The Effect of Risk Factors on Cost of Equity Estimation

Gregory L. Nagel
Mississippi State University

David R. Peterson
Florida State University

Robert S. Prati*
East Carolina University

Individual empirical tests conducted on cost of equity estimation methods increasingly have implied that sophisticated models are no more accurate than the simplest approaches. We test this implication by evaluating six commonly used methods for cost of equity estimation, contrasting them with a single factor model. Specifically, we explore whether adding risk factors to cost of equity models improves their predictive power both for individual firms and firm portfolios. When compared to the single factor model over horizons of one month to five years, we find mainstream models such as the CAPM and further evolved variants do not significantly reduce forecast error for portfolios. More importantly, the sophisticated models typically increase forecast error for individual firms. Either estimation errors cause the mainstream models to be rejected or the models are not well-specified and their assumptions need revision. Absent better models, we conclude ex ante estimation for cost of equity should utilize a simple single factor model for individual firms.

Introduction

The use of historical returns in estimating the cost of equity (K_E) long has been a tradition [Dimson, 1979; Fama and French, 1993 (FF93)]. While some recent

* We thank James M. Nelson for the insightful suggestion of extending the study to portfolios and for his advice on interpreting the results. Additional thanks to Yingmei Cheng for her econometric support that further aided the analysis of our results and to an anonymous referee at the *Quarterly Journal of Business and Economics* for valuable comments and suggestions. We appreciate the many helpful refinements of Gary Benesh and Don Nast; finally, we also appreciate the generosity of Ken French in supplying data through his website. A summer research grant from the College of Business at East Carolina University further provided support for this paper.

0747-5535/07/1600/0061/$2.50

authors, such as Claus and Thomas (2001), explore the use of analyst forecasts to compute a measure of expected returns or attempt other novel approaches, almost all of the widely used methods of estimating K_E involve historical returns. This can be easily verified by looking at most any corporate finance textbook (Brigham and Ehrhardt, 2002; Smart, Megginson, and Gitman, 2003; Brealey, Myers and Allen, 2006).

As determining K_E is a firm-specific or decision-specific task, however, it is important to examine the impact of various estimation techniques on individual firms. Elton (1999) argues that considerable empirical evidence exists to question whether historical returns should be used to estimate expected returns *imprimis*. The underlying problem with basing K_E estimates on historical returns seems to be the noise inherent to those returns. In fact, noise in mainstream models derived from historical returns has been sufficient in leading most studies to conclude that these models serve as poor estimators for K_E. This problem, already shown to exist with portfolios (Fama and French, 1997), is likely to be exacerbated when individual securities are considered. Nevertheless, many applications require *ex ante* estimates for K_E, and most estimation methods employ historical returns.

Ceteris paribus, gleaning the most accurate estimate possible remains the objective behind K_E models. After all, it is in the best interest of corporate managers to know which model provides superior accuracy. Considerable debate over which method should be used to estimate K_E, however, continues to dominate the literature on this topic. We seek to address this quagmire by examining the mainstream models for estimating K_E.

Fama and French (FF97) indicate that the capital asset pricing model (CAPM) is the most common choice for estimating K_E. Survey results reported by Bruner, Eades, Harris, and Higgins (1998) corroborate Fama and French's (1997) assertion, and Graham and Harvey (2001) confirm this once more with their own survey. An unanswered question remains, however: Even if the CAPM is the most commonly employed K_E model, could it be that other models offer superior forecasting?

We evaluate the forecasting ability of mainstream models, anticipating improvements as model complexity increases. Specifically, we investigate the improvement in realized forecast error gained by including risk premia and factor loadings. The forecasting errors of six well-known models are compared to a more naïve model for K_E, which assumes a fixed market risk premium and constrains beta to one. Lee, Myers, and Swaminathan (1999) (LMS) apply just such a model to large stocks (the Dow Thirty) as a part of their model for intrinsic firm valuation. In Lee, Myers, and Swaminathan (1999), K_E is represented as the sum of the risk-free rate (one-month Treasury bill) plus a risk premium averaged over many years. We refer to this as the LMS model and it represents our base model upon which each of the more sophisticated models successively are built. In their study of the Dow Thirty, Lee, Myers, and Swaminathan (1999) find that the mainstream CAPM and the FF93 three-factor

model do little to improve the ability to predict K_E for individual stocks over the baseline model, where they assume beta equals one for all stocks. We further develop these findings for a broad range of securities (and portfolios) by testing whether additional risk factors and factor loadings significantly reduce forecast errors when compared to the base model, as might be expected. In our evaluation of K_E models, it is important to clarify that our primary objective lies in finding the factors and factor loadings that provide statistically significant predictive power for estimating future returns.

The six mainstream models we investigate are:

- The CAPM,
- The Fama, and French (1993) three-factor model,
- The Carhart (1997) four-factor model, which builds on the Fama and French (1993) three-factor model by adding a momentum factor,
- A co-skewness model, which builds upon the CAPM model by adding a measure of risk associated with co-skewness,
- A co-kurtosis model, which builds upon the co-skewness model by adding a measure of risk associated with co-kurtosis, and
- A stock's own historical mean return.

Our findings are enlightening. For individual firms, the additional risk factors and factor loadings used by mainstream models fail to significantly improve forecasting performance relative to the LMS model for each holding period investigated (one-month to five years); rather, the more sophisticated models typically increase forecasting error. Our results hold with rare exception whether estimation error is analyzed by year, size, or book-to-market decile from 1953 to 1996. Moreover, as the holding period increases from one month to five years, the forecasting ability of the more complex models worsens with individual firms. The LMS forecasting error is less than that of mainstream models a majority of the time for one-month holding periods, increasing to at least 80 percent of the time for three-year holding periods, (and more than 60 percent for five-year holding periods). Further, for portfolios based on size, book-to-market, and industry groups, the mainstream models tested do not significantly improve upon the LMS model over horizons greater than one month.

In summary, we find the simplest model has the best forecasting ability among commonly-used models for K_E estimation. That is, the LMS model outperforms, generating the smallest forecast error. In the process, we show that the CAPM, the Fama and French (1993) three-factor model, the Carhart four-factor model, the co-skewness model, and the co-kurtosis model all fail to supersede the LMS model in forecasting future stock returns. Although previous studies, such as Fama and French (1997), Ghysels (1998) and Lee, Myers, and Swaminathan (1999), have established the poor performance of the first three of these models, our study is the first to consider individual firms. We also extend the literature by including (a) models that

have risk factors for co-skewness and co-kurtosis, (b) book-to-market and size groupings of stocks, (c) a momentum-based factor, and (d) returns from a variety of longer horizons. We thereby demonstrate the out-of-sample advantage of using a simple model to estimate K_E and illustrate the difficulty of implementing sophisticated asset pricing models based on historical returns due to underlying issues of estimation error.

Related Literature

The weaknesses of these mainstream methods for K_E estimation have not been left unquestioned in the financial literature. Empiricists routinely check these models for their ability to explain cross-sectional and time-series patterns in returns. For instance, Fama and French (1993) assert that the market return, size, and book-to-market equity factors explain most of the cross section of returns over time for industry portfolios. Yet, Fama and French (1997) show that their three-factor model and the CAPM's estimated forecasting standard error for portfolios of firms in the same industry is typically 3.0 percent per year. Both Fama and French (1997) and Pastor and Stambaugh (1999) find that uncertainty about factor premia and beta both contribute to weaknesses in models seeking an appropriate K_E. In fact, Fama and French (1997) conclude that "estimates of the cost of equity are distressingly imprecise" for portfolios and presume they are less precise for individual securities.

Building on literature originating with Kraus and Litzenberger (1976), Harvey and Siddique (2000) propose that three factors (book-to-market, size, and momentum) proxy for investors' preference for right-skewed returns distribution. They show that the CAPM model, augmented with a conditional skewness term, explains many aspects of the cross section of returns previously attributed to these three factors. Dittmar (2002) augments the CAPM model with terms for both skewness and kurtosis. He concludes that these two non-linear terms cause linear multi-factor models to lose their importance in explaining the cross section of returns. These results suggest that a model incorporating skewness and kurtosis may improve forecasting power; however, some authors argue the opposite may be true.

Ghysels (1998) proposes that the error in estimating a time-varying beta might outweigh the advantage gained by including it in the model, when compared to a model that assumes a constant beta. Similarly, it is possible that errors in estimating risk loadings as well as expected risk premia outweigh the advantage gained by including them in a model for estimating K_E. Baker and Wurgler (2000) even suggest that financing decisions may not be influenced so much by cost of capital estimation and traditionally viewed capital budgeting decision criteria. Rather, corporate financing decisions may be made as a function of market timing, with firms issuing relatively more equity than debt when the market is overvaluing firms as a whole (indicated by subsequent periods of significantly negative abnormal market returns) and vice-versa when the market may be undervaluing firms.

It is also worth mentioning that Gebhardt, Lee, and Swaminathan (2001) have commenced research in estimating an implied cost of capital, but their work has not yet been sufficiently developed to apply their findings to the empirical testing we conduct here. The initiatives undertaken by Fama and French (1997) and Lee, Martin, and Swaminathan suggest a simple model works best for applied K_E estimation in finance, yet several topics remain unexplored. Accordingly, we extend the basis of their work.

By directly comparing various K_E estimation models, we aim to help guide researchers and corporate managers who must select the best choice among them. First, we determine the forecasting error for all individual firms and for size, book-to-market, and industry portfolios. Size and book-to-market portfolio return prediction is important because some portfolio managers emphasize these characteristics in addition to industry portfolios. Second, we consistently use information available at or before the time the forecast is made to estimate K_E. This is also important in motivating the paper because forecasts are made by corporate managers in this manner. Third, Dittmar (2002) and Harvey and Siddique (2000) do not test the forecasting power of models incorporating skewness and kurtosis; we do. Fourth, where Fama and French (1997) look only at one-month horizons for portfolios (one-month predicted at T_0 compared to the one-month return three years later), we consider several horizons more common to firm investment, extending five years. Most importantly, where Fama and French (1997) analyze only industry portfolios, we also look at individual firms. This should be of interest to both practitioners and financial managers concerned with their own firm's K_E.

Hypothesis and Data

We employ six methods for K_E estimation: (a) the CAPM model developed by Sharpe (1964) and Lintner (1965), (b) the Fama and French (1993) (FF93) three-factor model, (c) the Carhart (1997) four-factor momentum model, (d) a version of the Harvey and Siddique (2000) model that incorporates skewness, (e) a model that incorporates kurtosis as suggested by the work of Dittmar (2002), (f) the stock's own historical average return and (g) a time-varying riskless rate plus the historical market risk premium, or the LMS model. These common models are delineated in Appendix A. The LMS model is a restricted model of the CAPM; the CAPM is a restricted model of the FF93 three-factor model as well as the models having terms for skewness and kurtosis; and the FF93 three-factor model is a restricted version of the four-factor Carhart model. As it is standard practice to test the validity of additional terms as they are added to a model, we test whether successively adding terms and coefficients to the LMS model leads to its rejection in favor of a more complex model. Thus, our null hypothesis:

H_0: The mean square error for the geometric average equity premium (LMS model) is less than or equal to the mean square error of the six alternative K_E forecast models.

The individual firm sample consists of all domestic non-financial common stock securities with prices greater than one dollar, as listed on the Center for Research in Security Prices (CRSP) monthly returns file from 1953[1] to 1996. Firm-specific CRSP stock return data are used to form estimates of K_E for a given horizon, using delisting returns when necessary. Size and book-to-market (B/M) portfolios are formed from CRSP monthly returns spanning from June 1953 to December 2001, and from annual book equity data from 1953 to 2001. Firms with book equity less than zero are excluded. Value-weighted decile portfolios based on both size and B/M are obtained from Kenneth French's website.[2] These portfolios are constructed using NYSE breakpoints. We also obtain value-weighted portfolios that break the market into ten industries as defined by Ken French's website.[3] The portfolios are rebalanced annually and include all NYSE, AMEX, and Nasdaq stocks for which data are available. Kenneth French's website also provides the values for all of the mimicking portfolios (R_m-R_f, SMB, HML, and UMD) and the one-month Treasury bill (required starting in April 1953 to proxy for the riskless security). For maturities of one year, three years, and five years, new-issue Treasury interest yields are obtained from the website of the Federal Reserve Bank of St. Louis,[4] also starting in April 1953.

We recognize that a complication can arise from overlapping returns, which might lead to overlap bias in the test statistics. That is, if a one-year horizon starts in January, then 11 of its monthly returns would overlap with a one-year horizon starting in February of the same year. We avoid this situation by commencing estimation of K_E for a given horizon at a particular year and month and then forcing subsequent estimations to avoid overlap. For instance, we start the three-year horizon estimate of K_E in January 1956. The next set of three-year horizon estimates for K_E is thus begun in January 1959. As a consequence of avoiding overlap, analysis results from three-year and five-year horizons are based on relatively fewer time periods and should thus be interpreted judiciously. Table 1 describes the annual sample size for horizons of one to five years.

Empirical Approach

Estimates of both beta and the expected return on the market are required for our first mainstream model, the CAPM. Consistent with Fama and French (1997), we

[1] We use data beginning in 1953 to avoid incorporating depression and World War II years, as many studies document different financial market characteristics in these years.
[2] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french
[3] Industry portfolio four digit Standard Industry Codes are listed on Kenneth French's website for each of the ten industry groups.
[4] The Federal Reserve Bank of St. Louis website is http://research.stlouisfed.org/index.html

define $R_{m,t}$ as the CRSP value-weighted monthly return index and estimate beta by using ordinary least squares (OLS) in the following time series regression equation:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i[R_{m,t} - R_{f,t}] + \varepsilon_{i,t} \quad (1)$$

where α_i is the estimated intercept term, β_i is the estimated slope, and $\varepsilon_{i,t}$ is a normally distributed mean-zero random error. Two samples are used for estimating the beta for firm i at time t. First, a five year rolling window of monthly data from month t-1 to month t-60 is used in a regression to find beta for firm i at time t. Second, a ten year rolling window from months t-1 to t-120 is used to estimate beta.

Once beta for firm i at time t is known, an estimate of the expected monthly equity risk premium, $E[R_{m,t}-R_{f,t}]$, is needed. As is commonly done, $E[R_{m,t} - R_{f,t}]$ is

Table 1–Individual Firm Sample Size and Lower Bounds on Cost of Equity (K_E)

Year	Sample size	1 Year horizon	3 Year horizon	5 Year horizon	Year	Sample size	1 Year horizon	3 Year horizon	5 Year horizon
1954	855	Y			1976	2056	Y		
1955	922	Y		Y	1977	1990	Y	Y	
1956	874	Y	Y		1978	1864	Y		
1957	851	Y			1979	1906	Y		
1958	856	Y			1980	1902	Y	Y	Y
1959	910	Y	Y		1981	1804	Y		
1960	922	Y		Y	1982	1779	Y		
1961	931	Y			1983	1831	Y	Y	
1962	934	Y	Y		1984	2227	Y		
1963	946	Y			1985	2325	Y		Y
1964	970	Y			1986	2726	Y	Y	
1965	1569	Y	Y	Y	1987	2751	Y		
1966	1639	Y			1988	2641	Y		
1967	1698	Y			1989	2738	Y	Y	
1968	1699	Y	Y		1990	2655	Y		Y
1969	1696	Y			1991	2495	Y		
1970	1696	Y		Y	1992	2664	Y	Y	
1971	1802	Y	Y		1993	3114	Y		
1972	1905	Y			1994	3457	Y		
1973	1896	Y			1995	3495	Y	Y	Y
1974	1839	Y	Y		1996	3883	Y		
1975	1921	Y		Y					

This table presents the sample size used to evaluate each of the models at one, three, and five-year horizons. For one-month horizons, all available firms are chosen monthly. The sample size for one-month horizon K_E estimates ranges from 859 firms on 1/1/1953 to 4,181 on 12/1/1996. All samples at one, three, and five-year horizons start in the month of January. For instance, one of the 15 three-year observation samples starts in January 1983. A "Y" (yes) indicates that a one-, three-, or five-year horizon analysis begins in January of the indicated year, with the stated number of data observations

estimated by the arithmetic average of the monthly risk premium from January 1945 until the month prior to the prediction at time t. (See for instance Brealey, Myers and Allen, 2006 or D'Mello and Shroff, 2000.) We start the average in 1945 to avoid using war year data. $R_{f,t}$ is set equal to the one-month Treasury bill rate observed at the beginning of month t. Then the expected value of the K_E is calculated from the CAPM. The estimates of $R_{m,t} - R_{f,t}$, $beta_i$, and $R_{f,t}$ described above are then substituted into CAPM in order to estimate the expected equity risk premium for firm i at time t. These approaches give a one-month expected return for firm i at time t. We obtain longer horizon forecasts by compounding the monthly expected return.

Our second model, the Fama and French (1993) three-factor model (FF3), states that a security's expected return depends upon the sensitivity of its return to the market risk premium, the SMB portfolio return, and the HML portfolio return. The slopes are estimated for firm i using:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i[R_{m,t} - R_{f,t}] + s_i SMB_t + h_i HML_t + \varepsilon_{i,t}. \quad (2)$$

As with the CAPM, the coefficients are estimated with rolling windows of both five and ten years. Once the estimated regression coefficients are found for firm i at time t, estimates of the expected risk premia are needed. The two approaches described for the CAPM are used again, this time to find the expected risk premia for the three factors. Next, $R_{f,t}$ is obtained using the same method as for the CAPM. The expected value for K_E is then calculated from the FF3 model.

Our third model, the Carhart (1997) model (Car4M), adds to the FF3 model in that each security's expected return depends not only upon the sensitivity of its return to the market return, the SMB portfolio return, and the HML portfolio return, but also on the UMD portfolio return as a measure of momentum. The slopes are estimated for firm i using:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i[R_{m,t} - R_{f,t}] + s_i SMB_t + h_i HML_t + u_i UMD_t + \varepsilon_{i,t}. \quad (3)$$

As with the FF3 model, we estimate the coefficients with rolling windows of five and ten years. When the four estimated regression coefficients are known for firm i at time t, estimates of the expected risk premia are once again needed. The two previous approaches are again used to find the expected risk premia for the four factors, and $R_{f,t}$ is again found in the same manner. The expected value of K_E for firm i at time t can then be calculated from the Car4M using these values.

Our fourth model (SKEW) is the CAPM augmented with a term for skewness. In this model, investor preference for a right-skewed distribution of returns is captured by lambda in the following model:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i\left[R_{m,t} - R_{f,t}\right] + \lambda_i\left[R_{m,t} - R_{f,t}\right]^2 + \varepsilon_{i,t}. \quad (4)$$

We follow the same estimation procedures described for the CAPM to obtain forecasts of expected returns for various horizons using the SKEW model.

Our fifth model (KURTOSIS) is the CAPM augmented with terms for both skewness and kurtosis. In this model, investor preference for a non-linear distribution of returns is captured by lambda on the skewness term and gamma on the kurtosis term in the following model:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i\left[R_{m,t} - R_{f,t}\right] + \lambda_i\left[R_{m,t} - R_{f,t}\right]^2 + \gamma_i\left[R_{m,t} - R_{f,t}\right]^3 + \varepsilon_{i,t} . \qquad (5)$$

We follow the estimation procedures described for the CAPM to obtain forecasts of expected returns for various horizons using the KURTOSIS model.

When we examined the sixth model, where the stock's own historical average return spawns an estimate of expected return, the resulting forecast error was significantly greater than that of any of the other models. This implies momentum may explain stock returns but cannot predict them. Accordingly, its further investigation was not warranted. We continue to use each of the remaining models, (CAPM, FF3, Car4M, SKEW, KURTOSIS, and LMS), to generate monthly estimates for each firm whenever data are available.

Finally, our base model, the LMS model, computes K_E by adding a riskless rate to an equity risk premium. The LMS model for expected return less the risk-free rate at time t is therefore denoted as:

$$E\left(R_{i,t}\right) - R_{f,t} = \left[\prod_{k=1}^{N_{t-1}}\left(R_{m,k} - R_{f,k}\right)\right]^{1/N_{t-1}} \qquad (6)$$

where k is the index value of time (in months) starting January 1945, $R_{m,k}$ is the return on the CRSP value weighted market portfolio at index time k, and N_{t-1} is the index value of time the month before an estimate of K_E is desired. The risk-free rate, $R_{f,k}$ is obtained in the same manner as with the CAPM. In this model the expected monthly equity risk premium is the geometric average excess return over the riskless rate for the value-weighted market portfolio from January 1945 to month t-1. We use the geometric, rather than arithmetic[5], excess return for this model because its estimate is compounded to compute expected returns over long horizons. Compounding the arithmetic average monthly return over long time periods, to obtain that period's return, consistently would give a higher return than is realized during the period, possibly leading to erroneous conclusions. Spanning the range of timeframes commonly encountered in both practical and academic work, we consider forward-looking investment horizons of one month, one year, three years, and five years.

We calculate a realized return for each investment horizon as the compounded monthly return over that horizon:

[5] In robustness tests we use the arithmetic average rather than the geometric average excess return; Table 3 shows that conclusions do not change.

$$\text{Realized horizon return}_{i,t} = \left[\prod_{k=t}^{T+t}(R_{i,k}+1)\right]-1 \quad (7)$$

where $R_{i,k}$ is the CRSP return for stock i at time t, and T is the number of months over the investment horizon. Whenever a company delists, we use the delisting return from CRSP as the last return for that firm. After the delisting date, we assume that the remaining money, if any, is reinvested in the CRSP value-weighted market index.

Significant Factors and Factor Loadings

The inclusion of additional terms in any model may lead to increased out-of-sample prediction errors due to estimation error and parameter instability. Accordingly, we cannot assume that factors with significant in-sample explanatory power also have out-of-sample predictive power. In our evaluation of K_E models, it is important to clarify that our objective lies in finding the factors and factor loadings that provide statistically significant predictive power for estimating future returns, not in-sample explanatory power.[6]

For single time periods, the procedure we use to find statistically significant predictive terms is based on paired differences in squared forecasting errors.[7] First, the squared forecasting error is computed for each horizon, model, and firm at time t. Then the following six differences are calculated:

1. Squared error of the CAPM less the squared error of the LMS model
2. Squared error of the FF3 model less the squared error of the LMS model
3. Squared error of the Car4M less the squared error of the LMS model
4. Squared error of the FF3 model less the squared error of the CAPM
5. Squared error of the Car4M less the squared error of the CAPM
6. Squared error of the Car4M less the squared error of the FF3.

In order, at each time t, the first three comparisons tell whether adding successive complexity to the LMS model significantly reduces forecasting error. If a significant reduction is not found in the first three comparisons (significant positive differences), then the factors and loadings captured by the CAPM, FF3, and the Car4M model are not significant, and LMS is shown to be the most parsimonious. If

[6] To find factors having significant in-sample explanatory power, a traditional test would involve estimating the full model (the Car4M) on the full sample (1953 to 2001), applying constraints to the factors and factor loadings, and then running the reduced models on the full sample. Statistical tests based on the sum-of-squared errors for the full and reduced models would then identify which factors and factor loadings are significant. As additional factors should always reduce modeling error, those which failed to provide significant additional explanatory power would simply not be included in the model imprimis.

[7] The forecast error for stock i at time t is defined to be the predicted return over the horizon less the actual return over the horizon. The squared error is the square of the forecast error for stock i at time t for a given horizon.

one model alone significantly reduces error compared to LMS, then we say there is information in its term(s). If two of the three models significantly reduce error compared to LMS, then we use either the fourth or fifth comparison to see whether the more complex model significantly reduces error compared to the less complex model. Either way, the more parsimonious model is identified. If all three models improve upon LMS, then we use the fourth and fifth comparison to see whether the FF3 or Car4M models improve upon the CAPM. If both improve upon the CAPM, then the sixth comparison is used to decide which model has significant terms. This logic identifies the factors and factor loadings that significantly improve out-of-sample forecasting ability at each time t. From these significance tests, we then compute the frequency with which each model is selected as the most parsimonious model at time t.

With multiple time periods, we use the same logic to identify the model that is most parsimonious for the full sample, not just for one time period. In this case, we find the averages of each of the six paired differences at each time t and determine the sign of these differences at each time t. Then a significance test is used to identify which model has significant terms and is the most parsimonious over time.

Our next step is determining significance. An analysis of the above six difference-distributions for single time periods shows, with probability exceeding 99.9 percent, they are non-normally distributed. This indicates we can simply apply a nonparametric sign test to determine when the differences are not equal to zero, thereby detecting significant differences. Similarly, when averaging in the case of multiple timeframes, the distributions are also often non-normally distributed. So we again employ the sign test. When we use a t-test for robustness, our conclusions do not change.

The above procedure identifies which of the four models (a) LMS, (b) CAPM, (c) FF3, or (d) Car4M, is the most parsimonious. We follow the procedure[8] just described to identify whether the LMS, CAPM, SKEW, or the KURTOSIS model is the most parsimonious in this group of models. If both procedures identify LMS as the most parsimonious, then LMS is the most parsimonious of the models.

An issue of highly-correlated returns potentially can arise in the test for multiple time periods. This may occur if overlapping samples are used, as mentioned earlier in the data section. There are two ways to address this issue. In testing for significant differences, if the correlation is estimated and used to correct the standard error estimate, then all sample information may be included. The correlation would increase the standard error, but the increased sample size would reduce it. Alternatively, non-overlapping samples could be used as described earlier. We apply the second approach to prevent error being introduced through estimates of correlation.

[8] The procedure is identical except that SKEW replaces FF3 and KURTOSIS replaces Car4M in the discussion.

Table 2–Model Results for Individual Firms by Year

Panel A: Test of Ho: MSE_t of LMS ≤ That of the CAPM, FF3, and Car4M Models

			The Percentage of the Time that Each Model is not Rejected as the Most Parsimonious at Time t			
Horizon	Sample Size	Winner	LMS	CAPM	FF3	Car4M
One Month	528	LMS	53.8	27.3	8.3	10.6
One Year	44	LMS	59.1	22.7	6.8	11.4
Three Years	15	LMS	86.7	13.3	0.0	0.0
Five Years	9	LMS	77.8	11.1	0.0	11.1

Panel B: Test of Ho: MSE_t of LMS ≤ That of the CAPM, SKEW, and KURTOSIS Models

			The Percentage of the Time that Each Model is not Rejected as the Most Parsimonious at Time t			
Horizon	Sample Size	Winner	LMS	CAPM	SKEW	KURTOSIS
One Month	528	LMS	58.1	23.7	5.3	12.9
One Year	44	LMS	65.9	20.5	0.0	13.6
Three Years	15	LMS	80.0	13.3	0.0	6.7
Five Years	9	LMS	66.7	0.0	22.2	11.1

Panel C

	The Percentage of Times in the Sample That the Sign of the Average Paired Difference between Two Models' MSE is Positive				
Horizon	CAPM - LMS	FF3 - LMS	Car4M - LMS	–SKEW - LMS	–KURTOSIS - LMS
One Month	51.7	54.2	55.7	50.9	54.4
One Year	63.6	68.2	70.5	63.6	68.2
Three Years	86.7	100.0	93.3	86.7	86.7
Five Years	66.7	88.9	88.9	66.7	77.8

Panel D

	Model Mean Squared Error					
Horizon	MSE of LMS	MSE of CAPM	MSE of FF3	MSE of Car4M	MSE of SKEW	MSE of KURTOSIS
One Month	0.01582	0.01582	0.01582	0.01586	0.01582	0.01584
One Year	0.33053	0.33380	0.33605	0.34340	0.33453	0.33779
Three Years	1.56058	1.62004	1.71558	1.87842	1.63184	1.72228
Five Years	8.05533	8.37748	10.52320	17.02684	8.53397	13.10573

In Panels A and B, a 5 percent significance level with a one sided hypothesis test is used to compare models at each time t in order to determine which is most parsimonious. The cost of equity (K_E) is estimated using six models (LMS, CAPM, FF3, Car4M, SKEW, KURTOSIS) for all firms at horizons of one month, one year, three years, and five years. Each model is used to estimate K_E for each firm monthly from 1953 to 1996. Table 1 identifies the years and months chosen for evaluation of each model at each horizon; a minimum of 855 firms are evaluated at each time t. Horizon is the holding period over which K_E is compared to realized returns. Forecasting error is the difference between the estimated K_E over the given horizon for firm i at time t and the realized return over the given horizon for firm i at time t. For each firm at a given date and horizon, the estimates of K_E are computed. Then the mean square forecasting error for all firms at time t (MSE_t) is found for each model. The average of MSE_t across all time periods for each model generates its MSE for each given horizon. The percentage of times that each model is not rejected as the most parsimonious at time t tests the hypothesis, *Ho: MSE_t of LMS is less than or equal to the alternative model MSE_t.* The significance level is 5 percent. The hypothesis test statistics are found by computing the paired differences in squared error for the model combinations described in the text. This table shows: (1) CAPM less LMS (2) FF3 less LMS, (3) Car4M less LMS, (4) SKEW less LMS, and (5) KURTOSIS less LMS. We apply the sign test to these differences, identifying the factors and factor loadings that significantly improve out-of-sample forecasting ability at each time t. From these significance

tests, we observe the frequency with which each model is selected as the most parsimonious model at time t based on statistical significance. The same logic is used to recognize the model that is most parsimonious for the full sample, over each given horizon; this is the reported as Winner. In this case, we find the averages of each of the paired differences at each time t and determine the sign of the difference at each time t. Then the sign test is used to identify which model has significant terms and is thus the most parsimonious over time. LMS is the one-month Treasury bill yield at time t plus the geometric average market risk premium from 1/1/1945 until the time, t, when an estimate of K_E is made. The market risk premium is computed as the difference between the return on the value-weighted market return of all NYSE, AMEX, NASDAQ stocks at time t less the one month Treasury bill yield at time t, or $R_{m,t} - R_{f,t}$. The CAPM estimate of the risk factor, beta, is found from the regression: $R_{i,t} - R_{f,t} = a_i + \beta_i[R_{m,t} - R_{f,t}]$. The beta is found by using a rolling regression from t-1 month to t-60 months. This estimate of beta is used to find the CAPM K_E as $E(R_{i,t})$ from $E(R_{i,t}) = R_{f,t} + \beta_{i,t}[E(R_{m,t}) - R_{f,t}]$, where $E(R_{i,t})$ is the expected return on stock i at time t, $R_{f,t}$ is the risk-free interest rate set equal to the one month Treasury bill rate at time t, and $E(R_{m,t}) - R_{f,t}$ is set equal to the average value of $R_{m,t} - R_{f,t}$ from 1/1/1945 to time t-1. The computations for the SKEW and KURTOSIS models are similar to the CAPM except that these models include terms for skewness and kurtosis, as described in the Empirical Approach section. The FF3 risk factors in $R_{i,t} - R_{f,t} = a_i + b_i[R_{m,t} - R_{f,t}] + s_i SMB_t + h_i HML_t + \varepsilon_{i,t}$ arise from a rolling regression of this equation over time t-1 month to t-60 months. SMB_t is the return on a portfolio that mimics the returns of a small stock portfolio minus the returns of a large stock portfolio. HML_t is the return on a portfolio that mimics the returns on a portfolio of low book-to-market equity stocks minus the returns on a portfolio of high book-to-market equity stocks. b_i, s_i, and h_i are the regression slopes on the market, size, and book-to-market risk factors, respectively, for firm i. These risk loading estimates are used to find the FF3 K_E as $E(R_{i,t})$ in $E(R_{i,t}) - R_{f,t} = b_{i,t}[E(R_{m,t}) - R_{f,t}] + s_{i,t}E[SMB_t] + h_{i,t}E[HML_t]$, where $E(SMB_t)$ and $E(HML_t)$ are set equal to the average of SMB_t and HML_t, respectively, from 1/1/1945 until time t-1 month. In $R_{i,t} - R_{f,t} = a_i + b_i[R_{m,t} - R_{f,t}] + s_i SMB_t + h_i HML_t + u_i UMD_t + \varepsilon_{i,t}$, (Car4M) the factor loadings come from using this equation in a rolling regression from time t-1 month to t-60 months. UMD_t is the return of a portfolio that mimics the returns on a portfolio of low momentum stocks minus the returns on a portfolio of high momentum stocks at time t. $u_{i,t}$ is the regression slope on the momentum factor, UMD_t, for firm i at time t. The factor loading estimates are used to find the Car4M K_E as $E(R_{i,t})$ in $E(R_{i,t}) - R_{f,t} = b_{i,t}[E(R_{m,t}) - R_{f,t}] + s_{i,t}E[SMB_t] + h_{i,t}E[HML_t] + u_{i,t}E[UMD_t]$. $E(UMD_t)$ is set equal to the average of UMD_t from 1/1/1945 until time t-1

Results for Individual Firms

Table 2 presents the results for individual firms by year. This table focuses on whether additional risk factors increase the explanatory power of models when compared to a single-factor, fixed-beta (beta = 1) geometric average return model (or the LMS model). Factor loadings here were estimated over 60 months with risk premia averaged from 1945 through one month before making the K_E estimate.[9] The data show clearly that additional risk factors and factor loadings used by mainstream models degrade forecasting ability relative to the LMS model, for all horizons.[10]

Panels A and B reports the frequency where a model is not rejected as the most parsimonious at time t, using a nonparametric sign-test with a one-sided significance

[9] The same information is computed both for factor loadings estimated over 120 months and for a series of rolling 60-month estimates of the risk factors. As the conclusions fail to change for any of the analyses performed, we do not present these results.

[10] We also use the absolute deviation to evaluate models in place of the squared deviation. Conclusions remain the same for both individual firms and for portfolios, so again we do not present these results.

level of five percent. Among the six models independently competing, the LMS model wins outright more than half of the time for all horizons. Further, as horizon increases, the percentage of the time when the LMS model is not rejected increases. In fact, at a three-year horizon, not only does LMS win at least 80 percent of the time, but CAPM takes second place in almost all remaining instances, with the FF3, Car4M, and SKEW models failing to win even once. A sign test indicates that none of the more complex models are able to significantly reduce forecasting error compared to the LMS model, for any horizon.

Panel C indicates the percentage of the time that a model has less forecasting error than its comparator model for a given stock. Here, when the average differences in mean squared error (MSE) for paired models are considered, the LMS model wins (the average difference between the paired models' MSE is positive) more than 50 percent of the time against each of the mainstream models, for all horizons.[11] Moreover, the proportion of observations where the LMS model wins increases again with horizon. For example, in 294 months of 528 (55.7 percent), the average paired difference in MSE at a one-month horizon for Car4M versus LMS (Car4M-LMS) is positive, indicating superiority of the LMS model. At a five-year horizon, 88.9 percent are greater.

Finally, Panel D demonstrates that MSE consistently increases with model complexity, at horizons of one to five years. Further, the MSE of the LMS model is less than that of the mainstream models for every horizon greater than one month. These results are consistent with the problem described by Ghysels (1998) who found that inaccurate estimation of conditional betas in a model negates their benefit. Overall then, the estimation of risk loadings and premia clearly introduces more error in the estimate of K_E than would occur by omitting them.

One problem may be that the empirical implementation of the theory is inadequate due to imprecision in estimation of factor loadings and risk premia. Although intuitively, the cost of equity should not be less than the yield on a riskless security, we nonetheless find that roughly fifteen percent of the K_E estimates fall short of their matched-horizon Treasury security. Investors obviously require higher returns for holding riskier equity assets; therefore, the U.S. Treasury security yield becomes a lower-bound on investor expectations. Yet, an upper-bound on K_E may prove more beneficial in reducing forecast errors because the distribution of K_E estimates is skewed with extreme positive values. These few extreme values likely contribute far more to forecasting error than a larger number of K_E estimates falling slightly below the lower-bound risk-free rate. Therefore, we create an upper-bound on K_E estimates

[11]As the more complex models fail to show a significant reduction in forecast error in the first three comparisons against the LMS model, the factors and loadings captured by the CAPM, FF3 and Car4M models are not significant, implying the last three of the six comparisons described in the previous section are unnecessary.

by winsorizing the top one percent of K_E estimates for each model, over all estimates made.[12]

In addition to robustness tests (discussed later), we show in Table 3 that bounding the K_E estimates does not change any of the conclusions presented thus far. While incorporating bounds results in smaller prediction errors for all models except LMS (its estimates are never outside the bounds), the reductions are insufficient to modify the conclusion. The LMS model still continues to win overwhelmingly. Thus, we determine that additional risk factors and loadings beyond the LMS model should not be included.[13]

Finally, while our results for individual firms by size and book-to-market deciles do not delineate quite as strongly as our results by year, they still unequivocally support LMS. The LMS model continues to win consistently over all horizons, in nearly all decile breakdowns. (In the interest of concision however, these results are available only upon request from the authors.) As such, we fail to reject the LMS model and conclude that grouping firms by size or book-to-market does not identify instances in which inclusion of additional risk factors and factor loadings improves model performance.

Results for Portfolios

Alternatives to the LMS model may provide superior predictive capabilities with portfolios because portfolio returns are expected to be more closely correlated with risk factors than are returns for individual firms. Further, averaging also may reduce security specific errors in model parameter estimates for portfolio returns. For these reasons, we analyze portfolio forecast errors using the same models as for individual firms, estimating model parameters in a manner similar to that used for individual securities. Though not as strong, our overall MSE results again lend the most support for the LMS model. That is, for all horizons greater than one month we fail to reject our null hypothesis.

Subsequent to the publication of Fama and French (1993), it became common protocol to examine portfolios grouped by size and book-to-market ratio, as well as by industry. Therefore, we first investigate whether alternative models capture more forecasting information than the LMS model, for portfolios grouped by size.[14] The

[12] When estimates for the cost of equity exceed the 99th percentile of all estimates, the 99th percentile value is substituted for the more extreme estimate. In this way, all observations are retained.

[13] As the bounds fail to change forecasting errors enough to reject the LMS model, we omit analysis of bounded K_E estimates for the remainder of this study.

[14] Value-weighted decile portfolios based on size are provided by Kenneth French at his website. These portfolios are formed from CRSP monthly returns from June 1926 to December 2001 and annual book equity data from 1926 to 2001 and are constructed using NYSE breakpoints. All firms on the NYSE, AMEX, and Nasdaq are included in the portfolios,

Table 3–Robustness Tests for Individual Firms Using All Models By Year

Panel A: Expected Return for the LMS Model is Computed Using the Arithmetic Average Equity Premium over the Treasury Bill Rate

	The Percentage of Times in the Sample that the Sign of the Average Paired Difference between Two Models' MSE is Positive					
Horizon	Winner	CAPM - LMS	FF3 - LMS	Car4M - LMS	SKEW - LMS	KURTOSIS - LMS
One Month	LMS	49.1	51.9	57.0	51.3	54.4
One Year	LMS	61.4	61.4	65.9	61.4	65.9
Three Years	LMS	73.3	80.0	86.7	80.0	73.3
Five Years	LMS	55.6	88.9	88.9	55.6	66.7

Panel B: Estimates of Expected Returns are Winsorized

	The Percentage of Times in the Sample that the Sign of the Average Paired Difference between Two Models' MSE is Positive					
Horizon	Winner	CAPM - LMS	FF3 - LMS	Car4M - LMS	SKEW - LMS	KURTOSIS - LMS
One Month	LMS	51.7	53.8	54.9	50.8	54.0
One Year	LMS	63.6	65.9	70.5	63.6	65.9
Three Years	LMS	86.7	100.0	86.7	86.7	86.7
Five Years	LMS	55.6	88.9	88.9	66.7	77.8

Panel C : Factor Loadings at Time "t" are Computed Using Information from the Prior Ten Years

	The Percentage of Times in the Sample that the Sign of the Average Paired Difference between Two Models' MSE is Positive					
Horizon	Winner	CAPM - LMS	FF3 - LMS	Car4M - LMS	SKEW - LMS	KURTOSIS - LMS
One Month	LMS	51.5	52.3	52.5	50.9	50.6
One Year	LMS	63.6	63.6	65.9	63.6	65.9
Three Years	LMS	80.0	100.0	80.0	80.0	80.0
Five Years	LMS	55.6	88.9	88.9	66.7	66.7

Panel D: The Risk Free Rate is Set Equal to a One-Month Yield that when Geometrically Compounded Equals the New Issue Ten Year Treasury Bond Yield

	The Percentage of Times in the Sample that the Sign of the Average Paired Difference between Two Models' MSE is Positive					
Horizon	Winner	CAPM - LMS	FF3 - LMS	Car4M - LMS	SKEW - LMS	KURTOSIS - LMS
One Month	LMS	49.1	51.9	57.0	50.9	54.4
One Year	LMS	61.4	61.4	65.9	63.6	68.2
Three Years	LMS	73.3	80.0	86.7	86.7	86.7
Five Years	LMS	55.6	88.9	88.9	66.7	77.8

Expected returns are computed as described in Table 2, with the exception noted in the header of each panel. The model that is most parsimonious at the 5 percent significance level for the full sample is the reported Winner; the hypothesis test to establish the "Winner" is described in Table 2. In this table we also report the percentage of times at each horizon that the sign of the average paired difference between the following model's mean squared error (MSE) is positive: (1) CAPM less LMS (2) FF3 less LMS, (3) Car4M less LMS, (4) SKEW less LMS, and (5) KURTOSIS less LMS

except for the omission of firms with book equity less than zero. In all cases, the portfolios are rebalanced yearly at the end of June. We use these portfolio returns from 1953 to 2001.

results presented in Table 4 indicate that the most commonly used alternative models (CAPM, FF3, and Car4M) fail to provide statistically significant information over and above LMS for size-based portfolios, at all horizons exceeding one month. The winning model is identified in the eighth row of each panel. Panel A shows that at a one-month horizon, the null hypothesis of LMS winning is rejected consistently. Of the ten deciles, the FF3 model wins five, CAPM three, and Car4M two. Thus, no discernable winner emerges at the one-month horizon. At all other horizons, however, (shown by row eight in Panels B through D) LMS wins every decile. In Appendix A, we show the same type of results for the LMS, CAPM, SKEW, and KURTOSIS models; again there is no discernable winner at the one month horizon, and the LMS model wins at all longer horizons. Therefore, we focus the following discussion on the more commonly used models shown in Table 4.

Unlike with individual firms, the frequency in which LMS wins outright with portfolios (the percentage of times shown in the first three rows of each panel) does not rise noticeably with increased horizon.[15] For LMS to win outright when compared to the more complex models, this figure must exceed 50 percent. At the one-month horizon (Panel A), LMS never wins outright, which may indicate the other models possess explanatory power. We investigate their MSEs accordingly.

While the eighth row of each panel presents the winning model (the most parsimonious model for the full sample), rows four through seven identify the MSE for each model. The last (tenth) row shows the difference found by subtracting the minimum MSE found among all four models from the winning model's MSE. For one month horizons, this difference is frequently positive, indicating a lower MSE for the other K_E models. Trading some bias in favor of very low variance, minimal MSEs are, of course, desirable when seeking to maximize the precision of a model's prediction. Thus, this begs the question of whether any tangible benefit can be derived by increasing model complexity when predicting size-based portfolio returns for a one-month horizon. Still, alternative model performance does improve statistically when portfolios are used, leading us to surmise that an alternative model has the potential to significantly improve forecasting ability at a one-month horizon. Nonetheless, we are unable to ascertain which alternative model should be used.

Beyond horizons of one month, we fail to reject LMS for all deciles at the five percent significance level. As many of the percentages in the first three rows of panels B through D fall below 50 percent however, LMS does not win outright with portfolios very often. The MSE of CAPM is frequently less than that of LMS, as shown in the first row of each panel. So, we again look to the MSEs. The MSE of the LMS model often exceeds the minimum MSE of the other models, especially for large firms (see the last row of Panels B, C, and D). For one year and longer horizons

[15] In some cases, it even declines. For example, the percentage of the time that the tenth decile MSE for CAPM is greater than the MSE for LMS is 45.1 percent at a one-month horizon (Panel A), but it falls to 33.3 percent by three and five years out (Panels C and D).

Table 4–Model Results for Portfolios Formed by Size Using NYSE Deciles

	S1 (Small)	S2	S3	S4	S5	S6	S7	S8	S9	S10 (Large)
Panel A: One Month Horizon; N = 528										
% CAPM MSE > LMS MSE	49.4	47.7	47.5	46.8	45.8	46.8	47.9	46.2	46.4	45.1
% FF3 MSE > LMS MSE	46.6	45.1	44.5	42.8	43.8	45.5	45.8	43.8	44.3	47.2
% Car4M MSE > LMS MSE	45.1	43.9	43.2	43.4	44.9	46.2	44.1	44.7	44.7	47.7
MSE Of LMS *100	0.338	0.311	0.299	0.282	0.257	0.239	0.227	0.214	0.187	0.161
MSE Of CAPM*100	0.335	0.309	0.297	0.280	0.255	0.237	0.225	0.212	0.185	0.158
MSE Of FF3*100	0.335	0.309	0.297	0.280	0.255	0.237	0.224	0.212	0.186	0.158
MSE Of Car4M*100	0.335	0.309	0.297	0.280	0.255	0.237	0.224	0.212	0.185	0.158
Winner	Car4M	FF3	FF3	FF3	CAPM	FF3	Car4M	CAPM	FF3	CAPM
(Winner MSE – LMS MSE) * 10,000	-0.239	-0.226	-0.225	-0.216	-0.204	-0.234	-0.219	-0.198	-0.189	-0.224
(Winner MSE – Min MSE) * 100,000	0.537	0.060	0.083	0.000	0.147	0.000	0.025	0.136	0.106	0.000
Panel B: One Year Horizon; N = 44										
% CAPM MSE > LMS MSE	50.0	45.5	43.2	45.5	43.2	56.8	50.0	50.0	50.0	47.7
% FF3 MSE > LMS MSE	54.5	43.2	40.9	40.9	38.6	52.3	47.7	50.0	54.5	50.0
% Car4M MSE > LMS MSE	40.9	47.7	45.5	45.5	45.5	50.0	50.0	54.5	43.2	50.0
MSE Of LMS *10	0.962	0.777	0.630	0.630	0.553	0.476	0.442	0.377	0.326	0.303
MSE Of CAPM*10	0.966	0.786	0.633	0.637	0.550	0.473	0.439	0.369	0.314	0.279
MSE Of FF3*10	0.962	0.776	0.623	0.625	0.544	0.464	0.432	0.367	0.315	0.281
MSE Of Car4M*10	0.953	0.783	0.619	0.628	0.547	0.465	0.432	0.369	0.314	0.282
Winner	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS
(Winner MSE – LMS MSE)	0	0	0	0	0	0	0	0	0	0
(Winner MSE – Min MSE) * 1,000	0.907	0.056	1.002	0.509	0.872	1.184	1.017	1.051	1.160	2.347
Panel C: Three Year Horizon; N = 15										
% CAPM MSE > LMS MSE	40.0	46.7	40.0	46.7	46.7	46.7	40.0	40.0	33.3	33.3
% FF3 MSE > LMS MSE	66.7	80.0	73.3	73.3	73.3	60.0	33.3	33.3	33.3	40.0
% Car4M MSE > LMS MSE	60.0	73.3	66.7	66.7	66.7	46.7	46.7	60.0	40.0	40.0
MSE of LMS	0.297	0.212	0.130	0.153	0.121	0.098	0.100	0.070	0.099	0.147
MSE of CAPM	0.313	0.224	0.146	0.169	0.125	0.098	0.097	0.065	0.084	0.130
MSE of FF3	0.308	0.215	0.138	0.157	0.121	0.092	0.089	0.063	0.085	0.135
MSE of Car4M	0.303	0.222	0.130	0.163	0.127	0.089	0.093	0.069	0.089	0.134
Winner	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS
(Winner MSE – LMS MSE)	0	0	0	0	0	0	0	0	0	0
(Winner MSE – Min MSE) * 10	0	0	0	0	0	0.089	0.107	0.075	0.145	0.166

Table 4 (cont.)–Model Results for Portfolios Formed by Size Using NYSE Deciles

	S1 (Small)	S2	S3	S4	S5	S6	S7	S8	S9	S10 (Large)
Panel D: Five Year Horizon; N = 9										
% CAPM MSE > LMS MSE	44.4	33.3	44.4	44.4	44.4	33.3	44.4	44.4	44.4	33.3
% FF3 MSE > LMS MSE	66.7	55.6	33.3	55.6	55.6	44.4	44.4	33.3	44.4	44.4
% Car4M MSE > LMS MSE	55.6	55.6	33.3	66.7	55.6	33.3	44.4	33.3	44.4	44.4
MSE of LMS	2.073	1.562	1.060	1.092	1.055	0.627	0.764	0.535	0.553	0.921
MSE of CAPM	2.201	1.571	1.115	1.107	1.065	0.587	0.687	0.452	0.417	0.770
MSE of FF3	2.247	1.569	1.107	1.107	1.040	0.558	0.652	0.426	0.413	0.805
MSE of Car4M	2.163	1.568	1.093	1.115	1.065	0.547	0.668	0.502	0.454	0.798
Winner	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS	LMS
(Winner MSE – LMS MSE)	0	0	0	0	0	0	0	0	0	0
(Winner MSE – Min MSE)	0.000	0.000	0.000	0.000	0.015	0.081	0.112	0.109	0.141	0.151

Value-weighted decile portfolios based on size are provided by Ken French at his website. These size portfolios are formed from CRSP monthly returns from June 1926 to December 2001 and annual book equity data from 1926 to 2001 and are constructed using NYSE breakpoints. Firms with book equity less than zero are not included. We use Ken French's size decile portfolio returns from 1953 to 2001. All firms on NYSE, AMEX, and NASDAQ are included in the portfolios. In all cases, the portfolios are rebalanced yearly. The four models (LMS, CAPM, FF3, and Car4M) for finding K_E are computed as defined in Table 2 using portfolio data. The sample is from 1953 to 1996. The sample selection for each horizon is shown in Table 1. Horizon is the holding period over which K_E is compared to realized returns. The null hypothesis is Ho: MSE_t of LMS is less than or equal to the alternative model MSE. The statistical test of Ho is defined in the text. The model that is most parsimonious at the 5 percent significance level for the full sample is the reported Winner. The mean square error (MSE) is defined in Table 2, as is the percentage of times in the sample that the sign of the average paired difference between two models' MSE is positive. These percentages are given here as (1) % CAPM MSE > LMS MSE, (2) % FF3 MSE > LMS MSE, and (3) % Car4M MSE > LMS MSE. Min MSE is the MSE of the model with the minimum MSE.

then, the alternative models show better forecasting ability for portfolios than for individual firms, (notably for large firm portfolios), but the results are not statistically significant.

Finally, although our results for portfolios analyzed by book-to-market and by industry do not delineate quite as strongly as our results by size, (just as with the results for individual firms by size and book-to-market), they still support LMS, increasing with horizon. By five years, the LMS model wins unequivocally in every industry and every book-to-market decile, with similar results for the SKEW and KURTOSIS models. (These results are available upon request from the authors.) Overall, for horizons one year and greater, we fail to reject the LMS model and conclude that grouping portfolios by size, industry, or book-to-market does not identify instances in which inclusion of additional risk factors and factor loadings improves model performance.

Robustness

Survey articles by Bruner, Eades, Harris, and Higgins (1998) and Graham and Harvey (2001) suggest that, in practice, a variety of proxies for the risk-free rate are used for the estimation of K_E. Yet while the most common industry practice employs the ten-year Treasury bond yield, theory asserts that the one-month Treasury bill yield should be used. We investigate the most commonly used models (CAPM, FF3, and Car4M) with alternative proxies at either end of the spectrum. Table 5 shows that the most common industry practice of using the ten-year Treasury bond yield leads to the minimum forecast MSE when compared to the other estimation procedures. Further, whereas Lee, Myers, and Swaminathan compute the risk premium from the arithmetic average, Ritter and Warr (2002) use the geometric average. Table 5 also shows that this change in estimation procedure does not affect our conclusions.

Additional robustness tests on all six models also are run in order to determine sensitivity to various inputs used. Table 3 shows that our conclusions are robust to (a) the use of the arithmetic average or geometric average version of the LMS model, (b) winsorizing unlikely estimates of expected returns as described in the results for individual firms section, (c) the time period (five or ten years) used to estimate factor loadings, and (d) estimating the risk free rate by using the one month Treasury bill rate or the ten year Treasury bond rate. All robustness checks demonstrate that each input change in the procedure fails to affect our conclusions.

It is also important to note that forecasting errors from the various models could cluster in time. Thus, conditions might exist under which a particular mainstream model (CAPM, FF3, Car4M, SKEW, or KURTOSIS) performs well in estimating a firm's K_E. To test for such clustering, we first compute the difference in forecast error between each mainstream model and the LMS model, each year, at a one-year horizon. We then apply the Kolmogorov-Smirnov one-sample test. No evidence of

Table 5–Robustness Tests for Individual Firms by Year Using Different Risk-Free Rates

Sample Size	Horizon	Rf1mth, LMS T-bill geo % LMS not Reject	Rf10T, LMS T-bill arithmetic % LMS not Reject	Rf10T, LMS T-bill geo % LMS not Reject	Rf10T, LMS 10T arithmetic % LMS not Reject	Rf10T, LMS 10T geo % LMS not Reject
528	0	53.8	52.5	52.1	53.4	53.0
44	1	59.1	59.1	56.8	54.5	63.6
15	3	86.7	60.0	80.0	73.3	66.7
9	5	77.8	44.4	66.7	55.6	55.6
	Horizon	MSE LMS	MSE LMS	MSE LMS	MSE LMS	MSE LMS
	0	0.0158	0.0158	0.0158	0.0158	0.0158
	1	0.3305	0.3300	0.3305	0.3293	0.3300
	3	1.5606	1.5582	1.5606	1.5558	1.5582
	5	8.0553	8.0431	8.0553	8.0283	8.0431

% LMS not Reject is the percentage of times that LMS is not rejected as the most parsimonious at time t. Sample size, Horizon, MSE, LMS, % LMS not Reject, and Winner are explained in Table 2. Rf1mth, LMS T-bill geo indicates that the three models (CAPM, FF3, and Car4M described in Table 2) are computed using the one-month Treasury bill yield as the risk-free rate. The LMS model (described in Table 2) risk premium is computed using the one-month Treasury bill yield as the risk-free rate. The risk premium is computed as the geometric average monthly risk premium. Rf10T, LMS, T-bill, and arithmetic indicate that the three models are computed using the one-month yield that when geometrically compounded equals the new issue ten year Treasury bond yield as the risk-free rate. The LMS risk premium is computed using the one-month Treasury bill yield as the risk-free rate. The risk premium is computed as the arithmetic average monthly risk premium. Rf10T, LMS T-bill geo indicates that the three models are computed using the one-month yield that when geometrically compounded equals the new issue ten year Treasury bond yield as the risk-free rate. The LMS risk premium is computed using the one-month Treasury bill yield as the risk-free rate. The risk premium is computed as the geometric average monthly risk premium. Rf10T, LMS 10T arithmetic indicates that all of the models described in Table 2 use a risk-free rate that is equal to the one month yield that when geometrically compounded equals the new issue 10 year Treasury bond yield. The LMS risk premium is computed as the arithmetic average monthly risk premium. Rf10T, LMS 10T geo indicates that all of the models described in Table 2 use a risk-free rate that is equal to the one month yield that when geometrically compounded equals the new issue ten year Treasury bond yield. The LMS risk premium is computed as the geometric average monthly risk premium

clustering over time appears, even when the significance level is relaxed to 20 percent.

Going one step further, we also examine how the LMS model performs against the mainstream models at the troughs and peaks of the seven business cycles that have occurred since 1953. During the seven peak years, LMS cannot be rejected as the most parsimonious model (winning three of seven). During the seven trough years, LMS wins outright (winning four of seven) and is less biased on average than any of the other models.

Summary and Conclusions

The models most widely used in forming K_E estimates are based on historical returns. When making direct comparisons of these models, we find the best ex ante estimation method available to financial managers is essentially the CAPM with beta restricted to one; that is, a naïve model where K_E equals the risk-premium added to the risk-free rate. Our results prove robust to different constructs of both the risk-free rate and the equity risk-premium. The more theoretically advanced models we reject consist of the CAPM, the Fama-French three-factor model, the Carhart four-factor model, and the CAPM augmented with both a skewness and a kurtosis term. Each of these models is a more sophisticated version of the naïve LMS model, yet none significantly improves upon the forecasting ability of the base LMS model. We conclude that forecast error caused by estimating factor loadings and expected risk premia in the more complex models exceeds the precision gained by including the risk factors. In other words, both parametric and nonparametric statistical tests show that increasing model complexity fails to significantly reduce forecast error.

The LMS model makes the trivial assumption that the same K_E is applicable to all stocks, at any one point in time, by assigning all stocks a beta of one. When factor loadings and risk premia are added to the base-model (the LMS model) to arrive at more sophisticated models, mean square forecasting errors increase for all holding periods we investigate (one month to five years) for individual firms. LMS generates the least forecast error followed in order by the CAPM, the CAPM augmented with a skewness term, the three-factor linear model, the CAPM augmented with both a skewness and a kurtosis term, and the four-factor linear model. With individual securities, therefore, forecast error most often increases with model complexity. These individual firm results also hold when firms are grouped by size or book-to-market.

With portfolios formed by size, industry, and book-to-market, we also find the more sophisticated models fail to provide significant improvements in forecasting accuracy over LMS, except marginally at one-month horizons. Further, while it is possible that the alternative models capture some information at a one-month horizon that LMS does not, no discernable pattern emerges to identify which model consistently offers the most improvement. Moreover, for all horizons beyond a month, the alternative models fail to show a significant improvement over LMS when applied to

portfolios of securities. This again suggests the theoretically more advanced models inherently possess either assumption (misspecification) or estimation errors.

The underlying problem of the mainstream K_E estimates may lie in the approaches used to find estimates for the risk factors and risk loadings that capture future expectations, but addressing those approaches goes beyond the scope of our work. While we certainly are not contending that all stocks are equivalently risky and should be assigned a beta of one, our overall results clearly demonstrate that either estimation errors cause the more sophisticated models to be rejected or the assumptions of the models need revision, both for individual securities and portfolios, at various horizons commonly considered in practice. We conclude that ex ante K_E estimation should utilize the simplest model for individual firms, until better models are developed.

As a main goal in financial modeling is helping managers to accurately estimate their expected K_E, the implications of this work are widespread. First, regardless of its simplicity, the LMS model outperforms more theoretically sophisticated models. This result underscores the inescapable fact that K_E estimates based on historical returns tend to be very noisy. This excessive noise remains present irrespective of the asset-pricing model employed. Our results indicate that the addition of risk factors to the most basic K_E estimation model increases noise sufficiently for the simple model to outperform the theoretically superior models. For individual firms, forecast error often increases, on average; and mainstream models do not significantly improve upon the simple model for portfolios either. This all serves to support the contention that the most widely-used models serve as poor estimators for K_E.

A second implication is the clear conclusion that managers and researchers should estimate the expected cost of equity capital for both individual firms and portfolios with a plainly trivial model, as opposed to a rigorous mainstream model. Despite the minimalism of the LMS model, our findings suggest that managers will simply achieve better results when using it or one similar.

Finally, our research indirectly supports Elton (1999), who argues that improved methods of estimating K_E are necessary, and any further statistical testing and development of models that rely on realized returns in order to estimate expected returns is unlikely to yield fruitful results. Decades of work refining econometric tests and developing conditional versions of asset pricing models has made little progress toward the ultimate goal of finding and presenting knowledge that improves financial management. Recent promising alternative approaches toward obtaining useful estimates such as forward-looking equity premiums and models incorporating behavioral aspects may one day provide a better solution, but the common use of additional risk factors today certainly does not seem to help. From the perspective of forecasting accuracy, financial managers should be indifferent among the various mainstream models, as none stand out as superior. A simple model works as well or better than the more advanced models, and in the end, may prove the best choice.

References

1. Baker, Malcolm, and Jeffrey Wurgler, "The Equity Share in New Issues and Aggregate Stock Returns," *Journal of Finance,* 55, no. 5 (October 2000), pp.2219-2257.

2. Brealey, Richard A., Stewart C. Myers, and Franklin Allen, *Principles of Corporate Finance*, 8th ed. (McGraw-Hill/Irwin, 2006).

3. Brigham, Eugene F., and Michael C. Ehrhardt, *Financial Management Theory and Practice*, 10th ed. (Thomson South-Western, 2002).

4. Bruner, Robert F., Kenneth M. Eades, Robert S. Harris, and Richard C. Higgins, "Best Practices in Estimating the Cost of Capital: Survey and Synthesis," *Financial Practice and Education,* 8, no. 1 (Spring/Summer 1998), pp. 13-28.

5. Carhart, Mark M., "On Persistence in Mutual Fund Performance," *Journal of Finance,* 52, no.1 (March 1997), pp. 57-82.

6. Claus, James, and Jacob Thomas, "Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings for Domestic and International Stock Markets," *Journal of Finance,* 56, no. 5 (October 2001), pp. 1629-1666.

7. Constantinides, G.M., "Rational Asset Prices," *Journal of Finance,* 57, no. 4 (August 2002), pp. 1567-1591.

8. D'Mello, Rajan and Pervin K. Shroff, "Equity Undervaluation and Decisions Related to Repurchase Tender Offers: An Empirical Investigation," *Journal of Finance,* 55, no. 5 (October 2000), pp. 2399-2424.

9. Dimson, Elroy, "Risk Measurement when Shares are Subject to Infrequent Trading," *Journal of Financial Economics,* 7, no. 2 (1979), pp. 197-226.

10. Dittmar, Robert F., "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns," *Journal of Finance* 57 (2002), pp. 369-403.

11. Elton, Edwin J., "Expected Returns, Realized Returns, and Asset Pricing Tests," *Journal of Finance,* 54, no. 4 (August 1999), pp.1199-1220.

12.Fama, Eugene F., and Kenneth R. French, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, no. 1 (1993), pp. 3-56.

13. Fama, Eugene F., and Kenneth R. French, "Industry Cost of Equity," *Journal of Financial Economics,* 43, no.2 (1997), pp. 153-193.

14. Gebhardt, William R., Charles M.C. Lee, and Bhaskaran Swaminathan, "Toward an Implied Cost of Capital," *Journal of Accounting Research,* 39, no.1 (June 2001), pp.135-176.

15. Ghysels, Eric, "On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt?" *Journal of Finance,* 53, no. 2 (April 1998), pp. 549-573.

16. Graham, John R., and Campbell R. Harvey, "The Theory and Practice of Corporate Finance: Evidence from the Field," *Journal of Financial Economics,* 60 (2001), pp. 187-243.

17. Harvey, Campbell R., and Akhtar Siddique, "Conditional Skewness in Asset Pricing Tests," *Journal of Finance* 55 (2000), pp. 1263-1295

18. Jegadeesh, Narasimhan, and Sheridan Titman, "Returns to Buying Winners and Selling Losers," *Journal of Finance,* 48, no. 1 (March 1993), pp. 65-91.

19. Kraus, Alan, and Robert H. Litzenberger, "Skewness Preference and the Valuation of Risk Assets," *Journal of Finance*, 31, no. 4 (September 1976), pp. 1085-1100.

20. Lee, Charles M., James Myers, and Bhaskaran Swaminathan, "What is the Intrinsic Value of the Dow," *Journal of Finance*, 54, no. 5 (October 1999), pp. 1693-1741.

21. Lintner, John, "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics,* 47, no. 1 (February 1965), pp. 13-37.

22. Pastor, Lubos, and Robert F. Stambaugh, "Costs of Equity Capital and Model Mispricing," *Journal of Finance,* 54, no.1 (February 1999), pp. 67-121.

23. Ritter, Jay R., and Richard S. Warr, "The Decline of Inflation and the Bull Market of 1982-1999," *Journal of Financial and Quantitative Analysis,* 37, no. 1 (March 2002), pp. 29-61.

24. Sharpe, William F., "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance,* 19, no.3 (September 1964), pp. 425-442.

25. Smart, Scott B., William L. Megginson, and Lawrence J. Gitman, *Corporate Finance*, 1st ed. (Thomson South-Western, 2003).

Appendix
Development of Models

The CAPM implies that the expected cost of equity less the risk-free rate for firm i at time t is given by:

$$E(R_{i,t}) - R_{f,t} = \beta_{i,t}[E(R_{m,t}) - R_{f,t}] \quad \text{(A1)}$$

where $E(R_{i,t})$ is the expected return on stock i at time t, $R_{f,t}$ is the risk-free interest rate[16] at time t, $E(R_{m,t})$ is the expected return on the value-weighted market portfolio at time t, and $\beta_{i,t}$ is the systematic risk of stock i at time t.

For the Fama and French (1993) three-factor time series model, the expected return less the risk-free rate at time t is:

$$E(R_{i,t}) - R_{f,t} = \beta_{i,t}[E(R_{m,t}) - R_{f,t}] + s_{i,t}E[SMB_t] + h_{i,t}E[HML_t] \quad \text{(A2)}$$

where $E(SMB_t)$ is the expected return of a portfolio that mimics the returns on a portfolio of small stocks minus the returns on a portfolio of large stocks, $E(HML_t)$ is the expected return of a portfolio that mimics the returns on a portfolio of high book-to-market equity stocks minus the returns on a portfolio of low book to market equity stocks, and $\beta_{i,t}, s_{i,t}$, and $h_{i,t}$ are the factor loadings on the market, size, and book-to-market risk factors, respectively, for firm i at time t. The Fama and French (1993) model captures much of the cross-sectional spread of returns.

Attempting to explain more of the return spread, Jegadeesh and Titman (1993) define momentum as an additional risk factor. To capture momentum, they create a four-factor model employing a UMD (up-minus-down) factor as the fourth factor. The UMD factor is a mimicking portfolio that captures one-year momentum in returns. Carhart (1997) details the construction of these factors.[17] The expected return less the risk-free rate at time t is:

$$\begin{aligned} E(R_{i,t}) - R_{f,t} = {} & \beta_{i,t}[E(R_{m,t}) - R_{f,t}] + s_{i,t}E[SMB_t] \\ & + h_{i,t}E[HML_t] + u_{i,t}E[UMD_t] \end{aligned} \quad \text{(A3)}$$

where $E(UMD_t)$ is the expected return of a portfolio that mimics the returns on a portfolio of low momentum stocks minus the returns on a portfolio of high

[16] Although Bruner, Eades, Harris, and Higgins (1998) find that the vast majority of corporate managers use a maturity of at least ten years for the riskless security, an abundance of theoretical literature indicates the one-month Treasury bill rate should be used to proxy for riskless asset returns.

[17] UMD is available on the Kenneth French website. According to Carhart (1997), UMD "is constructed as the equal-weight average of firms with the highest 30 percent eleven-month returns lagged one month minus the equal weight average of firms with the lowest 30 percent eleven–month returns lagged one month." The UMD portfolio includes all NYSE, AMEX, and NASDAQ stocks; it is reformed monthly.

momentum stocks at time t, and $u_{i,t}$ is the regression slope on the momentum factor, UMD_t, for firm i at time t.

The SKEW model is the CAPM augmented with a second order moment term. The SKEW model implies that the expected cost of equity less the risk-free rate for firm i at time t is given by:

$$E(R_{i,t}) - R_{f,t} = \beta_{i,t}\left[E(R_{m,t}) - R_{f,t}\right] + \lambda_{i,t}\left[E(R_{m,t}) - R_{f,t}\right]^2 \quad \text{(A4)}$$

where $E(R_{i,t})$ is the expected return on stock i at time t, $R_{f,t}$ is the risk-free interest rate at time t, $E(R_{m,t})$ is the expected return on the value-weighted market portfolio at time t, and $\beta_{i,t}$ is the systematic risk of stock i at time t.

The KURTOSIS model is the SKEW model augmented with a third order moment term. The KURTOSIS model implies that the expected cost of equity less the risk-free rate for firm i at time t is given by:

$$E(R_{i,t}) - R_{f,t} = \beta_{i,t}\left[E(R_{m,t}) - R_{f,t}\right] + \lambda_{i,t}\left[E(R_{m,t}) - R_{f,t}\right]^2 + \gamma_{i,t}\left[E(R_{m,t}) - R_{f,t}\right]^3 \quad \text{(A5)}$$

Ritter and Warr (2002) point out that the Fama and French (1993) estimates of K_E tend to be much more variable than the market risk premium; however, they find their results are not sensitive to the method of estimating K_E.[18] Therefore they, like Lee, Myers, and Swaminathan (1999), estimate K_E for all firms in their study, the Dow Thirty, as a time-varying riskless rate plus an historical market risk premium; in other words, every firm is assigned the same cost of equity at time t. The riskless rate used is the monthly Treasury bill rate. The equity risk premium is the geometric average excess return of the market portfolio over the one-month Treasury bill rate. The LMS model for expected return less the risk-free rate at time t is therefore denoted as:

$$E(R_{i,t}) - R_{f,t} = \left[\prod_{k=1}^{N_{t-1}} (R_{m,k} - R_{f,k})\right]^{1/N_{t-1}} \quad \text{(A6)}$$

where k is the index value of time (in months) starting January 1945, $R_{m,k}$ is the return on the market portfolio at index time k, and N_{t-1} is the index value of time the month before an estimate of K_E is desired. The LMS model implicitly assumes all stocks have a beta of one; thus, every stock has the same ex ante estimate of K_E, at any given time.

[18] Penman and Sougiannis (1998) find a similar insensitivity of their results to various estimates of the cost of equity.