Model Insight

The Barra US Equity Model (USE4)

Methodology Notes

Jose Menchero

D.J. Orr

Jun Wang

August 2011

Contents

1. Introduction

1.1. Model Highlights

This document describes the new methodologies that underpin the USE4 model. Our aim is to produce a document that is clear and concise, yet comprehensive as well. MSCI prides itself not only on setting the standard for excellence in factor risk modeling, but also on being the industry leader in model transparency.

This document is the complement to a companion document: *USE4 Empirical Notes*. Whereas the current document focuses on methodology, the *Empirical Notes* contain detailed information about Barra USE$ MUSE4 factor structure, extensive analysis on the explanatory power and statistical significance of the factors, and a systematic investigation into the forecasting accuracy of the model. The *Empirical Notes* also provide a thorough comparison with the USE3 model.

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 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. Both 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 the most accurate forecasts at a monthly prediction horizon. The USE4L model is designed for longer-term investors who are willing to trade some degree of accuracy for greater stability in risk forecasts.

1.2. Modern Portfolio Theory and Barra Risk Models: A Brief History

The pioneering work of Markowitz (1952) formally established the intrinsic tradeoff between risk and return. This paradigm provided the foundation upon which the modern theory of finance was built, and has proven so resilient that it has survived essentially intact for nearly 60 years. Almost as remarkable is the vigor with which the theory has been embraced by academics and practitioners alike.

The specific problem addressed by Markowitz was how to construct an efficient portfolio from a collection of risky assets. Markowitz defined an efficient portfolio as one that had the highest expected return for a given level of risk, which he measured as standard deviation of portfolio returns. Markowitz showed that the relevant risk of an asset is not its stand-alone volatility, but rather its contribution to *portfolio* risk. Thereafter, the concepts of risk and correlation became inseparable.

A plot of expected return versus volatility for the set of all efficient portfolios maps out a curve known as the efficient frontier. In order to construct the efficient frontier using the Markowitz prescription, an investor must provide expected returns and covariances for the universe of all investable assets. The Markowitz procedure identifies the optimal portfolio corresponding to the risk tolerance of any given investor.

Tobin (1958) took the Markowitz methodology and extended it in a very simple way that nonetheless had profound implications for portfolio management. By including cash in the universe of investable assets, Tobin showed that there existed a single portfolio on the efficient frontier that, when combined with cash, dominated all other portfolios. For any investor, therefore, the optimal portfolio would always consist of a combination of cash and the "super-efficient" portfolio. For instance, risk-averse investors may combine the super-efficient portfolio with a large cash position, whereas risk seekers would borrow cash to purchase more of the super-efficient portfolio. As a result, according to Tobin, the optimal investment strategy consists of two separate steps. The first is to determine the super-efficient portfolio. The second step is to determine the appropriate level of cash that matches the overall risk tolerance of the investor. This two-step investment process came to be known as the Tobin separation theorem.

The next major step in the development of Capital Market Theory was due to Sharpe (1964). By making certain assumptions (e.g., that all investors followed mean-variance preferences and agreed on the expected returns and covariances of all assets) Sharpe was able to show that the super-efficient portfolio was the market portfolio itself. Sharpe's theory, known as the Capital Asset Pricing Model, predicts that the expected return of an asset depends only on the expected return of the market and the beta of the asset relative to the market. In other words, within CAPM, the only "priced" factor is the market factor.

Using the CAPM framework, the return of any asset can be decomposed into a systematic component that is perfectly correlated with the market, and a residual component that is uncorrelated with the market. The CAPM predicts that the expected value of the residual return is zero. This does not preclude the possibility, however, of *correlations* among the residual returns. That is, even under the CAPM, there may be multiple sources of equity return co-movement, even if there is only one source of expected return.

Rosenberg (1974) was the first to develop multi-factor risk models to estimate the asset covariance matrix. This work was later extended by Rosenberg and Marathe (1975), who conducted a sweeping econometric analysis of multi-factor models. The intuition behind these models is that there exists a relatively parsimonious set of pervasive factors that drive asset returns. Returns that cannot be explained by the factors are deemed "stock specific" and are assumed to be uncorrelated.

Rosenberg founded Barra, which made widespread use of multi-factor risk models and dedicated itself to helping practitioners implement the theoretical insights of Markowitz, Tobin, Sharpe, and others. The first multi-factor risk model for the US market, dubbed the Barra USE1 Model, was released in 1975. That model was followed by the USE2 Model in 1985, and USE3 in 1997. Rapidly changing volatility levels during and after the Internet Bubble highlighted the need for more responsive risk models, and in 2002 the USE3 Model was upgraded to incorporate daily factor returns.

Another key step in developing the theoretical edifice of quantitative investing came with the publishing in 1995 of an influential book entitled *Active Portfolio Management,* written by Grinold and Kahn while at Barra. The widespread success of this book prompted a second edition by Grinold and Kahn (2000), and it serves today as an essential guidebook for many quantitative investment firms.

For modeling global portfolios, an important milestone came in 1989 with the development of the first Barra Global Equity Risk Model (GEM). This model was estimated via monthly cross-sectional regressions using countries, industries, and styles as explanatory factors, as described by Grinold, Rudd, and Stefek (1989).

GEM was followed by a second-generation Global Equity Risk Model, GEM2, as described by Menchero, Morozov, and Shepard (2008). GEM2 incorporated several advances over the previous model, such as improved estimation techniques, higher-frequency observations, and the introduction of the World factor to place countries and industries on an equal footing.

Barra also pioneered the use of integrated models, which combine the breadth of a global model with the detail of local single-country models. An innovative feature of this approach is that it assures consistency between the risk forecasts used by portfolio managers in the front office and risk managers in the middle office. The first-generation Barra Integrated Model (BIM) was introduced in 2002. The second-generation Barra Integrated Model, described by Shepard (2011), incorporated important advances in methodology, such as using the GEM2 model to estimate covariances among local factors and employing higher-frequency observations.

Barra risk models have long played an important role in applying the concepts of modern portfolio theory to solve practical investment problems. At MSCI, we are dedicated to continuing this proud tradition of developing industry-leading risk models. The release of the new Barra US Equity Model, USE4, marks only the latest step in this ongoing journey.

1.3 Forecasting Portfolio Risk with Factor Models

The asset covariance matrix is critical both for portfolio construction and for risk management purposes. A key challenge in estimating the asset covariance matrix lies in the sheer dimensionality of the problem. For instance, an active portfolio containing 2000 stocks requires more than *two million* independent elements. If the asset covariance matrix is computed naively — that is, by brute force — then the matrix is likely to be extremely ill-conditioned. This makes the asset covariance matrix highly susceptible to noise and spurious relationships that are unlikely to persist out-of-sample. For instance, if the number of time observations is less than the number of stocks (as would be typical for large portfolios), the matrix is said to be "rank deficient," meaning that it is possible to construct apparently riskless portfolios.

Factor risk models were developed to provide a more robust solution to this problem. Stock returns are attributed to a factor component that affects all stocks, and an idiosyncratic component that is unique to the particular stock. More specifically, the stock return is explained as

$$r_n = \sum_k X_{nk} f_k + u_n , \tag{1.1}$$

where X_{nk} is the exposure of stock n to factor k, f_k is the return to the factor, and u_n is the stock specific return.

Consider a portfolio with weights w_n, and return given by

$$R_P = \sum_n w_n r_n . \tag{1.2}$$

The portfolio factor exposures are given by the weighted average of the asset exposures, i.e.,

$$X_k^P = \sum_n w_n X_{nk} . \tag{1.3}$$

Therefore, the portfolio return can be expressed as

$$R_P = \sum_k X_k^P f_k + \sum_n w_n u_n . \tag{1.4}$$

Two key assumptions in factor risk modeling are that (a) the factor returns are uncorrelated with the specific returns and (b) the specific returns are uncorrelated among themselves. This allows the variance of the portfolio to be expressed as

$$\mathrm{var}(R_P) = \sum_{kl} X_k^P F_{kl} X_l^P + \sum_n w_n^2 \,\mathrm{var}(u_n) , \tag{1.5}$$

where F_{kl} is the predicted covariance between factors k and l. If, for example, the risk model contains 60 factors, then the factor covariance matrix contains less than two thousand independent elements. It is due to this tremendous reduction in the dimensionality of the problem that the factor model is able to filter out most of the noise and provide a robust measure of portfolio risk.

2. Factor Exposures

2.1 General Considerations

Sound factor structure represents a key pillar for producing a high-quality risk model. Factors represent broad drivers of asset return co-movement, segmenting portfolio risk and return into two distinct sources. The first source, due to factors, represents the systematic component. The second source represents the diversifiable component that cannot be explained by the factors, and is therefore deemed idiosyncratic or asset specific.

A key assumption underlying factor risk models is that the factors capture all systematic drivers of asset returns, thus implying that the specific returns are mutually uncorrelated. Therefore, it is essential that a high-quality factor structure explain as fully as possible the cross section of asset returns. A common pitfall in risk model construction is to omit important factors. A portfolio with exposures to these factors may be adversely impacted by “hidden” risk sources that spuriously appeared to be diversified.

Another important characteristic of a high-quality factor structure is parsimony, meaning that the systematic component of asset returns is explained with the fewest possible number of factors. Parsimonious factor structures tend to be more robust in capturing true underlying relationships. Indeed, another pitfall in risk model construction is to indiscriminately add weak or spurious factors, as this makes the model more susceptible to noise that can be detrimental to the portfolio construction process.

When investigating risk model factor structure, careful attention must also be paid to the statistical significance of the factor returns. In particular, the statistical significance of the factors should be persistent across time, and not due to isolated events that are unlikely to recur in the future. Careful and thorough statistical analysis helps ensure that spurious factors do not find their way into the model.

Stability is another characteristic of a high-quality factor structure. Stability means that typical stock exposures do not change drastically over short periods of time, which can be problematic from a risk management perspective. We define a *factor stability coefficient* as the cross-sectional correlation of factor exposures from one month to the next; this can be expressed as

$$\rho_{kt} = \frac{\sum_n v_n^t \left(X_{nk}^t - \bar{X}_k^t \right)\left(X_{nk}^{t+1} - \bar{X}_k^{t+1} \right)}{\sqrt{\sum_n v_n^t \left(X_{nk}^t - \bar{X}_k^t \right)^2}\sqrt{\sum_n v_n^t \left(X_{nk}^{t+1} - \bar{X}_k^{t+1} \right)^2}}, \tag{2.1}$$

where v_n^t is the regression weight of stock n at time t. As a rule of thumb, factor stability coefficients above 0.90 are considered desirable, whereas those below 0.80 are deemed too unstable for model inclusion.

Collinearity, another important consideration when researching and building a factor structure, occurs when a given factor can be approximately replicated by a combination of other factors. If the factor structure is excessively collinear, estimation errors in the regressions can become very large and the factor returns difficult to interpret. One measure of factor collinearity is given by the *Variance Inflation Factor* (VIF). This is computed by regressing a single factor against the other factors in the model:

$$X_{nk} = \sum_{k' \neq k} X_{nk'} b_{k'} + \varepsilon_{nk} . \tag{2.2}$$

The VIF is defined through the explanatory power of the regression, as measured by R-squared. More specifically:

$$VIF_k = \frac{1}{1 - R_k^2}. \tag{2.3}$$

A large value for VIF is indicative of excessive collinearity. As described by Menchero (2010), a factor rotation procedure known as orthogonalization can serve as an effective tool for both reducing collinearity and making the factors more intuitive. For instance, the Non-Linear Size factor, after orthogonalization, essentially captures the return differences between mid-cap stocks and the overall market.

Last, but not least, risk model factors should be intuitive. In other words, they must be transparent, easily interpretable, and consistent with investors' views about what these factors represent. For example, it is reasonable to demand that a broad value or growth index have positive exposure to a well-constructed value or growth factor. Further discussion of this issue can be found in the MSCI Research Insight, *Global Style Factors* (2010).

2.2. Data Quality and Outlier Treatment

The most difficult and time-consuming part of constructing an equity risk model lies in preparation of the input data. If the data inputs are of poor quality, the risk forecasts will likewise be poor, no matter how sophisticated the model. Assuring a high degree of data quality, therefore, is essential for building a reliable risk model.

In order to obtain the highest quality inputs, Barra risk models leverage the same data infrastructure used for the construction of MSCI Global Investable Market Indices. Data items such as raw descriptors, GICS® codes, country classifications, and clean daily stock returns are obtained from in-house sources that have already undergone extensive quality control.

No matter how stringent the data quality assurance process, we can never exclude the possibility of extreme outliers entering the data set. These extreme outliers may represent legitimate values or outright data errors. Either way, such observations must be carefully handled to prevent a few data points from having an undue impact on model estimation.

In USE4, we employ a multi-step algorithm to identify and treat outliers. The algorithm assigns each observation into one of three groups. The first group represents values so extreme that they are treated as potential data errors and removed from the estimation process. The second group represents values that are regarded as legitimate, but nonetheless so large that their impact on the model must be limited. We *trim* these observations to three standard deviations from the mean. The third group of observations, forming the bulk of the distribution, consists of values that are less than three standard deviations from the mean; these observations are left unadjusted.

The thorough data quality assurance process, coupled with the robust outlier algorithm, is designed to ensure that the input data are clean and reliable. There is still the issue, however, of dealing with *missing* data. Data could be missing either due to lack of availability or due to removal by the outlier algorithm.

If the underlying data required to calculate factor exposures are missing, then the factor exposures are generated using a data-replacement algorithm. In this case, we regress non-missing exposures against a subset of factors, with the slope coefficients being used to estimate the factor exposures for the stocks with missing data. The intuition is that stocks with similar attributes, such as industry membership or market capitalization, are likely to share similar exposures.

Note also that the data-replacement algorithm is applied at the factor level as opposed to the descriptor level. In other words, if a stock has data for some descriptors within a style factor, but not others, then the non-missing data will be used to compute the factor exposures. Only if *all* descriptors are absent does the replacement algorithm become active.

2.3. Style Exposures

Style factors constitute major drivers of equity cross-sectional returns. In fact, as shown by Menchero and Morozov (2011), there are periods when style factors explain a greater proportion of cross-sectional variation of global equity returns than either country factors or industry factors.

Style factors are constructed from descriptors, which represent financially intuitive stock attributes and serve as effective predictors of equity return covariance. Since the descriptors within a particular style factor are meant to capture the same underlying driver of returns, these descriptors tend to be significantly collinear. Combining these descriptors into a single style factor overcomes the collinearity problem, while also leading to a more parsimonious factor structure.

Descriptors are standardized to have a mean of 0 and a standard deviation of 1. In other words, if d_{nl}^{Raw} is the raw value of stock n for descriptor l, then the standardized descriptor value is given by

$$d_{nl} = \frac{d_{nl}^{Raw} - \mu_l}{\sigma_l}, \tag{2.4}$$

where μ_l is the cap-weighted mean of the descriptor (within the estimation universe), and σ_l is the equal-weighted standard deviation. We adopt the convention of standardizing using the cap-weighted mean so that a well-diversified cap-weighted portfolio has approximately zero exposure to all style factors. For the standard deviation, however, we use equal weights to prevent large-cap stocks from having an undue influence on the overall scale of the exposures.

Formally, descriptors are combined into style factors as follows:

$$X_{nk} = \sum_{l \in k} w_l \, d_{nl} , \tag{2.5}$$

where w_l is the descriptor weight, and the sum takes place over all descriptors within a particular style factor. Descriptor weights are determined using an optimization algorithm designed to maximize the explanatory power of the model. The final step of constructing the style factor is to re-standardize the exposures so that they have a cap-weighted mean of 0 and an equal-weighted standard deviation of 1.

2.4. Industry Factors

Industries represent another major class of risk model factors. When constructing industry factors, it is important that the factors reflect the return drivers of the local market. Using a single level of the GICS hierarchy, for example, fails to capture the unique characteristics of each financial market. Furthermore, this naive approach invariably leaves some systematic risk unaccounted for and suffers from the problem of thin industries and low statistical significance. To avert these shortcomings, Barra models employ customized factor structure to reflect the local characteristics of each market.

Identifying which industry factors to include in the model involves a combination of judgment and empirical analysis. Within each sector we drill into different levels of the GICS hierarchy. The basic criteria we use to guide industry factor selection are: (a) groupings of industries into factors must be economically intuitive, (b) industry factors should have a strong degree of statistical significance, (c) incorporating an additional industry factor should significantly increase the explanatory power of the model, and (d) thin industries (those with few assets) should be avoided.

2.5. Multiple-Industry Exposures

Within the context of an industry classification scheme such as GICS, the industry membership of a given stock is determined by analysts who scrutinize financial statements filed by the firm in question. By studying such filings, the analysts are able to determine the primary line of business for the firm and make an appropriate industry assignment. However, there are many instances where such analysis will reveal complications; for example, when a firm has substantial business activities in *multiple* industries. It is valuable to model *fractional* industry exposures in order to capture a more realistic representation of the underlying business activities of the stocks in a given portfolio.

Multiple-industry membership is modeled in USE4 by examining the impact of two key explanatory variables — Assets and Sales — on the market value of a given stock. More specifically, we regress the market capitalizations of the estimation universe stocks against their reported Assets within each industry,

$$M_n = \sum_k A_{nk} \beta_k^A + \varepsilon_n , \tag{2.6}$$

where M_n is the market capitalization of stock n, A_{nk} is the Assets of the stock within industry k, β_k^A is the industry beta, and ε_n is the residual. Note that β_k^A can be interpreted as the price-to-assets ratio of the industry. The industry exposures, using Assets as the explanatory variable, are given by the fraction of market capitalization explained by each industry,

$$X_{nk}^A = \frac{A_{nk} \beta_k^A}{\sum_k A_{nk} \beta_k^A} . \tag{2.7}$$

Note that the industry exposures in Equation 2.7, by construction, will sum to 1.

We also use Sales as an explanatory variable to estimate industry exposures X_{nk}^S. To produce the net USE4 multiple-industry exposures, we combine Assets exposures and Sales exposures according to the following weighting scheme:

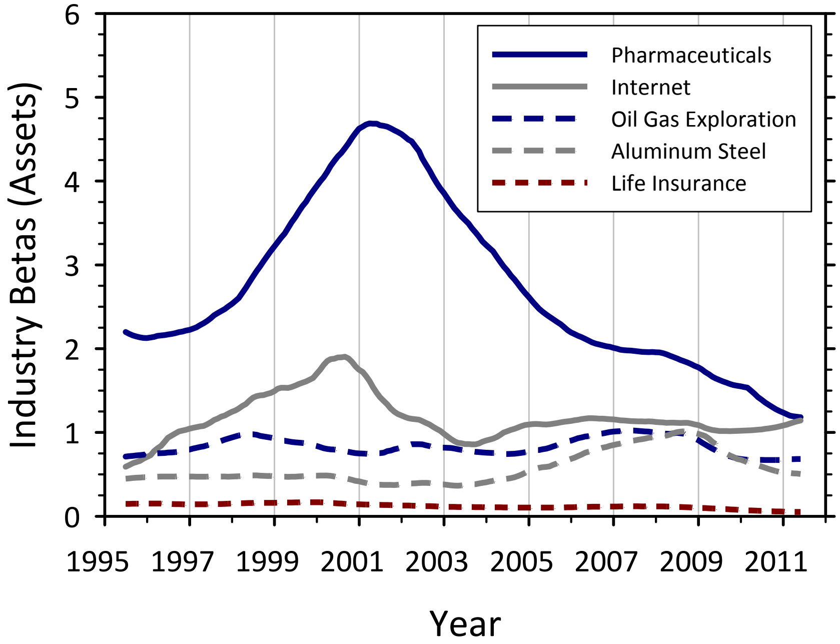
$$X_{nk} = 0.75 \cdot X_{nk}^A + 0.25 \cdot X_{nk}^S . \tag{2.8}$$

The maximum allowable number of multiple-industry exposures is limited to five. If a firm has more than five business segments, the top five are taken and the industry exposures are re-normalized to 1.

Intuitively, we expect that industry betas (e.g., price-to-assets ratios) may vary from one industry to the next. To illustrate this effect, in **Figure 2.1** we plot Asset industry betas versus time for Pharmaceuticals (with the largest average industry beta over the sample period), Life Insurance (which had the smallest average beta), and three other industries with intermediate betas. Clearly, a dollar of Assets in the Pharmaceuticals industry translated into a much higher market capitalization than a dollar of Assets in the Life Insurance industry.

Figure 2.1

Plot of Asset industry betas versus time for representative industries. A dollar of Assets in Pharmaceuticals translated into higher market capitalization than a dollar of Assets in Life Insurance.

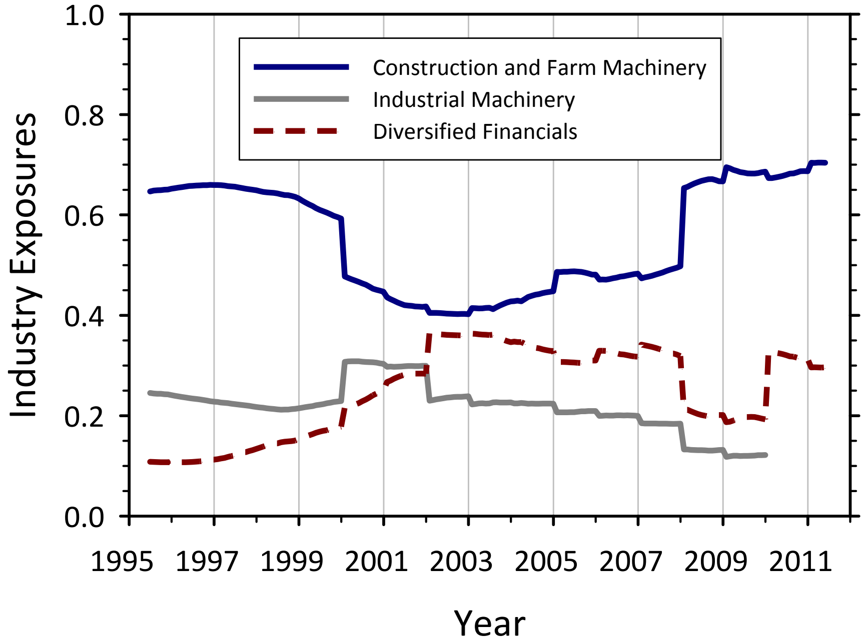


It is also important to note that most data sets for segment reporting utilize industry schemes that have a far higher level of granularity than GICS. For example, the North American Industry Classification System (NAICS) contains nearly ten times the number of segments as represented in GICS. Constructing multiple-industry exposures therefore requires an accurate mapping between NAICS and GICS, along with managing a list of legitimate exceptions. This is an activity for which MSCI is ideally suited. As a partner in the construction of GICS, our firm is able to leverage the same industry analysis experts who manage GICS in order to build the NAICS-to-GICS mapping that is central to the USE4 multiple-industry methodology. This produces a multiple-industry scheme that is not only *accurate*, but *consistent* with GICS as well.

As an illustrative example, we consider the multiple-industry exposures for Deere & Company, a leading producer of agricultural machinery. In **Figure 2.2**, we plot the exposures from 30-Jun-1995 to 31-May-2011. The largest exposure is consistently on the USE4 Construction and Farm Machinery, which also corresponds to the single-industry GICS assignment of the firm. However, Deere & Company also has sizeable allocations to Industrial Machinery and Diversified Financials. The jumps in multiple-industry exposures are attributable to changes in the financial statement reporting.

Figure 2.2

Multiple industry exposures for Deere & Company, plotted versus time.



3. Factor Returns

3.1. Country Factor

A significant advance in the USE4 methodology is explicit inclusion of the Country factor. This enhancement, which is analogous to the World factor in GEM2, is beneficial in two ways. First, it disentangles the market effect from the pure industry effect, thus providing greater insight into the sources of risk and return. Second, adding the Country factor allows for more accurate risk forecasts and more responsive correlations, as shown in the next section.

In the USE4 model, local excess returns r_n are explained by

$$r_n = f_c + \sum_i X_{ni} f_i + \sum_s X_{ns} f_s + u_n , \tag{3.1}$$

where f_c is the return of the Country factor, f_i is the return of industry factor i, f_s is the return of style factor s, X_{ni} and X_{ns} are the exposures of stock n to industry i and style s, respectively, and u_n is the specific return. Factor returns in USE4 are estimated using weighted least-squares regression, assuming that the variance of specific returns is inversely proportional to the square root of total market capitalization. This regression-weighting scheme reflects the empirical observation that the idiosyncratic risk of a stock decreases as the market capitalization of the firm increases.

In Equation 3.1, we see that every stock has an exposure of 1 to the Country factor. Country exposure, however, should not be confused with *market* exposure. Some stocks are clearly more sensitive to market movements than others. This effect is captured through the Beta factor. Note that the Beta factor is highly correlated (temporally, not cross-sectionally) with the Country factor. This reflects the observation that high-beta stocks tend to outperform low-beta stocks in an up-market, while underperforming in down-markets. From a risk perspective, therefore, it is important to consider the Country factor and the Beta factor jointly. For instance, in a fully invested portfolio holding high-beta stocks, the portfolio will have positive exposure to both factors, and the risk contributions will be additive. By contrast, in a fully invested portfolio of low-beta stocks, the risk contributions will partially cancel.

Note that adding the Country factor introduces an exact collinearity into the model. That is, for all stocks n, the sum of industry factor exposures is given by

$$\sum_i X_{ni} = 1 , \tag{3.2}$$

which is exactly the USE4 Country factor exposure. A constraint, therefore, must be applied to obtain a unique regression solution.

In USE4 we adopt the constraint that the cap-weighted industry factor returns sum to zero,

$$\sum_i w_i f_i = 0 , \tag{3.3}$$

where w_i is the capitalization weight of the estimation universe in industry i. The choice of constraint does not affect the regression fit or the explanatory power of the model, but does directly affect the interpretation of the factors. We select our constraint to provide the most intuitive interpretation of factor portfolios, as we now discuss.

As described by Menchero (2010), factor returns can be interpreted as the returns of pure factor portfolios. To understand the Country factor portfolio, consider the cap-weighted estimation universe, with holdings h_n^E. The return of this portfolio R_E can be attributed to the USE4 factors,

$$R_E = f_c + \sum_i w_i f_i + \sum_s X_s^E f_s + \sum_n h_n^E u_n , \tag{3.4}$$

where X_s^E is exposure of the cap-weighted estimation universe to style s. Note, however, by Equation 3.3, the sum over industries must equal zero. Similarly, since style factors are standardized to be cap-weighted mean zero (i.e., $X_s^E = 0$), the second summation in Equation 3.4 also vanishes. The final sum in Equation 3.4 corresponds to the specific return of a broadly diversified portfolio, and is therefore *approximately* zero (note that it would be *exactly* zero had we used regression weights instead of capitalization weights). Thus, to an excellent approximation, Equation 3.4 reduces to

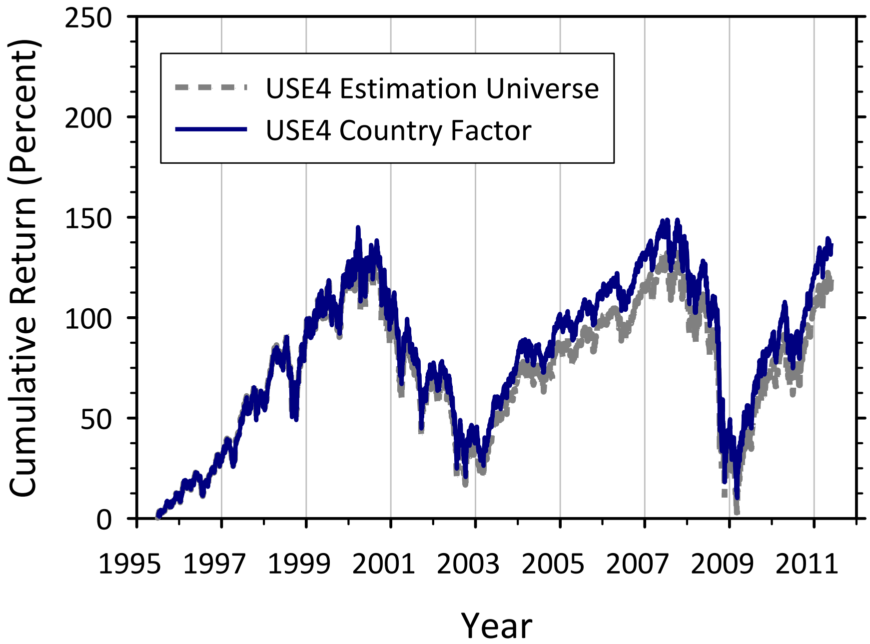
$$R_E \approx f_c . \tag{3.5}$$

In other words, the Country factor essentially represents the cap-weighted estimation universe.

In **Figure 3.1** we plot the cumulative return of the USE4 Country factor, and compare it to the cumulative local excess return of the USE4 estimation universe. The two portfolios track each other almost perfectly over the 16-year sample period, with a time-series correlation of approximately 99.9 percent, thus validating the interpretation of the Country factor.

Figure 3.1

Cumulative returns of the USE4 Country factor and the USE4 estimation universe (in excess of risk-free rate). The two portfolios track each other very closely over the 16-year sample window, with a time-series correlation of 99.9 percent.



The industry factors in Equation 3.4 can be interpreted as the returns of dollar-neutral portfolios that go 100 percent long the particular industry and go 100 percent short the Country factor. The industry factor portfolios also have zero exposure to all style factors. Therefore, industry factors capture the performance of the pure industry relative to the overall market, net of all style effects.

Similarly, style factors can be interpreted as the returns of dollar-neutral portfolios that have unit exposure to the style in question and zero exposures to all other factors. In particular, pure style factor portfolios have net zero weight in every industry. These portfolios allow us to measure the performance of a particular style without the confounding effects of other style or industry tilts.

3.2. Relation to Traditional Approach

Traditionally, single country models do not include the Country factor. In this case, stock returns are described as follows:

$$r_n = \sum_i X_{ni} \tilde{f}_i + \sum_s X_{ns} f_s + u_n . \tag{3.6}$$

There are two differences between Equation 3.1 and Equation 3.6. First, and most obvious, the former explicitly includes the Country factor, whereas the latter does not. Second, the industry factor returns are different in the two representations. The industry factor returns in Equation 3.6, however, can be easily related to those in Equation 3.1,

$$\tilde{f}_I = f_C + f_I . \tag{3.7}$$

In other words, the traditional industry factors are given by the sum of the Country factor and the corresponding dollar-neutral industry factor. This implies that the factor portfolio $\tilde{f}_i$ is 100 percent net long the particular industry and has net zero weight in all other industries.

Beyond the intuitive benefit of disentangling the market from the pure industries, including the Country factor provides advantages for risk-forecasting purposes as well. The underlying reason is that the covariance matrix, for reasons discussed in Section 4, uses different half-life parameters to estimate factor volatilities and factor correlations. Typically, the correlation half-life is significantly longer than the volatility half-life.

Consider the correlation $\tilde{\rho}_{ij}$ between two net-long industry factors, $\tilde{f}_i$ and $\tilde{f}_j$, as in Equation 3.7. If this is estimated in a model without the Country factor, the responsiveness of the correlation forecast is completely determined by the relatively slow correlation half-life. If, on the other hand, we estimate the correlation using a model that contains the Country factor, then the common exposure to the Country factor can act as a source return co-movement between $\tilde{f}_i$ and $\tilde{f}_j$. More specifically, the correlation can be expressed as

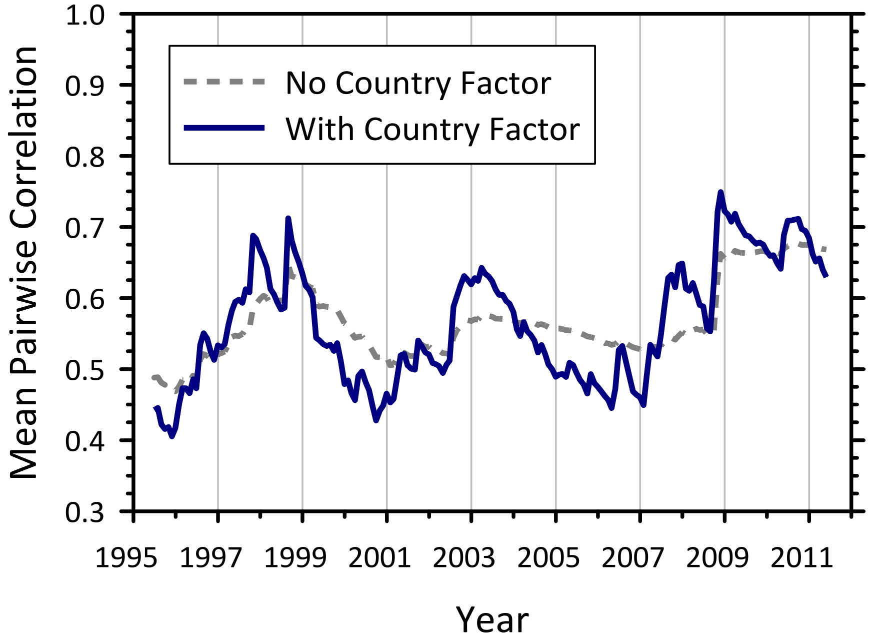
$$\tilde{\rho}_{ij} = \frac{\sigma_c^2 + \rho_{ci}\sigma_c\sigma_i + \rho_{cj}\sigma_c\sigma_j + \rho_{ij}\sigma_i\sigma_j}{\tilde{\sigma}_i\tilde{\sigma}_j} , \tag{3.8}$$

where $\tilde{\sigma}_i$ is the volatility of $\tilde{f}_i$, σ_c is the volatility of f_c, σ_i is the volatility of f_i, and ρ_{ij}, is the correlation between f_i and f_j. All estimates in Equation 3.8 are obtained from a model that includes the Country factor. Since the volatility of the Country factor is estimated with the more responsive volatility half-life, the correlation forecast $\tilde{\rho}_{ij}$ is likewise more responsive.

Intuitively, we know that during times of financial crisis, market volatility rises and industries become more correlated. In **Figure 3.2**, we plot the mean pair-wise correlation $\tilde{\rho}_{ij}$ for all industry pairs in the USE4 model. The dashed grey curve is obtained using a model without the Country factor, and the dark solid line is for a model with the Country factor. The average correlations are similar in magnitude, but the correlations are clearly much more responsive in the model with the Country factor.

Figure 3.2

Mean pair-wise correlation between net-long pure industry factors. Averages were computed across all 1770 possible pair-wise correlations for the 60 industry factors in USE4. Results were obtained using the USE4S covariance matrix parameters reported in Table 4.1. Common exposures to the Country factor allow industry correlations to be more responsive.



We must also verify that the correlations in the model including the Country factor are more accurate. For this purpose, we created long/short industry-pair portfolios for all possible combinations of USE4 industries. The returns to these portfolios are given by $\tilde{f}_i \quad \tilde{f}_j$, and have a large fraction of their risk deriving from the correlation term. To evaluate the accuracy of the risk forecasts, we computed standardized returns and Mean Rolling Absolute Deviations (MRAD) for these portfolios, as described in Appendix A. We found that including the Country factor reduced the MRAD statistic by about 80 bps on average during a roughly 16-year sample period (30-Jun-1995 to 31-May-2011).

4. Factor Covariance Matrix

4.1 Established Methods

The factor covariance matrix predicts the volatilities and correlations of the factors and thus represents a second key pillar for constructing a high-quality risk model. The USE4 factor covariance matrix builds upon methodologies used in the Barra Global Equity Model, GEM2, as described by Menchero, Morozov, and Shepard (2008). In this section, we provide a brief overview of these established methodologies. Innovations and enhancements to the factor covariance matrix methodology are described in Section 4.2 and Section 4.3.

Estimation of the USE4 factor covariance matrix follows a multi-step process. The first step is to compute the factor correlation matrix from the set of daily factor returns. We employ exponentially weighted averages, characterized by the factor correlation half-life parameter τ_ρ^F. This approach gives more weight to recent observations and is an effective method for dealing with data non-stationarity.

Special care must be exercised when selecting the correlation half-life. To ensure a well-conditioned correlation matrix, the half-life must be sufficiently long so that the effective number of observations T is significantly greater than the number of factors K. An ill-conditioned factor covariance matrix may be problematic for portfolio optimization purposes, since active portfolios can be constructed that exhibit spuriously low factor risk. On the other hand, if the correlation half-life is too long, then undue weight is placed on distant observations that have little relation to current market conditions. Finding the proper balance between these two limits is essential for producing a reliable correlation matrix.

The prediction horizon of the USE4 model is one month. The factor correlation matrix, however, is estimated from daily factor returns. We must therefore account for the possibility of serial correlation in factor returns, as these may affect risk forecasts over a longer horizon.

In USE4, we employ the Newey-West methodology (1987) to account for serial-correlation effects. A key parameter in this approach is the number of lags L_ρ^F over which the serial correlation is deemed important. For instance, $L_\rho^F = 2$ implies that the return of any factor may be correlated with the return of any other factor within a two-day time span.

Another complication in estimating the factor correlation matrix arises from the case of missing factor returns. In a global model, country holidays are one possible reason for missing factor returns. In single country models, missing factor returns may arise from using time series of differing lengths. For instance, the Internet factor may appear in the model only after the start date of the cross-sectional regressions.

We use the EM algorithm of Dempster (1977) to estimate the correlation matrix for the case of missing factor returns. This method employs an iterative procedure to estimate the correlation matrix. The EM algorithm also refines the correlation forecasts as new information flows into the model, while ensuring that the correlation matrix is positive semi-definite.

With the correlation matrix thus computed, the next step is to compute the factor volatilities. We use exponentially weighted averages, with half-life parameter τ_σ^F. In estimating factor volatilities, we also employ the Newey-West approach allowing factor returns to be auto-correlated over a number of lags L_σ^F. Next, we construct the initial covariance matrix by scaling in the factor volatilities. That is, the covariance between factors i and j is given by

$$F_{ij}^0 = \rho_{ij}\sigma_i\sigma_j \ , \tag{4.1}$$

where σ_i and σ_j are the factor volatilities and ρ_{ij} is their respective correlation.

Selecting the proper factor volatility half-life is an important modeling decision that involves analyzing the trade-off between accuracy and responsiveness on the one hand, and stability on the other. If the half-life is too long, the model gives undue weight to distant observations that have little to do with current market conditions. This leads to stable risk forecasts, but at the cost of reduced accuracy. By contrast, a short half-life makes the model more responsive, generally improving the risk forecasts, but at the cost of increased variability.

In general, the volatility half-life can be made considerably shorter than the correlation half-life. This is because sampling error for the diagonal elements has very little impact on the conditioning of the covariance matrix, whereas the effect is much larger for the off-diagonal elements. Of course, the volatility half-life cannot be made arbitrarily small; if it is too short, then sampling error can become so dominant that the risk forecasts are not only less stable, but also less accurate.

The parameters that are used to compute the USE4S and USE4L models are presented in **Table 4.1**. Note that the VRA (Volatility Regime Adjustment) half-life reported in Table 4.1 is described in Section 4.3.

Table 4.1
Factor covariance matrix parameters for the USE4 model. All values are reported in trading days.

Model	Factor Volatility Half-Life	Newey-West Volatility Lags	Factor Correlation Half-Life	Newey-West Correlation Lags	Factor VRA Half-Life
USE4S	84	5	504	2	42
USE4L	252	5	504	2	168

4.2 Optimization Bias Adjustment

In 1952, Markowitz (1952) established the mean-variance framework for constructing efficient portfolios. The required inputs include a set of expected asset returns and the asset covariance matrix, with the output being an optimal portfolio that has maximum expected return for a given level of risk. This framework forms the basis of modern portfolio theory and has led to a rich body of academic literature, as well as widespread adoption within the quantitative investment community.

Given the level of attention that the mean-variance framework has attracted, it is perhaps surprising that more than 40 years passed before it was empirically demonstrated (see Muller, 1993) that risk models have a systematic tendency to underpredict the risk of optimized portfolios. More recently, Shepard (2009) derived an analytic result for the magnitude of the bias. Under assumptions of normality, stationarity, and many assets (i.e., the large N limit), he found

$$\sigma_{true} \approx \frac{\sigma_{pred}}{1-\left(K/T\right)}, \tag{4.2}$$

where σ_{true} is the true volatility of the optimized portfolio, σ_{pred} is the predicted volatility from the risk model, K is the number of factors, and T is the effective number of observations used to compute the covariance matrix. The basic cause of the bias is estimation error; in essence, stocks that appear to be good hedges in-sample tend to be less effective at hedging risk out-of-sample.

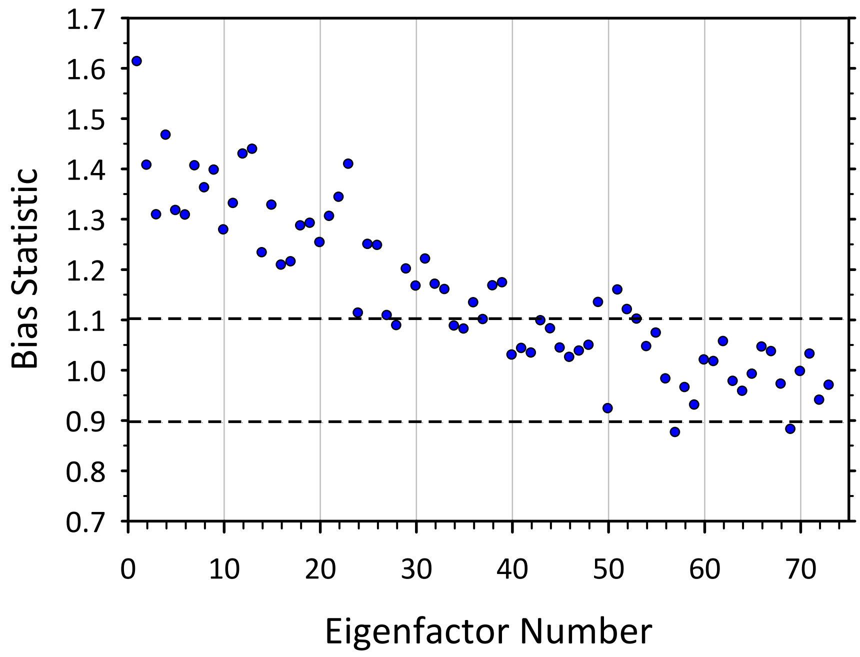
An important innovation in the USE4 model is to identify portfolios that reliably capture these biases, and to build corrections directly into the factor covariance matrix. This section provides a high-level overview of the methodology, with additional details given in Appendix B. Further analysis and discussion are provided by Menchero, Wang, and Orr (2011).

The underestimation of risk for optimized portfolios is closely linked with the concept of *eigenfactors*. Mathematically, eigenfactors are given by the eigenvectors of the factor covariance matrix; financially, they represent uncorrelated portfolios of pure factors.

In **Figure 4.1** we report optimization bias statistics computed using the factor covariance matrix given by Equation 4.1. As described in Appendix A, the bias statistic essentially represents the ratio of realized risk to predicted risk. The factor returns were taken from the USE4 model and the parameters used to construct F_{ij}^{0} were the same as for the USE4S model, as reported in Table 4.1. We standardize the eigenfactors to be unit norm and sort the eigenfactors from low volatility to high volatility. Figure 4.1 demonstrates a strong and direct relationship between the bias statistic of the eigenfactor and the eigenfactor number. More specifically, the lowest volatility eigenfactors have realized volatilities about 40 percent higher than their predicted volatilities and fall well outside the 95 percent confidence interval (indicated by the dashed horizontal lines). The larger eigenfactors, by contrast, fall mostly within the confidence interval.

Figure 4.1

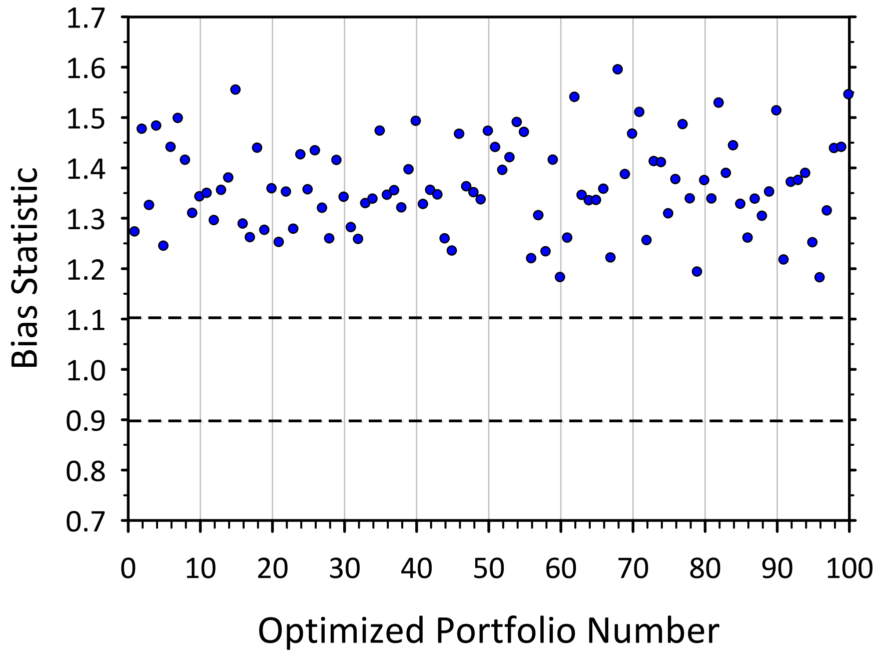
Bias statistics of eigenfactors using the unadjusted covariance matrix of Equation 4.1. Results were computed using the USE4S covariance matrix parameters reported in Table 4.1. The sample period comprised 191 months, from July 1995 through May 2011. The 95-percent confidence interval is indicated by the two dashed horizontal lines. While the large eigenfactors lie mostly within the confidence interval, the small eigenfactors fall well outside it.



While these results are intriguing, quantitative portfolio managers are more interested in optimized portfolios. To study these, we generated 100 random alpha signals at the factor level, drawn from a standard normal distribution. Each alpha signal was kept constant across time, but centered at the start of each month so that the cap-weighted alpha of the estimation universe was zero. We then constructed the minimum risk factor portfolio subject to the constraint that $\alpha = 1$. In **Figure 4.2**, we plot the bias statistics for the optimized factor portfolios. We observe the classic example of underprediction of risk of optimized portfolios, with all 100 bias statistics falling outside of the confidence interval.

Figure 4.2

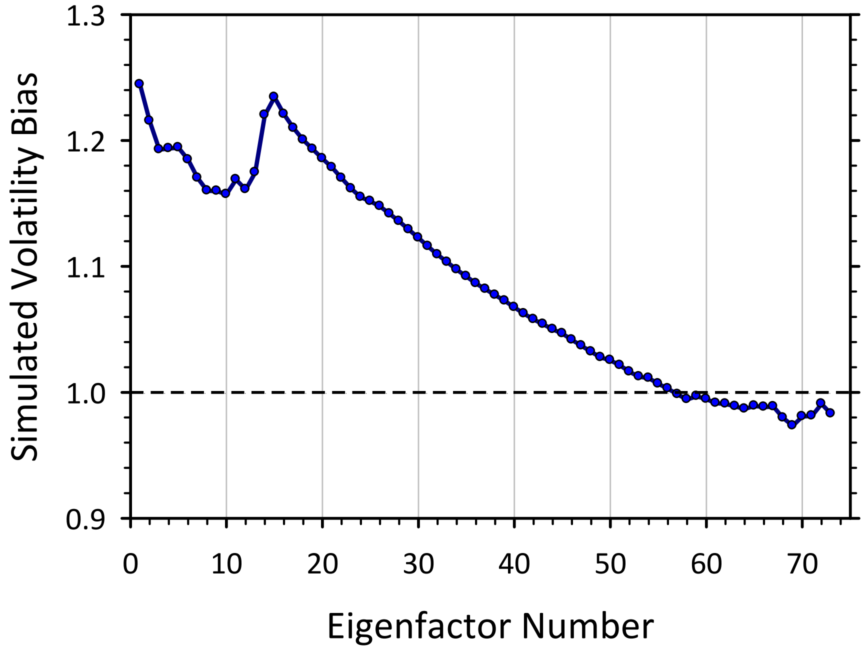
Bias statistics of 100 optimized factor portfolios using the unadjusted covariance matrix of Equation 4.1. Factor alphas were randomly drawn from a standard normal distribution, and held constant across time for each portfolio. Optimized factor portfolios were constructed at the start of each month by minimizing predicted factor risk subject to the $\alpha = 1$ constraint. Results were computed using the USE4S covariance matrix parameters reported in Table 4.1. The sample period comprised 191 months, from July 1995 through May 2011. The 95-percent confidence interval is indicated by the two dashed horizontal lines. All 100 observations fall outside the confidence interval, indicating underprediction of risk.



As described in Appendix B, we can estimate the magnitude of the empirical eigenfactor biases via Monte Carlo simulation. Although we cannot directly observe the true factor covariance matrix, we suppose for a moment that the sample factor covariance matrix represents the "truth." We then simulate histories of factor returns (using, for instance, the Cholesky decomposition) whose true distribution is drawn from the multivariate normal given by the sample factor covariance matrix. For each history, we compute a simulated factor covariance matrix. By comparing the predicted volatilities of the simulated eigenfactors with their "true" volatilities (i.e., using the sample factor covariance matrix), we determine the size of the simulated volatility bias.

In **Figure 4.3**, we plot the average simulated volatility bias over the entire sample period. Qualitatively, Figure 4.3 is very similar to Figure 4.1, indicating that the simulations are effective in capturing the biases. Quantitatively, however, we see that the realized biases are larger in magnitude than those observed in the simulations. Note that the simulations assume normality and stationarity, both of which are violated in practice. As described in Appendix B, slightly larger adjustments are required to fully remove the biases of the eigenfactors. We refer to this as the *scaled* adjustment, defined by Equation B8 of Appendix B. By contrast, Figure 4.3 was obtained using Equation B7; we refer to this as the *simulated* adjustment.

Figure 4.3
Mean simulated volatility bias, computed per Equation B7. Averages were computed from 30-Jun-1995 to 31-May-2011. Results were obtained using the USE4S parameters reported in Table 4.1.

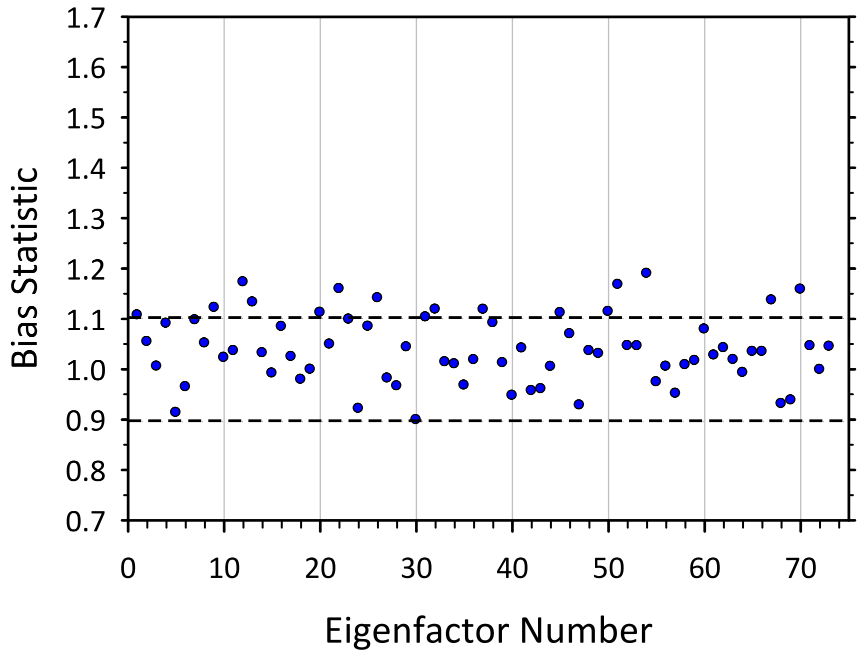


Since the sample factor covariance matrix and simulated factor covariance matrix share the same estimator, we assume that they also suffer from the same biases. We then adjust the eigenvariances of the sample factor covariance matrix. The final step, as described in Appendix B, is to rotate back to the original pure factor basis.

In **Figure 4.4** we show the bias statistics of the eigenfactors using the eigen-adjusted covariance matrix. The results were computed using the scaled adjustment of Equation B8. We see that most observations now fall squarely within the confidence interval, indicating unbiased risk forecasts. Furthermore, the relationship between the bias statistic and the eigenfactor number is now essentially flat, in sharp contrast to Figure 4.1.

Figure 4.4

Bias statistics of eigenfactors using the eigen-adjusted covariance matrix of Equation B10. Results were computed using the USE4S covariance matrix parameters reported in Table 4.1. The sample period comprised 191 months, from July 1995 through May 2011. Most observations fall within the 95-percent confidence interval.

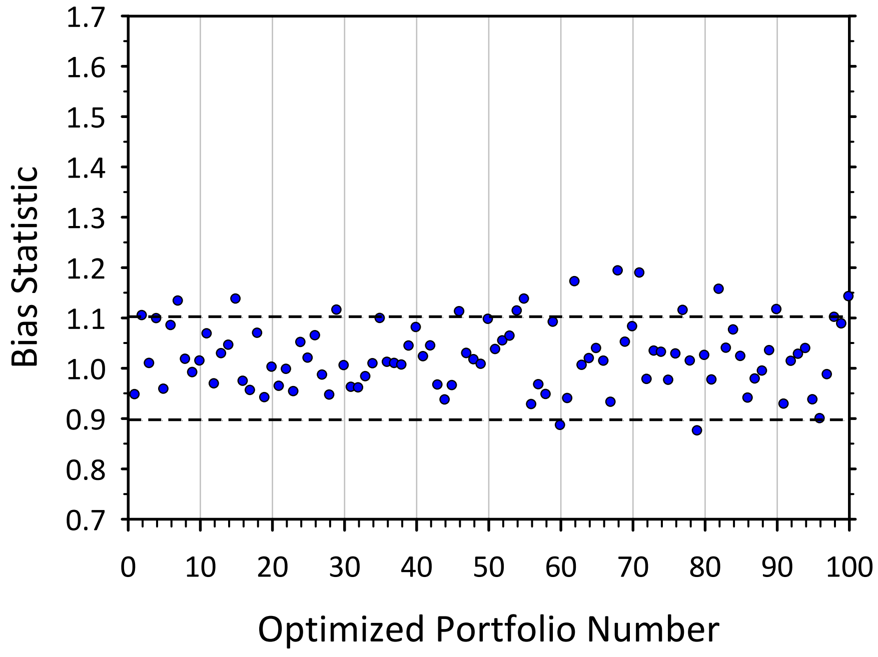


Again, while this result is encouraging, the primary interest of quantitative managers is not in eigenfactors, but rather in optimized portfolios. In **Figure 4.5** we report the bias statistics of the portfolios optimized using the same alpha signals as in Figure 4.2. The results were computed using the scaled adjustment of Equation B8. This demonstrates that the eigen-adjusted covariance matrix has essentially removed the biases of the optimized factor portfolios as well.

As discussed by Menchero, Wang, and Orr (2011), the Optimization Bias Adjustment may induce small biases for the pure factors. In particular, the factors with the smallest volatilities (i.e., the style factors) tend to have their volatilities adjusted slightly upward. To mitigate this effect, in the USE4 Model we employ the milder simulated adjustment defined by Equation B7. We find that this approach significantly reduces forecasting biases for optimized portfolios, while also ensuring that the bias statistics for the style factors fall mostly within the 95-percent confidence interval as expected.

Figure 4.5

Bias statistics of 100 optimized factor portfolios using the eigen-adjusted covariance matrix of Equation B10. Factor alphas were randomly drawn from a standard normal distribution, and held constant across time for each portfolio. Optimized factor portfolios were constructed at the start of each month by minimizing predicted factor risk subject to the $\alpha = 1$ constraint. Results were computed using the USE4S covariance matrix parameters reported in Table 4.1. The sample period comprised 191 months, from July 1995 through May 2011. Most observations fall within the 95-percent confidence interval.



4.3 Volatility Regime Adjustment

In the traditional approach to estimating factor volatilities, each factor is considered in isolation. That is, volatility estimates for a particular factor are based exclusively on the time series of returns to the individual factor.

Another significant innovation in the USE4 model is to use cross-sectional observations to produce timelier and more accurate forecasts of factor volatility. The basic intuition behind this method is that the returns and risk forecasts of the other factors provide additional information for refining the volatility estimates. In particular, if cross-sectional observations show that the model is consistently overforecasting or underforecasting risk over a recent period, then the volatilities of all factors can be collectively adjusted to remove this bias.

More specifically, let f_{kt} be the return to factor k on day t, and let σ_{kt} be the one-day volatility forecast for the factor at the start of the day. The standardized return of the factor is given by the ratio $\left(f_{kt}/\sigma_{kt}\right)$, and should have standard deviation close to 1 if the risk forecasts are accurate. Normally, as described in Appendix A, we compute the *time-series* standard deviation to investigate whether an individual factor is unbiased across time.

Alternatively, we can compute the *cross-sectional* standard deviation to investigate whether the factor volatility forecasts are collectively unbiased at a given point in time. We define the factor cross-sectional bias statistic B_t^F on day t as

$$B_t^F = \sqrt{\frac{1}{K}\sum_k \left(\frac{f_{kt}}{\sigma_{kt}}\right)^2} , \tag{4.3}$$

where K is the total number of factors. This quantity represents an *instantaneous* measure of factor risk bias. For instance, if the risk forecasts were too small on a particular day, then $B_t^F > 1$. By observing the cross-sectional bias statistics over time, we can determine the extent to which volatility forecasts should be adjusted to remove these biases.

We define the *factor volatility multiplier* λ_F as an exponentially weighted average

$$\lambda_F = \sqrt{\sum_t \left(B_t^F\right)^2 w_t} , \tag{4.4}$$

where w_t is an exponential weight with Volatility Regime Adjustment half-life τ_{VRA}^F. This parameter serves as the primary determinant of model responsiveness for factor risk. The Volatility Regime Adjustment forecasts are given by

$$\tilde{\sigma} \; [illegible] \; . \tag{4.5}$$

This is equivalent to multiplying the entire factor covariance matrix by a single number, λ_F^2. As a result, the Volatility Regime Adjustment has no effect on factor correlations.

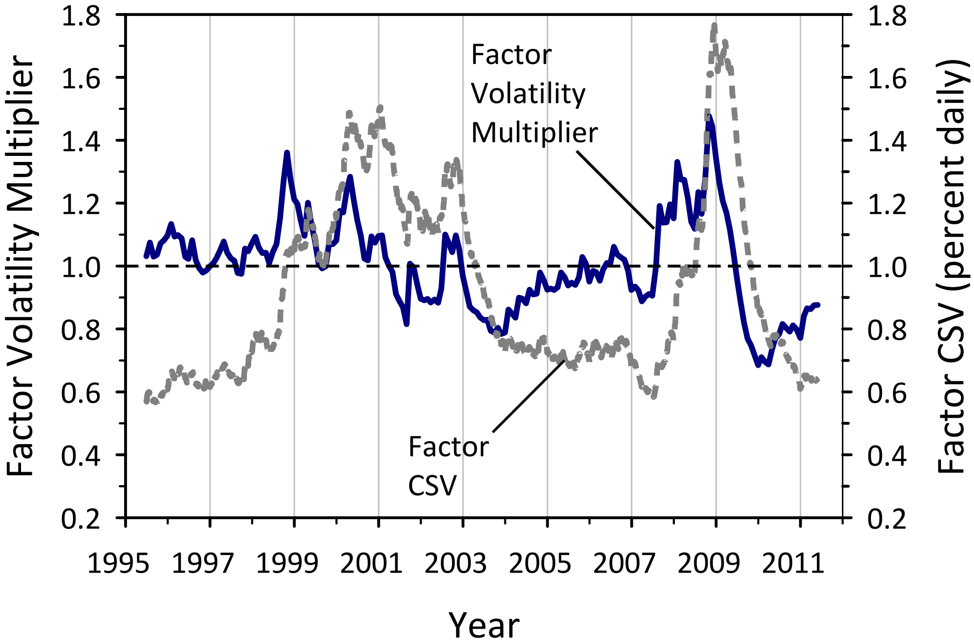
We define the factor cross-sectional volatility (CSV) on day t as

$$CSV_t^F = \sqrt{\frac{1}{K}\sum_k f_{kt}^2} . \tag{4.6}$$

In **Figure 4.6** we plot factor CSV, together with the factor volatility multiplier, λ_F. Results were computed using the USE4S model parameters, shown in Table 4.1.

Figure 4.6

Factor cross-sectional volatility (CSV) and factor volatility multiplier. Factor CSV levels are reported in daily percentage and smoothed using a 42-day half-life.



It is instructive to analyze the interplay between factor CSV levels and the factor volatility multiplier. From 1995-1998, we see that factor CSV gradually rose from about 60 bps to roughly 80 bps per day. During this same time, the volatility multiplier λ_F was only slightly greater than 1, indicating that the traditional time-series approach did quite well predicting factor volatility, needing only a slight upward adjustment.

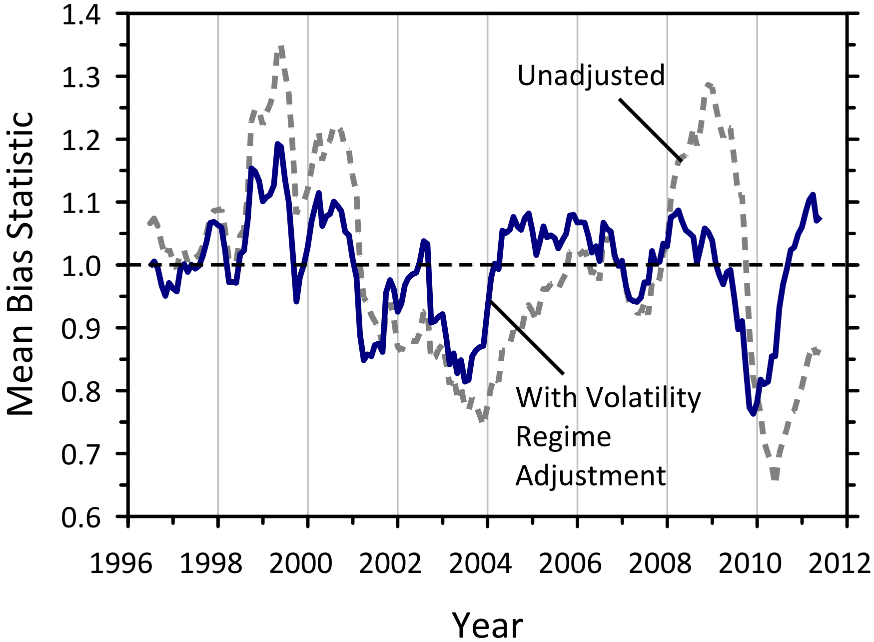
The Russian Default, in the summer of 1998, marks the first major crisis of the sample period. Over a very short window, we see that factor CSV spiked from about 70 bps to above 100 bps per day. The Volatility Regime Adjustment quickly detected the increased volatility levels, with λ_F rapidly adjusting to a value of nearly 1.4. Over the subsequent year, factor CSV stabilized in the range of 100-110 bps daily. During this period, the traditional time-series approach adjusted to the new volatility levels and λ_F decayed back to 1 by the fall of 1999.

Another interesting period to consider is 2005-2007, during which the factor CSV was quite stable at about 70 bps per day. During stable periods, the traditional approach works well, and again we see that the volatility multiplier was very close to 1. However, as the financial crisis unfolded, from August 2007 until the end of 2008, factor CSV increased dramatically to nearly 180 bps per day. During the entire financial crisis, the volatility multiplier was significantly greater than 1, reaching a peak of about 1.45 in late 2008. During 2009, as the financial crisis subsided, factor CSV plummeted back to pre-crisis levels. Again, the Volatility Regime Adjustment detected this sudden drop in volatility and λ_F fell to 0.7 by the end of 2009, thereby reducing the volatility forecast by 30 percent relative to the traditional approach.

We have seen that the factor volatility multiplier responds quickly and intuitively to market shocks in volatility. To evaluate the performance gain of the Volatility Regime Adjustment, we compute the mean of the rolling 12-month bias statistics for the factors over the sample period. The results are plotted in **Figure 4.7**, together with the corresponding results computed without the Volatility Regime Adjustment. Over the vast majority of the sample period, the Volatility Regime Adjustment produced mean bias statistics that were closer to 1. During times of extreme market stress, the outperformance was even greater. For instance, during the most severe part of the financial crisis in 2008, the Volatility Regime Adjustment provided mean bias statistics very close to 1. By contrast, without the methodology we significantly underpredicted risk. Similarly, in the wake of the financial crisis, the Volatility Regime Adjustment again produced bias statistics much closer to 1, whereas without the methodology we significantly overforecasted risk for an extended period of time.

Figure 4.7

Mean rolling 12-month bias statistics for USE4 factors, with and without the Volatility Regime Adjustment. Results were computed using the USE4S covariance matrix parameters reported in Table 4.1. The Volatility Regime Adjustment led to mean bias statistics closer to 1 over most of the sample period.



In summary, without the Volatility Regime Adjustment we observed the classic signature of underprediction of risk entering a crisis, followed by overprediction afterwards. The Volatility Regime Adjustment greatly reduced these non-stationarity biases.

5. Specific Risk

5.1 Established Methods

Accurate prediction of specific volatility constitutes the third essential pillar for producing a high-quality risk model. 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.

The specific returns are obtained from daily cross-sectional regressions,

$$r_{nt} = \sum_{k} X_{nk} f_{kt} + u_{nt} \ , \tag{5.1}$$

where r_{nt} is the return of stock n on day t, X_{nk} is the stock exposure to factor k, f_{kt} is factor return, and u_{nt} is the specific return of the stock. It is understood that although the factor exposure matrix varies slowly with time, the t subscript is suppressed for notational simplicity.

Asset-level specific risk is computed directly from the time series of specific returns,

$$\sigma_n^{TS} = C_n^{NW} \left[\sum_t w_t \left(u_{nt} - \bar{u}_n \right)^2 \right] , \tag{5.2}$$

where w_t is an exponential weight characterized by the specific volatility half-life parameter τ_σ^S, and C_n^{NW} is the Newey-West serial correlation adjustment for stock n. To estimate C_n^{NW}, we use an auto-correlation half-life parameter τ_ρ^S and allow specific returns to be correlated over a number of lags L_ρ^S.

While the time-series approach offers the advantage of estimating specific volatility directly, one challenge is that not all stocks are well-suited to such an estimation methodology. For instance, recently issued IPOs lack sufficient history to reliably estimate their specific risk. Similarly, thinly traded stocks often exhibit poor return characteristics such as consecutive days of zero returns followed by large returns on days when the stock eventually trades. Earnings announcements can also lead to extreme jumps in specific returns that make the time-series specific-risk forecasts unreliable.

For such stocks, the USE4 model follows the EUE3 treatment and blends the asset-level forecast with the fit value from a structural model. More specifically, we define a blending parameter γ_n that equals 1 if the stock has sufficiently few missing returns and the specific returns are not excessively fat tailed. For stocks that strongly violate either of these conditions, γ_n is set to zero. In general, γ_n varies from 0 to 1, although most stocks in USE4 have $\gamma_n = 1$.

To estimate the structural model, we take all stocks with $\gamma_n = 1$ and regress the logarithm of specific volatility forecasts against the factors in the model,

$$\ln\left(\sigma_n^{TS}\right) = \sum_k X_{nk} b_k + \varepsilon_n \ , \tag{5.3}$$

where b_k is the regression coefficient and ε_n is the residual. We perform this cross-sectional regression daily. The structural specific risk forecasts are then given by

$$\sigma_n^{STR} = E_0 \cdot \exp\left(\sum_k X_{nk} b_k \right) \ , \tag{5.4}$$

where E_0 is a number slightly greater than 1 that is required to remove the small bias arising from exponentiation of the residuals. The intuition behind the structural model is that stocks with similar characteristics are also likely to have similar specific volatilities.

The blended specific risk forecast is given by

$$\hat{\sigma}_n = \gamma_n \sigma_n^{TS} + \left(1-\gamma_n\right)\sigma_n^{STR} \ . \tag{5.5}$$

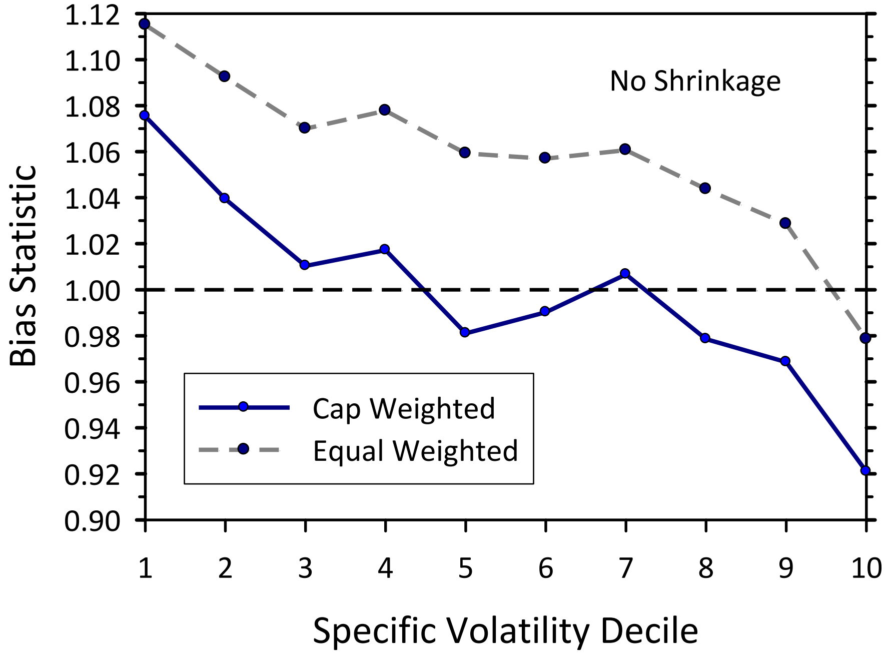
As most stocks in USE4 have $\gamma_n = 1$, they receive the full weight from the time-series component.

5.2 Bayesian Shrinkage

One potential problem with using a pure time-series approach is that specific volatilities may not fully persist out-of-sample. In particular, stocks with either extremely low or extremely high specific volatility forecasts tend to revert to the mean. To illustrate this effect, every month we segment stocks into deciles based on their specific risk forecasts as given by Equation 5.5. We then compute the bias statistics, as described in Appendix A, for each decile over the entire sample period. The results are plotted in **Figure 5.1**, and show that the bias statistics depend strongly on the volatility decile. More specifically, Equation 5.5 tends to underpredict the low-volatility stocks and overpredict those with high volatility. These biases can be problematic for portfolio construction purposes, since stocks that appear to have low specific volatility in-sample may exhibit higher volatility out-of-sample.

Figure 5.1

Bias statistics of specific returns grouped by specific risk deciles (sorted from low to high), without Bayesian shrinkage adjustments. Specific risk was computed per Equation 5.5. Both cap-weighted and equal-weighted results are reported. The sample period comprised 191 months, from July 1995 through May 2011. Without Bayesian shrinkage, note the tendency to overpredict high-volatility stocks and underpredict low-volatility stocks.



To remove this bias, we shrink our estimates toward the cap-weighted mean specific volatility for the size decile s_n to which the stock belongs. More precisely, the shrunk estimate σ_n^{SH} is given by

$$\sigma_n^{SH} = v_n \bar{\sigma}(s_n) + \left(1 - v_n\right)\hat{\sigma}_n \ , \tag{5.6}$$

where $\hat{\sigma}_n$ is the blended forecast (given by Equation 5.5) and v_n is the shrinkage intensity that determines the weight given to the Bayesian prior, also known as the shrinkage target,

$$\bar{\sigma}(s_n) = \sum_{n \in s_n} w_n \hat{\sigma}_n \ , \tag{5.7}$$

where w_n is the capitalization weight of stock n with respect to the size decile. The shrinkage intensity is given by

$$v_n = \frac{q\left|\hat{\sigma}_n - \bar{\sigma}(s_n)\right|}{\Delta_\sigma(s_n) + q\left|\hat{\sigma}_n - \bar{\sigma}(s_n)\right|} \ , \tag{5.8}$$

where q is an empirically determined shrinkage parameter and

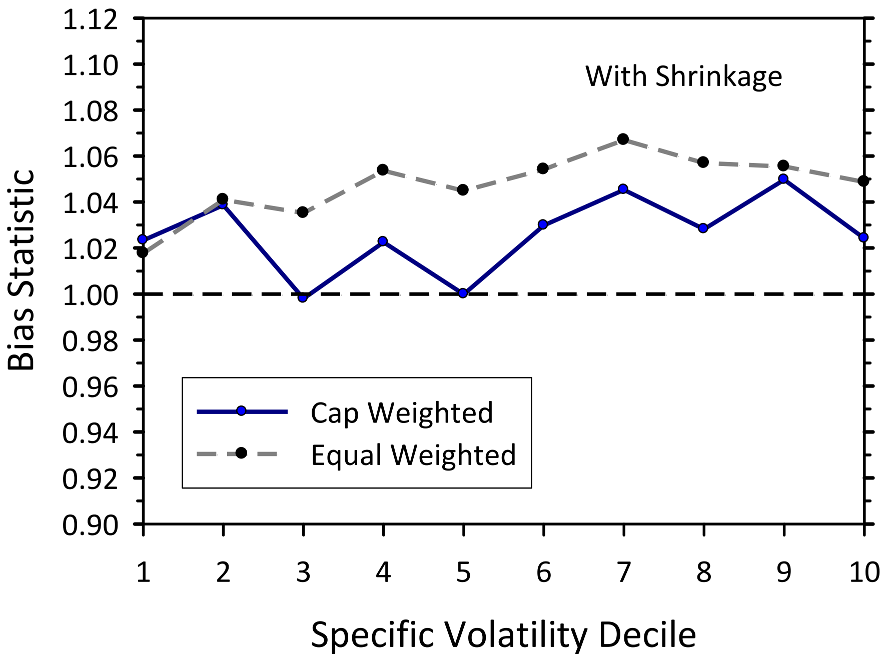
$$\Delta_{\sigma}(s_n) = \sqrt{\frac{1}{N(s_n)} \sum_{n \in s_n} \left(\hat{\sigma}_n - \bar{\sigma}(s_n) \right)^2} \tag{5.9}$$

is the standard deviation of specific risk forecasts within the size decile. The intuition behind this approach is straightforward: the more $\hat{\sigma}_n$ deviates from the mean, the greater the weight we assign to the Bayesian prior $\bar{\sigma}(s_n)$.

In **Figure 5.2** we show the bias statistics for the same specific risk deciles, now after applying the Bayesian shrinkage technique. We see that the bias statistics are closer to the ideal value of 1. Furthermore, the relationship is essentially flat, so that we do not observe significant biases associated with specific volatility levels.

Figure 5.2

Bias statistics of specific returns grouped by specific risk deciles (sorted from low to high), with Bayesian shrinkage adjustments. Specific risk was computed using Equation 5.6. Both cap-weighted and equal-weighted results are reported. The sample period comprised 191 months, from July 1995 through May 2011. The Bayesian shrinkage technique largely removes the biases associated with either extreme.



5.3 Volatility Regime Adjustment

The specific risk model follows a similar approach as for the factor covariance matrix and adjusts the volatility forecasts based on cross-sectional observations. Let u_{nt} be the specific return to stock n on day t, and let σ_{nt} be the one-day volatility forecast at the start of the day, obtained from Equation 5.6

after removing adjustments for serial correlation. We define the specific cross-sectional bias statistic on day t as the cap-weighted average

$$B_t^S = \sqrt{\sum_n w_{nt} \left(\frac{u_{nt}}{\sigma_{nt}} \right)^2} . \tag{5.10}$$

This quantity represents an *instantaneous* measure of specific risk bias. For instance, if the model underpredicts specific risk on day t, then $B_t^S > 1$.

We define the *specific volatility multiplier* by

$$\lambda_S = \sqrt{\sum_t \left(B_t^S \right)^2 w_t} , \tag{5.11}$$

where w_t is an exponential weight with a Volatility Regime Adjustment half-life parameter τ_{VRA}^S. In order to ensure that the specific risk model has consistent responsiveness with the factor risk model, we select $\tau_{VRA}^S = \tau_{VRA}^F$. The Volatility Regime Adjustment forecasts are then given by

$$\tilde{c}_{n} \qquad {}_{S} \quad {}^{H}_{n} , \tag{5.12}$$

where σ_n^{SH} is the specific risk forecast after Bayesian shrinkage (i.e., Equation 5.6).

We define the specific cross-sectional volatility (CSV) as

$$CSV_t^S = \sqrt{\sum_n w_{nt} u_{nt}^2} . \tag{5.13}$$

In **Figure 5.3** we plot the specific CSV, together with the specific volatility multiplier λ_S. Results were computed using the USE4S model parameters, reported in **Table 5.1**.

Figure 5.3

Specific cross-sectional volatility (CSV) and specific volatility multiplier. Specific CSV levels are reported in daily percentage and smoothed using a 42-day half-life.

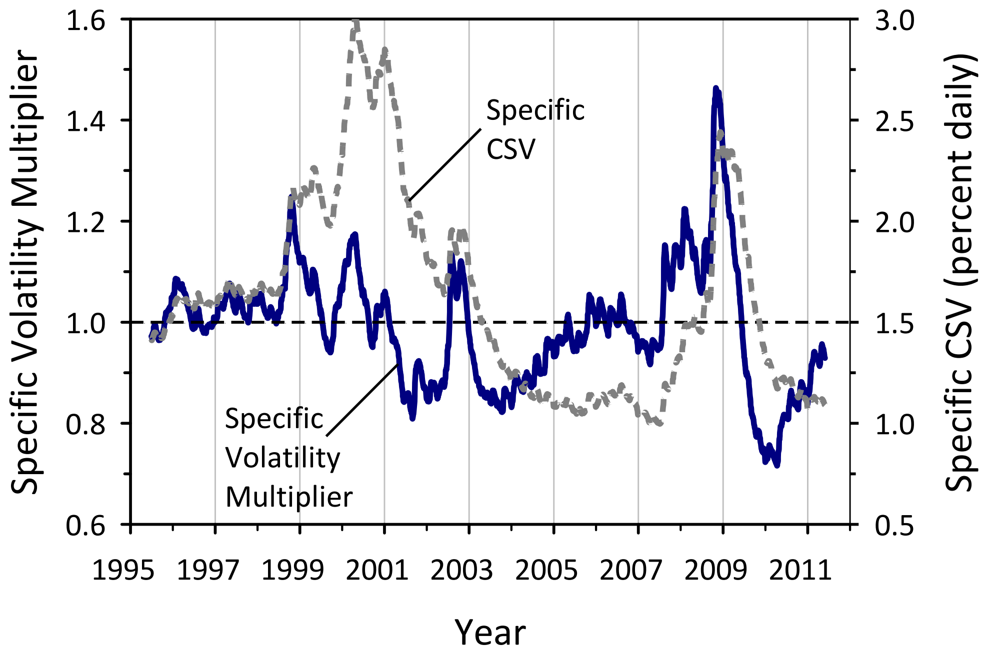


Table 5.1

Specific risk parameters for the USE4 model. Except for the dimensionless shrinkage parameter q, all values are reported in trading days.

Model	Specific Volatility Half-Life	Newey-West Auto-Corr. Lags	Newey-West Auto-Corr. Half-Life	Bayesian Shrinkage Parameter q	Specific VRA Half-Life
USE4S	84	5	252	0.1	42
USE4L	252	5	252	0.1	168

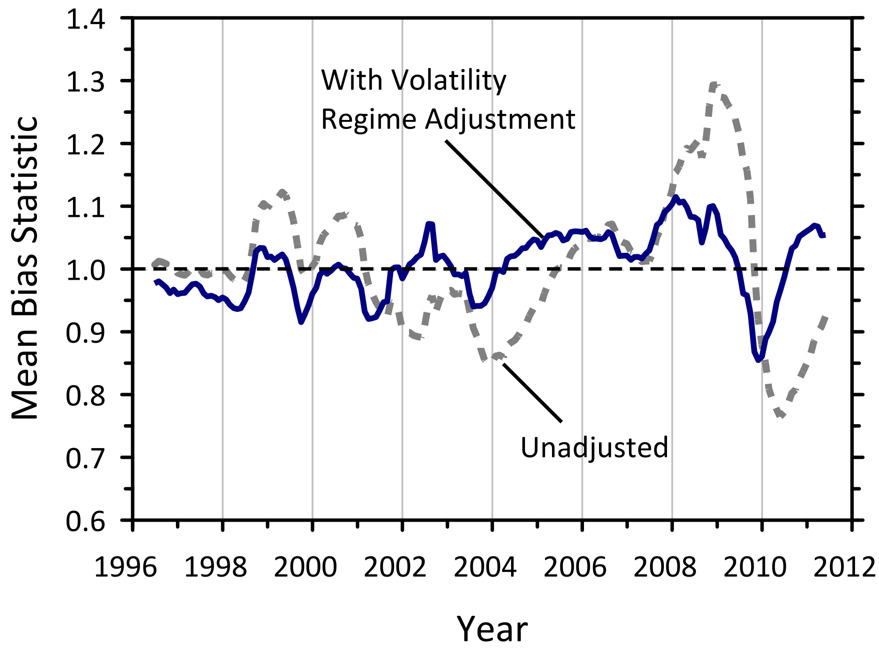
It is interesting to compare Figure 5.3 for specific risk with the corresponding result for factor risk, shown in Figure 4.6. Qualitatively, the plots are remarkably similar. For instance, specific CSV and factor CSV each exhibited major peaks during the Internet Bubble period of 2000 and the financial crisis of 2008. The differences, rather, are primarily in the details. For instance, factor CSV reached a maximum in 2008 whereas specific CSV reached an absolute peak in 2000.

The volatility multipliers also behave similarly. For example, during the stable period 2005-2007 immediately prior to the financial crisis, the specific volatility multiplier λ_S was very close to 1. At the height of the financial crisis, λ_S reached a peak of about 1.45 in late 2008, and hit a minimum near 0.75 at the end of 2009.

The specific volatility multiplier responds intuitively to market shocks. To demonstrate the performance gain of the Volatility Regime Adjustment, we compute the mean of the rolling 12-month bias statistics of the specific returns. The results are plotted in **Figure 5.4**, together with the corresponding results computed without the Volatility Regime Adjustment.

Figure 5.4

Mean rolling 12-month bias statistics for specific returns, with and without the Volatility Regime Adjustment. Results were equal weighted across the USE4 estimation universe and computed using the USE4S specific risk parameters reported in Table 5.1. The Volatility Regime Adjustment led to mean bias statistics closer to 1 over most of the sample period.



Similar to Figure 4.7, over most of the sample period we see a significant benefit in the accuracy of the risk forecasts, with the performance gain becoming even greater during times of financial stress. For instance, during the financial crisis, the Volatility Regime Adjustment produced bias statistics very close to 1, whereas without the methodology we observed the classic signature of underprediction during the crisis, followed by an extended period of overprediction after the crisis.

6. Conclusion

MSCI has a long and proud tradition in developing rigorous, innovative, and pioneering solutions for the investment management community. As the industry leader for portfolio construction and equity analytics, we are committed to the continual improvement of the models, tools, and products that we provide to our clients. We believe in the value of high-quality research and it remains a cornerstone of our approach to investment problem solving.

The new Barra US Equity Model, USE4, is the result of these research efforts in combination with extensive client consultations. The USE4 Model incorporates many methodological innovations designed to address long-standing problems in risk modeling. For instance, the Optimization Bias Adjustment addresses the issue of underestimation of risk of optimized portfolios, and leads to better conditioning of the covariance matrix. The Volatility Regime Adjustment calibrates volatility levels to current market levels and represents an important determinant of risk, especially during times of market turmoil. The introduction of a country factor leads to greater insight into the sources of portfolio risk, as well as providing timelier forecasts of industry correlations. Another enhancement is the use of a Bayesian adjustment technique to reduce biases in specific risk forecasts.

During consultations, our clients stressed the importance of factor structure. In particular, they strongly endorsed the value of our unique multiple industry exposure methodology. In response to strong client demand, the USE4 Model continues to provide multiple industry exposures. Furthermore, extensive research was performed to customize the industry structure to the local market. The USE4 Model also offers an enhanced set of style factors that have undergone comprehensive empirical and statistical testing. These rigorous procedures assure a robust factor structure with a high degree of explanatory power.

Building high-quality risk models is a challenging and fascinating endeavor. While the USE4 model marks a significant milestone in our research and development efforts, we remain committed to continuing our search for new and better ways to model risk.

Appendix A: Review of Bias Statistics

A1. 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,

$$b_{nt} = \frac{R_{nt}}{\sigma_{nt}}, \tag{A1}$$

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

$$B_n = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(b_{nt} - \bar{b}_n\right)^2}, \tag{A2}$$

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,

$$B_n \in \left[1 - \sqrt{2/T}, 1 + \sqrt{2/T}\right]. \tag{A3}$$

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 A1, 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.

A2. 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 A2. 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,

$$B_n^{\tau} = \sqrt{\frac{1}{11}\sum_{t=\tau}^{\tau+11}(b_{nt} - \bar{b}_n)^2} \text{ ,} \tag{A4}$$

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,

$$\bar{B}^{\tau} = \frac{1}{N}\sum_n B_n^{\tau} \text{ .} \tag{A5}$$

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

$$\mathrm{MRAD}^{\tau} = \frac{1}{N}\sum_n \left| B_n^{\tau} - 1 \right| . \tag{A6}$$

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 A2 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 A1

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.

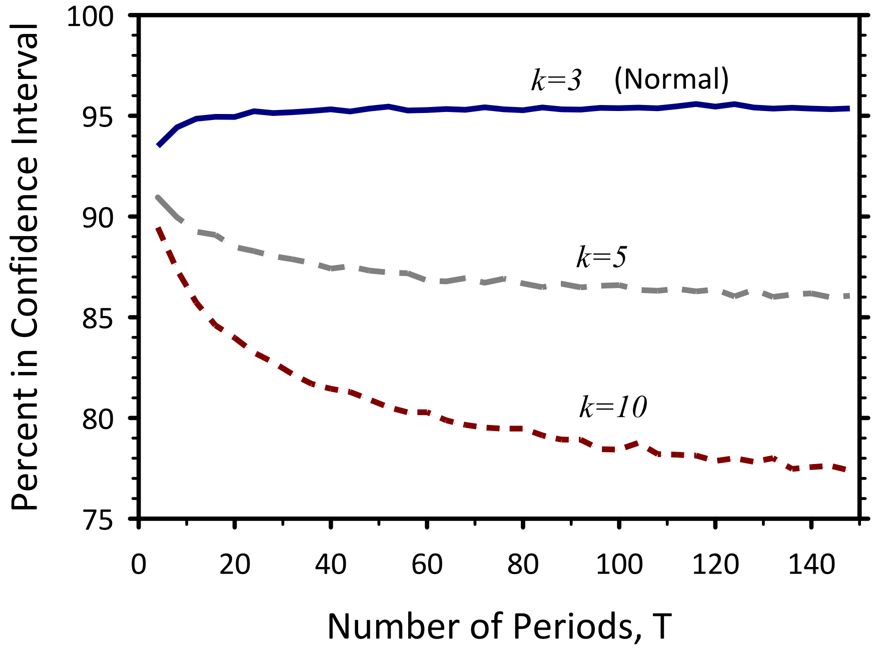
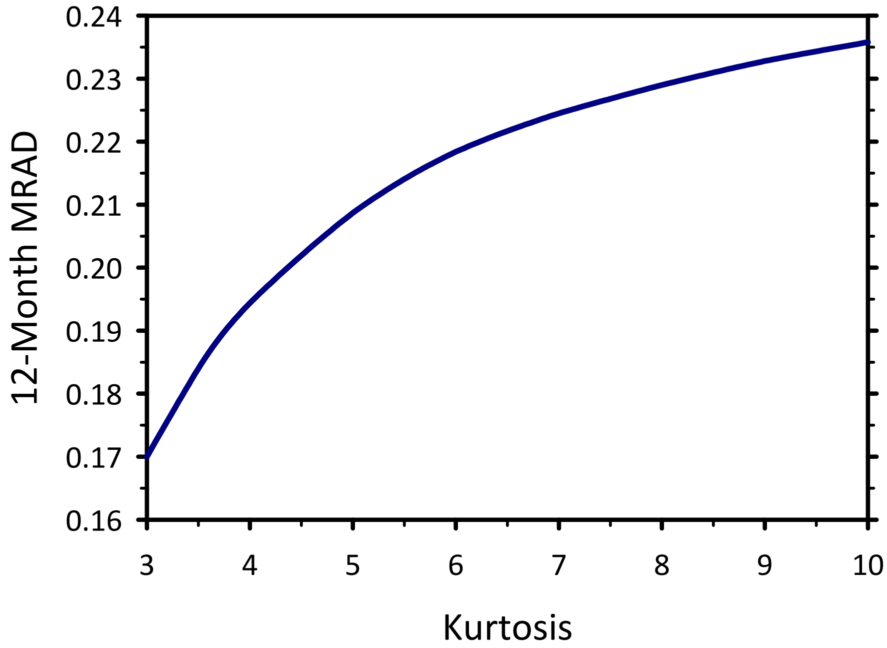


Figure A2

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



Appendix B. Optimization Bias Adjustment

Let $\mathbf{F}_0$ denote the $K\text{x}K$ sample factor covariance matrix (FCM),

$$\mathbf{F}_0 = \text{cov}(\mathbf{f},\mathbf{f}) \ , \tag{B1}$$

where $\mathbf{f}$ is the $K\text{x}T$ matrix of realized factor returns, K is the number of factors and T is the number of periods. The covariance matrix estimator is described in Section 4.1.

The sample FCM can be expressed in diagonal form as

$$\mathbf{D}_0 = \mathbf{U}_0'\mathbf{F}_0\mathbf{U}_0 \ , \tag{B2}$$

where $\mathbf{U}_0$ is the $K\text{x}K$ rotation matrix whose columns are given by the eigenvectors of $\mathbf{F}_0$. The j^{th} element of the k^{th} column of $\mathbf{U}_0$ gives the weight of pure factor j in eigenfactor k. The predicted variances of the eigenfactors are given by the diagonal elements of $\mathbf{D}_0$. The fact that $\mathbf{D}_0$ is diagonal indicates that the eigenfactors are mutually uncorrelated.

Although the true FCM is unobservable, we suppose for simulation purposes that the sample FCM $\mathbf{F}_0$ governs the "true" return-generating process. We generate a set of factor returns for simulation m as

$$\mathbf{f}_m = \mathbf{U}_0\mathbf{b}_m \ , \tag{B3}$$

where $\mathbf{b}_m$ is a $K\text{x}T$ matrix of simulated eigenfactor returns. The elements of row k of $\mathbf{b}_m$ are drawn from a random normal distribution with mean zero and variance given by the diagonal element $D_0(k)$ of matrix $\mathbf{D}_0$. It can be easily verified that the simulated returns in Equation B3 have a true FCM given by $\mathbf{F}_0$. Due to sampling error, however, the *estimated* FCM

$$\mathbf{F}_m = \text{cov}(\mathbf{f}_m,\mathbf{f}_m) \ , \tag{B4}$$

will differ from the true FCM $\mathbf{F}_0$. Nevertheless, $\mathbf{F}_m$ is unbiased in the sense that $E[\mathbf{F}_m] = \mathbf{F}_0$. We diagonalize the simulated FCM

$$\mathbf{D}_m = \mathbf{U}_m'\mathbf{F}_m\mathbf{U}_m \ , \tag{B5}$$

where $\mathbf{U}_m$ denotes the simulated eigenfactors with estimated variances given by the diagonal elements of $\mathbf{D}_m$, i.e., $D_m(k)$.

Since we know the true distribution that governs the simulated factor returns, we can compute the true FCM of the simulated eigenfactors,

$$\tilde{\mathrm{L}}_m \quad -_{m}-_{0}\mathbf{U}_m \ . \tag{B6}$$

Note that since $\mathbf{U}_m$ is not composed of the "true" eigenfactors, the matrix $\tilde{\Lambda}_m$ is not diagonal. Nevertheless, our current focus is on the diagonal elements of the matrix.

As shown in Figure 4.1, the predicted volatilities of the empirical eigenfactors are biased. We compute the *simulated* volatility biases according to

$$v(k) = \sqrt{\frac{1}{M}\sum_m \frac{\tilde{\Lambda}_m(k)}{D_m(k)}} \ , \tag{B7}$$

where M is the total number of simulations. The simulated volatility bias is computed daily, and the average over the entire sample period is plotted in Figure 4.3. As shown by Menchero, Wang, and Orr (2011), the simulated volatility bias is very stable over time.

Qualitatively, Figure 4.3 is in good agreement with Figure 4.1. Closer inspection, however, reveals that the simulations do not capture the full extent of the forecasting bias. Our simulations assume both normality and stationarity. Real financial data, of course, violate both of these assumptions. In practice, therefore, additional scaling is required to fully remove the biases of the eigenfactors.

To obtain the scaled version of $v(k)$, and hence fully remove the eigenfactor biases, two additional steps are required. First, we fit the simulated volatility bias $v(k)$ every day using a parabola. When estimating the fit, we assign zero weight to the first 15 eigenfactors, which has the effect of removing the sawtooth structure on the left side of Figure 4.3. The second step is to scale the fit values in proportion to their deviation from 1,

$$v_S(k) = a\left[v_P(k) - 1\right] + 1 \ , \tag{B8}$$

where $v_S(k)$ is the scaled value, $v_P(k)$ is obtained from the parabolic fit, and $a = 1.4$ is an empirically determined constant. Empirically, we find that this procedure is effective at removing the eigenfactor biases, as indicated in Figure 4.4. Furthermore, we confirmed that this procedure is robust over different sub-periods.

As discussed by Menchero, Wang, and Orr (2011), this procedure may induce small biases at the individual pure factor level. To mitigate this effect, in the USE4 Model we adopt the milder simulated adjustment given by Equation B7. We find that this approach significantly reduces the forecasting biases for optimized portfolios, while also ensuring that the bias statistics of the individual factors fall mostly within the 95-percent confidence interval as expected.

We now assume that the sample FCM $\mathbf{F}_0$, which uses the same covariance estimator as the simulated FCM $\mathbf{F}_m$, also suffers from the same biases. Let $\tilde{\Lambda}_0$ denote the diagonal FCM whose eigenvariances have been adjusted

$$\tilde{\Lambda}_0 \quad \cdot \; -_0 \ , \tag{B9}$$

where $\mathbf{v}^2$ is a diagonal matrix whose elements are given by $v^2(k)$. The FCM in Equation B9 is now rotated from the diagonal basis to the pure factor basis using the sample eigenfactors. That is,

$$\tilde{\mathbf{F}}_0 = \mathbf{U}_0 \tilde{\mathbf{D}}_0 \mathbf{U}_0', \tag{B10}$$

where $\tilde{\mathbf{F}}_0$ denotes the eigen-adjusted factor covariance matrix.

REFERENCES

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

Dempster, A.P., N.M. Laird, and D.B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society*, Vol. 39, No. 1: 1-38.

Grinold, R., and R. Kahn. 2000. *Active Portfolio Management*. New York: McGraw-Hill.

Grinold, R., A. Rudd, and D. Stefek, Fall 1989. "Global Factors: Fact or Fiction." *Journal of Portfolio Management*, vol. 16, no. 1: 79-88.

MSCI, Inc. 2010. "Global Style Factors." *MSCI Research Insight*.

Markowitz, Harry. 1952. "Portfolio Selection." *Journal of Finance*, vol. 7: 77-91.

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, 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.

Newey, W., and K. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, Vol. 55, No. 3: 703-708.

Rosenberg, B. March 1974. "Extra-Market Components of Covariance among Security Prices," *Journal of Financial and Quantitative Analysis*, vol. 9: 263-274.

Rosenberg, B., and V. Marathe, May 1975. "Tests of Capital Asset Pricing Hypotheses," *Research Program in Finance*, Working Paper no. 32.

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.

Shepard, Peter. 2011. "Barra Integrated Model." *Research Notes*.

Tobin, J. February 1958. "Liquidity Preference as a Behavior Toward Risk," *Review of Economic Studies*, vol. 67: 65-86.

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

Notice and Disclaimer

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.