SLl	– 29.0	– 10.9	0.0	– 26.6	– 35.0	– 14.2	– 25.3	– 32.8	– 12.5	– 25.4
SLh	– 13.1	0.0	0.0	– 17.0	0.0	– 31.7	– 20.8	– 11.0	– 36.7	– 22.4
SHl	113.5	0.0	0.0	125.1	34.1	67.2	123.3	85.8	107.1	121.6
SHh	28.4	77.5	0.0	1.0	65.6	6.7	0.0	27.8	2.3	0.0
BLl	0.0	– 34.3	0.0	0.0	0.0	15.1	0.0	0.0	0.0	0.0
BLh	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BHl	0.0	6.6	0.0	0.0	0.0	11.7	0.0	0.0	0.5	0.0
BHh	16.0	70.6	0.0	30.3	53.5	20.0	30.7	42.7	40.9	30.6
SMB	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HML	0.0	0.0	0.0	0.0	5.9	0.0	0.0	0.0	0.0	0.0
MKT	0.0	0.0	100.0	0.0	0.0	33.5	0.0	0.0	0.0	0.0
Cash	– 15.8	– 9.6	0.0	– 12.8	– 18.1	– 8.2	– 7.9	– 12.5	– 1.7	– 4.3
Mean	18.1	17.1	6.3	16.5	16.0	7.5	15.8	15.4	10.5	15.5
Std	21.0	21.6	15.0	19.2	20.5	15.7	18.2	18.7	16.7	17.6

Table 3 (*continued*)

	$\sigma_\alpha = 0$			$\sigma_\alpha = 1\%$			$\sigma_\alpha = 2\%$			$\sigma_\alpha = \infty$ (all)
	CB2	FF	CM	CB2	FF	CM	CB2	FF	CM	
					$c = 5$					
SLl	– 44.6	– 0.6	0.0	– 52.5	– 30.7	– 14.7	– 62.4	– 54.5	– 34.9	– 62.6
SLh	– 79.3	0.0	0.0	– 97.4	– 39.8	– 40.0	– 113.7	– 89.2	– 84.1	– 115.2
SHl	212.3	0.0	0.0	244.5	108.2	67.4	263.1	202.9	178.8	261.6
SHh	23.9	63.5	0.0	6.2	40.7	– 15.3	11.4	31.6	2.2	11.1
BLl	0.0	0.0	0.0	0.0	0.0	41.3	0.0	0.0	44.8	0.0
BLh	0.0	0.0	0.0	0.0	0.0	7.6	0.0	0.0	0.0	0.0
BHl	20.4	33.9	0.0	7.2	14.2	– 4.2	0.0	0.0	3.5	0.0
BHh	15.1	41.0	0.0	40.3	48.9	0.7	49.4	54.7	27.3	49.4
SMB	0.0	0.0	0.0	0.0	0.0	33.7	0.0	0.0	0.0	0.0
HML	52.2	180.5	0.0	25.9	108.8	80.9	0.0	33.6	62.1	0.0
MKT	0.0	0.0	100.0	0.0	0.0	79.5	0.0	0.0	0.0	0.0
Cash	– 47.8	– 37.8	0.0	– 48.3	– 41.5	– 22.5	– 47.7	– 45.5	– 37.8	– 44.3
Mean	28.6	25.5	6.3	28.0	24.3	8.4	28.5	26.1	15.6	28.6
Std	27.0	28.0	15.0	26.5	26.1	17.1	26.7	26.1	22.0	26.2
					$c = 10$					
SLl	– 40.8	0.0	0.0	– 47.8	– 23.0	– 15.5	– 58.5	– 47.4	– 35.9	– 69.5
SLh	– 118.1	0.0	0.0	– 138.7	– 66.2	– 42.0	– 156.9	– 123.3	– 97.9	– 170.8
SHl	278.9	0.0	0.0	296.5	155.8	61.1	322.9	265.8	170.8	345.1
SHh	0.0	0.0	0.0	0.0	0.0	– 30.9	0.0	0.0	– 28.8	0.0
BLl	74.1	0.0	0.0	79.8	63.9	46.9	74.0	78.1	99.0	69.3
BLh	– 27.1	0.0	0.0	0.0	12.3	9.5	0.0	0.0	17.8	0.0
BHl	21.2	0.0	0.0	0.0	0.0	– 9.9	0.0	0.0	– 7.0	0.0
BHh	0.0	0.0	0.0	0.0	5.5	– 0.7	8.0	14.7	2.7	16.9
SMB	0.0	63.2	0.0	0.0	13.2	63.6	0.0	0.0	65.9	0.0
HML	220.0	352.5	0.0	218.6	307.1	104.8	189.8	235.3	186.5	164.2
MKT	0.0	168.2	100.0	0.0	32.8	106.0	0.0	0.0	35.3	0.0
Cash	– 88.2	– 68.2	0.0	– 89.9	– 81.0	– 24.5	– 89.4	– 87.8	– 56.1	– 91.0
Mean	40.6	30.4	6.3	40.5	32.5	8.5	41.4	37.2	18.3	42.5
Std	35.2	32.8	15.0	35.0	33.3	17.3	35.5	34.5	25.1	36.0

					$c = \infty$					
SLl	− 45.2	0.0	0.0	− 54.2	− 29.0	− 15.5	− 63.5	− 53.4	− 38.5	− 72.7
SLh	− 142.8	0.0	0.0	− 168.9	− 81.7	− 42.0	− 189.3	− 150.5	− 104.4	− 204.7
SHl	256.0	0.0	0.0	244.1	106.6	61.1	259.8	196.3	151.9	272.7
SHh	− 106.4	0.0	0.0	− 190.2	− 77.4	− 30.9	− 199.1	− 142.6	− 76.6	− 200.4
BLl	163.3	0.0	0.0	247.0	102.1	46.9	255.1	188.1	116.5	261.5
BLh	− 101.6	0.0	0.0	8.2	23.3	9.5	41.3	42.9	23.5	62.9
BHl	− 12.2	0.0	0.0	− 22.9	− 23.4	− 9.9	− 41.3	− 43.1	− 24.6	− 59.4
BHh	− 46.7	0.0	0.0	− 24.9	− 4.5	− 0.7	− 16.2	− 8.3	− 1.6	− 10.3
SMB	191.6	63.2	0.0	350.2	195.7	63.6	379.7	308.1	158.0	398.2
HML	490.5	352.5	0.0	596.6	444.8	104.8	597.0	527.0	260.4	595.0
MKT	262.6	168.2	100.0	190.7	175.1	106.0	181.3	183.1	116.6	179.5
Cash	− 126.9	− 68.2	0.0	− 128.8	− 91.0	− 24.5	− 128.3	− 112.5	− 62.6	− 129.1
Mean	49.8	30.4	6.3	51.7	34.3	8.5	53.0	43.5	19.2	54.7
Std	42.0	32.8	15.0	42.7	34.8	17.3	43.3	39.2	26.1	44.0

the optimal portfolio has a mean of 54.7% and a standard deviation of 44.0%; it involves borrowing 129 and its largest position (of 595) is HML.

As in the previous section, the position-by-position analysis of the optimal allocations under the factor-based and characteristic-based models is complemented by an analysis of certainty-equivalent losses. Fig. 6 is the equivalent of Fig. 1, except that the characteristic-based model is now CB2 (instead of DT) and the investment universe is different (as explained earlier). With no investment restrictions ($c = \infty$) and a dogmatic belief in each model ($\sigma_\alpha = 0$), the certainty-equivalent loss for the FF portfolio is about 8% per year, quite similar to the corresponding loss in Fig. 1. As in Fig. 1, the loss for the CAPM portfolio is much higher, almost 21%. For $\sigma_\alpha = 3\%$, the FF loss is only 17 basis points, whereas the CAPM loss is 2.2%. CB2 is also closer to the FF model than to the CAPM when investment constraints are present ($c < \infty$). When $c = 2$ and σ_α is 2% or more, both the CAPM and the FF model are very close to CB2 (the losses are no more than 19 basis points).

Comparisons of CB2 and the CAPM under the predictive distribution from the FF model (an equivalent of Fig. 2, not reported to save space) lead to observations very similar to those made for Fig. 6. In general, optimal portfolios from the FF and CB2 models are quite similar when realistic investment constraints and modest mispricing uncertainty are incorporated. This result no doubt reflects an association between characteristics and expected returns present in the characteristic-based model as well as the FF model, as noted earlier, whereas that association is weaker under the CAPM due to the lower correlation between characteristics and market betas.

5. Incorporating model uncertainty

In Sections 3 and 4, optimal allocations are computed by combining data with prior beliefs centered at a particular pricing model. Such an approach is consistent with a common practice of first choosing the 'best' model, using judgment or model selection criteria, and then proceeding as if the selected model were the only one relevant. If the investor entertains several models a priori, this practice does not account for 'model uncertainty' in the selection process. For example, an investor who is uncertain as to whether expected returns are modeled better as risk-related or characteristic-related can potentially benefit from combining the implications of both models. It seems conceivable, given available empirical evidence, that one might be less than comfortable discarding one of the two models entirely. For example, Davis et al. (2000) find that a t-test rejects the null hypothesis that the FF model prices an equally weighted combination of characteristic-matched HML-beta spreads in the 1973–93 period, but the hypothesis cannot be rejected in the longer 1929–97 period. At the same time, the point estimate underlying the latter result does

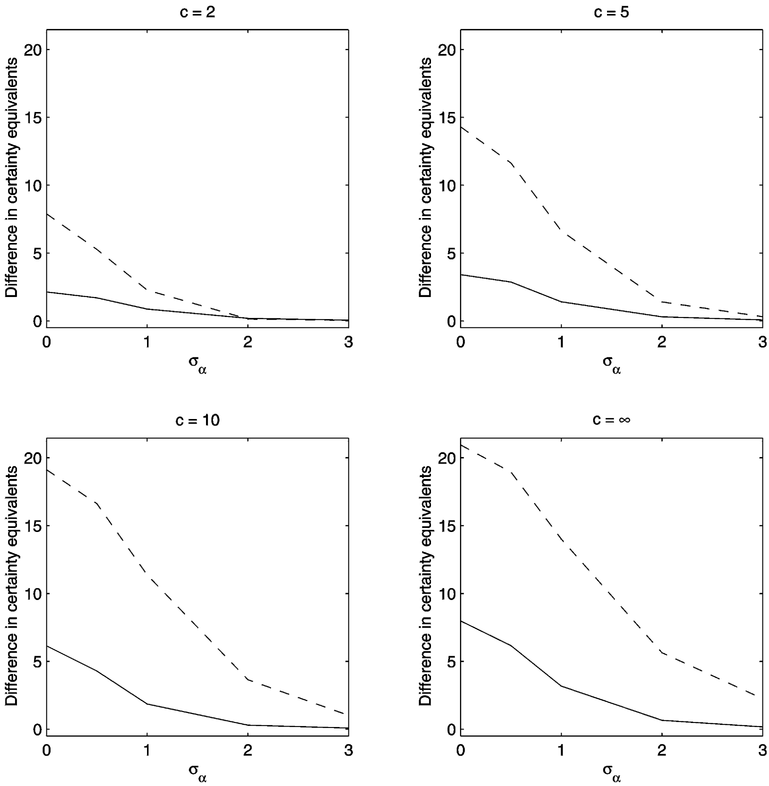

Fig. 6. Certainty-equivalent losses for other models' portfolios from the perspective of an investor who believes in the alternative characteristic-based model. The figure displays the certainty-equivalent loss (in % per year) for a mean-variance-optimizing investor whose prior beliefs are centered on the alternative characteristic-based model, with mispricing uncertainty σ_α, but who is forced to hold portfolios chosen by investors with the same degree of belief in either the Fama–French model (solid line) or the CAPM (dashed line). Investor risk aversion is set to $A = 2.83$. The maximum value of risky positions that can be established per dollar of wealth is denoted by c. Mispricing uncertainty, denoted by σ_α, is the prior standard deviation of the difference between each position's annualized expected payoff and the pricing model's exact implication, expressed as a percentage of initial position size.

have its sign in the direction of a characteristic-based alternative. Moreover, the authors report that the FF model, viewed as a null hypothesis, is formally rejected within a characteristic-sorted universe based on an F-test. Fortunately, investors need not limit their attention to only one of the two models.

Suppose the investor considers a universe of J models $M_1, \ldots, M_J$ of expected returns. Let $p(r_{T+1} \mid M_j, R)$ denote the predictive distribution of returns from model j, and let $P(M_j|R)$ denote the posterior probability of model j. The predictive distribution that accounts for model uncertainty is

$$p(r_{T+1}|R) = \sum_{j=1}^{J} p(r_{T+1} \mid M_j, R)\, P(M_j \mid R). \tag{28}$$

Optimal allocations are again computed from (7), but E and V are no longer replaced by E_j^* and V_j^*, the mean and covariance matrix of the predictive distribution from model j, but rather by E_M^* and V_M^*, the moments of the predictive distribution that accounts for model uncertainty. These moments are computed as follows (see Leamer, 1978):

$$E_M^* = \sum_{j=1}^{J} E_j^* P(M_j|R) \tag{29}$$

$$V_M^* = \sum_{j=1}^{J} V_j^* P(M_j|R) + \sum_{j=1}^{J} (E_j^* - E_M^*)(E_j^* - E_M^*)' P(M_j|R). \tag{30}$$

The predictive mean is an average of the predictive means from the J models, weighted by the model probabilities. The predictive covariance matrix has two components. The first is a weighted average of the predictive covariance matrices from the J models, and the second is the covariance matrix of the predictive means across models. The posterior model probabilities $P(M_j \mid R)$ are computed as

$$P(M_j|R) = \frac{P(M_j)p(R \mid M_j)}{\sum_{j=1}^{J} P(M_j)p(R \mid M_j)}, \tag{31}$$

where $P(M_j)$ denotes the prior probability assigned to model j before observing the data, and $p(R \mid M_j)$ denotes the so-called marginal likelihood of model j. The marginal likelihood is computed as

$$p(R \mid M_j) = \int p(\theta_j|M_j)\, p(R \mid \theta_j,\, M_j)\, \mathrm{d}\theta_j, \tag{32}$$

where θ_j denotes the parameters, $p(\theta_j|M_j)$ denotes the prior distribution, and $p(R|\theta_j,\, M_j)$ denotes the likelihood function, all from model j.

The calculation of posterior model probabilities is beyond the scope of this study. Such a task would have to address several issues. First, recall that the priors for parameters not involving a pricing restriction are specified as noninformative or even diffuse, which are in some sense completely noninformative. Diffuse priors are improper in that they are not integrable over the parameter space. With improper priors, posterior model probabilities can be computed only in some special cases, as discussed in Kass and Raftery (1995). Second, the marginal likelihood is often quite sensitive to the choice of a noninformative prior distribution, more so than is the posterior distribution typically used in estimation. A prior distribution can be made essentially noninformative but still proper by specifying a large prior variance, but the latter can generally assume a wide range of values. For example, doubling an already large variance keeps the prior noninformative, hence having little effect on the posterior distribution, but the change can greatly affect the marginal likelihood. Therefore, a study that computes marginal likelihoods should analyze their sensitivity to specifications of noninformative priors. Finally, prior model probabilities should also be thoughtfully specified. For example, suppose the universe of models includes one with strong theoretical motivation (e.g., the CAPM) as well as one that is partially motivated by observing data that either overlap or are correlated with the sample R. It might be reasonable to assign a higher prior probability to the former model.

In Sections 3 and 4, we assume that an investor's prior beliefs center on a particular pricing model, say the DT model, and this assumption corresponds to assigning a posterior probability of one to that model in the context described above. In those sections, the effect of model uncertainty is assessed by computing the loss to an investor who is forced to hold the allocation that would be optimal if that investor assigned a probability of one to a different model, say the FF model. Such a scenario measures the maximum effect of model uncertainty in that two-model universe. The loss from accepting the FF allocation is smaller when the probability assigned to the DT model is less than one. Recall that even the maximum effect of model uncertainty is quite small with modest mispricing uncertainty and realistic investment constraints.

6. Conclusions

This study compares asset pricing models from the perspective of investors who center their prior beliefs on the models and then update those beliefs with data for the 1963–97 period. The pricing models considered include the Sharpe–Lintner CAPM, the three-factor model of Fama and French (1993), and the characteristic-based model of Daniel and Titman (1997). An investor who has dogmatic beliefs in a given pricing model perceives a large certainty-equivalent loss if forced to hold the portfolio chosen by an investor with equally strong beliefs in another model. The largest such losses occur when the second portfolio is chosen by a CAPM believer, but the differences between the Fama–French and Daniel–Titman models are also large when judged by this metric.

The differences described above are reduced by a consideration of two issues confronting investors. First, even an investor who prefers a given model is unlikely to believe it to be completely accurate. The investor's prior beliefs more likely include some degree of mispricing uncertainty. Second, most investors face margin requirements to at least some degree. Both of these issues, especially the second, diminish the importance of differences among the pricing models from an investment perspective. In fact, we find that considering such issues can virtually eliminate any differences between the Fama–French and Daniel–Titman models, even though these models reflect fundamentally different views about the economic determinants of expected returns. It is noteworthy that these models lead to similar portfolio choices within investment universes constructed to exploit differences between the models.

This study is not intended to assist investors in choosing one pricing model over another. One might instead view our results as questioning the economic importance of deliberating such a choice. Moreover, a rational Bayesian investor who is uncertain about which model to use will generally use them all, weighted by posterior model probabilities. Computing such probabilities is beyond the intended scope of this study but offers a direction for future research.

Finally, the single-period mean-variance framework provides only one of many investment perspectives from which pricing models might be compared empirically. For example, the differences between risk-based and characteristic-based models could be more important for investors who optimize multiperiod objective functions. Alternatives to the i.i.d. stochastic setting, adopted here for tractability, could further enrich an empirical comparison of models from a multiperiod investment perspective. Such a perspective would be more consistent with that of the representative investor in Merton's (1973) intertemporal version of a risk-based model. (An investor in the present study is not assumed to be representative, insofar as different assumptions about behavior and objectives might be required to derive a particular pricing model.) Such issues present additional directions for future research.

Appendix

We provide here the methods for obtaining the first two moments, E^* and V^*, of the predictive distribution of returns, $p(r_{T+1}|R)$, for each of the pricing models in Sections 3 and 4. As shown below, these predictive moments can be computed directly from the first and second posterior moments of each model's parameters.

A.1. Risk-based models

Define $Y = (r_{1,1}, \dots, r_{1,T})'$, $X = (r_{2,1}, \dots, r_{2,T})'$, and $Z = (\iota_T \;\; X)$, where ι_T denotes a T-vector of ones. Also define the $(k+1) \times m$ matrix $A = (\alpha \;\; B)'$, and let $a = \mathrm{vec}(A)$. For the T observations $t = 1, \dots, T$, the regression model in Eq. (15)

can be written as

$$Y = ZA + U, \quad \mathrm{vec}(U) \sim \mathrm{N}(0, \Sigma \otimes I_T), \tag{A.1}$$

where $U = (u_1, \ldots, u_T)'$. The matrix $R = (Y\ X)$ contains the entire sample. Define the statistics $\hat{A} = (Z'Z)^{-1}Z'Y$, $\hat{a} = \mathrm{vec}(\hat{A})$, $\hat{\Sigma} = (Y - Z\hat{A})'(Y - Z\hat{A})/T$, $\hat{E}_2 = X'\iota_T/T$, and $\hat{V}_{22} = (X - \iota_T\hat{E}_2')'(X - \iota_T\hat{E}_2')/T$. (Recall that θ denotes all of the parameters in a given model.) The likelihood function can be factored as

$$p(R|\theta) = p(Y|\theta, X)\ p(X|\theta), \tag{A.2}$$

where

$$\begin{aligned} p(Y|\theta, X) &\propto |\Sigma|^{-T/2} \exp\{ -\tfrac{1}{2}\mathrm{tr}(Y - ZA)'(Y - ZA)\Sigma^{-1}\} \\ &\propto |\Sigma|^{-T/2} \exp\left\{ -\frac{T}{2}\mathrm{tr}\hat{\Sigma}\Sigma^{-1} - \frac{1}{2}\mathrm{tr}(A - \hat{A})'Z'Z(A - \hat{A})\Sigma^{-1} \right\} \\ &\propto |\Sigma|^{-T/2} \exp\left\{ -\frac{T}{2}\mathrm{tr}\hat{\Sigma}\Sigma^{-1} - \frac{1}{2}(a - \hat{a})'(\Sigma^{-1} \otimes Z'Z)(a - \hat{a}) \right\} \end{aligned} \tag{A.3}$$

and

$$\begin{aligned} p(X|\theta) &\propto |V_{22}|^{-T/2} \exp\{ -\tfrac{1}{2}\mathrm{tr}(X - \iota_T E_2')'(X - \iota_T E_2')V_{22}^{-1}\} \\ &\propto |V_{22}|^{-T/2} \exp\left\{ -\frac{T}{2}\mathrm{tr}\hat{V}_{22}V_{22}^{-1} - \frac{T}{2}\mathrm{tr}(E_2 - \hat{E}_2)(E_2 - \hat{E}_2)'V_{22}^{-1} \right\}. \end{aligned} \tag{A.4}$$

The joint prior distribution of all parameters is

$$p(\theta) = p(\alpha|\Sigma)\ p(\Sigma)\ p(B)\ p(E_2)\ p(V_{22}), \tag{A.5}$$

where

$$p(\alpha|\Sigma) \propto |\Sigma|^{-1/2} \exp\left\{ -\frac{1}{2}\alpha'\left(\frac{\sigma_\alpha^2}{s^2}\Sigma\right)^{-1}\alpha \right\}, \tag{A.6}$$

$$p(\Sigma) \propto |\Sigma|^{-(v+m+1)/2} \exp\{ -\tfrac{1}{2}\mathrm{tr}H\Sigma^{-1}\}, \tag{A.7}$$

$$p(B) \propto 1, \tag{A.8}$$

$$p(E_2) \propto 1, \tag{A.9}$$

$$p(V_{22}) \propto |V_{22}|^{-(k+1)/2}. \tag{A.10}$$

The priors of B, E_2, and V_{22} are diffuse. The prior of Σ is inverted Wishart with a small number of degrees of freedom, so that it is essentially noninformative. The prior on α given Σ is normal and centered at the pricing restriction. Note that

$$\alpha'\left(\frac{\sigma_\alpha^2}{s^2}\Sigma\right)^{-1}\alpha = a'(\Sigma^{-1}\otimes D)a, \tag{A.11}$$

where D is a $(k+1)\times(k+1)$ matrix whose $(1, 1)$ element is s^2/σ_α^2 and all other elements are zero.

Combining the likelihood in Eqs. (A.2)–(A.4) with the prior in Eqs. (A.5)–(A.10) yields the posterior distribution of θ:

$$p(\theta|R) \propto p(R|\theta)\ p(\theta). \tag{A.12}$$

Both the likelihood and the prior can be factored into two parts, one that involves the regression parameters (a, Σ) and another that involves the benchmark moments (E_2, V_{22}). As a result, the posterior distribution also splits into two parts. The joint posterior of the regression parameters is

$$\begin{aligned} p(a, \Sigma|R) \propto |\Sigma|^{-(k+1)/2}&\exp\{-\tfrac{1}{2}[a'(\Sigma^{-1}\otimes D)a + (a-\hat{a})' \\ &\times(\Sigma^{-1}\otimes Z'Z)(a-\hat{a})]\}|\Sigma|^{-(T+v+m-k+1)/2} \\ &\times\exp\{-\tfrac{1}{2}\mathrm{tr}(H + T\hat{\Sigma})\Sigma^{-1}\}. \end{aligned} \tag{A.13}$$

Let $F = D + Z'Z$, $Q = Z'(I_T - ZF^{-1}Z')Z$. Completing the square on a yields

$$\begin{aligned} p(a, \Sigma|R) &\propto |\Sigma|^{-(k+1)/2}\exp\{-\tfrac{1}{2}(a-\tilde{a})'(\Sigma^{-1}\otimes F)(a-\tilde{a})\} \\ &\times|\Sigma|^{-(T+v+m-k+1)/2}\exp\{-\tfrac{1}{2}\mathrm{tr}(H + T\hat{\Sigma} + \hat{A}'Q\hat{A})\Sigma^{-1}\}, \end{aligned} \tag{A.14}$$

where

$$\tilde{a} = (I_m\otimes F^{-1}Z'Z)\hat{a}. \tag{A.15}$$

It follows from (A.14) that

$$a|\Sigma, R \sim \mathrm{N}(\tilde{a},\ \Sigma\otimes F^{-1}), \tag{A.16}$$

$$\Sigma^{-1}|R \sim W(T+v-k,\ (H + T\hat{\Sigma} + \hat{A}'Q\hat{A})^{-1}). \tag{A.17}$$

Therefore,

$$\tilde{a} = \mathrm{E}(a|R) = (I_m\otimes F^{-1}Z'Z)\hat{a}, \tag{A.18}$$

$$\tilde{\Sigma} = \mathrm{E}(\Sigma|R) = \frac{1}{T+v-m-k-1}(H + T\hat{\Sigma} + \hat{A}'Q\hat{A}), \tag{A.19}$$

$$\mathrm{Var}(a|R) = \tilde{\Sigma}\otimes F^{-1}. \tag{A.20}$$

We denote posterior means using tildes for the remainder of the Appendix.

The joint posterior of the benchmark moments is

$$p(E_2, V_{22}|X) \propto |V_{22}|^{-(T+k+1)/2} \times \exp\left\{ -\frac{T}{2}\mathrm{tr}\hat{V}_{22}V_{22}^{-1} - \frac{T}{2}\mathrm{tr}(E_2 - \hat{E}_2)(E_2 - \hat{E}_2)'V_{22}^{-1} \right\}. \tag{A.21}$$

It follows that

$$E_2|V_{22}, R \sim \mathrm{N}\left(\hat{E}_2, \frac{1}{T}V_{22}\right), \tag{A.22}$$

$$V_{22}^{-1}|R \sim W(T-1, (T\hat{V}_{22})^{-1}). \tag{A.23}$$

Therefore,

$$\tilde{E}_2 = \mathrm{E}(E_2|R) = \hat{E}_2, \tag{A.24}$$

$$\tilde{V}_{22} = \mathrm{E}(V_{22}|R) = \frac{T}{T-k-2}\hat{V}_{22}, \tag{A.25}$$

$$\mathrm{Var}(E_2|R) = \frac{1}{T-k-2}\hat{V}_{22}. \tag{A.26}$$

Note that the predictive mean obeys the relation,

$$E^* = \mathrm{E}(r_{T+1}|R) = \mathrm{E}(\mathrm{E}(r_{T+1}|\theta, R)|R) = \mathrm{E}(E|R) = \tilde{E}. \tag{A.27}$$

Since B and E_2 are independent in the posterior, the mean of the predictive distribution is

$$E^* = \tilde{E} = \mathrm{E}\left(\begin{matrix}\alpha + BE_2 \\ E_2\end{matrix}\middle| R\right) = \begin{pmatrix}\tilde{\alpha} + \tilde{B}\tilde{E}_2 \\ \tilde{E}_2\end{pmatrix}, \tag{A.28}$$

where $\tilde{\alpha}$ and $\tilde{B}$ are obtained from Eq. (A.18) using $\tilde{a} = \mathrm{vec}((\tilde{\alpha}\ \tilde{B})')$.

Partition the predictive covariance matrix as

$$V^* = \begin{bmatrix} V_{11}^* & V_{12}^* \\ V_{21}^* & V_{22}^* \end{bmatrix}. \tag{A.29}$$

The first submatrix, V_{11}^*, can be represented in terms of its (i, j) element. Denote the i-th element of $r_{1,T+1}$ as $y_{i,T+1}$, the ith element of α as α_i, the ith element of u_{T+1} as $u_{i,T+1}$, the ith row of B as b_i', the ith column of A as a_i, and the (i, j)

element of Σ as $\sigma_{i,j}$. Note that

$$y_{i,T+1} = \alpha_i + b_i' r_{2,T+1} + u_{i,T+1} \tag{A.30}$$

$$= [1 \; r_{2,T+1}'] a_i + u_{i,T+1}, \tag{A.31}$$

since $a_i = (\alpha_i \; b_i')'$. The predictive covariance between $y_{i,T+1}$ and $y_{j,T+1}$, the (i, j) element of V_{11}^*, can be obtained using the decomposition

$$\mathrm{Cov}(y_{i,T+1}, y_{j,T+1}|R) = \mathrm{E}(\mathrm{Cov}(y_{i,T+1}, y_{j,T+1}|a, R)|R) + \mathrm{Cov}(\mathrm{E}(y_{i,T+1}|a, R), \mathrm{E}(y_{j,T+1}|a, R)|R). \tag{A.32}$$

To compute the first term in (A.32), observe using (A.30) that

$$\mathrm{Cov}(y_{i,T+1}, y_{j,T+1}|a, R) = b_i' V_{22}^* b_j + \tilde{\sigma}_{i,j}, \tag{A.33}$$

since, given the assumed properties of $u_{i,T+1}$, $\mathrm{Cov}(r_{2,T+1}, u_{i,T+1}|R) = 0$ for all i and

$$\mathrm{Cov}(u_{i,T+1}, u_{j,T+1}|R) = \mathrm{E}(\mathrm{Cov}(u_{i,T+1}, u_{j,T+1}|\theta, R)|R) = \tilde{\sigma}_{i,j}. \tag{A.34}$$

Taking the expectation of (A.33) with respect to a gives

$$\mathrm{E}(\mathrm{Cov}(y_{i,T+1}, y_{j,T+1}|a, R)|R) = \tilde{b}_i' V_{22}^* \tilde{b}_j + \mathrm{tr}[V_{22}^* \mathrm{Cov}(b_i, b_j'|R)] + \tilde{\sigma}_{i,j}. \tag{A.35}$$

To compute the second term in (A.32), observe using (A.31) that

$$\mathrm{E}(y_{i,T+1}|a, R) = [1 \; \tilde{E}_2'] a_i, \tag{A.36}$$

so

$$\mathrm{Cov}(\mathrm{E}(y_{i,T+1}|a, R), \mathrm{E}(y_{j,T+1}|a, R)|R) = [1 \; \tilde{E}_2'] \mathrm{Cov}(a_i, a_j'|R) [1 \; \tilde{E}_2']'. \tag{A.37}$$

Note that $\mathrm{Cov}(b_i, b_j'|R)$ and $\mathrm{Cov}(a_i, a_j'|R)$ are submatrices of $\mathrm{Var}(a|R)$ in (A.20).

To compute the remaining submatrices in (A.29), observe that

$$V^* = \mathrm{Var}(r_{T+1}|R) = \mathrm{E}(V|R) + \mathrm{Var}(E|R) = \tilde{V} + \mathrm{Var}(E|R), \tag{A.38}$$

which follows from the law of iterated expectations and the variance decomposition rule. Applying this decomposition to the lower-right submatrix gives

$$V_{22}^* = \tilde{V}_{22} + \mathrm{Var}(E_2|R), \tag{A.39}$$

and applying it to the off-diagonal submatrices gives

$$V_{12}^* = V_{21}^{*'} = \mathrm{E}(BV_{22}|R) + \mathrm{Cov}(\alpha + BE_2, E_2'|R) = \tilde{B}\tilde{V}_{22} + \tilde{B}\mathrm{Var}(E_2|R). \tag{A.40}$$

A.2. The first characteristic-based model

The likelihood function and the prior on (B, Σ, E_2, V_{22}) are the same as in the factor-based model presented in Section A.1. The only difference from the factor-based model is in the prior for α:

$$p(\alpha|B, E_2) \propto \exp\left\{ -\frac{1}{2\sigma_\alpha^2}(\alpha + BE_2)'(\alpha + BE_2)\right\}. \tag{A.41}$$

The conditional prior on α is normal and centered at the pricing restriction. Note that

$$\begin{aligned}
\frac{1}{\sigma_\alpha^2}(\alpha + BE_2)'(\alpha + BE_2) &= \frac{1}{\sigma_\alpha^2}(1\ \ E_2')\begin{pmatrix}\alpha' \\ B'\end{pmatrix}(\alpha\ \ B)\begin{pmatrix}1 \\ E_2\end{pmatrix} \\
&= \frac{1}{\sigma_\alpha^2}\mathrm{tr}(1\ \ E_2')AA'\begin{pmatrix}1 \\ E_2\end{pmatrix} \\
&= \mathrm{tr}\, A'\Psi A\left(\frac{1}{\sigma_\alpha^2}I_m\right) \\
&= a'\left(\frac{1}{\sigma_\alpha^2}I_m \otimes \Psi\right)a, \qquad \text{(A.42)}
\end{aligned}$$

where

$$\Psi = \begin{pmatrix} 1 & E_2' \\ E_2 & E_2E_2' \end{pmatrix}. \tag{A.43}$$

The full conditional posterior distribution of a is

$$\begin{aligned}
p(a|\cdot) &\propto \exp\left\{ -\frac{1}{2}\left[a'\left(\frac{1}{\sigma_\alpha^2}I_m \otimes \Psi\right)a + (a - \hat{a})'(\Sigma^{-1} \otimes Z'Z)(a - \hat{a})\right]\right\} \\
&\propto \exp\{ -\tfrac{1}{2}(a - \bar{a})'G(a - \bar{a})\}, \qquad \text{(A.44)}
\end{aligned}$$

where

$$G = \left(\frac{1}{\sigma_\alpha^2}I_m \otimes \Psi\right) + (\Sigma^{-1} \otimes Z'Z), \tag{A.45}$$

$$\bar{a} = G^{-1}(\Sigma^{-1} \otimes Z'Z)\hat{a}. \tag{A.46}$$

Hence, the full conditional posterior of a is a normal distribution:

$$a|\cdot \sim \mathrm{N}(\bar{a}, G^{-1}). \tag{A.47}$$

The full conditional posterior distribution of Σ is

$$p(\Sigma|\cdot) \propto |\Sigma|^{-(T+\nu+m+1)/2} \exp\{-\tfrac{1}{2}\mathrm{tr}[(Y-ZA)'(Y-ZA)+H]\Sigma^{-1}\}. \tag{A.48}$$

Hence, the full conditional posterior of Σ is an inverted Wishart distribution:

$$\Sigma^{-1}|\cdot \sim W(T+\nu, [(Y-ZA)'(Y-ZA)+H]^{-1}). \tag{A.49}$$

The full conditional posterior distribution of E_2 is

$$\begin{aligned} p(E_2|\cdot) \propto \exp\{ & -\tfrac{1}{2}[E_2'B'BE_2 + 2E_2'B'\alpha + \mathrm{tr}E_2\iota_T'\iota_T E_2' V_{22}^{-1} \\ & - 2E_2' V_{22}^{-1} X'\iota_T]\} \\ \propto & \exp\{-\tfrac{1}{2}(E_2 - \bar{E}_2)'P(E_2 - \bar{E}_2)\}, \end{aligned} \tag{A.50}$$

where

$$P = B'B + TV_{22}^{-1}, \tag{A.51}$$

$$\bar{E}_2 = P^{-1}(TV_{22}^{-1}\hat{E}_2 - B'\alpha). \tag{A.52}$$

Hence, the full conditional posterior of E_2 is a normal distribution:

$$E_2|\cdot \sim \mathrm{N}(\bar{E}_2, P^{-1}). \tag{A.53}$$

The full conditional posterior distribution of V_{22} is

$$p(V_{22}|\cdot) \propto |V_{22}|^{-(T+k+1)/2} \exp\{-\tfrac{1}{2}\mathrm{tr}(X-\iota_T E_2')'(X-\iota_T E_2')V_{22}^{-1}\}. \tag{A.54}$$

Hence, the full conditional posterior of V_{22} is an inverted Wishart distribution:

$$V_{22}^{-1}|\cdot \sim W(T, [(X-\iota_T E_2')'(X-\iota_T E_2')]^{-1}). \tag{A.55}$$

Posterior draws of (a, Σ, E_2, V_{22}) can be obtained using Gibbs sampling (see Casella and George, 1992). A chain of draws is constructed by making repeated draws from the full conditional distributions in (A.47), (A.49), (A.53), and (A.55). After an initial burn-in stage, these draws are taken from the joint posterior distribution $p(a, \Sigma, E_2, V_{22}|R)$. The predictive moments are computed as in Eqs. (A.28) through (A.40), where the required posterior moments in those equations are computed using the draws from the Gibbs chain.

A.3. The alternative characteristic-based model

As before, r_t is assumed to be normally distributed with mean E and covariance matrix V. Unlike the previous models, however, this model is not recast in a regression framework. The likelihood function is therefore

$$p(R|\theta) \propto |V|^{-T/2} \exp\{-\tfrac{1}{2}\mathrm{tr}(R - \iota_T E')'(R - \iota_T E')V^{-1}\}. \tag{A.56}$$

The set of parameters θ is now redefined as $\theta = (E, V, \gamma)$, where γ is defined below.

The prior on θ is

$$p(E, V, \gamma) = p(E|\gamma)p(V)p(\gamma), \tag{A.57}$$

where

$$p(E|\gamma) \propto \exp\left\{-\frac{1}{2\sigma_\alpha^2}(E - C\gamma)'(E - C\gamma)\right\}, \tag{A.58}$$

$$p(V) \propto |V|^{-(n+1)/2}, \tag{A.59}$$

$$p(\gamma) \propto \exp\left\{-\frac{1}{2\sigma_\gamma^2}(\gamma - \bar{\gamma})'(\gamma - \bar{\gamma})\right\}. \tag{A.60}$$

The prior on V is diffuse. The prior on γ is normal with a large variance σ_γ^2, so that it is noninformative. The normal prior on E given γ is centered at the pricing restriction.

The full conditional posterior distribution of E is

$$\begin{aligned} p(E|\cdot) &\propto \exp\left\{-\frac{1}{2}\left[(E - C\gamma)'(\sigma_\alpha^2 I_n)^{-1}(E - C\gamma) \right.\right. \\ &\qquad \left.\left. + (E - \hat{E})'\left(\frac{1}{T}V\right)^{-1}(E - \hat{E})\right]\right\} \\ &\propto \exp\{-\tfrac{1}{2}(E - E^c)'(V_E^c)^{-1}(E - E^c)\}, \end{aligned} \tag{A.61}$$

where

$$\hat{E} = R'\iota_t / T, \tag{A.62}$$

$$V_E^c = \left(\frac{1}{\sigma_\alpha^2} I_n + TV^{-1}\right)^{-1}, \tag{A.63}$$

$$E^c = V_E^c\left(\frac{1}{\sigma_\alpha^2} C\gamma + TV^{-1}\hat{E}\right). \tag{A.64}$$

Hence, the full conditional posterior of E is a normal distribution:

$$E|\cdot \sim \mathrm{N}(E^c, V_E^c). \tag{A.65}$$

The full conditional posterior distribution of V is

$$p(V|\cdot) \propto |V|^{-(T+n+1)/2}\exp\{-\tfrac{1}{2}\mathrm{tr}(R-\iota_T E')'(R-\iota_T E')V^{-1}\}. \tag{A.66}$$

Hence, the full conditional posterior of V is an inverted Wishart distribution:

$$V^{-1}|\cdot \sim W(T, [(R-\iota_T E')'(R-\iota_T E')]^{-1}). \tag{A.67}$$

The full conditional posterior distribution of γ is

$$p(\gamma|\cdot) \propto \exp\{-\tfrac{1}{2}[(E-C\gamma)'(\sigma_\alpha^2 I_n)^{-1}(E-C\gamma)+(\gamma-\bar{\gamma})'(\sigma_\gamma^2 I_L)^{-1}(\gamma-\bar{\gamma})]\}$$
$$\propto \exp\{-\tfrac{1}{2}(\gamma-\gamma^c)'(V_\gamma^c)^{-1}(\gamma-\gamma^c)\}, \tag{A.68}$$

where

$$V_\gamma^c = \left(\frac{1}{\sigma_\alpha^2}C'C + \frac{1}{\sigma_\gamma^2}I_L\right)^{-1}, \tag{A.69}$$

$$\gamma^c = V_\gamma^c\left(\frac{1}{\sigma_\alpha^2}C'E + \frac{1}{\sigma_\gamma^2}\bar{\gamma}\right). \tag{A.70}$$

Hence, the full conditional posterior of γ is a normal distribution:

$$\gamma|\cdot \sim \mathrm{N}(\gamma^c, V_\gamma^c). \tag{A.71}$$

Posterior draws of the parameters are again obtained using Gibbs sampling. A chain of draws is constructed by making repeated draws from the full conditional distributions in (A.65), (A.67), and (A.71). After an initial burn-in stage, these draws are taken from the joint posterior distribution $p(E, V, \gamma|R)$. The posterior mean of V and the posterior mean and variance of E are computed using the posterior draws. The predictive moments E^* and V^* are then computed using (A.27) and (A.38).

When $\sigma_\alpha = 0$, $E = C\gamma$. In this special case, the full conditional posteriors are derived for a reduced set of parameters (γ, V). Whereas V is drawn in the same manner as before, γ is now drawn from the following normal distribution:

$$\gamma|\cdot \sim \mathrm{N}(\gamma^0, V_\gamma^0), \tag{A.72}$$

where

$$V_\gamma^0 = \left(C'\left(\frac{V}{T}\right)^{-1}C + \frac{1}{\sigma_\gamma^2}I_L\right)^{-1}, \tag{A.73}$$

$$\gamma^0 = V_\gamma^0\left[C'\left(\frac{V}{T}\right)^{-1}\hat{E} + \frac{1}{\sigma_\gamma^2}\bar{\gamma}\right]. \tag{A.74}$$

References

Anderson, T.W., 1984. An Introduction to Multivariate Statistical Analysis.. Wiley, New York.

Black, F., Litterman, R., 1991. Asset allocation: combining investor views with market equilibrium. Journal of Fixed Income 1 (2), 7–18.

Black, F., Litterman, R., 1992. Global portfolio optimization. Financial Analysts Journal 48 (5), 28–43.

Brown, S.J., 1979. Optimal portfolio choice under uncertainty: a Bayesian approach In Bawa, V.S., Brown, S.J., Klein, R.W. (Eds.), Estimation Risk and Optimal Portfolio Choice. North-Holland, Amsterdam.

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

Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. Journal of Finance 52, 1–33.

Davis, J., Fama, E.F., French, K.R., 2000. Characteristics, covariances, and average returns: 1929–1997. Journal of Finance 55, 389–406.

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

Grinblatt, M., Titman, S., 1987. The relation between mean-variance efficiency and arbitrage pricing. Journal of Business 60, 97–112.

Huberman, G., Kandel, S., Stambaugh, R.F., 1987. Mimicking portfolios and exact arbitrage pricing. Journal of Finance 42, 1–9.

Ingersoll, J., 1984. Some results in the theory of arbitrage pricing. Journal of Finance 39, 1021–1039.

Jobson, J.D., Korkie, B., 1985. Some tests of linear asset pricing with multivariate normality. Canadian Journal of Administrative Sciences 2, 114–138.

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

Kandel, S., Stambaugh, R.F., 1996. On the predictability of stock returns: an asset allocation perspective. Journal of Finance 51, 385–424.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of the American Statistical Association 90, 773–795.

Klein, R.W., Bawa, V.S., 1976. The effect of estimation risk on optimal portfolio choice. Journal of Financial Economics 3, 215–231.

Leamer, E.E., 1978. Specification Searches: Ad Hoc Inference with Non-Experimental Data.. Wiley, New York.

Lintner, J., 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.C., 1995. Multifactor models do not explain the CAPM. Journal of Financial Economics 38, 3–28.

MacKinlay, A.C., Pástor, L., 2000. Asset pricing models: implications for expected returns and portfolio selection. Review of Financial Studies, in press.

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

Pástor, L., 2000. Portfolio selection and asset pricing models. Journal of Finance 55, 179–223.

Pástor, L., Stambaugh, R.F., 1999. Costs of equity capital and model mispricing. Journal of Finance 54, 67–121.

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

Shanken, J., 1987. A Bayesian approach to testing portfolio efficiency. Journal of Financial Economics 19, 195–215.

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

Wang, Z., 1998. Efficiency loss and constraints on portfolio holdings. Journal of Financial Economics 48, 359–375.

Zellner, A., Chetty, V.K., 1965. Prediction and decision problems in regression models from the Bayesian point of view. Journal of the American Statistical Association 60, 608–616.