MSCI

Model Insight

The Barra US Equity Model (USE4)

Empirical Notes

Yang Liu

Jose Menchero

D. J. Orr

Jun Wang

September 2011

Contents

1. Introduction

1.1. Model Highlights

This document provides empirical results and analysis for the new Barra US Equity Model (USE4). These notes include extensive information on factor structure, commentary on the performance of select factors, an analysis of the explanatory power of the model, and an examination of the statistical significance of the factors. Furthermore, these notes also include a thorough side-by-side comparison of the forecasting accuracy of the USE4 Model and the USE3 Model, its predecessor. The methodological details underpinning the USE4 Model may be found in the companion document: *USE4 Methodology Notes*, described by Menchero, Orr, and Wang (2011).

Briefly, the main advances of USE4 are:

- An innovative Optimization Bias Adjustment that improves risk forecasts for optimized portfolios by reducing the effects of sampling error on the factor covariance matrix
- A Volatility Regime Adjustment designed to calibrate factor volatilities and specific risk forecasts to current market levels
- The introduction of a country factor to separate the pure industry effect from the overall market and provide timelier correlation forecasts
- A new specific risk model based on daily asset-level specific returns
- A Bayesian adjustment technique to reduce specific risk biases due to sampling error
- A uniform responsiveness for factor and specific components, providing greater stability in sources of portfolio risk
- A set of multiple industry exposures based on the Global Industry Classification Standard (GICS®)
- An independent validation of production code through a double-blind development process to assure consistency and fidelity between research code and production code
- A daily update for all components of the model

The USE4 Model is offered in short-term (USE4S) and long-term (USE4L) versions. The two versions have identical factor exposures and factor returns, but differ in their factor covariance matrices and specific risk forecasts. The USE4S Model is designed to be more responsive and provide more accurate forecasts at a monthly prediction horizon. The USE4L model is designed for longer-term investors willing to trade some degree of accuracy for greater stability in risk forecasts.

2. Methodology Highlights

2.1. Optimization Bias Adjustment

One significant bias of risk models is the tendency to underpredict the risk of optimized portfolios, as demonstrated empirically by Muller (1993). More recently, Shepard (2009) derived an analytic result for the magnitude of the bias, showing that the underforecasting becomes increasingly severe as the number of factors grows relative to the number of time periods used to estimate the factor covariance matrix. The basic source of this bias is estimation error. Namely, spurious correlations may cause certain stocks to appear as good hedges in-sample, while these hedges fail to perform as effectively out-of-sample.

An important innovation in the USE4 Model is the identification of portfolios that capture these biases and to devise a procedure for correcting these biases directly within the factor covariance matrix. As shown by Menchero, Wang, and Orr (2011), the *eigenfactors* of the sample covariance matrix are systematically biased. More specifically, the sample covariance matrix tends to tends to underpredict the risk of low-volatility eigenfactors, while overpredicting the risk of high-volatility eigenfactors. Furthermore, removing the biases of the eigenfactors essentially removes the biases of optimized portfolios.

In the context of the USE4 Model, eigenfactors represent portfolios of the original pure factors. The eigenfactor portfolios, however, are special in the sense that they are mutually uncorrelated. Also note that the number of eigenfactors equals the number of pure factors within the model.

As described in the *USE4 Methodology Notes*, we estimate the biases of the eigenfactors by Monte Carlo simulation. We then adjust the predicted volatilities of the eigenfactors to correct for these biases. This procedure has the benefit of building the corrections directly into the factor covariance matrix, while fully preserving the meaning and intuition of the pure factors.

2.2. Volatility Regime Adjustment

Another major source of risk model bias is due to the fact that volatilities are not stable over time, a characteristic known as *non-stationarity*. Since risk models must look backward to make predictions about the future, they exhibit a tendency to underpredict risk in times of rising volatility, and to overpredict risk in times of falling volatility.

Another important innovation in the USE4 Model is the introduction of a Volatility Regime Adjustment for estimating factor volatilities. As described in the *USE4 Methodology Notes*, the Volatility Regime Adjustment relies on the notion of a cross-sectional bias statistic, which may be interpreted as an *instantaneous* measure of risk model bias for that particular day. By taking a weighted average of this quantity over a suitable interval, the non-stationarity bias can be significantly reduced.

Just as factor volatilities are not stable across time, the same holds for specific risk. In the USE4 Model, we apply the same Volatility Regime Adjustment technique for specific risk. We estimate the adjustment by computing the cross-sectional bias statistic for the specific returns.

2.3. Country Factor

Traditionally, single country models (e.g., USE3) have included industry and style factors, but no Country factor. An important improvement with the USE4 Model is to explicitly include the Country factor, which is analogous to the World factor in the Barra Global Equity Model (GEM2), as described by Menchero, Morozov, and Shepard (2008, 2010).

One significant benefit of the Country factor is the insight and intuition that it affords. For instance, as discussed in the *USE4 Methodology Notes*, the USE4 Country factor portfolio can be cleanly interpreted as the cap-weighted country portfolio. Furthermore, the Country factor disentangles the pure industry effect from the overall market effect, thus providing a cleaner interpretation of the industry factors.

Without the Country factor, industry factors represent portfolios that are 100 percent net long the particular industry, with zero net weight in every other industry. With the Country factor, by contrast, industry factors represent *dollar-neutral* portfolios that are 100 percent long the industry and 100 percent short the Country factor; that is, industry performance is measured net of the market.

Dollar-neutral industry factor portfolios are important from an attribution perspective. For instance, suppose that a portfolio manager is overweight an industry that *underperforms* the market, but which nonetheless has a *positive* return. Clearly, overweighting an underperforming industry *detracts* from performance. If the industry factors are represented by net-long portfolios, however, an attribution analysis would spuriously show that overweighting the underperforming industry contributed *positively* to performance. This non-intuitive result is resolved by introducing the Country factor, which makes the industry factor portfolios dollar-neutral and thereby produces the intuitive result that overweighting an underperforming industry detracts from performance. Including the Country factor also resolves other problematic issues in risk attribution, as described by Davis and Menchero (2011).

Another benefit of the Country factor pertains to improvements in risk forecasting. Intuitively and empirically, we know that industries tend to become more highly correlated in times of financial crisis. As shown in the *USE4 Methodology Notes*, the Country factor is able to capture these changes in industry correlation in a timelier fashion. The underlying mechanism for this effect is that net-long industry portfolios have common exposure to the Country factor, and when the volatility of the Country factor rises during times of market stress, it explains the increased correlations for the industries.

2.4. Specific Risk Model with Bayesian Shrinkage

The USE4 specific risk model builds upon methodological advances introduced with the European Equity Model (EUE3), as described by Briner, Smith, and Ward (2009). The EUE3 model utilizes daily observations to provide timely estimates of specific risk directly from the time series of specific returns. A significant benefit of this approach is that specific risk is estimated individually for every stock, thus reflecting the idiosyncratic nature of this risk source.

A potential shortcoming of the pure time-series approach is that specific volatilities may not fully persist out-of-sample. In fact, as shown in the *USE4 Methodology Notes*, there is a tendency for time-series volatility forecasts to overpredict the specific risk of high-volatility stocks, and underpredict the risk of low-volatility stocks.

To reduce these biases, we apply a Bayesian shrinkage technique. We segment stocks into deciles based on their market capitalization. Within each size bucket, we compute the mean and standard deviation of the specific risk forecasts. We then pull or “shrink” the volatility forecast to the mean within the size decile, with the shrinkage intensity increasing with the number of standard deviations away from the mean.

3. Factor Structure Overview

3.1. Estimation Universe

The coverage universe is the set of all securities for which the model provides risk forecasts. The estimation universe, by contrast, is the subset of stocks that is used to actually estimate the model. Judicious selection of the estimation universe is an important part of building a sound risk model. The estimation universe must be broad enough to accurately represent the investment opportunity set of investors, without being so broad as to include illiquid stocks that may introduce spurious return relationships into the model. Furthermore, the estimation universe must be sufficiently stable to ensure that factor exposures are well behaved across time. *Representation, liquidity, and stability*, therefore, are the three primary issues that must be addressed when selecting a risk model estimation universe.

A well-constructed equity index must address these very same issues, and therefore serves as an excellent basis for the estimation universe. The USE4 estimation universe utilizes the MSCI USA *Investable Markets Index* (USA IMI), which aims to reflect the full breadth of investment opportunities within the US market by targeting 99 percent of the float-adjusted market capitalization. The MSCI index construction methodology applies innovative rules designed to achieve index stability, while reflecting the evolving equity markets in a timely fashion. Moreover, liquidity screening rules are applied to ensure that only investable stocks that meet the index methodological requirements are included for index membership.

3.2. Industry Factors

Industries are important variables for explaining the sources of equity return co-movement. One of the strengths of the USE4 Model is that it uses the Global Industry Classification Standard (GICS®) for the industry factor structure. The GICS scheme is hierarchical, with 10 sectors at the top level, 24 industry groups at the next level, followed with increasing granularity at the industry and sub-industry levels. GICS applies a consistent global methodology to classify stocks based on careful evaluation of the firm's business model and economic operating environment. GICS codes are reviewed annually to ensure that the classifications are timely and accurate.

It is important that the industry factor structure for each country reflect the unique characteristics of the local market. For instance, some countries may require fine industry detail in some sectors, while a coarser structure may be appropriate for other sectors. When building Barra risk models, special care is taken in customizing the industry factor structure to the local market. Within each sector, we analyze which combinations of industries and sub-industries best reflect the market structure, while also considering the economic intuition and explanatory power of such groupings.

The result of this investigative process is the set of 60 USE4 industry factors, presented in **Table 3.1**. Industries that qualify as factors tend to exhibit high volatility and have significant weight. We find that this relatively parsimonious set of factors captures most of the in-sample *R*-squared explained by the full set of sub-industries (the finest level of granularity in the GICS hierarchy), but with a much higher degree of statistical significance. Also reported in Table 3.1 are the average weights (from 30-Jun-1995 to 31-May-2011) and end-of-period industry weights.

Table 3.1

USE4 Industry Factors. Weights were determined within the USE4 estimation universe using total market capitalization. Averages were computed over the sample period (30-Jun-1995 to 31-May-2011).

GICS Sector	USE4 Code	USE4 Industry Factor Name	Average Weight	31-May-2011 Weight
Energy	1	Oil and Gas Drilling	0.50	0.40
	2	Oil and Gas Equipment and Services	1.11	1.92
	3	Oil Gas and Consumable Fuels	3.12	4.03
	4	Oil and Gas Exploration and Production	2.14	4.15
Materials	5	Chemicals	1.43	1.71
	6	Specialty Chemicals	0.73	0.90
	7	Construction Materials	0.09	0.06
	8	Containers and Packaging	0.31	0.33
	9	Aluminum Steel	0.39	0.42
	10	Precious Metals Gold Mining	0.46	0.81
	11	Paper and Forest Products	0.33	0.20
Industrials	12	Aerospace and Defense	1.49	1.89
	13	Building Products	0.24	0.20
	14	Construction and Engineering	0.20	0.33
	15	Electrical Equipment	1.02	0.86
	16	Industrial Conglomerates	1.13	0.69
	17	Construction and Farm Machinery	0.48	0.92
	18	Industrial Machinery	1.18	1.38
	19	Trading Companies and Distributors	0.57	0.58
	20	Commercial and Professional Services	2.44	1.74
	21	Transportation Air Freight and Marine	0.79	0.72
	22	Airlines	0.24	0.21
	23	Road and Rail	0.87	1.12
Consumer Discretionary	24	Automobiles and Components	0.72	0.94
	25	Household Durables (non-Homebuilding)	0.57	0.49
	26	Homebuilding	0.20	0.15
	27	Leisure Products Textiles Apparel and Luxury	0.80	0.91
	28	Hotels Leisure and Consumer Services	1.05	1.26
	29	Restaurants	0.82	1.15
	30	Media	4.07	3.28
	31	Distributors Multiline Retail	1.28	0.86
	32	Internet and Catalog Retail	0.85	1.49
	33	Apparel and Textiles	0.93	1.03
	34	Specialty Retail	1.56	1.93
	35	Specialty Stores	0.48	0.33

Table 3.1 (cont.)

GICS Sector	USE4 Code	USE4 Industry Factor Name	Average Weight	31-May-2011 Weight
Consumer Staples	36	Food and Staples Retailing	1.66	1.81
	37	Beverages Tobacco	2.81	3.36
	38	Food Products	1.95	1.90
	39	Household and Personal Products	2.09	1.95
Health Care	40	Health Care Equipment and Technology	2.07	2.47
	41	Health Care Providers (non-HMO)	1.34	1.20
	42	Managed Health Care	0.79	0.93
	43	Biotechnology Life Sciences	1.85	2.15
	44	Pharmaceuticals	6.41	4.68
Financials	45	Banks	6.12	3.00
	46	Diversified Financials	7.02	6.75
	47	Insurance Brokers and Reinsurance	3.20	2.41
	48	Life Health and Multi-line Insurance	1.30	1.03
	49	Real Estate	2.00	3.00
Information Technology	50	Internet Software and IT Services	2.64	4.17
	51	Software	4.42	4.74
	52	Communications Equipment	3.11	2.17
	53	Computers Electronics	3.82	4.34
	54	Semiconductor Equipment	0.45	0.39
	55	Semiconductors	2.75	2.39
Telecom	56	Diversified Telecommunication Services	3.40	1.77
	57	Wireless Telecommunication Services	1.00	0.89
Utilities	58	Electric Utilities	2.07	1.66
	59	Gas Utilities	0.46	0.37
	60	Multi-Utilities Water Utilities Power	0.67	1.08

In **Table 3.2**, we report the underlying GICS codes that map to each of the USE4 industry factors. This table helps illustrate the customization that takes place within each sector. Taking the Health Care sector as an example, we see that this sector is divided into five industry factors. The first three factors are derived from GICS Industry Group 3510. Two industries within this group (351010 and 351030) are combined to form a single risk factor: Health Care Equipment and Technology. The other industry (351020) is divided into two risk factors: (a) Managed Health Care (based on sub-industry 35102030), and (b) Health Care Providers (based on three sub-industries). The second Industry Group within Health Care (3520) is divided into two risk factors; this is accomplished by splitting off the Pharmaceuticals industry (352020) from Biotechnology (352010) and Life Sciences (352030). In each case, the industry structure is guided by a combination of financial intuition and empirical analysis.

Table 3.2

Mapping of USE4 industry factors to GICS codes.

Code	USE4 Industry Factor Name	GICS Codes
1	Oil and Gas Drilling	10101010
2	Oil and Gas Equipment and Services	10101020
3	Oil Gas and Consumable Fuels	10102010, 10102030, 10102040, 10102050
4	Oil and Gas Exploration and Production	10102020
5	Chemicals	15101010, 15101020, 15101030, 15101040
6	Specialty Chemicals	15101050
7	Construction Materials	151020
8	Containers and Packaging	151030
9	Aluminum Steel	15104010, 15104050
10	Precious Metals Gold Mining	15104020, 15104030, 15104040
11	Paper and Forest Products	151050
12	Aerospace and Defense	201010
13	Building Products	201020
14	Construction and Engineering	201030
15	Electrical Equipment	201040
16	Industrial Conglomerates	201050
17	Construction and Farm Machinery	20106010
18	Industrial Machinery	20106020
19	Trading Companies and Distributors	201070
20	Commercial and Professional Services	2020
21	Transportation Air Freight and Marine	203010, 203030, 203050
22	Airlines	203020
23	Road and Rail	203040
24	Automobiles and Components	2510
25	Household Durables (non-Homebuilding)	25201010, 25201020, 25201040, 25201050
26	Homebuilding	25201030
27	Leisure Products Textiles Apparel and Luxury	252020, 252030
28	Hotels Leisure and Consumer Services	25301010, 25301020, 25301030, 253020
29	Restaurants	25301040
30	Media	2540
31	Distributors Multiline Retail	255010, 255030
32	Internet and Catalog Retail	255020
33	Apparel and Textiles	25504010
34	Specialty Retail	25504020, 25504030, 25504050, 25504060
35	Specialty Stores	25504040

Table 3.2 (cont.)

Code	USE4 Industry Factor Name	GICS Codes
36	Food and Staples Retailing	3010
37	Beverages Tobacco	302010, 302030
38	Food Products	302020
39	Household and Personal Products	3030
40	Health Care Equipment and Technology	351010, 351030
41	Health Care Providers (non-HMO)	35102010, 35102015, 35102020
42	Managed Health Care	35102030
43	Biotechnology Life Sciences	352010, 352030
44	Pharmaceuticals	352020
45	Banks	4010
46	Diversified Financials	4020
47	Insurance Brokers and Reinsurance	40301010, 40301040, 40301050
48	Life Health and Multi-line Insurance	40301020, 40301030
49	Real Estate	4040
50	Internet Software and IT Services	451010, 45102010, 45102020
51	Software	451030
52	Communications Equipment	452010
53	Computers Electronics	452020, 452030, 452040
54	Semiconductor Equipment	45205010, 45301010
55	Semiconductors	45205020, 45301020
56	Diversified Telecommunication Services	501010
57	Wireless Telecommunication Services	501020
58	Electric Utilities	551010
59	Gas Utilities	551020
60	Multi-Utilities Water Utilities Power	551030, 551040, 551050

In **Table 3.3** we report the largest firm within each industry, as well as the total market capitalization as of 31-May-2011. The largest firm was Exxon Mobil, with a market capitalization exceeding $420 billion. The next three largest stocks, Apple, Microsoft, and IBM, were all within the Information Technology sector.

Table 3.3
Largest stock within each industry as of 31-May-2011. Market capitalizations are reported in billions of US dollars.

Code	USE4 Industry Factor Name	Largest Stock (May 31, 2011)
1	Oil and Gas Drilling	NOBLE CORPORATION BAAR ($10.6)
2	Oil and Gas Equipment and Services	SCHLUMBERGER LTD ($117.0)
3	Oil Gas and Consumable Fuels	EXXON MOBIL CORP ($420.9)
4	Oil and Gas Exploration and Production	APACHE CORP ($47.6)
5	Chemicals	DU PONT E I DE NEMOURS & CO ($48.7)
6	Specialty Chemicals	LYONDELLBASELL INDUSTRIES N V ($24.8)
7	Construction Materials	VULCAN MATLS CO ($5.2)
8	Containers and Packaging	BALL CORP ($7.0)
9	Aluminum Steel	ALCOA INC ($17.2)
10	Precious Metals Gold Mining	FREEPORT-MCM GLD ($48.6)
11	Paper and Forest Products	INTL PAPER CO ($13.7)
12	Aerospace and Defense	UNITED TECHNOLOGIES CORP ($81.0)
13	Building Products	MASCO CORP ($5.1)
14	Construction and Engineering	FLUOR CORP NEW ($12.3)
15	Electrical Equipment	EMERSON ELEC CO ($41.1)
16	Industrial Conglomerates	GENERAL ELECTRIC CO ($209.3)
17	Construction and Farm Machinery	CATERPILLAR INC DEL ($67.2)
18	Industrial Machinery	DANAHER CORP DEL ($35.7)
19	Trading Companies and Distributors	GRAINGER W W INC ($10.4)
20	Commercial and Professional Services	WASTE MGMT INC DEL ($18.5)
21	Transportation Air Freight and Marine	UNITED PARCEL SERVICE INC ($53.6)
22	Airlines	SOUTHWEST AIRLS CO ($9.4)
23	Road and Rail	UNION PAC CORP ($51.8)
24	Automobiles and Components	FORD MTR CO DEL ($50.8)
25	Household Durables (non-Homebuilding)	STANLEY BLACK & DECKER INC ($12.3)
26	Homebuilding	NVR INC ($4.2)
27	Leisure Products Textiles Apparel and Luxury	NIKE INC ($32.8)
28	Hotels Leisure and Consumer Services	LAS VEGAS SANDS CORP ($28.4)
29	Restaurants	MCDONALDS CORP ($86.1)
30	Media	WALT DISNEY CO ($78.8)
31	Distributors Multiline Retail	TARGET CORP ($35.1)
32	Internet and Catalog Retail	AMAZON COM INC ($88.3)
33	Apparel and Textiles	TJX COS INC NEW ($21.0)
34	Specialty Retail	HOME DEPOT INC ($60.4)
35	Specialty Stores	STAPLES INC ($12.2)

Table 3.3 (cont.)

Code	USE4 Industry Factor Name	Largest Stock (May 31, 2011)
36	Food and Staples Retailing	WAL MART STORES INC ($196.7)
37	Beverages Tobacco	COCA COLA CO ($155.1)
38	Food Products	KRAFT FOODS INC ($61.1)
39	Household and Personal Products	PROCTER & GAMBLE CO ($187.5)
40	Health Care Equipment and Technology	MEDTRONIC INC ($44.0)
41	Health Care Providers (non-HMO)	EXPRESS SCRIPTS INC ($31.3)
42	Managed Health Care	UNITEDHEALTH GROUP INC ($53.8)
43	Biotechnology Life Sciences	AMGEN INC ($57.2)
44	Pharmaceuticals	JOHNSON & JOHNSON ($184.8)
45	Banks	WELLS FARGO & CO NEW ($148.9)
46	Diversified Financials	JPMORGAN CHASE & CO ($169.0)
47	Insurance Brokers and Reinsurance	BERKSHIRE HATHAWAY [B] ($82.1)
48	Life Health and Multi-line Insurance	AMERICAN INTL GROUP INC ($54.1)
49	Real Estate	SIMON PPTY GROUP INC NEW ($34.6)
50	Internet Software and IT Services	INTERNATIONAL BUSINESS MACHS ($209.9)
51	Software	MICROSOFT CORP ($214.0)
52	Communications Equipment	QUALCOMM INC ($94.8)
53	Computers Electronics	APPLE INC ($319.1)
54	Semiconductor Equipment	APPLIED MATLS INC ($18.4)
55	Semiconductors	INTEL CORP ($125.6)
56	Diversified Telecommunication Services	AT&T INC ($186.5)
57	Wireless Telecommunication Services	AMERICAN TOWER CORP ($22.1)
58	Electric Utilities	SOUTHERN CO ($33.3)
59	Gas Utilities	ONEOK INC NEW ($7.6)
60	Multi-Utilities Water Utilities Power	DOMINION RES INC VA NEW ($27.7)

3.3. Multiple Industry Exposures

The USE4 Model assigns multiple industry exposures to stocks based on the firm's business segment reporting and an analysis of two explanatory variables: Assets and Sales. As described in the *USE4 Methodology Notes*, we estimate multiple industry exposures by first computing slope coefficients, or "industry betas," that represent the price-to-assets ratios and price-to-sales ratios for the industries. The industry betas are computed by separately regressing the market capitalizations of the firms against their Assets and Sales within the primary industry. From the business segment reporting, we then determine a breakdown of the firm's Assets of Sales across industries. By combining these Assets and Sales with the corresponding price ratios of the industries, we obtain an estimate of the market capitalization of the firm explained by each industry. The fraction of a firm's total market capitalization explained by each industry gives the multiple industry exposure. Note that the maximum number of industry exposures is limited to five; by construction, these exposures add to 1.

In **Table 3.4** we report average industry betas for Assets and Sales. For Sales, the lowest industry betas are found in Airlines, Automobiles, and Food & Staples Retailing, indicating that these industries have high revenues in proportion to their market capitalization. In contrast, the highest industry betas are found within Biotechnology Life Sciences, Software, and Pharmaceuticals. For Assets, the lowest industry betas are in the Financials sector, consistent with the high leverage used by these firms.

Table 3.4

Industry betas for Assets and Sales. Results were averaged over the sample history (30-Jun-1995 to 31-May-2011). The final column reports the weight of stocks with multiple industry exposures as of 31-May-2011.

USE4 Code	USE4 Industry Factor Name	Assets Beta	Sales Beta	Multi-Ind Weight
1	Oil and Gas Drilling	1.25	2.74	55.52
2	Oil and Gas Equipment and Services	1.38	1.43	80.25
3	Oil Gas and Consumable Fuels	0.88	0.60	90.61
4	Oil and Gas Exploration and Production	0.83	2.02	54.43
5	Chemicals	0.87	0.92	98.67
6	Specialty Chemicals	0.95	0.91	66.71
7	Construction Materials	0.93	1.25	100.00
8	Containers and Packaging	0.42	0.57	74.37
9	Aluminum Steel	0.57	0.47	77.20
10	Precious Metals Gold Mining	0.82	1.31	88.26
11	Paper and Forest Products	0.51	0.62	83.38
12	Aerospace and Defense	0.83	0.81	96.14
13	Building Products	0.71	0.51	94.49
14	Construction and Engineering	0.72	0.39	76.44
15	Electrical Equipment	1.24	1.16	88.69
16	Industrial Conglomerates	0.67	1.71	99.62
17	Construction and Farm Machinery	0.62	0.57	78.35
18	Industrial Machinery	0.99	0.91	88.82
19	Trading Companies and Distributors	0.96	0.53	76.91
20	Commercial and Professional Services	0.92	0.77	67.48
21	Transportation Air Freight and Marine	0.86	0.69	75.70
22	Airlines	0.22	0.23	10.98
23	Road and Rail	0.55	0.77	44.55
24	Automobiles and Components	0.26	0.22	89.85
25	Household Durables (non-Homebuilding)	0.80	0.65	54.48
26	Homebuilding	0.47	0.41	64.60
27	Leisure Products Textiles Apparel and Luxury	0.92	0.77	48.51
28	Hotels Leisure and Consumer Services	0.69	1.20	62.66
29	Restaurants	1.42	1.02	22.91
30	Media	0.76	1.87	53.16
31	Distributors Multiline Retail	0.68	0.48	45.07
32	Internet and Catalog Retail	1.26	1.04	37.74
33	Apparel and Textiles	1.45	0.71	73.71
34	Specialty Retail	1.53	0.72	55.19
35	Specialty Stores	0.68	0.43	37.18

Table 3.4 (cont.)

USE4 Code	USE4 Industry Factor Name	Assets Beta	Sales Beta	Multi-Ind Weight
36	Food and Staples Retailing	0.99	0.35	58.58
37	Beverages Tobacco	1.08	1.28	27.01
38	Food Products	0.84	0.59	75.13
39	Household and Personal Products	1.94	1.92	87.76
40	Health Care Equipment and Technology	1.69	2.15	54.99
41	Health Care Providers (non-HMO)	0.86	0.62	69.87
42	Managed Health Care	0.58	0.57	48.87
43	Biotechnology Life Sciences	2.48	5.53	22.63
44	Pharmaceuticals	2.73	3.71	43.71
45	Banks	0.15	2.06	29.94
46	Diversified Financials	0.14	1.60	53.15
47	Insurance Brokers and Reinsurance	0.36	1.35	49.44
48	Life Health and Multi-line Insurance	0.12	1.35	80.85
49	Real Estate	0.57	3.25	41.13
50	Internet Software and IT Services	1.17	1.20	30.94
51	Software	2.69	3.74	13.86
52	Communications Equipment	1.55	1.64	50.30
53	Computers Electronics	1.10	0.91	27.52
54	Semiconductor Equipment	1.86	2.25	51.06
55	Semiconductors	2.06	2.91	5.36
56	Diversified Telecommunication Services	0.80	1.52	84.09
57	Wireless Telecommunication Services	0.79	1.90	91.83
58	Electric Utilities	0.41	1.01	97.56
59	Gas Utilities	0.50	0.72	95.70
60	Multi-Utilities Water Utilities Power	0.47	0.87	75.71
Average		0.96	1.27	62.99

Also reported in Table 3.4 is the weight of stocks with multiple industry exposure as of 31-May-2011. Airlines and Semiconductors have the lowest weights, suggesting that these industries are the least diversified in terms of their business activities. The average multiple-industry weight is nearly 63 percent, showing that multiple-industry exposures are well represented in USE4.

3.4. Style Factors

Investment style represents another major source of systematic risk for equity portfolios. Style factors are constructed from financially intuitive stock attributes called *descriptors*, which serve as effective predictors of equity return covariance.

In order to facilitate comparison across style factors, the factors are standardized to have a cap-weighted mean of 0 and an equal-weighted standard deviation of 1. The cap-weighted estimation universe, therefore, is *style neutral*.

The USE4 Model contains 12 style factors. The factors are described in Appendix A, together with descriptor definitions and descriptor weights. Here we provide a brief qualitative description of the factors:

- The *Beta* factor is typically the most important style factor. It captures market risk that cannot be explained by the Country factor. We compute Beta by time-series regression of excess stock returns against the cap-weighted estimation universe, as described in Appendix A. To better understand how Beta relates to the Country factor, consider a fully invested long-only portfolio that is tilted toward high-beta stocks. Intuitively, this portfolio has greater market risk than a portfolio with a beta of 1. This additional market risk is captured through positive exposure to the Beta factor. Since the time-series correlation between the Country factor and the Beta factor is typically very high, these two sources of risk are additive in this example. If, by contrast, the portfolio were invested primarily in low-beta stocks, then the risk from the Beta and the Country factors would have been partially offset, as expected.
- The *Momentum* factor is often the second strongest factor in the model, although sometimes it may surpass Beta in importance. Momentum differentiates stocks based on their performance over the trailing 6-12 months. When computing Momentum exposures we exclude the last month (21 days) of returns in order to avoid the effects of short-term reversal.
- The *Siz*e factor represents another strong source of equity return covariance, and captures return differences between large-cap stocks and small-cap stocks. We measure Size by the log of market capitalization.
- The *Earnings Yield* factor describes return differences based on a company's earnings relative to its price. Earnings Yield is considered by many investors to be a strong value signal. The most important descriptor in this factor is the analyst-predicted 12-month forward earnings-to-price ratio.
- The *Residual Volatility* factor is composed of three descriptors: (a) the volatility of daily excess returns, (b) the volatility of daily residual returns, and (c) the cumulative range of the stock over the last 12 months. Since these descriptors tend to be highly collinear with the Beta factor, the Residual Volatility factor is orthogonalized with respect to the Beta factor, as described by Menchero (2010).
- The *Growth* factor differentiates stocks based on their prospects for sales or earnings growth. The most important descriptor in this factor is the analyst predicted long-term earnings growth. Other descriptors include sales and earnings growth over the trailing five years.
- The *Dividend Yield* factor explains return differences attributable to dividend payouts of the firm. This factor is defined by the trailing 12-month dividend divided by the current price.
- The *Book-to-Price* factor is also considered by some to be an indicator of value. This factor is given by the last reported book value of common equity divided by current market capitalization.
- The *Leverage* factor captures return differences between high-leverage and low-leverage stocks. The descriptors within this style factor include market leverage, book leverage, and debt-to-assets ratio.
- The *Liquidity* factor describes return differences due to relative trading activity. The descriptors for this factor are based on the fraction of total shares outstanding that trade over a recent window.
- The *Non-Linear Size* (NLS) factor captures non-linearities in the payoff to the Size factor across the market-cap spectrum. This factor is based on a single raw descriptor: the cube of the Size exposure. However, since this raw descriptor is highly collinear with the Size factor, it is orthogonalized with respect to Size. This procedure does not affect the fit of the model, but does mitigate the confounding effects of collinearity, while preserving an intuitive meaning for the Size factor. As

described by Menchero (2010), the NLS factor roughly captures the risk of a "barbell portfolio" that is long mid-cap stocks and short small-cap and large-cap stocks.

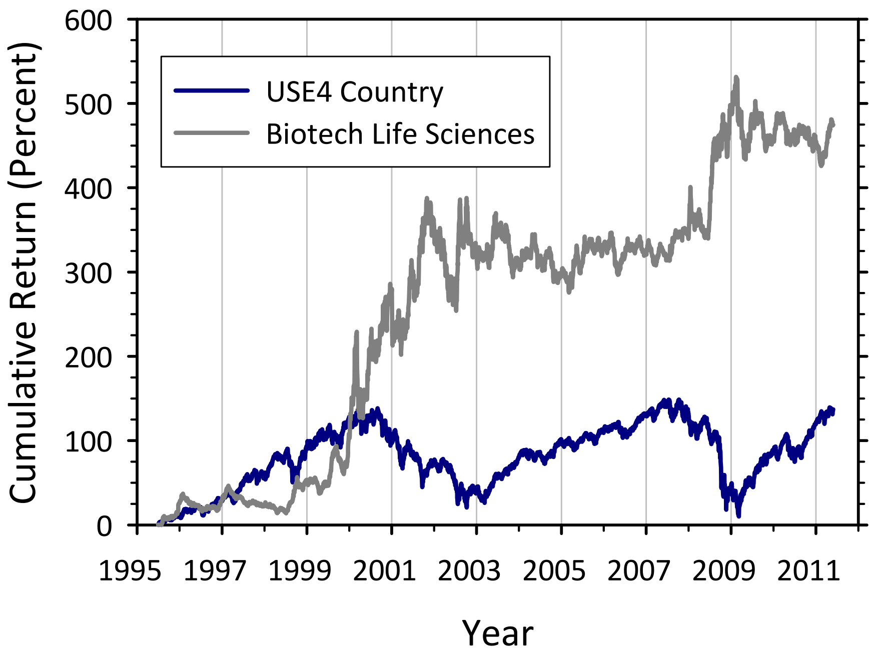
- The *Non-Linear Beta* (NLB) factor captures non-linearities in the payoff to the Beta factor. Similar to the NLS factor, we first cube the Beta factor and then orthogonalize it with respect to the Beta. Roughly speaking, the NLB factor captures the risk of holding a "barbell portfolio" that is long stocks with average betas (i.e., close to 1) and short stocks with betas that deviate strongly from the mean.

3.5. Performance of Select Factors

It is helpful to consider the performance of individual factors. In **Figure 3.1**, we report cumulative returns to the USE4 Country factor. As described in the *USE4 Methodology Notes*, the Country factor return essentially represents the excess return (i.e., above the risk-free rate) of the cap-weighted country portfolio. Figure 3.1 clearly illustrates the main features of the US equity market over the last 16 years. For instance, the three bull markets of the sample period are clearly visible, as is the bear market after the Internet Bubble, and the market crash of 2008.

Figure 3.1

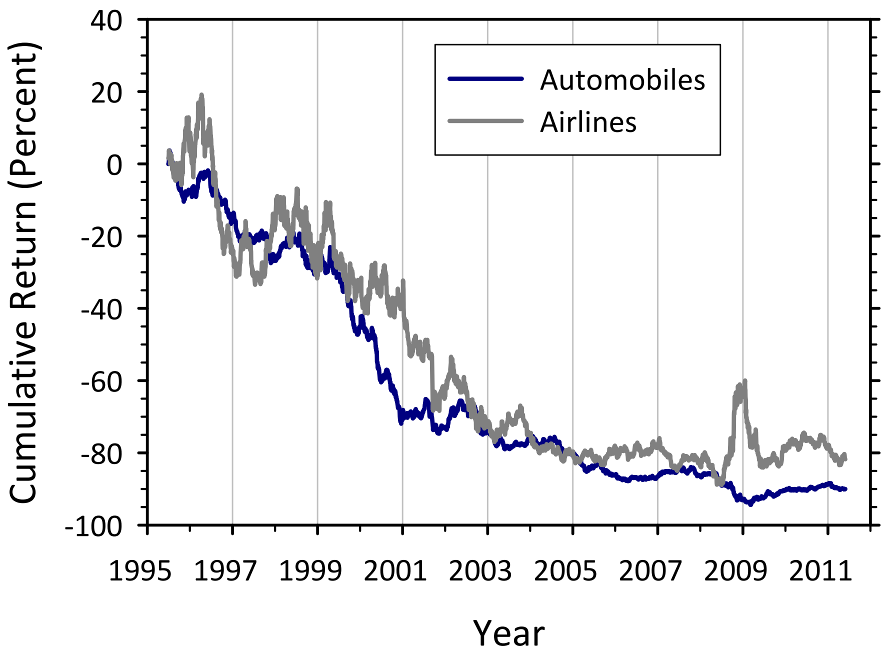
Cumulative returns of USE4 Country factor and Biotechnology Life Sciences factor.



As discussed in the *USE4 Methodology Notes*, industry factor returns represent the performance of the pure industry relative to the overall market, net of all style effects. In other words, the pure industry factor portfolio is dollar neutral and has zero exposure to every style. In Figure 3.1 we report the cumulative return of the Biotechnology Life Sciences factor. We see that this factor performed extremely well over the sample period, particularly during 2000-2002.

In **Figure 3.2** we report the cumulative returns for the Automobiles and Airlines factors, which suffered from poor performance. Over the roughly 16-year sample period, the Airlines factor lost about 80 percent. The Automobile factor fared even worse, declining by roughly 90 percent over the entire sample period.

Figure 3.2
Cumulative returns of Automobiles factor and Airlines factor.

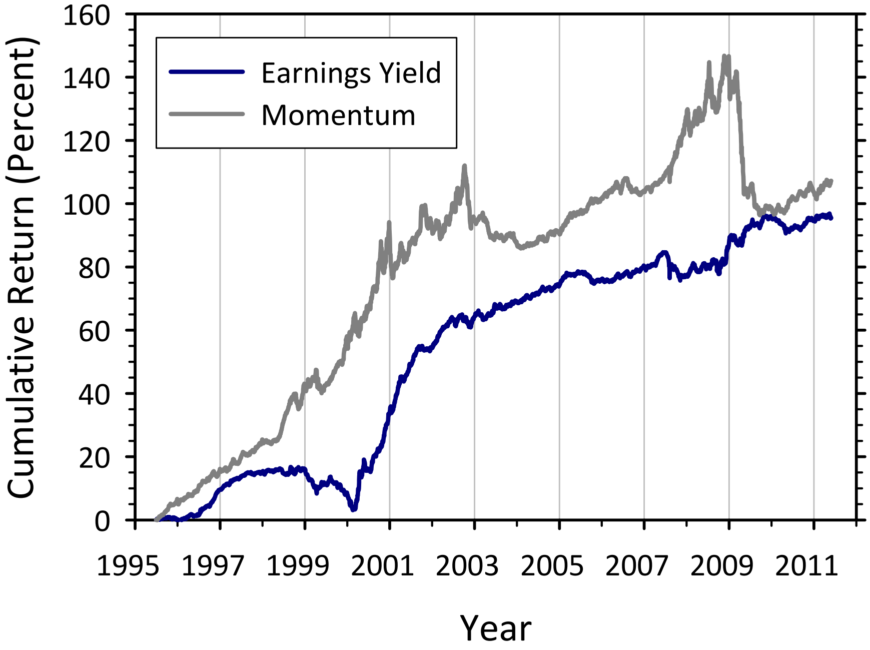


Style factor returns represent the returns of pure factor portfolios that have exposure only to the style in question. In other words, they have net zero weight in every industry, and have zero exposure to every other style factor. A more detailed discussion of pure factor portfolios is provided by Menchero (2010).

In **Figure 3.3**, we report the cumulative returns to the Earnings Yield and Momentum factors, which represent two common strategies often used by quantitative investors. Overall, Earnings Yield performed very well over the past 16 years, consistent with the notion of a "value premium." As described by Basu (1977), this reflects the tendency of stocks that are priced low relative to fundamentals to outperform. A notable exception, however, occurred during the Internet Bubble in 1999, when Earnings Yield performed poorly. The Quant Meltdown is also visible as a "downward blip" in August 2007.

Figure 3.3

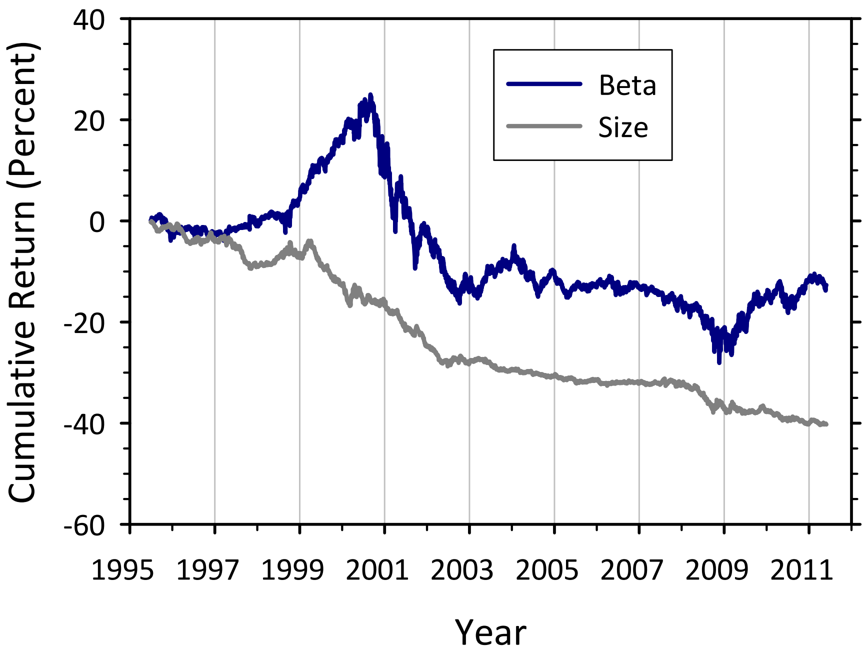
Cumulative returns of Earnings Yield and Momentum factors.



Momentum also performed well over the last 16 years, consistent with the empirical observation noted by Jegadeesh and Titman (1993) that stocks with strong performance over the previous 6-12 months continue to outperform. There were, however, two major periods of underperformance. The first occurred in late 2002, which coincided with the market rebound following the 2000-2002 bear market. The second major downturn for the Momentum factor began in March 2009, which again coincided with a market recovery — this time after the crash of 2008.

In **Figure 3.4** we report cumulative returns to the Beta and Size factors. Qualitatively, the Beta factor showed the same main features as the Country factor in Figure 3.1, consistent with the high correlation between these factors. A key difference, however, is that the Beta factor exhibited a significant negative drift with respect to the Country factor. For instance, the Beta factor did not participate in the market rally of 2003-2007. As a result, the cumulative return to the Beta factor was actually negative over the last 16 years, although the Country factor itself was up by over 100 percent. This result is inconsistent with the Capital Asset Pricing Model (CAPM) described by Sharpe (1964), which states that high-beta stocks are expected outperform low-beta stocks. Nevertheless, our results are consistent with those of Fama and French (1992), who find that there does not appear to be a return premium associated with Beta.

Figure 3.4
Cumulative returns of Beta and Size factors.



The Size factor also exhibited a consistent downward drift over the sample period. In fact, with the exception of 1998, the Size factor had negative returns every year of the 16-year sample. This negative performance is consistent with the notion of a "small-cap premium," as described by Banz (1981).

4. Model Characteristics and Properties

4.1. Country and Industry Factors

One requirement of a high-quality factor structure is that the factor returns be statistically significant. This helps prevent weak or noisy factors from finding their way into the model. We measure statistical significance by the t-statistic of the factor return. Assuming normality, absolute t-statistics greater than 2 are considered significant at the 95-percent confidence level. In other words, if the factor truly had no explanatory power (e.g., pure noise), then by chance we would observe $|t|>2$ about 5 percent of the time.

In **Table 4.1** we report mean absolute t-statistics for the USE4 Country factor and industry factors, as well as the percentage of observations with $|t|>2$. Note that the t-statistics reported in Table 4.1 were computed using *monthly* cross-sectional regressions, even though we run daily cross-sectional regressions for purposes of constructing the factor covariance matrix. This distinction is important, because what is ultimately relevant is the explanatory power of the factors at the prediction horizon of the model.

From Table 4.1, we see that the Country factor was by far the strongest factor during the sample period. On average, it had an absolute t-statistic of 15.33, and was significant nearly 96 percent of the months. Table 4.1 also shows that the strongest industry factors tended to be clustered in the Energy and Information Technology sectors. Across all factors, the t-statistics were significant in 40.2 percent of the observations. This reflects, on the whole, the high degree of statistical significance for the factors.

Table 4.1

Industry factor summary statistics. The first two columns pertain to *t*-statistics, and were computed using monthly cross-sectional regressions. The last four columns were computed based on daily factor returns. The sample period is from 30-Jun-1995 to 31-May-2011 (191 months of returns).

Factor Name	Average Absolute *t*-stat	Percent Observ. |*t*|>2	Annual. Factor Return	Annual. Factor Volatility	Factor Sharpe Ratio	Correl. with ESTU
Country Factor	15.33	95.8	5.57	20.25	0.28	0.999
Oil and Gas Drilling	3.14	60.4	5.12	32.96	0.16	-0.04
Oil and Gas Equipment and Services	3.44	60.9	4.08	27.68	0.15	-0.03
Oil Gas and Consumable Fuels	2.56	51.6	0.88	17.76	0.05	-0.05
Oil and Gas Exploration and Production	4.06	70.3	2.38	22.17	0.11	-0.04
Chemicals	1.71	32.8	-0.88	15.64	-0.06	0.00
Specialty Chemicals	1.16	18.2	-5.69	10.85	-0.52	0.08
Construction Materials	0.94	10.4	-6.16	20.09	-0.31	0.02
Containers and Packaging	1.18	18.8	-5.26	14.93	-0.35	0.00
Aluminum Steel	2.24	44.3	-7.78	21.20	-0.37	0.07
Precious Metals Gold Mining	2.20	44.8	6.40	26.53	0.24	-0.10
Paper and Forest Products	1.58	30.7	-8.08	20.78	-0.39	0.04
Aerospace and Defense	1.97	40.6	-1.47	14.33	-0.10	-0.09
Building Products	1.16	15.6	-6.31	14.57	-0.43	0.04
Construction and Engineering	1.27	19.3	-2.01	15.78	-0.13	0.04
Electrical Equipment	1.23	20.3	-2.21	10.51	-0.21	0.14
Industrial Conglomerates	1.13	16.1	-7.50	32.93	-0.23	0.01
Construction and Farm Machinery	1.60	27.6	-0.16	16.58	-0.01	0.08
Industrial Machinery	1.42	26.0	-2.95	9.67	-0.30	0.12
Trading Companies and Distributors	1.11	12.5	-6.36	11.93	-0.53	0.04
Commercial and Professional Services	1.52	27.6	-2.91	6.27	-0.46	-0.05
Transportation Air Freight and Marine	1.25	21.4	-1.03	13.85	-0.07	0.03
Airlines	2.44	51.6	-10.14	32.84	-0.31	0.00
Road and Rail	1.63	32.8	-0.02	14.46	0.00	0.03
Automobiles and Components	2.07	41.7	-13.48	16.49	-0.82	0.04
Household Durables (non-Homebuilding)	1.20	16.1	-5.72	11.62	-0.49	0.03
Homebuilding	2.23	47.4	-2.35	25.87	-0.09	0.06
Leisure Products Textiles Apparel and Luxury	1.65	31.3	-5.06	11.21	-0.45	-0.01
Hotels Leisure and Consumer Services	2.06	46.4	-3.45	12.57	-0.27	-0.04
Restaurants	1.87	39.1	-0.68	13.21	-0.05	-0.07
Media	2.19	46.4	-0.57	8.88	-0.06	-0.05
Distributors Multiline Retail	2.17	48.4	-1.23	16.93	-0.07	-0.02
Internet and Catalog Retail	1.93	38.0	1.51	22.03	0.07	0.01
Apparel and Textiles	2.52	52.6	3.03	18.64	0.16	-0.01
Specialty Retail	2.11	48.4	3.43	17.39	0.20	0.02
Specialty Stores	1.73	30.2	-6.16	13.71	-0.45	0.00

Table 4.1 (cont.)

Factor Name	Average Absolute *t*-stat	Percent Observ. |*t*|>2	Annual. Factor Return	Annual. Factor Volatility	Factor Sharpe Ratio	Correl. with ESTU
Food and Staples Retailing	1.70	33.9	-1.97	11.55	-0.17	-0.16
Beverages Tobacco	1.61	30.7	3.22	12.21	0.26	-0.23
Food Products	1.58	31.3	-0.52	9.40	-0.06	-0.23
Household and Personal Products	1.46	26.6	1.78	12.33	0.14	-0.14
Health Care Equipment and Technology	1.81	37.5	3.17	9.28	0.34	-0.17
Health Care Providers (non-HMO)	2.19	43.8	-0.21	12.06	-0.02	-0.15
Managed Health Care	2.40	50.0	2.41	22.29	0.11	-0.10
Biotechnology Life Sciences	3.08	56.3	11.61	17.20	0.67	-0.04
Pharmaceuticals	2.54	54.7	5.67	12.27	0.46	-0.14
Banks	3.51	62.5	-6.00	13.47	-0.45	0.08
Diversified Financials	2.64	54.7	0.02	12.66	0.00	0.19
Insurance Brokers and Reinsurance	1.95	41.1	-4.21	10.29	-0.41	-0.04
Life Health and Multi-line Insurance	1.54	29.2	-10.12	18.76	-0.54	0.14
Real Estate	2.23	44.8	-4.56	13.64	-0.33	0.05
Internet Software and IT Services	2.18	44.3	-0.74	12.03	-0.06	0.08
Software	2.50	51.6	5.24	11.62	0.45	0.09
Communications Equipment	2.83	56.8	5.34	14.37	0.37	0.10
Computers Electronics	2.90	56.8	2.15	11.56	0.19	0.08
Semiconductor Equipment	3.12	60.4	-0.30	28.25	-0.01	0.10
Semiconductors	4.14	71.9	5.94	21.51	0.28	0.09
Diversified Telecommunication Services	2.30	46.9	-4.74	14.89	-0.32	-0.06
Wireless Telecommunication Services	1.89	35.4	3.79	19.41	0.20	-0.04
Electric Utilities	2.34	51.0	-5.53	14.06	-0.39	-0.15
Gas Utilities	1.15	17.2	-5.48	11.65	-0.47	-0.07
Multi-Utilities Water Utilities Power	1.61	28.1	-9.87	16.29	-0.61	-0.12
Average	2.27	40.2		16.33		

Also reported in Table 4.1 are the returns, volatilities, and Sharpe ratios for the factors, during the sample period. These quantities were computed using *daily* factor returns and stated on an annualized basis. The Country factor had an annualized return of 5.57 percent and a volatility of 20.25 percent, leading to a Sharpe ratio of 0.28 over the roughly 16-year sample period. The best-performing industries tended to be concentrated in the Energy, Health Care, and Information Technology sectors, whereas Industrials, Financials, and Utilities generally underperformed. Also note that factors within the Energy sector tended to be quite volatile, while those in Consumer Staples tended toward low volatility.

Table 4.1 also reports the correlations of the daily factor returns with the estimation universe. Particularly noteworthy is the 99.9 percent correlation between the Country factor and the estimation universe, indicating the essential equivalence of the two. By contrast, most industry factors, being dollar-neutral portfolios, had relatively small correlations with the estimation universe. Industry factors within the Consumer Staples, Health Care, and Utilities tended to have small negative correlations, whereas those within the Information Technology sector were slightly positive. It is important to stress that these correlations represent averages over the entire sample period. Within different sub-periods or market regimes, the correlations may deviate significantly from these reported values.

4.2. Style Factors

In **Table 4.2**, we report summary statistics for the USE4 style factors, during the sample period. The sample is broken up into two roughly equal sub-periods. Note that the statistical significance of the style factors, on the whole, was slightly greater than that for the industry factors. As measured by volatility and t-statistics, the strongest factors were generally Beta, Momentum, and Size — although Earnings Yield, Residual Volatility, and Non-Linear Size also exhibited considerable strength. In the first sample period (30-Jun-1995 to 30-Jun-2003), Momentum and Earnings Yield performed extremely well, while Residual Volatility and Size performed poorly. The relative performance of these factors also persisted for the second sample period (30-Jun-2003 to 31-May-2011), although the returns and Sharpe ratios were smaller in magnitude. Also noteworthy is the high statistical significance of the Growth factor during the first sample period, which spanned the Internet Bubble.

Most style factors, being dollar-neutral portfolios, had relatively small correlation with the estimation universe. The glaring exception is the Beta factor, which had a correlation of 0.77 in the first sample period and 0.90 during the second. The Residual Volatility factor also had a sizeable positive correlation with the estimation universe during both sub-periods. Again, it is important to stress that the correlations reported in Table 4.2 represent averages, and that the actual correlations in different market regimes may deviate from these reported values.

Also reported in Table 4.2 is the factor stability coefficient, described in the *USE4 Methodology Notes*. Briefly, this coefficient is computed as the cross-sectional correlation of factor exposures from one month to the next. Although there is no strict lower limit for what is considered acceptable, a useful rule of thumb is that values below 0.80 are regarded as too unstable for model inclusion, while those above 0.90 have desirable stability characteristics. From Table 4.2, we see that the average factor stability coefficient for style factors was 0.96 during both sample periods.

Table 4.2 also reports the Variance Inflation Factor (VIF). As explained in the *USE4 Methodology Notes*, VIF measures the degree of collinearity among the factors. Excessive collinearity can lead to increased estimation error in the factor returns and non-intuitive correlations among factors. Although there exists no strict upper bound, VIF scores above 5 are generally considered problematic. As shown in Table 4.2, all USE4 style factors were below this level during both sample periods.

Table 4.2

Style factor summary statistics. The first two columns pertain to *t*-statistics, and were computed using monthly cross-sectional regressions. The next four columns were computed based on daily factor returns. The factor stability coefficient and Variance Inflation Factor were computed on monthly data using square root of market-cap weighting. The entire sample period, comprising 191 months (30-Jun-1995 to 31-May-2011), is divided into two sub-periods.

A. 30-Jun-1995 to 30-Jun-2003 (96 months)

Factor Name	Average Absolute *t*-stat	Percent Observ. |*t*|>2	Annual. Factor Return	Annual. Factor Volatility	Factor Sharpe Ratio	Correl. with ESTU	Factor Stability Coeff.	Variance Inflation Factor
Beta	3.96	64.9	-1.34	8.95	-0.15	0.77	0.96	3.74
Momentum	4.33	71.1	8.34	4.46	1.87	0.04	0.91	2.14
Size	3.60	70.1	-4.06	3.42	-1.19	0.27	1.00	4.38
Earnings Yield	2.86	55.7	6.68	3.05	2.19	-0.17	0.97	2.21
Residual Volatility	3.06	57.7	-6.08	4.16	-1.46	0.44	0.96	3.50
Growth	2.17	50.5	1.34	2.51	0.53	0.23	0.98	2.19
Dividend Yield	1.37	23.7	1.53	2.08	0.74	0.12	0.99	2.93
Book-to-price	1.33	21.6	0.58	1.99	0.29	-0.01	0.97	2.29
Leverage	1.46	27.8	-0.53	1.66	-0.32	-0.01	0.99	1.72
Liquidity	1.91	36.1	2.91	2.89	1.00	0.33	0.97	2.48
Non-linear Size	2.23	52.6	-0.12	3.40	-0.03	0.21	0.98	1.42
Non-linear Beta	1.90	32.0	-0.02	2.50	-0.01	0.06	0.85	1.33
Average	2.52	46.99	0.77	3.42	0.29	0.19	0.96	2.53

B. 30-Jun-2003 to 31-May-2011 (95 months)

Factor Name	Average Absolute *t*-stat	Percent Observ. |*t*|>2	Annual. Factor Return	Annual. Factor Volatility	Factor Sharpe Ratio	Correl. with ESTU	Factor Stability Coeff.	Variance Inflation Factor
Beta	4.02	66.3	-0.35	8.36	-0.04	0.90	0.97	3.46
Momentum	3.50	64.2	1.12	3.69	0.30	-0.18	0.89	1.88
Size	2.85	53.7	-2.29	3.09	-0.74	0.06	1.00	3.56
Earnings Yield	2.11	44.2	1.95	2.17	0.90	0.04	0.97	1.90
Residual Volatility	2.35	45.3	-2.69	3.42	-0.79	0.50	0.95	2.88
Growth	1.53	21.1	0.53	1.36	0.39	0.00	0.95	1.77
Dividend Yield	1.59	30.5	-0.09	1.77	-0.05	0.10	0.99	2.64
Book-to-price	1.70	33.7	-0.06	1.88	-0.03	0.19	0.98	1.90
Leverage	1.92	34.7	-0.87	2.07	-0.42	0.10	0.99	1.90
Liquidity	1.47	28.4	0.36	2.09	0.17	-0.11	0.98	2.30
Non-linear Size	2.00	47.4	1.37	2.70	0.51	0.02	0.99	1.52
Non-linear Beta	1.59	35.8	-0.33	1.53	-0.22	0.05	0.91	1.47
Average	2.22	42.11	-0.11	2.84	0.00	0.14	0.96	2.26

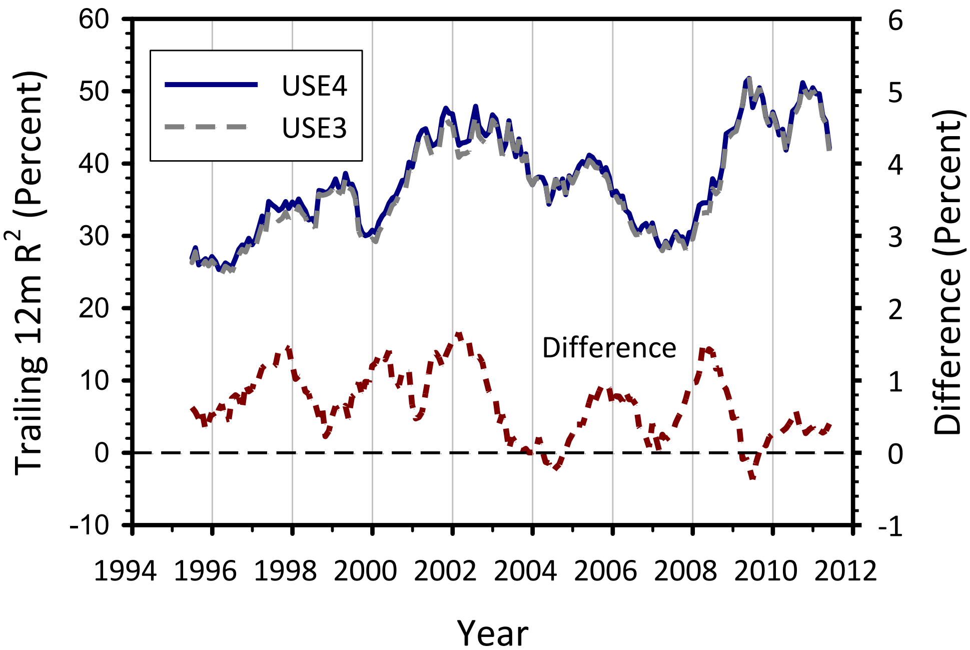
4.3. Explanatory Power

The explanatory power of the factors, as measured by *R*-squared, is a key metric of model quality. The value of *R*-squared, however, can depend sensitively on the regression weighting scheme, the estimation universe, and the time period under consideration. Caution must be exercised, therefore, when comparing *R*-squared values across different models. Nevertheless, if each of these variables is carefully controlled, then a meaningful apples-to-apples comparison between models is possible.

In **Figure 4.1**, we report the trailing 12-month *R*-squared for the USE3 and USE4 models. In order to ensure a fair comparison, the estimation universe (MSCI USA IMI Index) and regression weighting scheme (square root of market capitalization) were identical for the two sets of regressions. Clearly, both models track each other closely in terms of explanatory power. The *R*-squared of both models varied from just under 30 percent in the mid 1990s to slightly above 50 percent in 2009.

Figure 4.1

Trailing 12-month total *R*-squared for USE4 and USE3 models. Results were computed based on monthly cross-sectional regressions using a common estimation universe (MSCI USA IMI) and regression weighting scheme (square root of market capitalization). The difference is also plotted, with the scale indicated on the right axis.



To facilitate a detailed comparison, we also plot in Figure 4.1 the *difference* in explanatory power between the two models. Over the entire sample period, the USE4 Model outperformed the USE3 Model by an average of 66 bps in *R*-squared. Moreover, the increased explanatory power was persistent across time, with only two brief periods (in 2004 and 2009) when the USE3 Model slightly outperformed the USE4 Model.

In **Table 4.3**, we investigate the sources of the increased explanatory power of USE4 factors. The average *R*-squared for USE3 was 37.25 percent over the sample period. We also performed monthly regressions by combining USE4 industry factors with USE3 styles; this led to an increase in *R*-squared of

56 bps over the USE3 model. The full USE4 model had an *R*-squared of 37.91 percent over the sample period (i.e., 66 bps above USE3). Therefore, while both industries and styles contributed to the increased explanatory power, the majority can be attributed to the USE4 industries.

Table 4.3

Explanatory power of USE4 factors versus USE3. Results are averages of monthly *R*-squared values taken over the sample period 30-Jun-1995 to 31-May-2011 (191 months). The first row corresponds to the USE3 factor structure, which led to an *R*-squared of 37.25 percent. Substituting USE4 industries for USE3 industries increased the explanatory power by 56 bps. The bottom row shows the effect of replacing USE3 styles with USE4 styles, which boosted the *R*-squared by an additional 10 bps. Over the entire sample period, the full USE4 model had 66 bps greater explanatory power than USE3.

Industries	Styles	R^2(percent)	ΔR^2(bps)
USE3	USE3	37.25	0
USE4	USE3	37.81	56
USE4	USE4	37.91	66

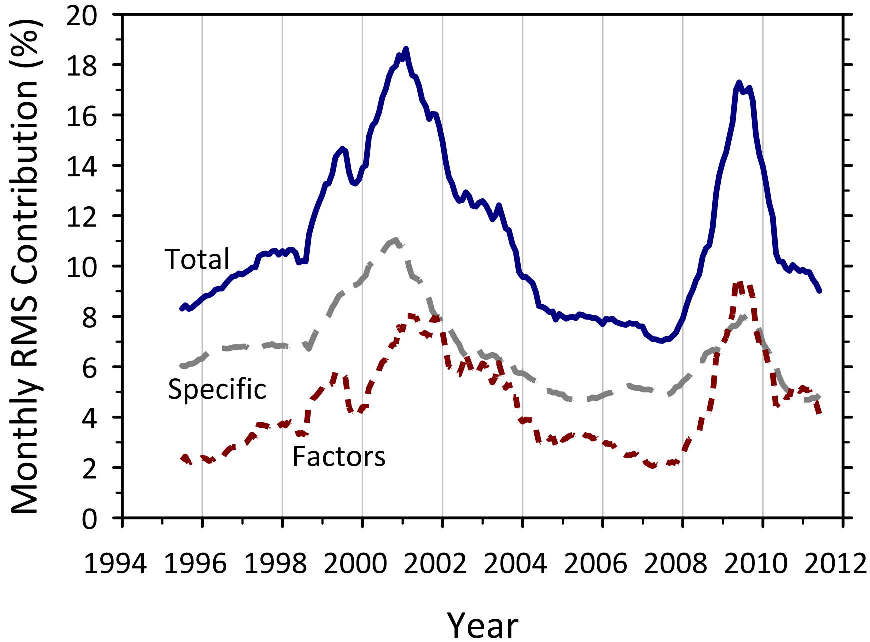
4.4. Cross-Sectional Dispersion

It is informative to study the cross-sectional dispersion of monthly stock returns. As discussed by Menchero and Morozov (2011), dispersion can be measured in one of two ways. The first is by cross-sectional volatility (CSV), which measures the dispersion relative to the *mean* return. The second way is by root mean square (RMS) return, which measures the dispersion relative to *zero* return. The main difference between the two is that the Country factor makes no contribution to CSV, whereas it does contribute to RMS levels.

In **Figure 4.2**, we plot the trailing 12-month total RMS return. The two most prominent features corresponded to the Internet Bubble (peak monthly RMS of 18 percent) and the financial crisis (peak RMS of 17 percent). Note that the Internet Bubble peak was much broader — the buildup and aftermath spanned several years — whereas the financial crisis peak was relatively short in duration.

Figure 4.2

Total monthly cross-sectional dispersion as measured by root mean square (RMS) return. Also displayed are the stock-specific and factor contributions. Lines were smoothed using 12-month moving averages.

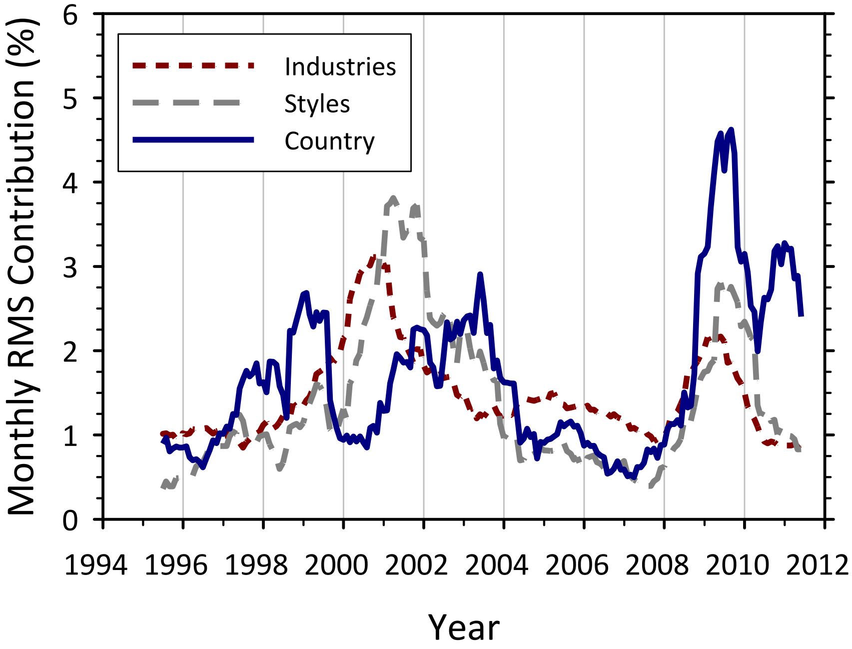


As discussed by Menchero and Morozov (2011), and shown in Appendix B, the RMS return can be decomposed and attributed to individual factors or groups of factors. Figure 4.2 shows the net RMS contributions from factors and stock-specific sources. During most of the sample period, the stock-specific contribution dominated. An important exception occurred during the financial crisis, when factors became the main driver of equity returns.

In **Figure 4.3**, we further decompose the factor RMS contributions into contributions from the Country factor, industries, and styles. We see that all three sources were of comparable importance in explaining the cross section of RMS returns. However, the relative importance of these factors varied over time. For instance, the Country factor was the largest contributor to RMS return from 1997-1999, and again during 2009-2011. Industries, by contrast, dominated in 2000, and from 2004-2008, whereas style factors dominated during the period immediately following the Internet Bubble.

Figure 4.3

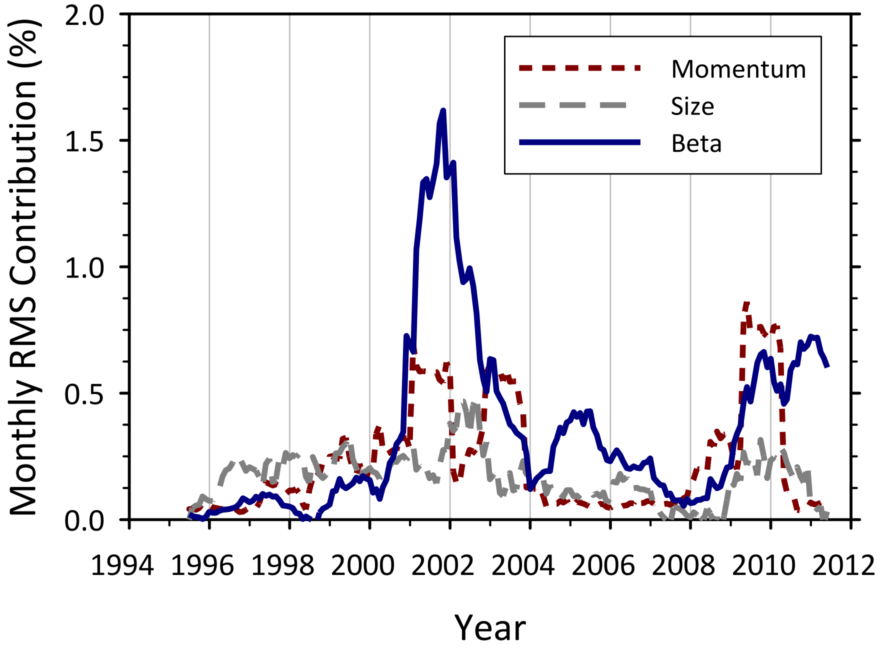
Contributions to monthly root mean square (RMS) return from Country factor, industries, and styles. Lines were smoothed using 12-month moving averages. All three sources are important contributors to cross-sectional dispersion, and each dominates over different sub-periods within the history.



In **Figure 4.4** we report RMS contributions from the Beta, Momentum, and Size factors. Particularly noteworthy is the large peak attributed to Beta from 2001-2003. The Size factor was the largest contributor from 1996-1998, although it contributed less than 25 bps per month. The Momentum factor dominated in 2009, following the rebound from the market crash of 2008.

Figure 4.4

Contributions to monthly root mean square (RMS) return from Beta, Momentum, and Size factors. Lines were smoothed using 12-month moving averages.

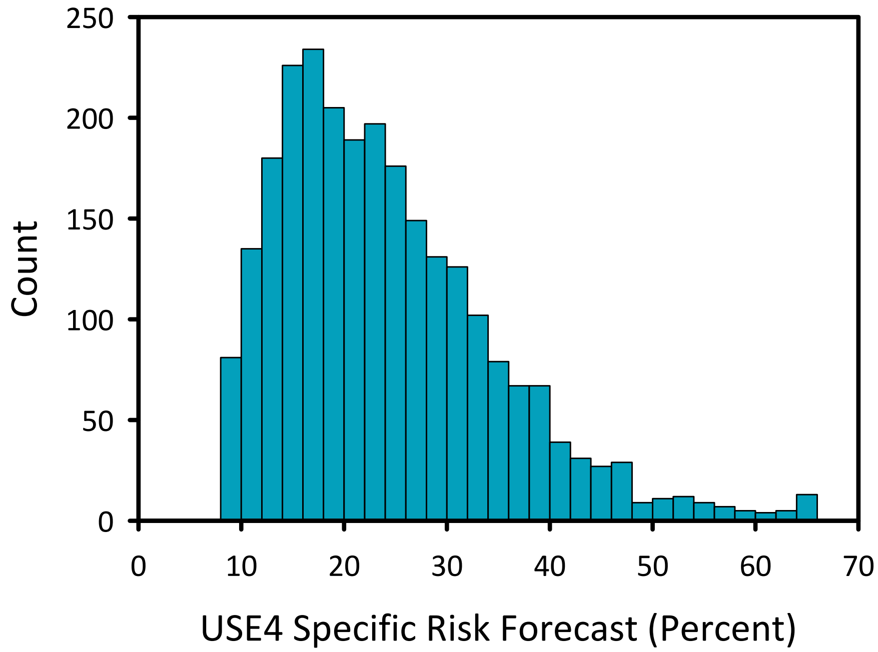


4.5. Specific Risk

The distribution of specific volatilities is an important characteristic to examine. In **Figure 4.5** we plot the histogram of USE4S specific risk forecasts for analysis date 31-May-2011. Most stocks had specific risk forecasts within the range 10-45 percent, although the most volatile stocks had forecasts exceeding 60 percent. The mean specific risk forecast on 31-May-2011 was about 25 percent.

Figure 4.5

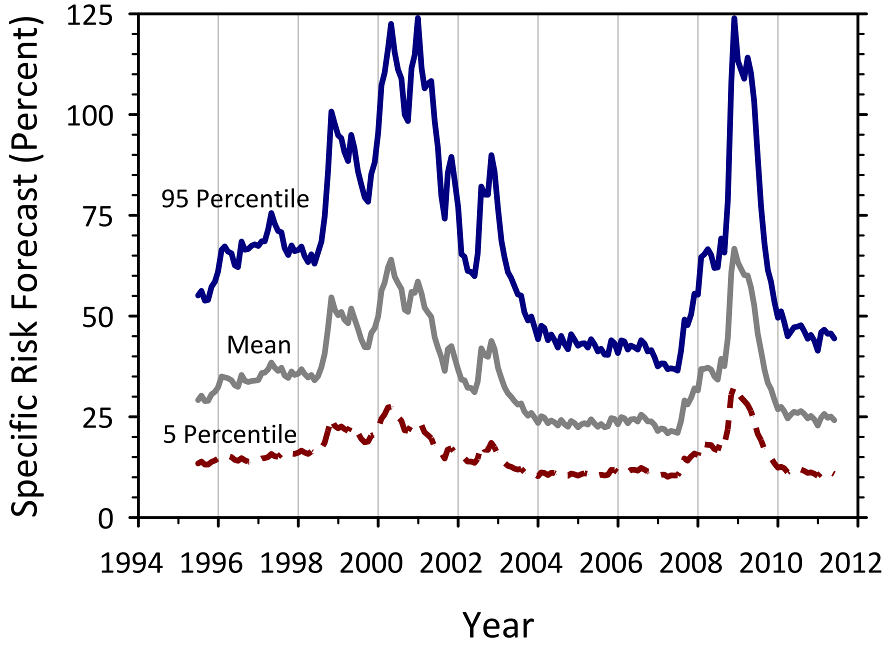
Histogram of USE4S specific-risk forecasts as of 31-May-2011.



It is also interesting to study how the distribution of specific risk varied over time. In **Figure 4.6**, we plot the 5-percentile, mean, and 95-percentile values for the specific risk distribution. We see that the 5-percentile specific volatility historically ranged from about 10-25 percent, with the maximum occurring in late 2008. The mean specific risk varied within the range of 25-65 percent. Again, the peak happened in late 2008, although the Internet Bubble period saw comparable levels of specific risk. The 95-percentile specific volatility ranged from a low of 40 percent in 2007, to highs in excess of 120 percent during the Internet Bubble and financial crisis.

Figure 4.6

Specific risk levels versus time for USE4S.



5. Forecasting Accuracy

5.1. Overview of Testing Methodology

In this section, we describe our methodology for evaluating and comparing the accuracy of risk model forecasts. We aim for a systematic and quantitative approach, yet one that is also visually intuitive.

The foundation of our approach rests on the bias statistic, described in Appendix C. Conceptually, the bias statistic is an out-of-sample measure that represents the ratio of realized risk to predicted risk. The ideal bias statistic for perfect risk forecasts should be close to 1. However, even for perfect risk forecasts, the bias statistic will never be *exactly* 1 due to sampling error. Nevertheless, we may define a confidence interval that is expected to contain 95 percent of the observations under the hypothesis of perfect risk forecasts. If the bias statistic falls outside of the confidence interval, we infer that the risk forecast was not accurate.

When determining the size of the confidence interval, standard practice is to assume that returns are normally distributed. In reality, however, stock returns tend to have fat tails (i.e., positive excess kurtosis). As shown in Appendix C, fewer than 95 percent of the observations are expected to fall within the standard confidence interval when kurtosis is taken into account.

We are interested in testing the full sample period, comprising 191 months from July 1995 through May 2011. One potential shortcoming of the bias statistic is that over long windows, we may have sub-periods of overforecasting and underforecasting, yet obtain a bias statistic close to 1 over the entire window. In other words, forecasting errors may cancel out over the long term, even though the risk forecasts may be poor over shorter periods. For a portfolio manager who may be devastated by a single year of poor performance, it is small consolation knowing that a risk forecast is good *on average*.

For this reason, we focus on 12-month rolling windows. By plotting the mean rolling 12-month bias statistic across time for a collection of portfolios, we quickly visualize the magnitude of the average biases and can judge whether they were persistent or regime-dependent.

It is not enough, however, knowing the average bias statistic. We must also understand the extremes. We also compute, therefore, the 5-percentile (P5) and 95-percentile (P95) bias statistics across time. Assuming normally distributed returns and perfect risk forecasts, on average 5 percent of the rolling 12-month bias statistics will fall below 0.66 by pure chance. Therefore, if the P5 bias statistic falls significantly below this level, we infer that we are likely overpredicting the risk of at least some of the portfolios with bias statistics below 0.66. Similarly, if the P95 bias statistic lies well above 1.34, we infer that we are underpredicting the risk of some portfolios with bias statistics above 1.34. It is worth pointing out, however, that if we relax the normality assumption and allow for fat-tailed distributions, then for perfect risk forecasts the P5 bias statistic tends to fall below 0.66, and the P95 value generally lies above 1.34.

Another measure that provides insight into the accuracy of risk forecasts is the *mean rolling absolute deviation*, or MRAD. As described in Appendix C, this is computed by averaging the absolute deviation of the bias statistics from 1 for a collection of portfolios. Conceptually, MRAD penalizes any deviation from the ideal bias statistic of 1, whether due to overforecasting or underforecasting.

Assuming normally distributed returns and perfect risk forecasts, the expected value of MRAD is 0.17. Real financial returns, of course, tend to have fat tails. For example, the monthly standardized pure factor returns for the USE4S Model had a kurtosis of approximately 4.0 over the 16-year sample period.

For the broadly diversified active portfolios described in Figure 5.4 below, the mean kurtosis level was roughly 3.5. As shown in Appendix C, kurtosis levels within the range of 3.5 to 4.0 lead to MRAD values of approximately 0.19 for perfect risk forecasts. When comparing MRAD values across two models, it is crucial to keep in mind the lower bound of MRAD. For instance, assuming a 0.19 lower bound, reducing MRAD from 0.23 to 0.21 constitutes a 50 percent reduction in *excess* MRAD.

It is also important to recognize that MRAD is a *statistical* measure. As such, by pure chance the MRAD may dip below the level of 0.17. Indeed, consider a portfolio that has been overforecast for many months, leading to a bias statistic less than 1. Eventually, the risk model may begin underforecasting the risk of that same portfolio. When the transition from overforecasting to underforecasting occurs, the bias statistic must necessarily cross through 1, thereby producing an MRAD value close to zero. For a *large collection* of portfolios, however, it is highly improbable that the bias statistics of all portfolios will cross through 1 simultaneously. Consequently, for a sufficiently diverse set of portfolios, the MRAD is unlikely to dip significantly below 0.17 for any sustained period of time.

Our testing approach therefore relies principally on these four measures: the mean bias statistic, the P5 and P95 bias statistics, and the MRAD. All are computed and plotted on a rolling 12-month basis. These plots allow us to quickly evaluate the accuracy of risk forecasts in a visually intuitive manner.

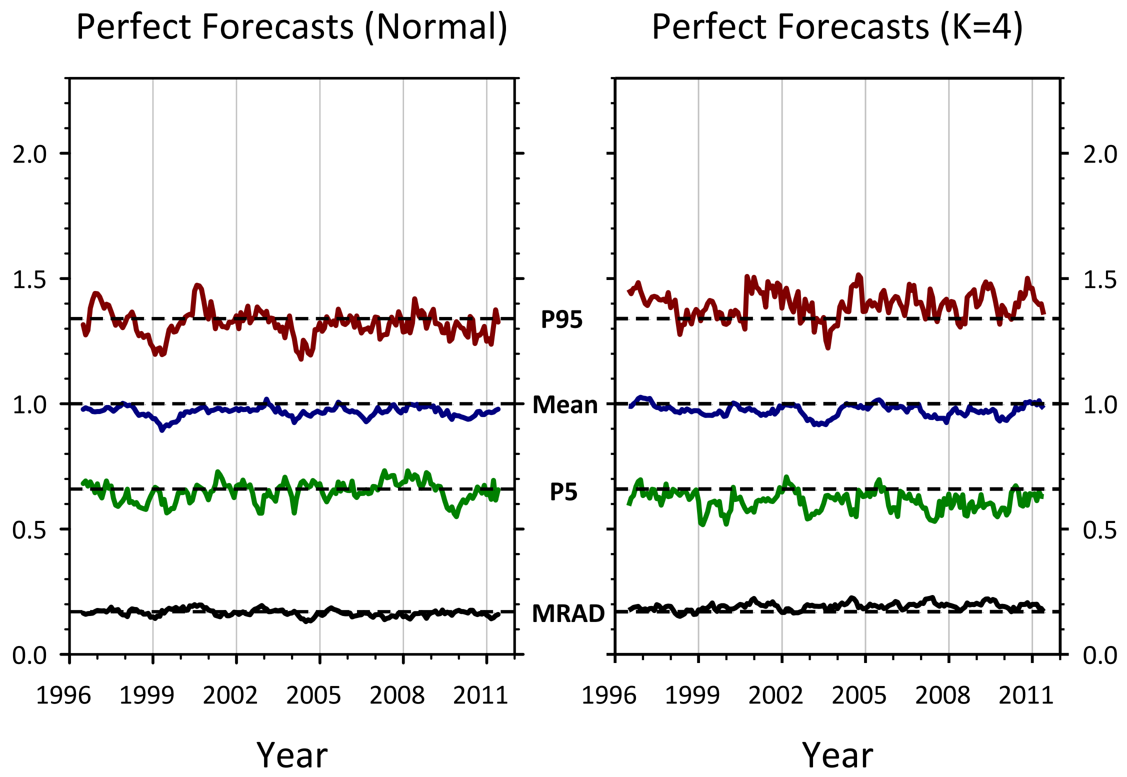
In order to develop a better understanding for how these measures behave in the ideal case of stationary returns and perfect risk forecasts, we perform two separate simulations for 100 sets of returns over 191 months (representing July 1995 through May 2011). In the first simulation, the returns were drawn from a standard normal distribution. In the second simulation, the returns were drawn from a *t*-distribution with standard deviation of 1 and kurtosis of 4. In all simulations, the predicted volatilities were equal to 1 (i.e., perfect risk forecasts).

In **Figure 5.1** we plot MRAD and bias statistics for the mean, P5 and P95 levels. The dashed horizontal lines represent the ideal positions of the curves for the case of perfect forecasts and normal distributions. On the left panel (normal distribution), we see that the realized curves indeed lie close to their ideal positions. In particular, the MRAD is closely centered at the 0.17 level. Note that the degree of "noise" in the lines depends on the number of portfolios in the sample. That is, the more portfolios that we use, the smaller the observed variability.

On the right panel of Figure 5.1 we plot MRAD and bias statistics for perfect risk forecasts and a kurtosis of 4. The effect of higher kurtosis is to increase the frequency of observations with bias statistics above 1.34 or below 0.66. In this case, the mean of the P5 line is shifted down to 0.61, whereas the P95 line moves upward to a mean of 1.40. This has the effect of increasing MRAD to approximately 0.19.

Figure 5.1

Simulated results for the rolling 12-month mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The left plot is for 100 portfolios with simulated returns drawn from a normal distribution with standard deviation of 1. The right plot is for 100 portfolios with simulated returns drawn from a *t*-distribution with a standard deviation of 1 and a kurtosis of 4. The risk forecasts in each case were perfect (i.e., predicted volatility of 1). The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



It is worth reiterating that Figure 5.1 represents the idealized case of perfect risk forecasts and stationary returns. In reality, risk forecasts are never perfect and returns are not stationary. Nevertheless, Figure 5.1 serves as a useful baseline for understanding the empirical backtesting results that follow.

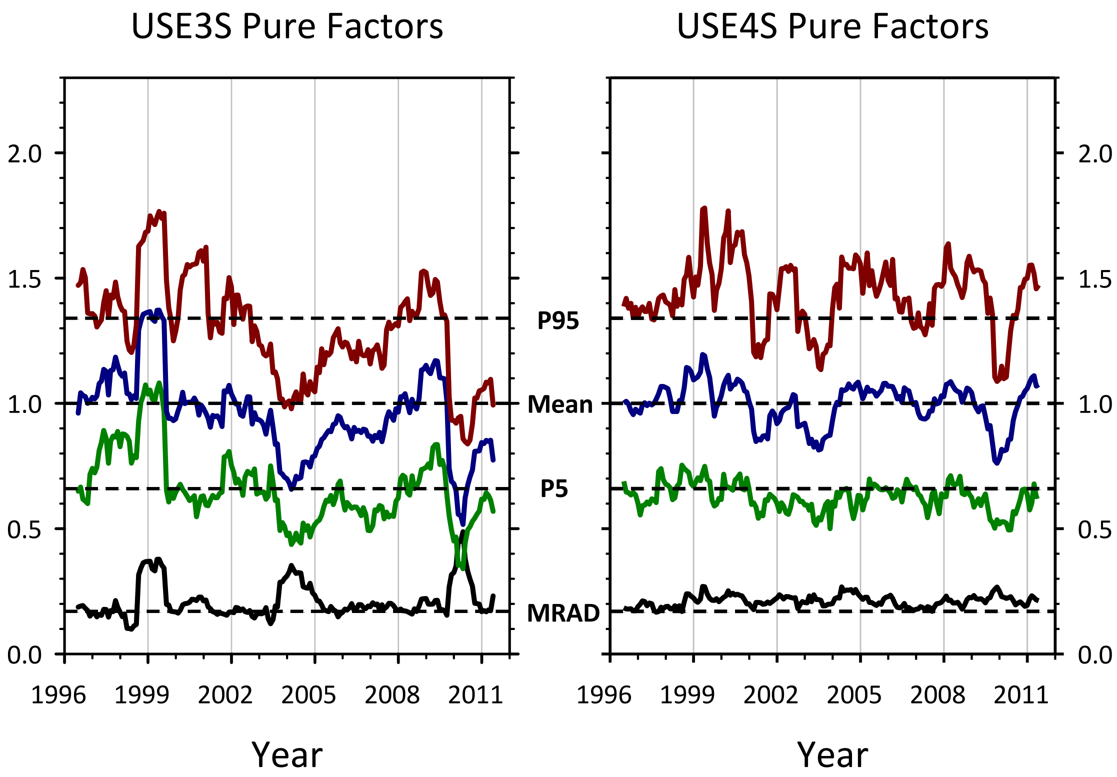
5.2. Backtesting Results

In this section we perform side-by-side comparisons for the USE3 and USE4 Models. We plot MRAD and bias statistics for a variety of test portfolios using both short-horizon and long-horizon models. The analysis period is approximately 16 years, running from July 1995 through May 2011. Rolling 12-month quantities are therefore plotted starting in July 1996.

In **Figure 5.2** we report MRAD and bias statistics for the USE3S and USE4S pure factors. We see that the MRAD and bias statistics were more stable for USE4S and centered more closely to their ideal positions. Both models overpredicted risk in early 2010, but the mean bias statistics for USE4S were significantly closer to 1. Also, note that USE3S tended to overpredict risk in 2004, whereas USE4S had bias statistics close to 1.

Figure 5.2

Comparison of USE3S Model and USE4S Model for pure factors. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.

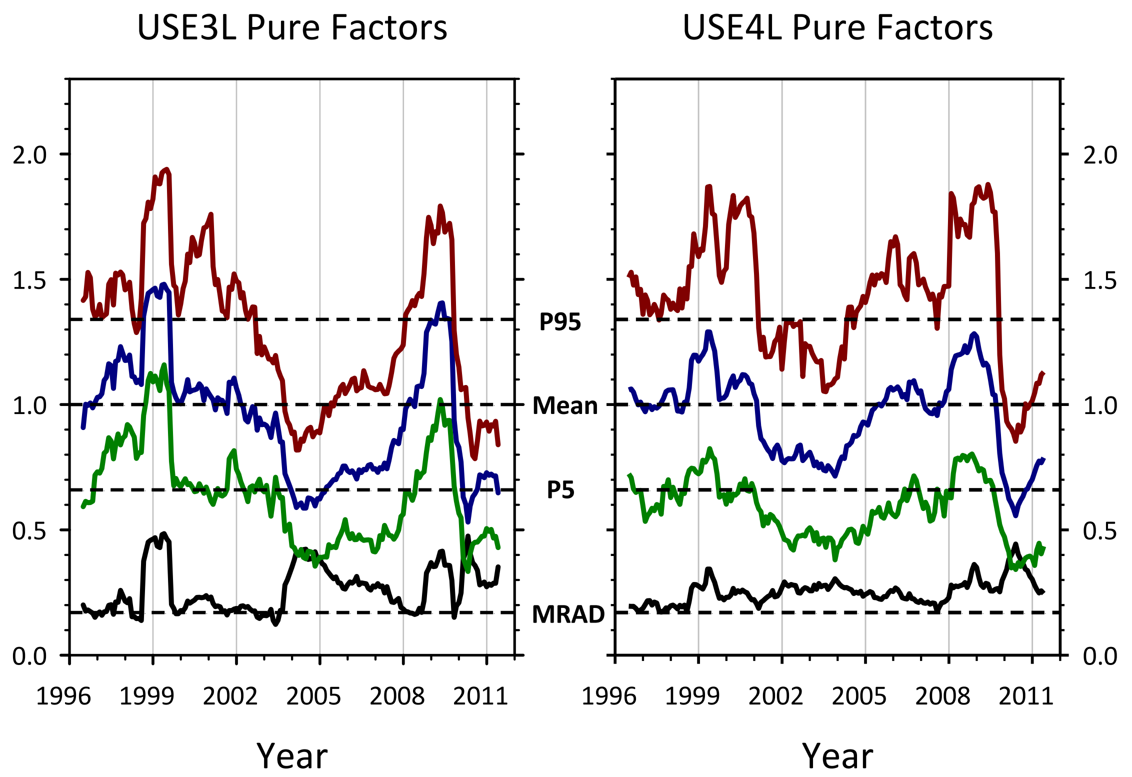


It is important to stress, however, that pure factors do not provide for a strict apples-to-apples comparison, since most of the USE3 factors correspond to *net-long* industry portfolios, whereas the USE4S industry factors represent *dollar-neutral* portfolios. This explains, for example, why the USE3S bias statistics in Figure 5.2 spiked in response to the Russian Default of August 1998, whereas this feature is largely absent from the USE4S results.

In **Figure 5.3** we report MRAD and bias statistics for the USE3L and USE4L pure factors. Again, although this is not a strict apples-to-apples comparison, the exercise is nonetheless informative. We see that USE3L overpredicted risk from 2003-2008. The USE4L Model also overpredicted risk for part of this period, but by a smaller margin. Both models tended to underpredict risk entering the 2008 financial crisis and to overpredict as the crisis subsided.

Figure 5.3

Comparison of USE3L Model and USE4L Model for pure factors. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.

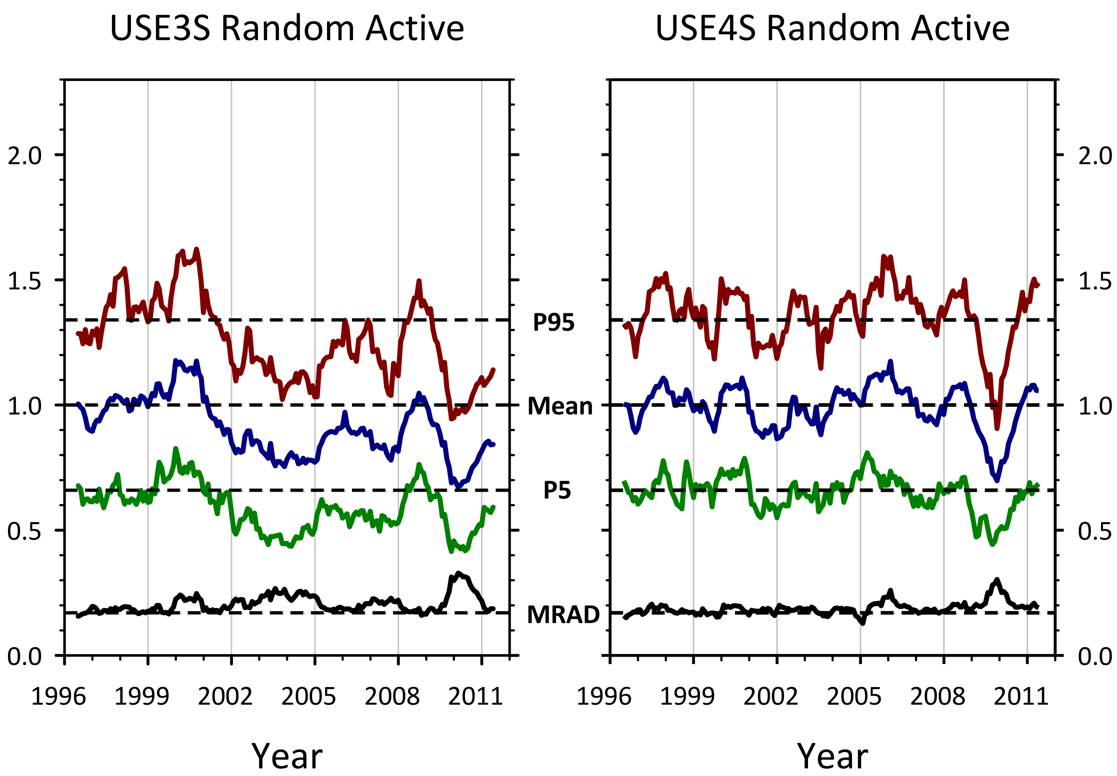


It is also instructive to compare the USE4L results in Figure 5.3 with the corresponding results for USE4S in Figure 5.2. For the more responsive USE4S Model, we see that the MRAD and bias statistics were consistently closer to their ideal values.

In **Figure 5.4** we plot MRAD and bias statistics for 100 random active portfolios using the short-horizon models. The random active portfolios were constructed by going long 500 cap-weighted randomly selected stocks and shorting the cap-weighted USE4 estimation universe. Since identical portfolios were used for the two models, the comparison is apples-to-apples. The USE3S Model tended to overpredict risk during much of the sample period, but especially during 2002-2005 and from late 2009 to mid 2011. For the USE4S model, by contrast, the average bias statistics were close to 1 and the MRAD was near the ideal value for most of the sample period. The notable exception was in late 2009, when USE4S also tended to overpredict risk. Observe, however, that the Volatility Regime Adjustment quickly corrected for the overprediction bias, so that by late 2010 the average bias statistics were again close to 1.

Figure 5.4

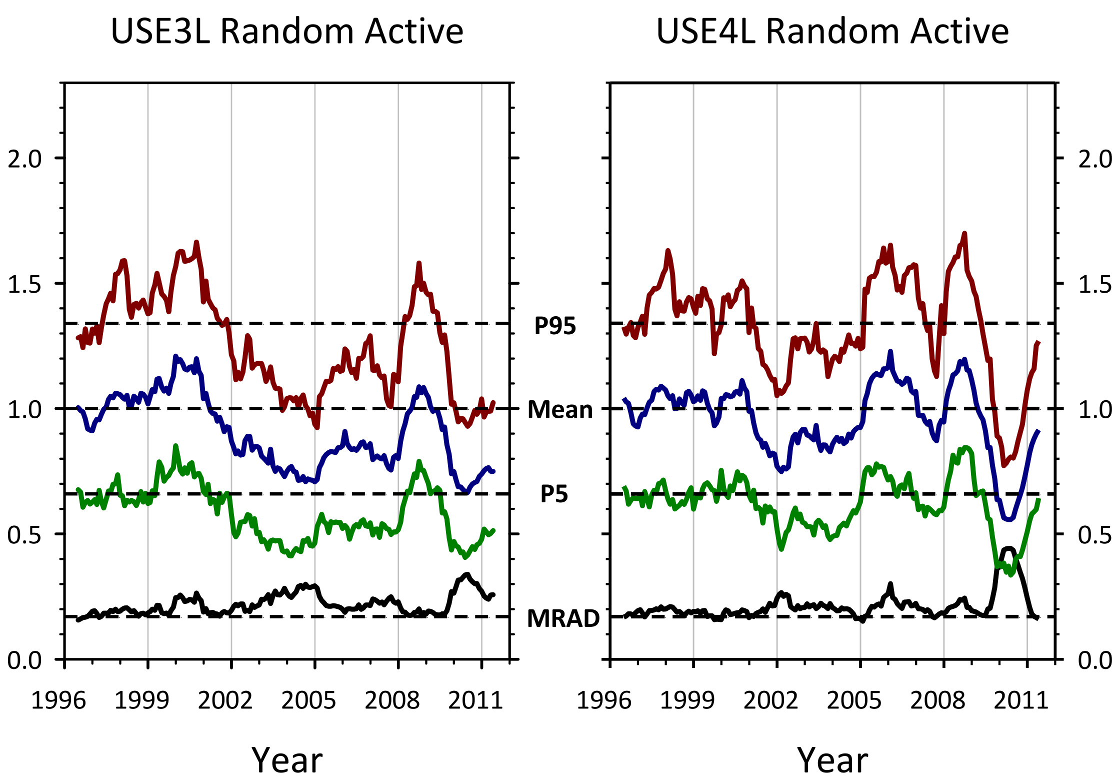
Comparison of USE3S Model and USE4S Model for 100 random active portfolios. The portfolios were generated by randomly selecting 500 stocks, which were then capitalization weighted. These portfolios were then run against the estimation universe to form dollar-neutral active portfolios. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.5** we plot MRAD and bias statistics for the same 100 random active portfolios, except now using the long-horizon models. The USE3L model tended to overpredict risk during much of the sample period. The USE4L model produced bias statistics closer to 1 over most of the sample period, with the notable exception of 2010, when the USE4L significantly overpredicted risk. Even in this case, however, the Volatility Regime Adjustment helped bring the risk forecasts back into line, with the mean bias statistics again close to 1 by May 2011.

Figure 5.5

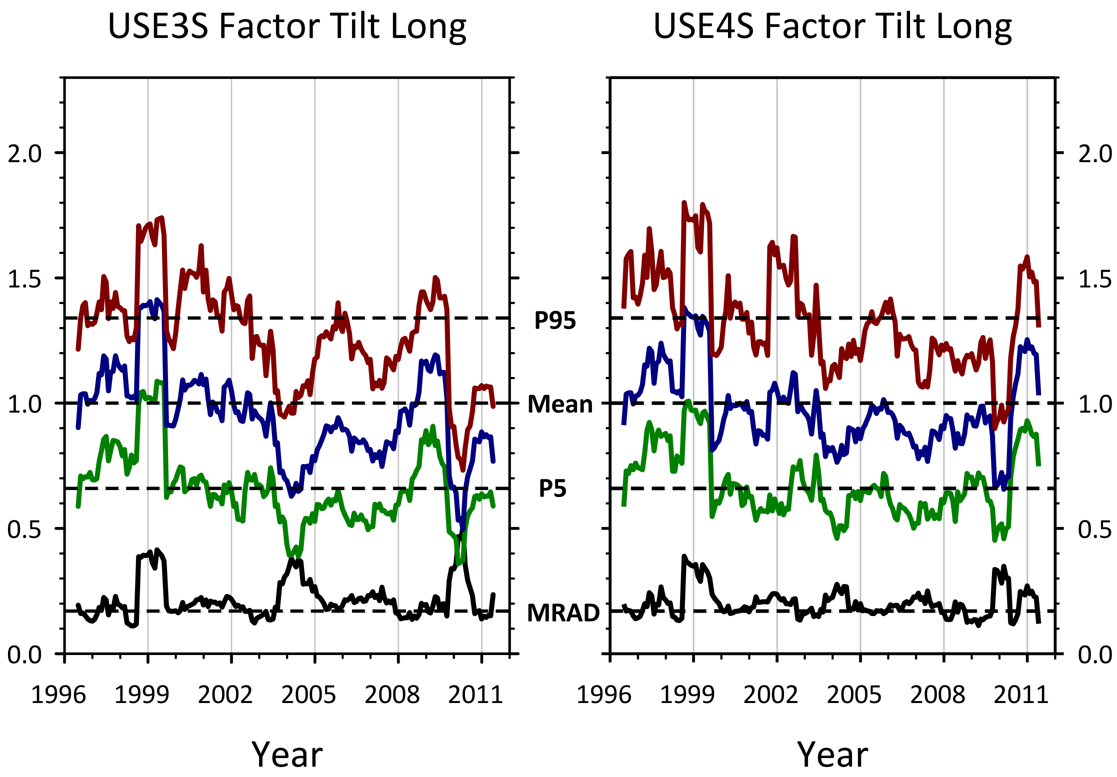
Comparison of USE3L Model and USE4L Model for 100 random active portfolios. The portfolios were generated by randomly selecting 500 stocks, which were then capitalization weighted. These portfolios were then run against the estimation universe to form dollar-neutral active portfolios. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.6** we plot MRAD and bias statistics for 163 long-only factor-tilt portfolios using the short-horizon models. The portfolios were constructed by cap-weighting the 55 USE3 industries, the 60 USE4 industries, and the top and bottom quintiles for the 12 USE3 styles and the 12 USE4 styles. The MRAD curve for USE3S had three distinct features. The first was the Russian Default of August 1998. This represents an outlier event that no risk model could have reliably predicted based on information available at the end of July, 1998. The second feature corresponded to overprediction during 2004. This feature is prominent in USE3 but largely absent from USE4. This reflects the Volatility Regime Adjustment quickly calibrating to the downward-trending volatility of that period. The third main feature occurred in 2010, when USE3S again overpredicted risk. Although USE4S also overpredicted risk immediately following the financial crisis, the degree of overprediction was much smaller.

Figure 5.6

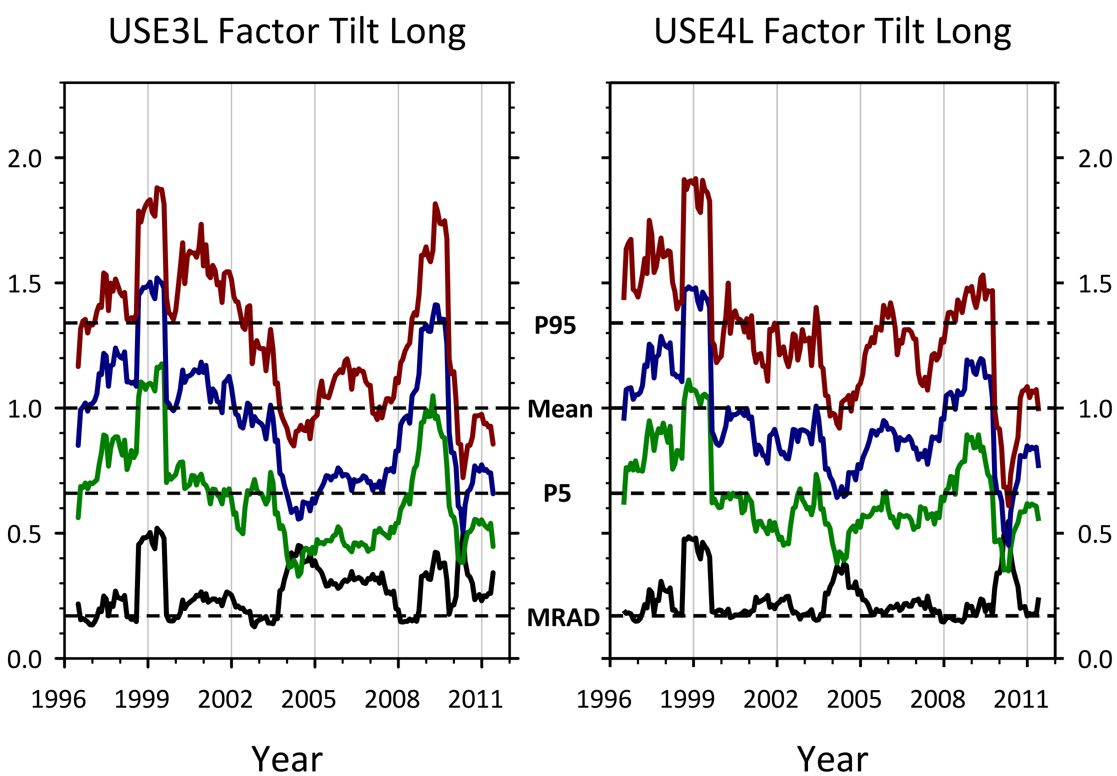
Comparison of USE3S Model and USE4S Model for industry and style-tilt long-only portfolios. For each model, we construct industry portfolios by cap-weighting all stocks with primary exposure to a particular industry factor. We also construct two portfolios for each style factor by cap-weighting the top and bottom quintile stocks. This leads to 163 long-only portfolios. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.7** we plot MRAD and bias statistics for the same 163 long-only factor-tilt portfolios as in Figure 5.6, except now using the long-horizon models. The USE3L Model overpredicted risk by a wide margin from 2004-2008. Entering the financial crisis of 2008, it underpredicted risk, followed by overprediction as the crisis subsided. The USE4L Model exhibited the same main features, except that the magnitudes of the biases were considerably smaller. That is, the Volatility Regime Adjustment helped to reduce the dramatic swings of overprediction and underprediction.

Figure 5.7

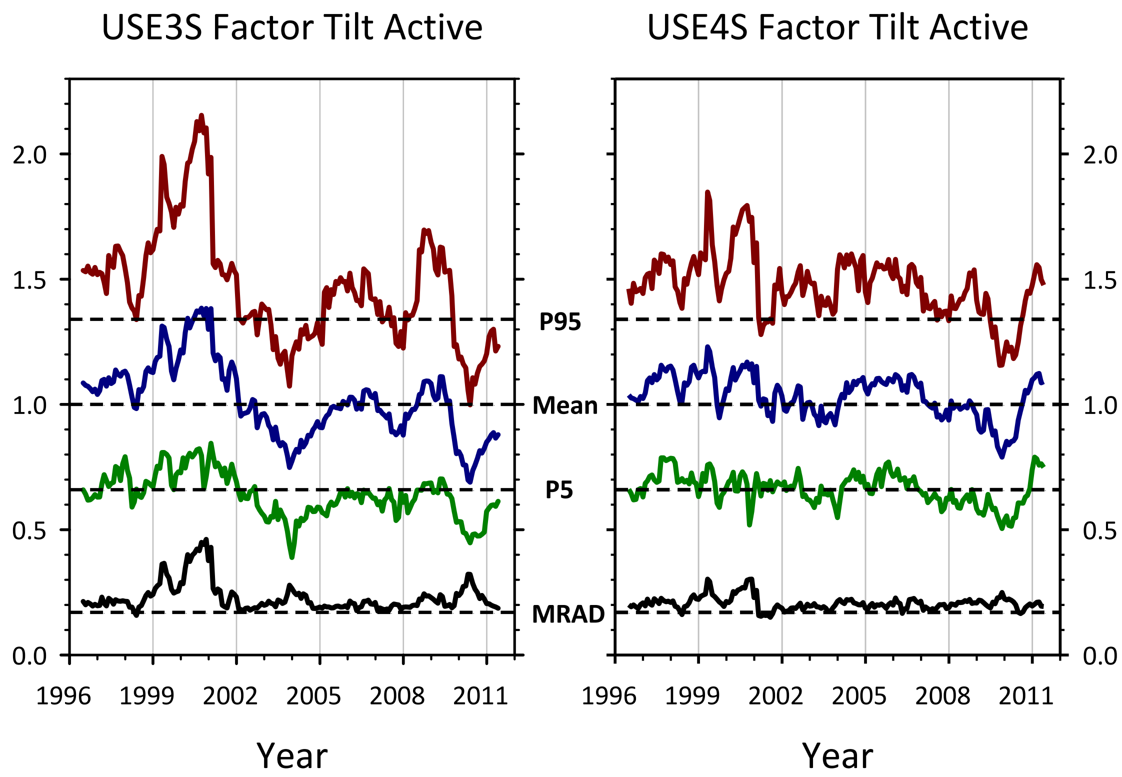
Comparison of USE3L Model and USE4L Model for industry and style-tilt long-only portfolios. For each model, we construct industry portfolios by cap-weighting all stocks with primary exposure to a particular industry factor. We also construct two portfolios for each style factor by cap-weighting the top and bottom quintile stocks. This leads to 163 portfolios. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.8** we plot MRAD and bias statistics for the 163 active factor-tilt portfolios using the short-horizon models. The portfolios were constructed by going long the factor-tilt portfolios of Figure 5.6 and shorting the USE4 estimation universe. The USE3S Model tended to significantly underpredict the risk of these portfolios during the Internet Bubble, followed by overprediction in 2003-2004. The USE3S Model also significantly overpredicted risk in the wake of the financial crisis. The USE4S Model, by contrast, produced MRAD and bias statistics that were very stable and close to their ideal values over the entire sample period. In particular, the underforecasting during the Internet Bubble was very mild, as was the overforecasting following the financial crisis.

Figure 5.8

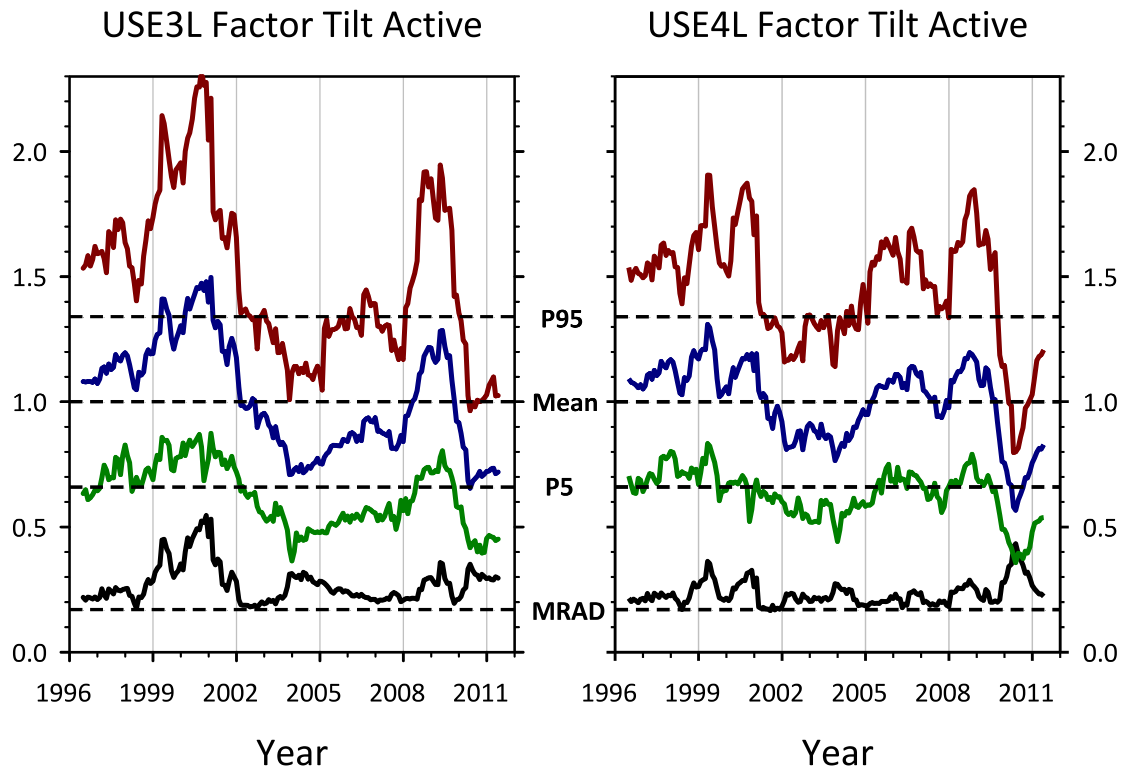
Comparison of USE3S Model and USE4S Model for industry and style-tilt active portfolios. We form the portfolios by going long the 163 factor-tilt portfolios described in Figure 5.6 and going short the USE4 estimation universe. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.9** we plot MRAD and bias statistics for the same 163 active factor-tilt portfolios, except now using the long-horizon models. The USE3L Model significantly underpredicted the risk of these portfolios during the Internet Bubble, followed by a five-year period of overprediction as volatility levels subsided. Entering the financial crisis, the USE3S Model underpredicted risk, followed by overprediction in the aftermath of the crisis. The USE4L Model generally gave more accurate risk forecasts throughout the sample period, with the exception of a brief period in 2010 during which the model significantly overpredicted risk.

Figure 5.9

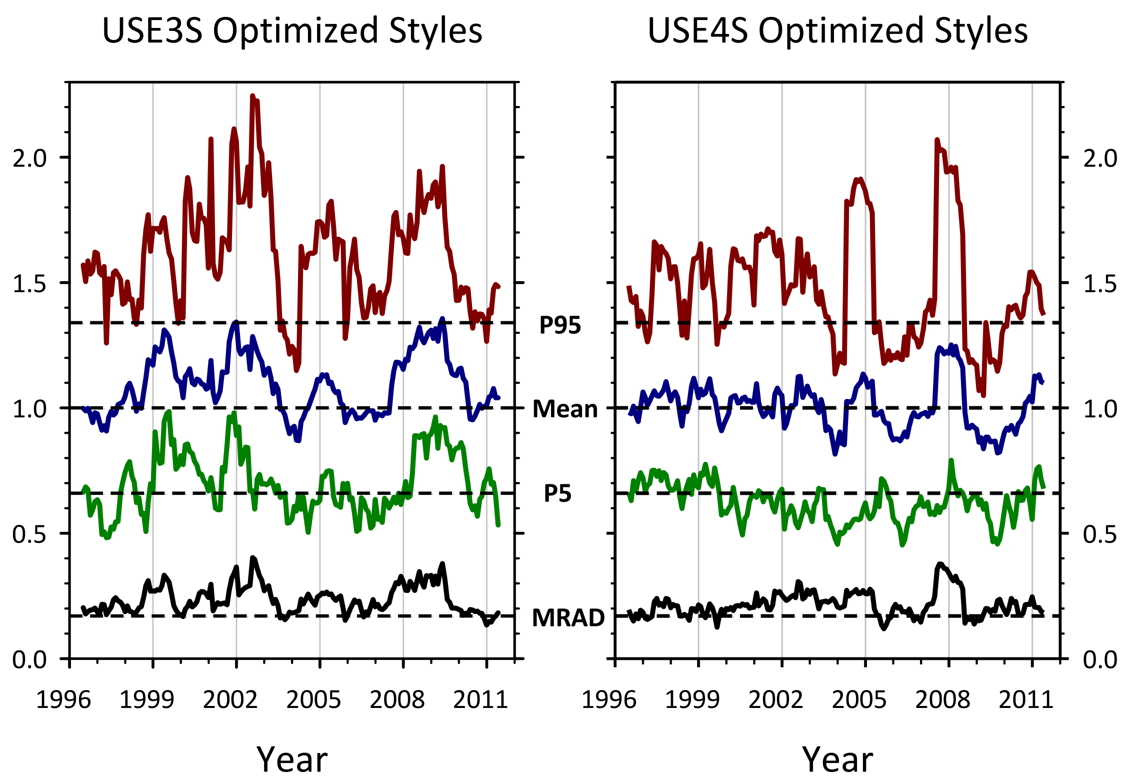
Comparison of USE3L Model and USE4L Model for industry and style-tilt active portfolios. We form the portfolios by going long the 163 factor-tilt portfolios described in Figure 5.6 and going short the USE4 estimation universe. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.10** we plot MRAD and bias statistics for 24 optimized style-tilt portfolios using the short-horizon models. The 24 USE3S optimized portfolios were constructed by using the 12 USE3 style factors as “alpha signals” and then forming the minimum volatility portfolio (with alpha equal to 1) for two draws of 500 randomly selected stocks. The 24 USE4S optimized portfolios were constructed similarly, except using the 12 USE4 style factors as the “alpha signals.” The mean bias statistics for the USE3S model were greater than 1 over most of the sample period, indicating underprediction of risk for these optimized portfolios. By contrast, the mean bias statistics for USE4S were much closer to 1 on average, indicating that the Optimization Bias Adjustment was effective at reducing the underforecasting biases for these optimized portfolios. It is also interesting to note that the Quant Meltdown in August 2007 is clearly visible for the USE4S Model.

Figure 5.10

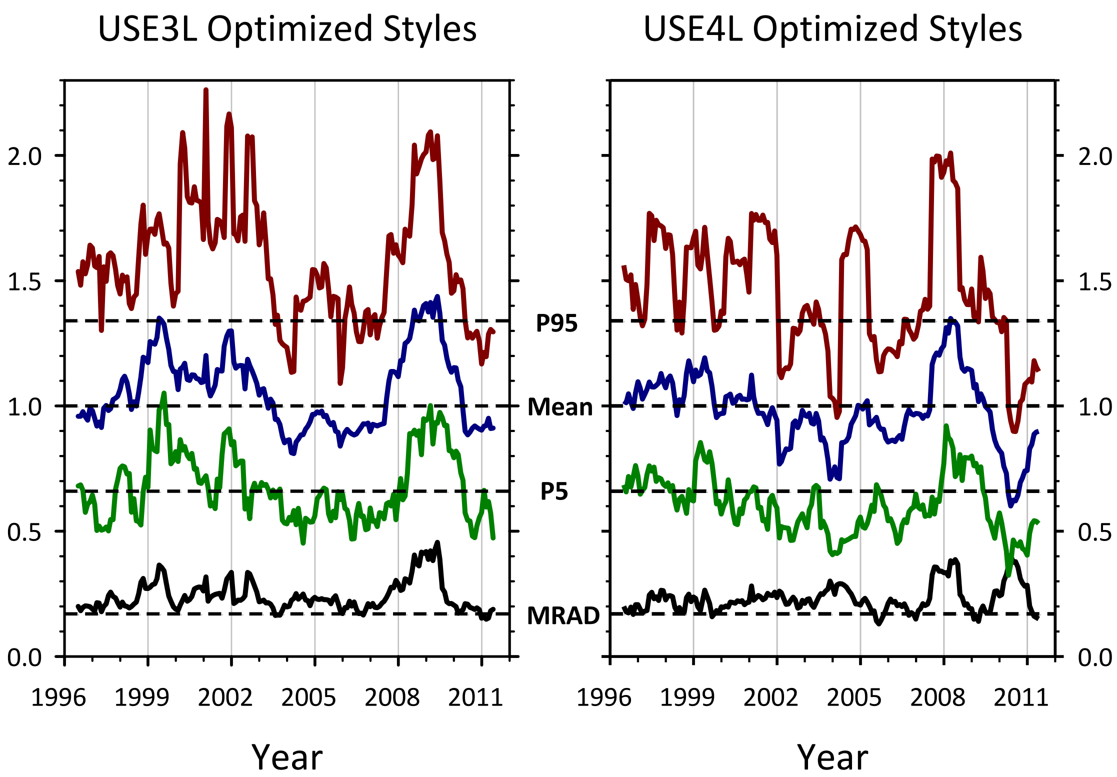
Comparison of USE3S Model and USE4S Model for 24 optimized portfolios. The portfolios were constructed by minimizing risk subject to the unit alpha constraint, where the alpha signals were taken from the 12 style factors of each model. Two sets of optimizations were performed using 500 randomly selected stocks for each style factor. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.11** we plot MRAD and bias statistics for 24 optimized style-tilt portfolios constructed in the same fashion as in Figure 5.10, except now using the long-horizon models. The bias statistics for the USE3L Model were shifted upward throughout most of the sample period, indicating underprediction of risk for these optimized portfolios. The mean bias statistics for the USE4L Model were closer to 1 on average, although there was a tendency to overpredict risk in 2010.

Figure 5.11

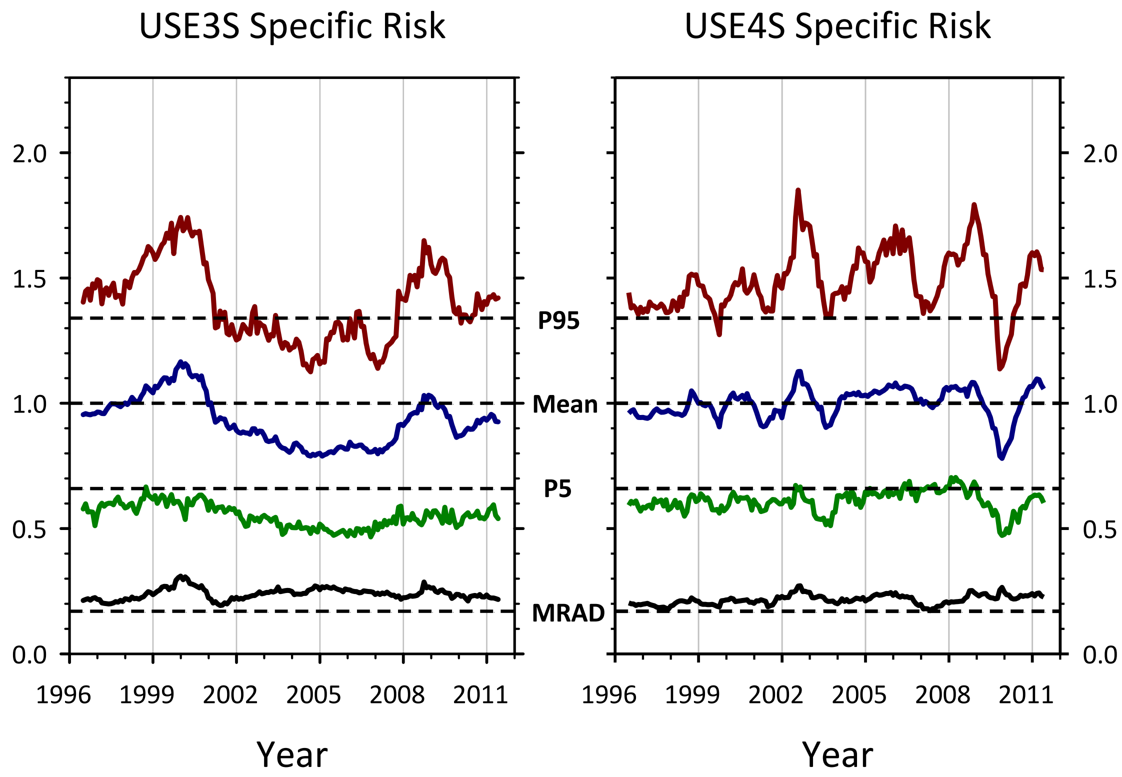
Comparison of USE3L model and USE4L model for optimized portfolios. The portfolios were constructed by minimizing risk subject to the unit alpha constraint, where the alpha signals were taken from the 12 style factors of each model. Two sets of optimizations were performed using 500 randomly selected stocks for each style factor. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.12** we plot cap-weighted MRAD and bias statistics for the specific returns of all stocks in the USE4 estimation universe using the short-horizon models. With the exception of the Internet Bubble, the USE3S Model tended to overpredict specific risk over the sample period. By contrast, the USE4S specific risk model had mean bias statistics very close to 1 over the sample period. A notable exception was in late 2009, when the USE4S Model tended to overforecast risk slightly. Nonetheless, the overforecasting period was rather short-lived, as the Volatility Regime Adjustment again produced bias statistics close to 1 by late 2010.

Figure 5.12

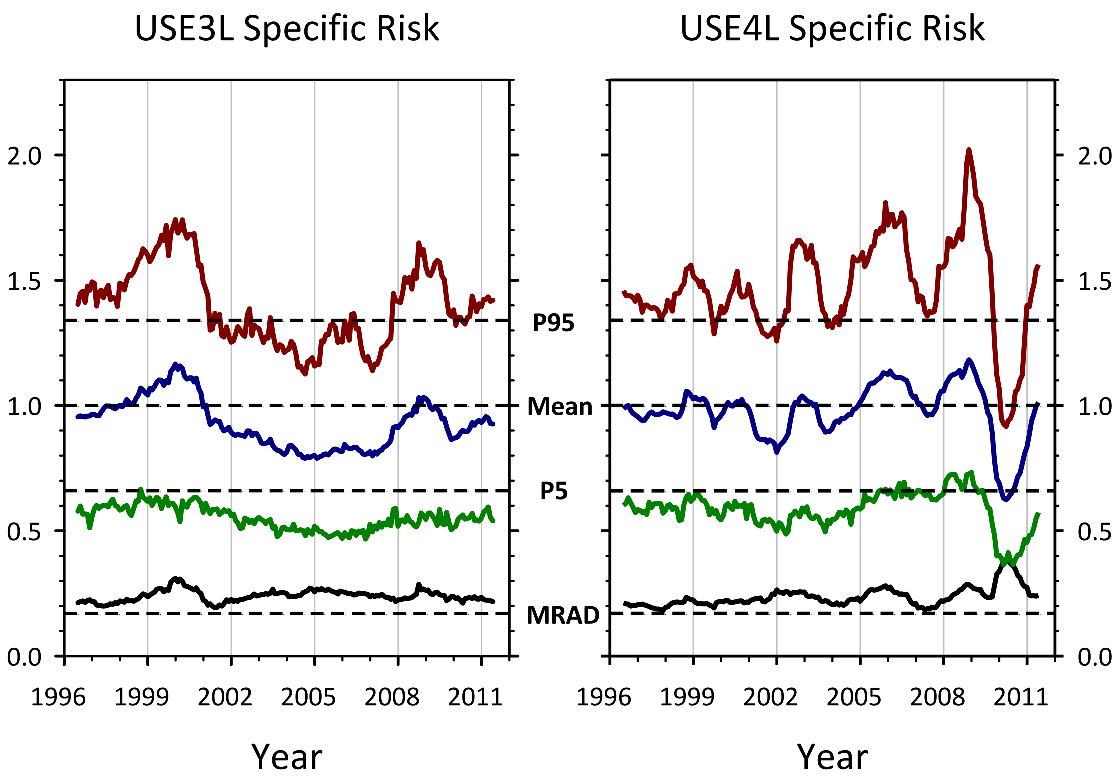
Comparison of USE3S Model and USE4S Model for specific risk. Results were capitalization weighted. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In **Figure 5.13** we plot cap-weighted MRAD and bias statistics for the same specific returns as in Figure 5.12, except now using the long-horizon models. Note that the USE3 Model uses the same specific risk forecasts for both the short-horizon and long-horizon versions. The mean bias statistics for USE4L were close to 1 over most of the sample period, with the exception of late 2009 and early 2010, when the model overpredicted risk. Again, however, the overprediction was fairly short-lived, as the Volatility Regime Adjustment produced bias statistics close to 1 by early 2011.

Figure 5.13

Comparison of USE3L Model and USE4L Model for specific risk. Results were capitalization weighted. Each plot reports rolling 12-month values for the mean bias statistic, the 5-percentile and 95-percentile bias statistics, and MRAD. The dashed horizontal lines indicate the ideal positions assuming normally distributed returns and perfect risk forecasts.



In Table 5.1 we present summary MRAD and mean bias statistic numbers for the USE3 and USE4 Models for the test cases presented in Figures 5.2 to 5.13. In virtually every case, the mean bias statistics for the USE4 Model were closer to 1 compared with the USE3 Model. Furthermore, for every class of portfolios, the USE4 Model produced more accurate risk forecasts as measured by the MRAD statistic. The average outperformance by the MRAD measure, coincidentally, was 174 bps for both the short-horizon and long-horizon models. If we exclude pure factors, which represent an apples-to-oranges comparison, then the MRAD reduction is greater than 200 bps for USE4S and nearly 200 bps for USE4L. If we take 0.19 as the lower bound of MRAD for perfect risk forecasts and realistic levels of kurtosis, then the USE4S Model represents a reduction in excess MRAD of more than 50 percent relative to the USE3S Model.

Table 5.1

Summary of mean bias statistics and MRAD for USE3 and USE4 Models.

Figures	MRAD (USE3S)	Mean B (USE3S)	MRAD (USE4S)	Mean B (USE4S)	MRAD Diff (bp)	(S-Model) Portfolio Type
5.2	0.2111	0.96	0.2109	1.00	2	Pure Factors
5.4	0.2070	0.91	0.1884	0.99	186	Random Active
5.6	0.2186	0.96	0.2025	0.98	161	Factor Tilts Long
5.8	0.2305	1.02	0.2060	1.03	245	Factor Tilts Active
5.10	0.2365	1.09	0.2151	1.00	214	Optimized Styles
5.12	0.2399	0.93	0.2162	1.00	237	Specific Risk
Average	0.2239	0.98	0.2065	1.00	174	
Figures	MRAD (USE3L)	Mean B (USE3L)	MRAD (USE4L)	Mean B (USE4L)	MRAD Diff (bp)	(L-Model) Portfolio Type
5.3	0.2630	0.95	0.2542	0.96	88	Pure Factors
5.5	0.2199	0.90	0.2136	0.95	63	Random Active
5.7	0.2716	0.96	0.2358	0.95	358	Factor Tilts Long
5.9	0.2732	1.02	0.2338	1.00	394	Factor Tilts Active
5.11	0.2409	1.06	0.2305	0.98	104	Optimized Styles
5.13	0.2399	0.93	0.2364	0.97	35	Specific Risk
Average	0.2514	0.97	0.2341	0.97	174	

6. Conclusion

The new Barra US Equity Model (USE4) is the result of extensive research efforts in combination with client consultations. The USE4 Model incorporates many methodological innovations and advances designed to address long-standing problems in risk modeling. For instance, the Optimization Bias Adjustment addresses the issue of underestimation of risk for optimized portfolios, and leads to better conditioning of the covariance matrix. The Volatility Regime Adjustment calibrates volatilities to current market levels and represents a key determinant of risk forecasts, especially during times of market turmoil. The introduction of the Country factor leads to more intuitive attribution of portfolio risk and return, while also providing timelier forecasts of industry correlations. Another enhancement is the use of a Bayesian adjustment technique which aims to reduce biases in specific risk forecasts.

This document provides a thorough empirical analysis of the USE4 Model. The factor structure is described in transparency and detail, for both industries and styles. The performance of select factors is presented and discussed. Key metrics are reported at the individual factor level, including statistical significance, performance, volatility, and correlation.

We also analyze the explanatory power of the USE4 Model, and compare it with the USE3 Model. We find that the USE4 Model had consistently higher explanatory power. We attribute most of this increase to the new USE4 industry factors, although the USE4 style factors also contributed to the improved performance. In addition, we study the contributions to cross-sectional dispersion from the Country factor, industries, and styles. We find that each category of factors was of comparable importance in explaining the observed cross-sectional dispersion of equity returns.

Lastly, we systematically compare the forecasting accuracy of the USE4S and USE4L Models versus their USE3 counterparts over a roughly 16-year backtesting window. We consider several classes of portfolios, including pure factors, random active portfolios, factor-tilt portfolios (both long-only and dollar-neutral), and optimized portfolios. We also compare the accuracy of specific risk forecasts between the two models. For every portfolio type considered, we find that the USE4S and USE4L Models provided more accurate risk forecasts than their USE3 counterparts during the sample period.

Appendix A: Descriptors by Style Factor

Beta

Definition: *1.0 BETA*

BETA *Beta* (β)

Computed as the slope coefficient in a time-series regression of excess stock return, $r_t - r_{ft}$, against the cap-weighted excess return of the estimation universe R_t,

$$r_t - r_{ft} = \alpha + \beta R_t + e_t . \quad \text{(A1)}$$

The regression coefficients are estimated over the trailing 252 trading days of returns with a half-life of 63 trading days.

Momentum

Definition: *1.0 · RSTR*

RSTR *Relative strength*

Computed as the sum of excess log returns over the trailing $T = 504$ trading days with a lag of $L = 21$ trading days,

$$RSTR = \sum_{t=L}^{T+L} w_t \left[\ln(1 + r_t) - \ln(1 + r_{ft}) \right], \quad \text{(A2)}$$

where r_t is the stock return on day t, r_{ft} is the risk-free return, and w_t is an exponential weight with a half-life of 126 trading days.

Size

Definition: *1.0 · LNCAP*

LNCAP *Log of market cap*

Given by the logarithm of the total market capitalization of the firm.

Earnings Yield

Definition: *0.75 · EPFWD + 0.15 · CETOP + 0.10 · ETOP*

EPFWD *Predicted earnings-to-price ratio*

Given by the 12-month forward-looking earnings divided by the current market capitalization. Forward-looking earnings are defined as a weighted average between the average analyst-predicted earnings for the current and next fiscal years.

CETOP *Cash earnings-to-price ratio*

Given by the trailing 12-month cash earnings divided by current price.

ETOP *Trailing earnings-to-price ratio*

Given by the trailing 12-month earnings divided by the current market capitalization. Trailing earnings are defined as the last reported fiscal-year earnings plus the difference between current interim figure and the comparative interim figure from the previous year.

Residual Volatility

Definition: *0.75 · DASTD + 0.15 · CMRA + 0.10 · HSIGMA*

DASTD *Daily standard deviation*

Computed as the volatility of daily excess returns over the past 252 trading days with a half-life of 42 trading days.

CMRA *Cumulative range*

This descriptor differentiates stocks that have experienced wide swings over the last 12 months from those that have traded within a narrow range. Let $Z(T)$ be the cumulative excess log return over the past T months, with each month defined as the previous 21 trading days

$$Z(T) = \sum_{\tau=1}^{T} \left[\ln(1+r_{\tau}) - \ln(1+r_{f\tau}) \right], \qquad \text{(A3)}$$

where r_{τ} is the stock return for month τ (compounded over 21 days), and $r_{f\tau}$ is the risk-free return. The cumulative range is given by

$$CMRA = \ln(1+Z_{\max}) - \ln(1+Z_{\min}), \qquad \text{(A4)}$$

where $Z_{\max} = \max\{Z(T)\}$, $Z_{\min} = \min\{Z(T)\}$, and $T = 1,\ldots,12$.

HSIGMA *Historical sigma* (σ)

Computed as the volatility of residual returns in Equation A1,

$$\sigma = \mathrm{std}(e_t). \qquad \text{(A5)}$$

The volatility is estimated over the trailing 252 trading days of returns with a half-life of 63 trading days.

Note: The Residual Volatility factor is orthogonalized to Beta to reduce collinearity.

Growth

Definition: *0.70 · EGRLF + 0.20 · EGRO + 0.10 · SGRO*

EGRLF *Long-term predicted earnings growth*

Long-term (3-5 years) earnings growth forecasted by analysts.

EGRO *Earnings growth (trailing five years)*

Annual reported earnings per share are regressed against time over the past five fiscal years. The slope coefficient is then divided by the average annual earnings per share to obtain the earnings growth.

SGRO *Sales growth (trailing five years)*

Annual reported sales per share are regressed against time over the past five fiscal years. The slope coefficient is then divided by the average annual sales per share to obtain the sales growth.

Dividend Yield

Definition: *1.0 · YILD*

YILD *Dividend-to-price ratio*

Given as the trailing 12-month dividend per share divided by current price.

Book-to-Price

Definition: *1.0 · BTOP*

BTOP *Book-to-price ratio*

Last reported book value of common equity divided by current market capitalization.

Leverage

Definition: *0.75 · MLEV + 0.15 · DTOA + 0.10 · BLEV*

MLEV *Market leverage*

Computed as

$$MLEV = \frac{ME + PE + LD}{ME}, \tag{A6}$$

where ME is the market value of common equity on the last trading day, PE is the most recent book value of preferred equity, and LD is the most recent book value of long-term debt.

DTOA *Debt-to-assets*

Computed as

$$DTOA = \frac{TD}{TA}, \tag{A7}$$

where TD is the book value of total debt (long-term debt and current liabilities), and TA is most recent book value of total assets.

BLEV *Book leverage*

Computed as

$$BLEV = \frac{BE + PE + LD}{BE}, \tag{A8}$$

where BE is the most recent book value of common equity, PE is the most recent book value of preferred equity, and LD is the most recent book value of long-term debt.

Liquidity

Definition: *0.35 · STOM + 0.35 · STOQ + 0.30 · STOA*

STOM — *Share turnover, one month*

Computed as the log of the sum of daily turnover during the previous 21 trading days,

$$STOM = \ln\left(\sum_{t=1}^{21} \frac{V_t}{S_t}\right), \tag{A9}$$

where V_t is the trading volume on day t, and S_t is the number of shares outstanding.

STOQ — *Average share turnover, trailing 3 months*

Let $STOM_\tau$ be the share turnover for month τ, with each month consisting of 21 trading days. The quarterly share turnover is defined by

$$STOQ = \ln\left[\frac{1}{T}\sum_{\tau=1}^{T} \exp\left(STOM_\tau\right)\right], \tag{A10}$$

where $T = 3$ months.

STOA — *Average share turnover, trailing 12 months*

Let $STOM_\tau$ be the share turnover for month τ, with each month consisting of 21 trading days. The annual share turnover is defined by

$$STOA = \ln\left[\frac{1}{T}\sum_{\tau=1}^{T} \exp\left(STOM_\tau\right)\right], \tag{A11}$$

where $T = 12$ months.

Non-linear Size

Definition: *1.0 · NLSIZE*

NLSIZE — *Cube of Size*

First, the standardized Size exposure (i.e., log of market cap) is cubed. The resulting factor is then orthogonalized to the Size factor on a regression-weighted basis. Finally, the factor is winsorized and standardized.

Non-linear Beta

Definition: *1.0 · NLBETA*

NLBETA — *Cube of Beta*

First, the standardized Beta exposure is cubed. The resulting factor is then orthogonalized to the Beta factor on a regression-weighted basis. Finally, the factor is winsorized and standardized.

Appendix B: Decomposing RMS Returns

We decompose excess stock returns r_n into a systematic component, due to factors, and a stock-specific component u_n. The factor returns f_k are estimated each period by cross-sectional regression

$$r_n = \sum_k X_{nk} f_k + u_n \ , \tag{B1}$$

where X_{nk} is the exposure of stock n to factor k. The specific returns are assumed to be uncorrelated with one another as well as to the other factors.

The total *R*-squared of a regression measures the cross-sectional variation explained by the factors,

$$R_T^2 = 1 - \frac{\sum_n v_n u_n^2}{\sum_n v_n r_n^2} \ , \tag{B2}$$

where v_n is the regression weight of stock n (proportional to square-root of market capitalization). The root mean square (RMS) return, computed as

$$RMS = \sqrt{\sum_n v_n r_n^2} \, , \tag{B3}$$

measures the cross-sectional dispersion from zero return. As described by Menchero and Morozov (2011), the RMS return can be exactly decomposed into the return sources of Equation B1 using a cross-sectional version of the *x-sigma-rho* formula,

$$RMS = \sum_k f_k \sigma(X_k) \rho(X_k, r) + \sigma(u) \rho(u, r) \ , \tag{B4}$$

where $\sigma(X_k)$ is the RMS dispersion of factor k, and $\rho(X_k, r)$ is the cross-sectional correlation between factor k and the asset returns. The last term in Equation B4 represents the contribution to RMS coming from stock-specific sources.

Appendix C: Review of Bias Statistics

C1. Single-Window Bias Statistics

A commonly used measure to assess a risk model's accuracy is the bias statistic. Conceptually, the bias statistic represents the ratio of realized risk to forecast risk.

Let R_{nt} be the return to portfolio n over period t, and let σ_{nt} be the beginning-of-period volatility forecast. Assuming perfect forecasts, the *standardized* return,

$$b_{nt} = \frac{R_{nt}}{\sigma_{nt}}, \tag{C1}$$

has an expected standard deviation of 1. The bias statistic for portfolio n is the *realized* standard deviation of standardized returns,

$$B_n = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(b_{nt} - \bar{b}_n\right)^2}, \tag{C2}$$

where T is the number of periods in the observation window.

Assuming normally distributed returns and perfect risk forecasts, for sufficiently large T the bias statistic B_n is approximately normally distributed about 1, and roughly 95 percent of the observations fall within the confidence interval,

$$B_n \in \left[1 - \sqrt{2/T}, 1 + \sqrt{2/T}\right]. \tag{C3}$$

If B_n falls outside this interval, we reject the null hypothesis that the risk forecast was accurate.

If returns are not normally distributed, however, then fewer than 95 percent of the observations will fall within the confidence interval, even for perfect risk forecasts. In **Figure C1**, we show simulated results for the percentage of observations actually falling within this interval, plotted versus observation window length T, for several values of kurtosis k.

For the normal case (kurtosis $k = 3$), except for the smallest values of T, the confidence interval indeed captures about 95 percent of the observations. As the kurtosis increases, however, the percentage falling within the interval drops significantly. For instance, at a kurtosis level of 5, only 86 percent of bias statistics fall inside the confidence interval for an observation window of 120 periods.

C2. Rolling-Window Bias Statistics

The purpose of bias-statistic testing is to assess the accuracy of risk forecasts, typically over a long sample period. Let T be the length of the observation window, which corresponds to the number of months in the sample period. One possibility is to select the entire sample period as a single window, and to compute the bias statistic as in Equation C2. This would be a good approach if financial data were stationary, as sampling error is reduced by increasing the length of the window. In reality, however, financial data are not stationary. It is possible to significantly overpredict risk for some years, and underpredict it for others, while ending up with a bias statistic close to 1.

Often, a more relevant question is to study the accuracy of risk forecasts over 12-month periods. For this purpose, we define the rolling 12-month bias statistic for portfolio n,

$$B_n^\tau = \sqrt{\frac{1}{11}\sum_{t=\tau}^{\tau+11}(b_{nt} - \bar{b}_n)^2}\ , \tag{C4}$$

Where τ denotes the first month of the 12-month window. The 12-month windows are rolled forward one month at a time until reaching the end of the observation window. If T is the number of periods in the observation window, then each portfolio will have $T-11$ (overlapping) 12-month windows.

It is useful to consider, for a collection of N portfolios, the mean of the rolling 12-month bias statistics,

$$\bar{B}^\tau = \frac{1}{N}\sum_n B_n^\tau\ . \tag{C5}$$

We also define $B^\tau(5\%)$ and $B^\tau(95\%)$ to be the 5-percentile and 95-percentile values for the rolling 12-month bias statistics at a given point in time. Assuming normal distributions and perfect risk forecasts, these values should be centered about 0.66 and 1.34, respectively. Plotting these quantities versus time for different classes of portfolios is a visually powerful way of understanding the predictive accuracy of the risk model.

Another useful measure to consider is the 12-month *mean rolling absolute deviation* (MRAD), defined as

$$\mathrm{MRAD}^\tau = \frac{1}{N}\sum_n \left| B_n^\tau - 1 \right|. \tag{C6}$$

This penalizes every deviation away from the ideal bias statistic of 1. Smaller MRAD numbers, of course, are preferable to larger ones. A lower limit for this statistic can be obtained by assuming the ideal case of normally distributed returns and perfect risk forecasts, which leads to an expected value of 0.17 for the 12-month MRAD.

It is interesting to consider how MRAD depends on kurtosis levels. In **Figure C2** we report simulated results for 12-month MRAD assuming perfect risk forecasts. For normally distributed returns, as discussed, the expected MRAD value is 0.17. At higher kurtosis levels, however, the expected MRAD for perfect forecasts increases significantly. For instance, even at moderate kurtosis levels in the range of 3.5 to 4.0, the 12-month MRAD for perfect risk forecasts rises to approximately 0.19.

Figure C1

Percent of observations falling within the confidence interval $1 \pm \sqrt{2/T}$, where T is the number of periods in the observation window. Results were simulated using a normal distribution $k = 3$, and using a *t*-distribution with kurtosis values $k = 5$ and $k = 10$. The standard deviations were equal to 1 in all cases.

For the normal distribution, the percentage of observations inside the confidence interval quickly approaches 95 percent. As kurtosis is increased, however, the proportion within the confidence interval declines considerably.

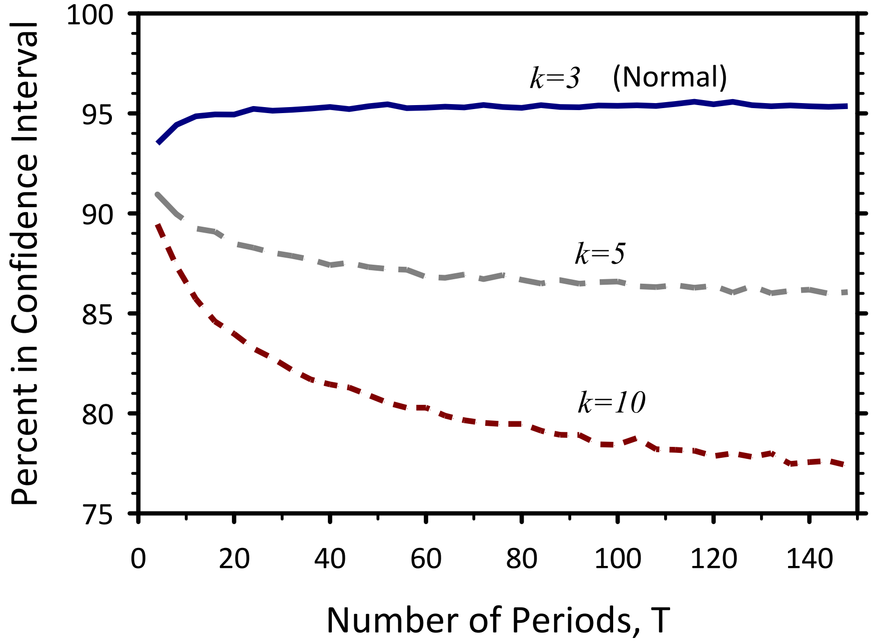
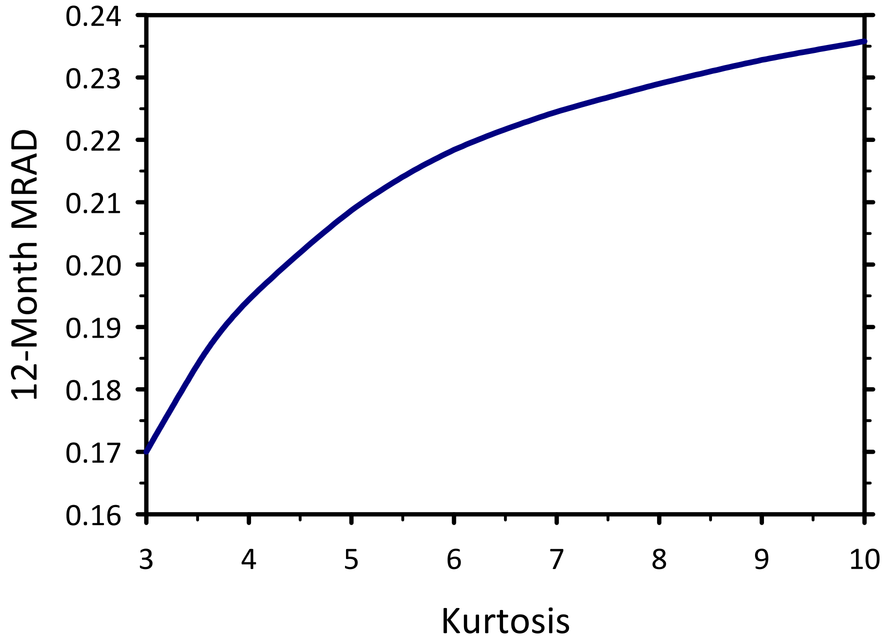


Figure C2

Plot of 12-month MRAD versus kurtosis levels for perfect risk forecasts. Results were simulated using a *t*-distribution.



REFERENCES

Banz, R., 1981. “The Relationship between Return and Market Value of Common Stock,” *Journal of Financial Economics*, vol. 9: 3-18.

Basu, S., 1977. “Investment Performance of Common Stocks in Relation to their Price-Earnings Ratio: A Test of the Efficient Market Hypothesis,” *Journal of Finance*, vol. 32, June, 663-682.

Briner, Beat, Rachael Smith, and Paul Ward. 2009. “The Barra European Equity Model (EUE3),” *Research Notes*.

Davis, Ben, and Jose Menchero. 2011. “Pitfalls in Risk Attribution.” *MSCI Research Insight*.

Fama, E., and R. French. 1992. “The Cross-Section of Expected Stock Returns,” *Journal of Finance*, vol. 47. no. 2: 427-465.

Jegadeesh, N., and S. Titman, 1993. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency,” *Journal of Finance*, vol. 48: 65-92.

Menchero, Jose. 2010. “The Characteristics of Factor Portfolios.” *Journal of Performance Measurement*, vol. 15, no. 1 (Fall): 52-62.

Menchero, Jose, Andrei Morozov, and Peter Shepard. 2008. “The Barra Global Equity Model (GEM2).” *Research Notes*.

Menchero, J., Morozov, A., and P. Shepard. 2010. “Global Equity Risk Modeling,” in J. Guerard, Ed., *The Handbook of Portfolio Construction: Contemporary Applications of Markowitz Techniques*, (New York: Springer), 439-480.

Menchero, Jose, and Andrei Morozov. 2011. “Decomposing Global Equity Cross-Sectional Volatility.” *Financial Analysts Journal*, vol. 67, no. 5 (September/October): to appear.

Menchero, Jose, Jun Wang, and D.J. Orr. “Optimization Bias Adjustment—Improving Risk Forecasts Using Eigen-Adjusted Covariance Matrices.” MSCI Research Insight, May 2011.

Muller, Peter. 1993. “Financial Optimization.” *Cambridge University Press* (Zenios): 80-98.

Sharpe, W.F., September 1964. “Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk,” *Journal of Finance*, vol. 19, no. 3: 425-442.

Shepard, Peter. 2009. “Second Order Risk.” *Working paper, http://arxiv.org/abs/0908.2455v1*.

Client Service Information is Available 24 Hours a Day

clientservice@msci.com

Americas

Americas	1.888.588.4567 (toll free)
Atlanta	+ 1.404.551.3212
Boston	+ 1.617.532.0920
Chicago	+ 1.312.675.0545
Montreal	+ 1.514.847.7506
Monterrey	+ 52.81.1253.4020
New York	+ 1.212.804.3901
San Francisco	+ 1.415.836.8800
Sao Paulo	+ 55.11.3706.1360
Stamford	+1.203.325.5630
Toronto	+ 1.416.628.1007

Europe, Middle East & Africa

Cape Town	+ 27.21.673.0100
Frankfurt	+ 49.69.133.859.00
Geneva	+ 41.22.817.9777
London	+ 44.20.7618.2222
Milan	+ 39.02.5849.0415
Paris	0800.91.59.17 (toll free)

Asia Pacific

China North	10800.852.1032 (toll free)
China South	10800.152.1032 (toll free)
Hong Kong	+ 852.2844.9333
Seoul	798.8521.3392 (toll free)
Singapore	800.852.3749 (toll free)
Sydney	+ 61.2.9033.9333
Tokyo	+ 81.3.5226.8222

About MSCI

MSCI Inc. is a leading provider of investment decision support tools to investors globally, including asset managers, banks, hedge funds and pension funds. MSCI products and services include indices, portfolio risk and performance analytics, and governance tools.

The company's flagship product offerings are: the MSCI indices which include over 148,000 daily indices covering more than 70 countries; Barra portfolio risk and performance analytics covering global equity and fixed income markets; RiskMetrics market and credit risk analytics; ISS governance research and outsourced proxy voting and reporting services; FEA valuation models and risk management software for the energy and commodities markets; and CFRA forensic accounting risk research, legal/regulatory risk assessment, and due-diligence. MSCI is headquartered in New York, with research and commercial offices around the world.