A Bloomberg Professional Services Offering

Bloomberg Terminal

Portfolio Risk and Analytics

MAC3 Global Equity Risk Model

Methodology Notes

Jose Menchero
Ercument Cahan
Lei Ji
Peng Li
Marty O'Brien

October 2020

Bloomberg

Risk analysis has transitioned from being a "check-the-box" portfolio management function into an essential tool for risk management, portfolio construction, and performance enhancement. Identifying the sources of risk and return—and controlling the portfolio exposure to them—enables managers to implement their views cleanly and avoid taking unwanted risks.

Risk models are called upon to serve the dual responsibilities of risk management and portfolio construction across a wide array of investment opportunities and geographies. Factor models provide a parsimonious framework to capture the underlying drivers of portfolio risk and return and understand their interrelationships. However, the large number of factors required to capture the nuances of different investment strategies across different markets makes accurate estimation of their joint properties challenging. This challenge is further complicated by the fact that factor properties may change over time—sometimes abruptly when a market crisis occurs.

The Bloomberg MAC3 GRM suite of equity risk models introduced in this paper incorporates a number of advanced techniques that result in superior performance across portfolio types, geographies and investment styles. Drawing on deep expertise and novel research, the MAC3 Model produces accurate risk forecasts to satisfy its risk management responsibilities, while also providing a reliable tool for advanced portfolio construction techniques, including optimization. The MAC3 Model offers different risk estimates calibrated for different investment horizons, correctly accounting for serial correlation of returns as well as mean reversion of risk. The model monitors the risk state of the market and responds quickly when conditions change, but scales the responsiveness according to the model horizon.

All models are estimated and updated on a daily basis to ensure that they fully reflect the latest available information from the financial markets. A set of enhanced model quality assurance processes is run prior to model publication to prevent data and calculation errors. Moreover, model performance metrics are collected and reviewed on a quarterly basis as part of our model governance process.

We would like to acknowledge the large number of our colleagues from global data, engineering, operations, sales and research teams that have contributed to the development and implementation of the MAC3 GRM models. The launch of these models reflects Bloomberg's ongoing commitment to research and innovation, so that we may provide our clients with the industry-leading set of investment tools and portfolio analytics.

Jose Menchero
Head of Portfolio Analytics Research

Anthony Lazanas
Head of Portfolio, Index & ESG Research

Bloomberg

Table of Contents

Section (1): Model Overview

(A) Model Highlights

This whitepaper describes the equity models contained within the equity portion of the MAC3 Global Risk Model (MAC3 GRM), which represents the third-generation Bloomberg multi-asset-class risk model. The MAC3 GRM is the successor to the MAC2 Model, which was launched in 2016.

The MAC3 Model is the result of an intensive, multi-year research endeavor to create the most advanced suite of factor models in the industry. As such, the MAC3 Model incorporates a multitude of methodology advances and model innovations, as highlighted below.

Term Structure of Risk. The MAC3 Model uses mixed-frequency (MXF) estimation to predict risk at different horizons. The MXF approach combines information from both high-frequency (daily) and low-frequency (e.g., weekly or monthly) observations to obtain more precise estimates of portfolio risk at different horizons. Each model is offered with six distinct prediction horizons: daily, weekly, monthly, quarterly, annual, and long-term. These models serve distinct investment use cases. For instance, the daily or weekly models represent attractive choices for risk managers with relatively short horizons. By contrast, the monthly, quarterly, and annual models typically appeal more to portfolio managers with longer investment horizons. Finally, the long-term model is designed for asset owners making strategic asset-allocation decisions.

Daily Model Updates. All components of the MAC3 Model are estimated and updated on a daily basis. These model components include: (a) factor exposures, (b) factor returns, (c) factor covariance matrices, and (d) specific risk forecasts. The daily production process ensures that the models reflect the most up-to-date information from the financial markets.

Industry and Country Beta Factors. The MAC3 equity models move beyond the naïve assumption that all stocks in a given industry/country have the same unit exposure to the corresponding factor. We explicitly recognize that some stocks are more sensitive to industry and country factor movements than others. We estimate these sensitivities (betas) by time-series regression. We find that using industry and country betas leads to increased explanatory power, while also mitigating spurious correlations between specific returns and factor returns.

Improved and Expanded Style Factors. The MAC3 equity models contain a common set of 14 style factors, compared with 10 styles in the previous model (MAC2). One major improvement in factor structure was to take the MAC2 volatility factor and split it into two style factors in MAC3: market beta and residual volatility. These two factors capture different elements of risk, and both act as strong risk factors with high explanatory power. In a similar fashion, the MAC2 value factor was split into two factors in MAC3: earnings yield and valuation (based on a set of price multiples). Another factor introduced by the MAC3 model is the mid-cap factor, which is designed to capture the unique characteristics of this market segment. The final enhancement was to include long-term reversal as a style factor.

Satellite Country Factors. The MAC3 equity models introduce the notion of "satellite" country factors. Stocks in satellite countries are exposed to all of the factors in the local model, but they are not included in the estimation universe used to calculate factor returns. In other words, the only factor portfolio in which satellite-country stocks have non-zero weight is the satellite factor itself. This treatment allows the model to have a dedicated factor to capture the unique risk associated with the country, without having the country compromise the integrity of the estimation universe.

Finite-Sample Adjustment. A long-standing problem in the traditional approach for building factor models is that factor risk and specific risk have not been properly disentangled. More specifically, the traditional approach "double counts" the specific risk in the factor portfolios, leading to inflated estimates of factor volatility and deflated estimates of specific risk. This effect is a direct consequence of having a finite number of stocks in the estimation universe. In the MAC3 Model, we implement an innovative solution to this problem, leading to a better decomposition of factor/specific risk and more accurate volatility forecasts.

Cross-Sectional Volatility (CSV) Adjustment. The MAC3 Model utilizes an innovative algorithm to quickly detect changing market dynamics and adjust the volatility forecasts accordingly. In particular, the algorithm uses cross-sectional observations to detect "instantaneous" biases in the risk forecasts. This information is then used in a "feedback loop" to adjust predicted volatilities and reduce biases. This technique greatly mitigates underforecasting of risk heading into a financial crisis, and allows predicted volatility to decline rapidly after the crisis subsides.

New Specific Risk Model. The MAC3 Model incorporates major improvements in the specific risk forecasts. Similar to the MAC3 factor covariance matrix, the specific risk model uses daily return observations to forecast specific risk (also known as non-factor risk) at multiple horizons. The specific risk model contains two components: a time-series estimate based on realized specific returns, and a structural model based on factor exposures. The structural model allows for accurate specific risk forecasts of recently issued IPOs, with little or no return history.

New BICS Industry Classification. The MAC3 Model utilizes the latest Bloomberg Industry Classification System (BICS), officially launched in 2020. The BICS scheme is hierarchical in nature, with 11 sectors at the top level (Level 1), followed by 20 industry groups (Level 2), 59 industries (Level 3), and 214 sub-industries (Level 4). The new BICS structure has been updated to better reflect the current landscape of the global equity markets.

Improved Regression-Weighting Scheme. Traditional equity factor models use square-root-of-market-cap for the regression weights. Econometrically, however, the optimal regression weights are proportional to inverse residual variance, since this maximizes the precision (i.e., minimizes noise) of the factor-return estimates. The MAC3 equity models apply inverse residual variance for the regression weights. Empirically, we find that these weights significantly reduce noise in the factor returns and mitigate spurious correlations between factor returns and specific returns.

Imputation Algorithm. The MAC3 Model incorporates a sophisticated imputation algorithm to handle factors with missing returns. Factor returns may be missing due to either short histories or national holidays. In either case, the missing factor returns must be imputed so that the factor covariance matrix can be properly estimated.

Independent Validation of Model Code. In order to minimize the potential for implementation errors and coding bugs, we applied *two* rounds of double-blind code reconciliation in the development of the MAC3

Model. The first round was applied during the research phase of the project, in which the research team developed two versions of the model, written in two different programming languages. The two research codes were reconciled by feeding in the same input data and iterating until the two models produced identical model outputs (to extremely high precision). In the second round, we reconciled the research code with the production code, which was written by an independent team of software engineers in a different programming language. This rigorous verification process—though tedious and time consuming—is crucial to ensure the highest quality standards in model implementation.

Robust Estimates of Factor Correlations. The MAC3 Model adopts the same advanced technique for estimating factor correlations that was first developed for the launch of the MAC2 Model in 2016. This method assures that the estimated factor correlations can be reliably used for both risk forecasting and portfolio optimization. Our research indicates that the conventional method used in the industry tends to systematically underforecast factor correlations, which can lead to serious biases in risk forecasts.

(B) Local, Global, and Integrated Models

The MAC3 equity suite consists of the MAC3 Global Equity Model and 13 local equity models: Asia Pacific (APAC), Australia, Canada, China A-shares, Europe, Europe/ME/Africa (EMEA), India, Japan, Korea, Latin America, South Africa, UK, and US. Users who select the MAC3 Global Equity Model will analyze their portfolio through the lens of a global market factor, plus factors that represent countries, global industries, and global styles.

An attractive feature of the Global Equity Model is that it provides comprehensive coverage of global equities under a single parsimonious factor structure. This makes it useful to gain a high-level insight into the drivers of a global equity portfolio. For instance, the aggregate impact of momentum on the portfolio can be quickly ascertained by analyzing the risk and return contributions of a single global momentum factor.

Parsimony, however, also comes with drawbacks. The return of the global momentum factor effectively represents an average of local momentum factors across the globe. However, momentum does not move in lock-step across different markets. For example, suppose that momentum performs well in Europe, but performs poorly across the rest of the globe. In this case, a European portfolio manager with a positive tilt on momentum would rightfully obtain a positive return contribution from the momentum factor in the Europe model, but would erroneously obtain a *negative* contribution using the global model.

For this reason, investors with portfolios concentrated in local equity markets typically prefer to use a local model that has been specifically calibrated to best reflect their investment universe. The advantage of this approach is that local factors better describe the actual behavior of the factors in a particular local market. The main disadvantage of local models is that they do not offer comprehensive coverage spanning the globe.

The third option is to select the MAC3 Integrated[1] Model, which effectively represents an aggregation of all 13 local models. In this case, US stocks would have exposure to US local factors, while Japanese stocks would

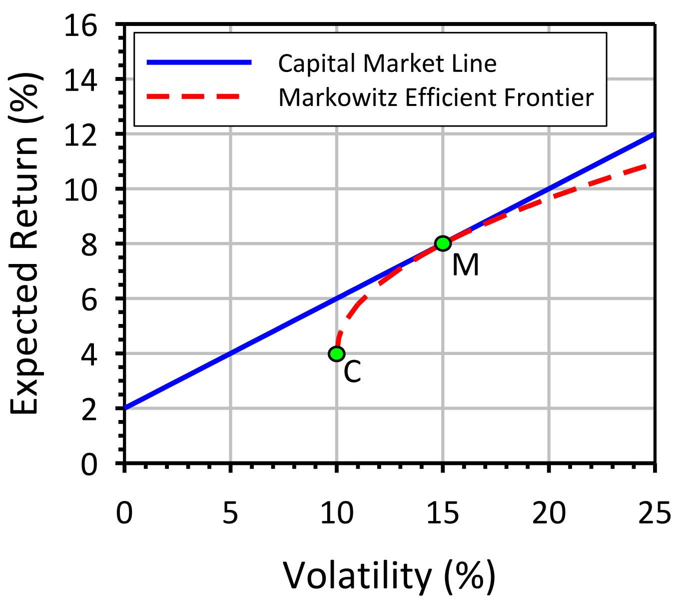
[1] Previously known as the "regional" model in the MAC2 parlance.

have exposure to Japan local factors. The advantage of the MAC3 Integrated Model is that it provides the best of both worlds: namely, it offers the same comprehensive coverage as the MAC3 Global Equity Model, while retaining the same granular factor structure of the local models.

(C) Brief History of Modern Portfolio Theory and Factor Models

Modern Portfolio Theory (MPT) originated with the pioneering work of Markowitz (1952), who introduced mean-variance optimization (MVO) as a portfolio-construction technique. According to this theory, rational investors seek to hold an *efficient* portfolio, defined as one having the highest expected return for a given level of risk. The risk measure used by Markowitz was the standard deviation of portfolio returns (i.e., volatility).

Figure 1.1. Capital Market Line and the Markowitz Efficient Frontier, assuming a 2% risk-free rate and a 6% equity risk premium.



A plot of expected portfolio return versus portfolio volatility, applied to the set of all fully invested efficient portfolios, maps out a concave curve called the *efficient frontier*, as shown by the dashed red line in Figure 1.1. Note that Portfolio C represents the minimum-volatility fully invested portfolio, while Portfolio M represents the fully invested portfolio with maximum expected return per unit of risk. Since each investor in principle has a different risk tolerance, it follows that each investor would select a different portfolio along the efficient frontier.

Tobin (1958) extended MPT in a simple—yet crucial—manner by introducing the risk-free instrument as an investable asset. With this seemingly trivial extension, Tobin showed that the efficient frontier collapsed into a straight line (known as the Capital Market Line), as indicated by the solid blue line in Figure 1.1. In this case, the set of all efficient portfolios was given by a weighted combination between just *two* portfolios: (a) the risk-free asset, and (b) the unique *super-efficient* portfolio M that had the maximum return/risk ratio of all Markowitz portfolios along the efficient frontier. This concept came to be known as the *two-fund separation theorem*.

The next step in the development of MPT was due to Sharpe (1964), who developed the Capital Asset Pricing Model (CAPM). By making several assumptions, such as that all investors were rational, built Markowitz-efficient portfolios, and had homogeneous expectations, Sharpe was able to show that the "super-efficient" portfolio M was the market portfolio itself. This ultimately gave rise to the *theory of market efficiency*.

A central prediction of the CAPM is that the expected return r_{nt} (in excess of the risk-free rate) of any asset is given by the product of two terms: (a) the asset's beta relative to the market portfolio, and (b) the expected excess return R_{Mt} of the market portfolio. The asset beta in this case represents the slope coefficient β_n in a time-series regression,

$$r_{nt} = \beta_n R_{Mt} + e_{nt} \ . \tag{1C.1}$$

The CAPM asserts that the residuals e_{nt} are mean zero.

While the CAPM represents a beautiful and unassailable edifice if one accepts its basic premises, it has not held up well under empirical scrutiny. In the years that followed, academics identified many "pricing anomalies" that appeared to violate the CAPM. For instance, Basu (1977) showed that stocks with high earnings-to-price ratios (i.e., value stocks) seemed to have higher returns than those predicted by CAPM. Later, Banz (1981) documented the small-cap premium, in which small-cap stocks tended to outperform large-cap stocks, again in apparent violation of the CAPM. Another example of a pricing anomaly was given by the momentum factor, as described by Jegadeesh and Titman (1993). Perhaps the most critical study was due to Fama and French (1992), who showed that there does not appear to be a return premium associated with market beta, in direct violation of the most basic prediction of CAPM.

The existence of pricing anomalies provided a theoretical justification for active portfolio management. If an investor were only able to identify the factors that truly explained expected returns—not an easy task—they would be able to build a customized efficient portfolio with higher risk-adjusted return (in expectation) than the market portfolio.

Building a customized efficient portfolio involves two basic steps. First, the investor must define the expected returns of each asset, known as the "alpha" signal. Second, the investor must apply Markowitz MVO to construct the optimal portfolio. Note that MVO requires three basic inputs: (a) the alpha signal, (b) the set of portfolio constraints, and (c) the asset covariance matrix.

Early attempts by practitioners to apply Markowitz MVO ended in disappointment. Michaud (1989) documented many of these shortcomings, and coined the phrase that "optimizers are error maximizers."

In fact, Markowitz optimization contains two potential pitfalls. First, since nobody knows the "true alpha," any *forecast* alpha signal will necessarily contain some error. In the extreme case that the forecast alpha signal contains no useful information, the investor cannot expect to outperform the market and should simply invest in the market portfolio.

The second basic pitfall in MVO is that the "true" asset covariance matrix is not known. Hence, we must instead use an *estimated* covariance matrix. The easiest approach is to compute it directly using the textbook definition of covariance, which results in the so-called *sample* covariance matrix. Of course, any estimated matrix will contain estimation error. Menchero and Ji (2019) showed that sampling error in the sample covariance matrix leads to several detrimental effects, including: (a) inferior risk-adjusted performance, (b) underforecasting risk of optimized portfolios, and (c) increased portfolio leverage and turnover. However, Menchero and Ji also showed that if the detrimental effect of "noise" could be mitigated, MVO clearly leads to superior risk-adjusted performance. More specifically, the portfolio has higher expected return per unit of risk. Hence, the key lesson here is that *to build a reliable portfolio using MVO, one must first create a robust estimate of the asset covariance matrix*.

The fundamental reason why sample covariance matrices are not reliable for MVO is that the number of assets N is typically large compared with the number of time periods T used to estimate the covariance matrix. For instance, the Bloomberg US 3000 Index contains roughly $N \approx 3,000$ stocks. If five years of daily returns are used to estimate the covariance matrix, then the number of trading days is approximately $T \approx 1,250$. A fundamental result from linear algebra is that whenever $N > T$, the sample covariance matrix becomes "rank deficient," which implies the existence of "riskless" portfolios composed entirely of risky assets. Although this is simply a mathematical artifact that does not reflect reality, MVO "locks into" these spurious risk-free portfolios, which invariably leads to disappointing out-of-sample performance.

Rosenberg (1974) devised a brilliant solution to this problem, and in the process revolutionized the practice of modern finance. Rosenberg posited that asset returns r_n are driven by a parsimonious set of K factors,

$$r_n = \sum_{k=1}^{K} X_{nk} f_k + u_n \quad , \tag{1C.2}$$

where X_{nk} is the exposure of asset n to factor k, f_k is the return of the factor, and u_n is the unexplained residual, also known as the specific return, the non-factor return, or the idiosyncratic return. Rosenberg assumed that the idiosyncratic returns were mutually uncorrelated and also uncorrelated with the factors.

The efficacy of factor models lies in their "dimensional reduction." Instead of directly estimating the NxN sample covariance matrix, which is unreliable for portfolio construction, Rosenberg estimated the much smaller KxK *factor* covariance matrix. The number of factors K in a typical local model may be several dozen, whereas the number of time periods used to estimate the factor covariance matrix may be several hundred. As a result, $T \gg K$, leading to a well-conditioned factor covariance matrix. Given the factor exposure matrix, the factor covariance matrix, and the specific risk forecast of each asset, the full NxN asset covariance matrix could be reliably estimated for portfolio optimization. This key advance made Markowitz optimization applicable *in practice*, not just in theory.

Of course, factor models bring another added benefit to the table. Namely, they provide great insight into the drivers of portfolio risk and return. Rather than analyzing the portfolio through the lens of individual stocks, the investor could focus on a much smaller number of factors that "really matter." This makes factor models an indispensable tool for portfolio analysis, as described in the next section.

The next major step in the development of factor models came with the introduction of *integrated* models. The basic idea behind integrated models is to combine a set of local models. For example, an equity model suite is typically comprised of multiple local models (e.g., US, Japan, Europe, etc.) that collectively span all countries across the globe. The US stocks would be exposed to US local factors, whereas Japanese stocks would be exposed to Japan local factors, and so on. Since there may be a dozen or more local equity models, each of which contains several dozen factors, the number of equity factors alone may easily reach into the hundreds. The equity models are then combined with other asset classes, such as fixed income, commodities, and currencies. The total number of factors in an integrated multi-asset-class (MAC) risk model may therefore easily exceed 2,000.

Hence, the original problem faced by Rosenberg (i.e., too many stocks with too few time periods) now reappears with integrated models. Namely, the number of factors is too large relative to the number of time periods to directly estimate the factor covariance matrix in a reliable fashion. Some kind of dimensional-reduction technique must be applied, i.e., a "factor model of factors."

The most common technique used to solve this problem was developed by Shepard (2007/2008). In this approach—known as the time-series method—one tries to identify time series of global factor returns (known as "core" factors) with the aim of explaining the correlations of local factors. The "loadings" of the local factors on the core factors are estimated by time-series regression, from which the method derives its name. Given the factor loadings and the covariance matrix of core factors, the correlations of the local factors can be estimated.

The major pitfall of this technique, as shown by Menchero and Ji (2020), is that it tends to systematically *underestimate* factor correlations, effectively "shrinking" them towards zero. The degree of underestimation may be quite large —especially across asset classes. For instance, Menchero and Ji found that the correlations between equity factors and fixed income factors were systematically and persistently underestimated by about 70%. In other words, a factor correlation of 0.50 may be estimated to be only 0.15 using the time-series method. The basic reason for underestimation is that the core factor returns cannot be properly identified, especially when trying to explain correlations across different asset classes.

Menchero and Li (2020) conducted a systematic analysis on the relationship between risk forecasting errors and underestimation of correlations. In particular, they studied the dependency of mean-squared error (MSE) on three quantities: (a) pair-wise correlation, (b) sampling error, and (c) shrinkage intensity. They found that for factor pairs with significant correlations (e.g., 0.50), the MSE increases dramatically with increasing shrinkage intensity.

A major advance in the estimation of MAC risk models came with the launch of the second-generation Bloomberg multi-asset-class model (MAC2) in 2016. The MAC2 Model[2] utilized an innovative technique for estimating factor correlations. As described by Menchero and Ji (2020), the MAC2 methodology blended the sample factor correlation matrix with the PCA factor correlation matrix (derived from Principle Component Analysis). The resulting correlation matrix greatly mitigates the large biases found in the time-

[2] Note that the Bloomberg first-generation MAC model (MAC1), available prior to 2016, utilized the time-series method.

series approach, while still preserving a well-conditioned matrix that is reliable for portfolio construction.

This whitepaper describes the MAC3 methodology for estimating volatilities and correlations. The MAC3 Model leverages the innovative technique introduced with the MAC2 Model, namely blending with the PCA correlation matrix to produce a well-conditioned factor covariance matrix that can be reliably used for portfolio construction. One important difference between MAC2 and MAC3 is that the MAC2 Model used weekly factor-return observations to estimate the factor covariance matrix, whereas MAC3 uses daily factor return observations. It is well known that high-frequency observations contain valuable information to improve the accuracy of risk forecasts. The MAC3 Model uses these high-frequency observations to predict a term structure of risk. While the MAC2 Model effectively provided a single forecast horizon (based on weekly observations), the MAC3 Model provides multiple forecasting horizons: daily, weekly, monthly, quarterly, annual, and long-term.

(D) Analyzing Portfolio Risk and Return with Factor Models

The factor model in Equation (1C.2) can be conveniently expressed in matrix form,

$$\mathbf{r} = \mathbf{X}\mathbf{f} + \mathbf{u} \ , \tag{1D.1}$$

where $\mathbf{r}$ is an Nx1 vector of local excess stock returns (i.e., the local stock return minus the risk-free rate), $\mathbf{X}$ is an NxK matrix of factor exposures, $\mathbf{f}$ is a Kx1 vector of factor returns, and $\mathbf{u}$ is an Nx1 vector of unexplained residuals (specific returns).

Let $\mathbf{F} = \mathrm{cov}(\mathbf{f}, \mathbf{f})$ denote the KxK factor covariance matrix, and let $\boldsymbol{\Delta} = \mathrm{cov}(\mathbf{u}, \mathbf{u})$ be NxN specific variance matrix. The basic structure of the model assumes that the specific returns are uncorrelated with the factor returns, i.e., $\mathrm{cov}(\mathbf{f}, \mathbf{u}) = \mathbf{0}$. Moreover, since specific returns are mutually uncorrelated, this implies that $\boldsymbol{\Delta}$ is a diagonal matrix[3] with diagonal entries given by the specific variance of the stock. Under this model structure, the NxN asset covariance matrix $\boldsymbol{\Omega}$ can be written as

$$\boldsymbol{\Omega} = \mathbf{X}\mathbf{F}\mathbf{X}' + \boldsymbol{\Delta} \ , \tag{1D.2}$$

where the first term $\mathbf{X}\mathbf{F}\mathbf{X}'$ represents the factor component, while the second term $\boldsymbol{\Delta}$ gives the (specific) component.

Next, we show how to use factor models to attribute portfolio return. Consider a portfolio with an Nx1 weight vector $\mathbf{w}_P$. The portfolio return $R_P = \mathbf{w}_P'\mathbf{r}$ can be expressed as

$$R_P = \mathbf{X}_P'\mathbf{f} + \mathbf{w}_P'\mathbf{u} \ , \tag{1D.3}$$

where $\mathbf{X}_P' \equiv \mathbf{w}_P'\mathbf{X}$ is a 1xK vector of portfolio factor exposures. A manager earns positive return contributions by having positive exposure to factors with positive returns, or negative exposures to factors with negative returns. Similarly, a manager earns positive idiosyncratic return contributions by overweighting stocks with

[3] The only exception to this rule occurs for *linked* securities, whose specific correlations are non-zero.

positive specific returns, or by underweighting stocks with negative specific returns.

The portfolio variance is given by $\sigma_P^2 = \mathbf{w}'_P \mathbf{\Omega} \mathbf{w}_P$, which can be rewritten as

$$\sigma_P^2 = \mathbf{X}'_P \mathbf{F} \mathbf{X}_P + \mathbf{w}'_P \mathbf{\Delta} \mathbf{w}_P \quad . \tag{1D.4}$$

The first term represents the variance contribution from factors, whereas the second term gives the variance contribution from the specific component. From Equation (1D.4), it is clear that all asset correlations are derived from exposures to common factors.

Section (2): Factor Exposure Matrix

(A) Estimation Universe

The first step in building a risk model is to define the estimation universe (ESTU). The ESTU forms the basis both for standardizing factor exposures as well as estimating factor returns.

The ESTU logic is designed with four main objectives. First, the ESTU for a particular country should be representative of the broad market portfolio for that country. Second, the ESTU should be stable, without large influxes or outfluxes of stocks in a short period of time. Third, stocks in the ESTU should trade with sufficient frequency to prevent thinly traded stocks from entering the universe. Fourth, the rules should be simple and transparent.

To determine the list of eligible securities for the ESTU of a given country, we apply six filters to the full universe of securities in the Bloomberg database. The first filter is the Bloomberg Country of Risk field, which must match the country for which the ESTU is being constructed.

The second filter is Security Type. Each country has an allowable set of security types. All countries allow Common Stocks and Real Estate Investment Trusts (REIT's) as eligible security types; for many countries (e.g., US), these are the only two security types allowed. Other countries, such as Canada, Australia, and Brazil, also allow Unit Trusts and Stapled Securities[4]. For some emerging market countries, such as Russia and Argentina, Depository Receipts are the most liquid and most important securities. The same is true for frontier markets. Hence, such countries permit Depository Receipts to be included in the ESTU.

Filter 3 removes duplicate securities. To apply Filter 3, we first identify all linked securities (based on a common Bloomberg ID035 security identifier) that have passed the first two filters. Two common examples of linked securities are: (1) different share classes of the same stock, and (2) depository receipts. From the list of linked securities, we retain only the unique Bloomberg Fundamentals Ticker (DT110), while removing all other linked securities. The Bloomberg Fundamentals Ticker can be thought of as the "primary" share, and the data from this ticker are used for purposes of computing exposures to technical and fundamental factors.

---

[4] Stapled securities are financial instruments consisting of two or more securities that are contractually bound to be sold as a single indivisible unit.

Filter 4 sets a minimum allowable price threshold (measured in US dollars) for each country. For the US and Canada, this is set to $1.00 (USD) per share. Other countries may have a lower price threshold. To reduce potential instabilities in the ESTU caused by stocks fluctuating near the price threshold, we smooth the price using exponentially weighted moving averages (EWMA).

Filter 5 specifies that to be eligible for inclusion in the ESTU, a security must have a non-missing industry classification. We employ the Bloomberg Industry Classification System (BICS) to determine the industry factors.

Filter 6 sets a lower bound on the allowable frequency of missing/zero returns. For each stock n, for every day t, we define an indicator function δ_{nt}, whose value is equal to 1 if the stock has non-zero return (measured in the trading currency), and is equal to zero if the stock return is zero or missing.

Next, we define the return-availability measure a_{nt} for stock n at the end of day t as the EWMA average of the indicator function,

$$a_{nt} = \sum_{\tau=1}^{t} \lambda_{\tau} \delta_{n\tau} \quad , \tag{2A.1}$$

where λ_{τ} is the EWMA weight assigned to day τ. For each country c, we apply a unique threshold A_c for the minimum allowable value of the return availability. More specifically, if $a_{nt} < A_c$, the security is excluded from the ESTU. For highly developed markets, such as the US, the threshold may be quite stringent (e.g., $A_c = 0.80$), whereas for emerging or frontier markets the threshold is more liberal (e.g., $A_c = 0.50$).

Stocks that pass all six filters are eligible for inclusion in the ESTU. In order to ensure that the ESTU is not dominated by a large number micro-cap stocks, we also apply a minimum market-cap threshold for each country. At the start of each day, for each country, we rank-order all stocks that have passed the six ESTU filters by their market capitalization. We continue down the list until we capture a fraction X_C of the total market cap. For instance, if $X_C = 0.97$, the ESTU captures 97% of the total market capitalization of the eligible universe within each country.

In some markets, certain sectors may be dominated by small-cap stocks, few of which may exceed the X_C threshold. To mitigate the effect of thin sectors, we rank order by market capitalization all eligible stocks within each sector, retaining all stocks whose cumulative weight exceeds a sector threshold of X_S. For instance, if $X_S = 0.90$, the ESTU captures 90% of the market capitalization of each sector. The final ESTU is given by the set of all stocks that pass the X_C country threshold, *augmented* by the set of all stocks needed to satisfy the X_S sector threshold.

In rare instances, we may find that the above ESTU logic may exclude a stock that is deemed important in a particular market, or include a stock that doesn't belong. To allow full flexibility in setting the ESTU, we apply and maintain *inclusion* and *exclusion* lists for each country. Stocks on the inclusion list are forced into the ESTU, whereas stocks on the exclusion list are forcibly removed.

The main reason to maintain inclusion and exclusion lists is to properly handle dual-listed companies. Dual-listed companies represent corporate structures whereby the two corporations operate as a single business unit, but retain separate legal identities for tax purposes.

As an illustration, the US model contains only a single stock (Carnival Corporation) on the exclusion list, and no stocks on the inclusion list. Note that Carnival Corporation is a dual-listed company, with one share that trades in New York, and another share that trades in London. These two securities have virtually identical return streams. Nonetheless, they have different Bloomberg ID035 security identifiers, since they are legally distinct entities. To avoid "double counting" the stock, the share that trades in London is excluded.

(B) Outlier Algorithm

Style factors are formed from raw descriptors, which are then standardized into z-scores, as described in the next section. Some common examples of raw descriptors include: stock momentum, size, market beta, and book-to-price ratio.

Raw descriptors may contain extreme outliers that in turn distort the distribution. These outliers must be trimmed so that they do not have an undue impact on either the factor returns or the factor exposures. The typical approach for identifying outliers is to measure how many standard deviations the observation lies away from the mean. This must be applied with caution, however, since the standard deviation itself is subject to distortion from extreme outliers.

Our approach is to first trim extreme outliers using the *robust standard deviation*, which is insensitive to extreme outliers. Let y_{nk} denote the raw value of descriptor k for asset n. The robust standard deviation is defined as

$$\sigma_k^R = 1.4826 \cdot \text{median}\left| y_{nk} - \theta_k \right| , \qquad \text{(2B.1)}$$

where θ_k is the median value of y_{nk}. In other words, the robust standard deviation measures the median deviation away from the median value. Note that the robust standard deviation converges to the conventional standard deviation when the descriptor is normally distributed.

We define extreme outliers as observations that lie more than z_R *robust* standard deviations away from the median. These observations are trimmed at $\pm z_R \sigma_k^R$ with respect to the median value. The value for z_R is established for each descriptor with the objective of catching the most egregious outliers, while not trimming an excessive number of points.

With the most egregious outliers thus trimmed, the conventional standard deviation can now be reliably computed. The second step of the outlier algorithm is to further trim raw descriptors z_C *conventional* standard deviations away from the equal-weighted mean.

Although Bloomberg offers excellent coverage for fundamental data, occasionally we may find a stock that has missing values for one or more raw descriptors. Our approach for filling missing fundamental data is guided by the empirical observation that stocks in the same industry tend to have similar raw descriptor values. For instance, stocks in the financial sector tend to have above average earnings-to-price (E/P) ratios, whereas stocks in the technology sector tend to have below-average E/P ratios. Hence, we fill missing fundamental data with the equal-weighted industry mean of the raw descriptor.

(C) Standardization of Style Factors

Let $\tilde{y}_{nk}^{c}$ denote the final trimmed raw value of stock n in country c for descriptor k. We standardize the descriptor into a z-score

$$\tilde{x}_{nk}^{c} = \frac{\tilde{y}_{nk}^{c} - \mu_{k}^{c}}{\sigma_{k}^{c}} , \tag{2C.1}$$

where μ_{k}^{c} is the equal-weighted mean of $\tilde{y}_{nk}^{c}$ across all stocks in the country, and σ_{k}^{c} is the equal-weighted standard deviation.

The final standardized raw descriptor (x_{nk}^{c}) is obtained by rigidly shifting the distribution so that the *cap-weighted* market portfolio has a standardized exposure of exactly zero,

$$x_{nk}^{c} = \tilde{x}_{nk}^{c} - \lambda_{k}^{c} , \tag{2C.2}$$

where λ_{k}^{c} is the cap-weighted mean of $\tilde{x}_{nk}^{c}$. The standardization convention in Equation (2C.2) ensures that the market portfolio is *style neutral*.

Some factors are composed of multiple descriptors, while other factors may contain only a single descriptor. As with individual descriptors, composite factors are also standardized to be cap-weighted mean zero and equal-weighted unit standard deviation.

The first step in forming the composite factor is to combine the underlying descriptors using descriptor weights v_k,

$$\tilde{X}_{n}^{c} = \sum_{k} v_k x_{nk}^{c} , \tag{2C.3}$$

where $\tilde{X}_{n}^{c}$ is the weighted average exposure of stock n (in country c) to the composite factor. The exposures $\tilde{X}_{n}^{c}$ are cap-weighted mean zero, but do not have unit standard deviation. The final step in forming the standardized z-score for the composite factor is to scale the exposures to be unit standard deviation. Specifically, the final standardized exposure to the composite factor is given by

$$X_{n}^{c} = \frac{\tilde{X}_{n}^{c}}{\tilde{\sigma}_{c}}, \tag{2C.4}$$

where $\tilde{\sigma}_{c}$ is the equal-weighted standard deviation of $\tilde{X}_{n}^{c}$. Note that the standardized composite factor exposures are cap-weighted mean zero and equal-weighted unit standard deviation within each country c.

For most style factors, the difference between the largest exposure and the smallest exposure is typically around 6. However, this does not imply that the exposures are distributed within the range ±3. In particular, when large-cap stocks cluster on the extremes, the distribution will be shifted up or down to maintain a cap-weighted mean of zero, as we show in the next section. Moreover, some factors occasionally have maximum differences in exposure slightly greater than 6. This may occur for composite factors, since a stock that scores high on all descriptors will have an even higher score for the composite factor.

(D) Style Factor Definitions

Style factors represent important sources of risk and return that are related to known characteristics of the stock. Most style factors are well documented in the academic literature and have been used by practitioners for many years.

As described by Menchero and Lee (2015), style factors can be categorized into *alpha* factors and *risk* factors. Alpha factors are useful as investment signals, as they are related to the expected returns of the assets. In the academic literature, they are referred to as the "pricing anomalies," as they appear to violate the CAPM. One common technique for identifying alpha factors is to sort stocks into quintiles by rank-ordering them according to some stock-level characteristic, such as E/P ratio. If the style factor explains alpha, then we expect that a dollar-neutral portfolio formed by going long the top quintile and short the bottom quintile would exhibit a positive return.

Risk factors are useful for identifying sources of portfolio volatility or systematic return comovement. Suppose we sort a universe of 1000 stocks into quintiles based on market beta, and we go long the top quintile (high-beta stocks), while shorting the bottom quintile (low-beta stocks). Since high-beta stocks are much more sensitive to market movements than their low-beta counterparts, the volatility of the resulting portfolio would be much higher than that of a portfolio formed by randomly selecting 200 stocks to go long and 200 stocks to go short. This argument shows that market beta is a strong risk factor.

Of course, some style factors can serve both as good risk factors *and* good alpha factors. For instance, in many countries, stock momentum is useful both for predicting expected returns and portfolio volatility.

It should be noted that the primary aim of a risk model is to predict portfolio volatility. Hence, while most practitioners agree that the number of independent alpha factors may be small enough to be counted on one hand, risk models generally contain a far greater number of factors. For example, it is well known that industry and country membership are important predictors of portfolio volatility and correlation, although they typically are not considered as sources of return premia (i.e., alpha factors).

Market Beta. According to the CAPM, the only "priced" factor (i.e., with a return premium) is market beta. This result rests ultimately on the assumption that the market portfolio is mean-variance efficient. According to this view, high-beta stocks should have higher expected returns than low-beta stocks. Empirically, however, this central prediction of the CAPM appears to be violated, as described by Fama and French (1992). Nonetheless, even if market beta does not have a return premium associated with it, it still represents a very strong risk factor.

We compute raw market betas by regressing local excess stock returns r_{nt} against the local excess return R_{Mt} of the cap-weighted market portfolio (ESTU),

$$r_{nt} = \beta_n R_{Mt} + e_{nt} \ . \tag{2D.1}$$

Local excess stock returns are defined relative to the risk-free rate. We estimate market betas using weighted least-squares regression, which assigns more weight to recent observations. The half-life (HL) parameter is selected to maximize the explanatory power of the factor.

Special care must be taken to reliably estimate market betas. First, time-series betas are sensitive to return outliers. Hence, prior to estimating market betas, we apply return trimming to both the individual stock returns as well as to the ESTU return.

Second, trading asynchronicity and serial correlation in daily returns may cause a material impact on estimated betas, especially in multi-country models that span multiple time zones. To handle such effects, we use *L*-day rolling windows to estimate market betas. The value of *L* is selected to maximize the explanatory power of the factor.

Third, special treatment is required for IPOs and stocks with a high frequency of missing/zero returns. For such stocks, we shrink the estimated market betas toward the cap-weighted mean beta of the country in question. When the market beta factor is standardized, this ensures that a recent IPO will have a standardized exposure of zero. As the return availability increases, we slowly assign more weight to the estimated beta.

Residual Volatility. Empirically, it has been observed that high-volatility stocks have tended to underperform their low-volatility peers, even after controlling for other factors. For instance, Ang *et al*. (2009) studied 23 developed markets and found that, on average, the high-volatility quintile underperformed the low-volatility quintile by 1.31% per month, even after controlling for the market beta, size, and value factors.

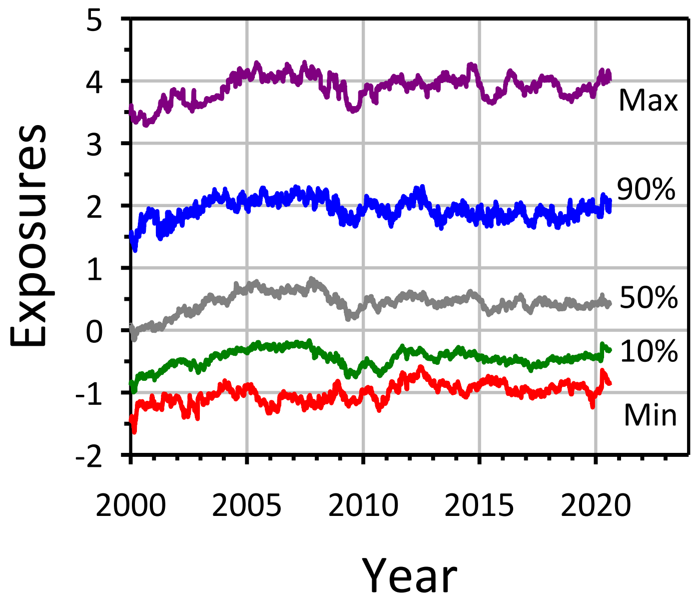
We estimate residual volatility from a time-series regression of local excess stock returns against the local excess return of the ESTU, as in Equation (2D.1). The raw value of the factor exposure is given by the EWMA volatility of the residual returns,

$$\sigma_n^R = \left(\sum_t v_t e_{nt}^2 \right)^{1/2} , \tag{2D.2}$$

where v_t is the EWMA weight assigned to day t. The HL parameter is calibrated to maximize the explanatory power of the factor over a one-month horizon.

Example. In Figure 2.1, we plot the distribution of exposures for the residual volatility factor in the UK Equity Risk Model. Large-cap stocks tend to have negative exposure to residual volatility, while the far more numerous small-cap stocks tend to have positive exposure. Since the distribution is cap-weighted mean zero, the minimum exposure tends to be close to -1, while the maximum exposure is typically around +4.

Figure 2.1. Distribution of residual volatility exposures for the MAC3 UK Equity Model. For each date, we plot the minimum, 10-percentile, median, 90-percentile, and maximum exposure.



Momentum. There is considerable evidence that past performance may be an indicator of future returns. For instance, Jegadeesh (1990) documented the short-term reversal effect, in which stocks that have performed well over the previous month tended to perform poorly over the subsequent month.

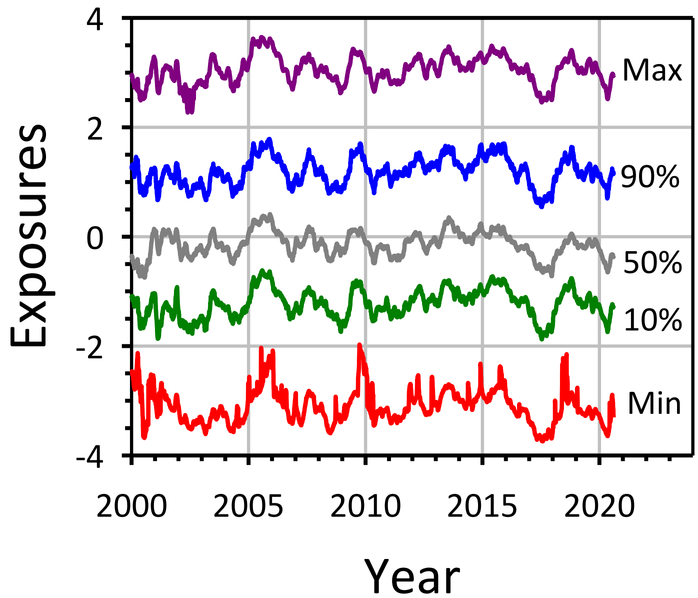
Another example is given by the momentum effect, which was described by Jegadeesh and Titman (1993). They showed that stocks which have performed well over the last 6-12 months (the "winners") tended to outperform stocks that have performed poorly (the "losers") over the same time period. A possible behavioral explanation for momentum is that investors tend to *under-react* to the flow of new information, which results in a delayed price response.

We compute the raw momentum descriptor by taking the weighted average of trailing log returns,

$$M_n^t = \sum_{l=1}^{L} w_{t-l} \ln\left(1 + r_n^{t-l}\right) , \tag{2D.3}$$

where w_{t-l} is the weight assigned to day $(t-l)$, r_n^{t-l} is the local stock return for the day, and L is the total number of days used to compute the raw momentum score.

Figure 2.2. Distribution of momentum exposures for the MAC3 Korea Equity Model. For each date, we plot the minimum, 10-percentile, median, 90-percentile, and maximum exposure.



The typical definition of momentum uses trailing 12-month returns, with a one-month lag. Note that only 12 months ($L = 252$) of data are used, with the most recent month discarded. Hence, the raw score is really based on an 11-month cumulative return.

We deviate slightly from the conventional weighting scheme. To remove the negative returns associated with short-term reversal, we assign zero weight to the first two weeks (the 10 most-recent trading days). Next, we use linearly increasing weights for the next 21 trading days, followed by constant weights over the next 200 trading days. Finally, over the last 21 trading days, we use linearly declining weights. The linear ramps mitigate spurious jumps in exposure when large returns suddenly enter or exit the window. Momentum is a single-descriptor factor.

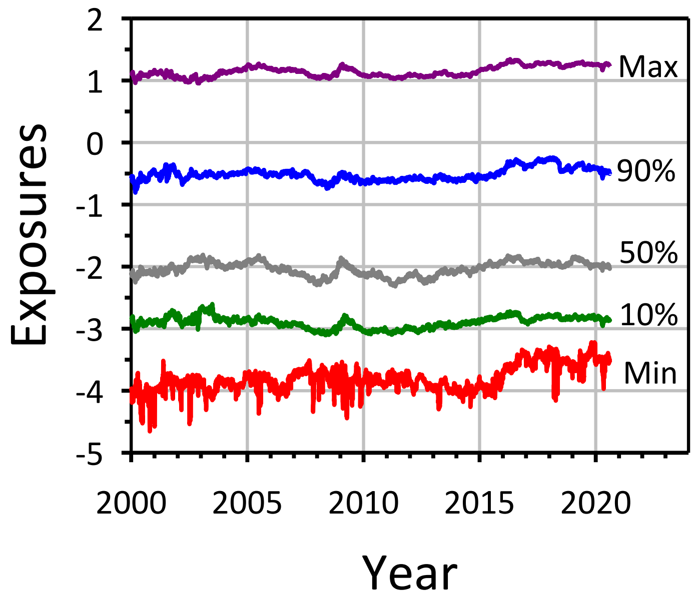
Special treatment is required of IPOs. On the IPO launch date, the stock lacks a return history, so the raw momentum cannot be computed. In this case, we set the raw momentum equal to the cap-weighted mean raw momentum of all ESTU stocks within the particular country. Since standardized factors are cap-weighted mean zero, this implies zero standardized exposure to momentum. After acquiring sufficient return history, we begin to blend the raw momentum of the IPO with the cap-weighted mean of the ESTU. As we continue acquiring more history, we assign increasing weight to the raw momentum. After acquiring 12 months of returns, we assign full weight to the raw momentum, as shown in Equation (2D.3).

Example. In Figure 2.2, we plot the distribution of momentum exposures for the MAC3 Korea Equity Model. In contrast to the residual volatility factor, the momentum exposures are quite symmetric in their distribution. In particular, the minimum exposure is roughly -3, the median exposure is close to zero, and the maximum exposure is approximately +3.

Size. The size effect was first documented by Banz (1981), who found empirical evidence to support the view of a "small-cap premium." More specifically, he found that small-cap stocks tended to outperform large-cap stocks on a risk-adjusted basis.

The distribution of market capitalization typically spans several orders of magnitude. To obtain a well-behaved distribution, we define the size factor as the log of market capitalization (CUR_MKT_CAP). If the small-cap premium holds, we expect the size pure factor portfolio to exhibit negative drift.

Figure 2.3. Distribution of size exposures for the MAC3 Australia Equity Model. For each date, we plot the minimum, 10-percentile, median, 90-percentile, and maximum exposure.



Example. In Figure 2.3, we plot the distribution of size exposures for the Australia Equity Model. In this case, large-cap stocks have positive exposure, while mid-cap and small-cap stocks have negative exposure. Since the distribution is cap-weighted mean zero, the entire distribution is shifted downward. As a result, while the maximum size exposure is typically slightly greater than +1, the minimum exposure is approximately -4. Note that the median exposure is roughly -2.

Mid-cap. The size factor alone may not fully capture the component of stock return associated with market capitalization. In particular, the return of the mid-cap segment may deviate from that explained by the size factor alone.

To capture this effect, we use the following raw descriptor for the mid-cap factor:

$$y_n = \exp\left[-a\left(x_n - \tilde{x}\right)^2\right] , \tag{2D.4}$$

where x_n is the standardized size exposure, and $\tilde{x}$ and a are empirically determined parameters. When the raw descriptor is standardized as a z-score, we find that mid-cap stocks have positive exposure to the factor, while large-cap stocks and small-cap stocks have negative exposures. Hence, when mid-cap stocks outperform the other segments of the market, the return to the mid-cap factor is typically positive.

Earnings Yield. Value investors have long held the view that stocks which are inexpensive relative to their fundamentals should outperform over the long run. One of the leading value signals is the earnings-to-price ratio. Basu (1977) studied this effect, showing that stocks with high earnings-to-price ratios have tended to outperform those with low ratios.

Our earnings yield factor is composed of two descriptors. The first descriptor is the historical earnings-to-price ratio (HEP), which is given by the trailing 12-month net income (NET_INCOME) divided by total market capitalization (CUR_MKT_CAP).

The second descriptor is the 12-month forward earnings-to-price ratio (FEP). This descriptor is computed using Bloomberg Estimates (BEST_EPS), which represent the mean earnings-per-share predicted by analysts for a given fiscal year. The FEP descriptor is found by taking a weighted average between the current fiscal year earnings forecast (B_1) and the next fiscal year forecast (B_2) ,

$$FEP = \frac{w_1 B_1 + (1 - w_1) B_2}{P} \quad , \tag{2D.5}$$

where w_1 is the weight assigned to the current fiscal year, and P is the last available price-per-share (PX_LAST).

For example, consider a company that has a fiscal year coinciding with the calendar year. Suppose that our analysis date is 31-Mar-2015. Since there are nine months remaining in the fiscal year, we assign weight of 0.75 to the current fiscal year and 0.25 to the next fiscal year. Six months later (30-Sep-2015), we have three months remaining in the fiscal year, and we assign weight 0.25 to the current fiscal year (2015) and 0.75 to the next fiscal year (2016).

Valuation. Earnings yield is not the only signal that can be associated with "value." Other widely used value metrics include book-to-price (B2P), sales-to-price (S2P), and cash-flow-to-price (CFP) ratios. We use all three descriptors to construct the valuation factor.

The B2P descriptor is found by taking the book value of the company (TOT_COMMON_EQY) and dividing by the market capitalization (CUR_MKT_CAP). The S2P descriptor is given by total sales (SALES_REV_TURN) divided by market capitalization. Finally, the CFP descriptor is the cash flow from operations (CF_CASH_FROM_OPER) divided by market capitalization.

Dividend Yield. The dividend-to-price ratio is another characteristic that sometimes falls under the "value" umbrella. For instance, Naranjo *et al*. (1998) found evidence of a positive relationship between stock returns and dividend yield.

We use indicated dividend yield (DIVIDEND_INDICATED_ YIELD) to construct the dividend yield factor. The indicated dividend yield represents the most recently announced dividend (annualized) divided by the current market price.

Long-Term Reversal. Empirically, De Bondt and Thaler (1985) demonstrated that stocks which performed poorly over the last 3-5 years tended to outperform over subsequent periods. This effect is known as long-term reversal (LTR). A possible behavioral explanation for this effect is that investors may over-react to a long string of positive or negative news. Once investors realize their "mistake," the stock price reverts.

The LTR factor is constructed in a similar fashion to momentum, as in Equation (2D.3), except that the two factors employ different weighting schemes. More specifically, to remove the effect of momentum, LTR assigns zero weight to the most recent 12 months. Over the 13th month, the weights linearly increase. For months 14-47, the weights are held constant. Finally, over the 48th month, the weights linearly decline to zero.

For IPOs, we follow the same treatment as for momentum. Namely, as of the IPO launch date, we set the raw LTR equal to the cap-weighted mean of the ESTU for the particular country, which implies a standardized factor exposure of zero. As the IPO builds up sufficient return history, we begin blending the raw LTR with the cap-weighted mean LTR. After acquiring four years of return history, we assign 100% weight to the raw LTR.

Liquidity. The liquidity factor is composed of three descriptors: (a) share turnover, (b) bid-ask spread, and (c) the modified Amihud measure. For some descriptors, as explained below, we flip the sign so that positive standardized exposure is associated with higher liquidity.

One widely used measure of liquidity is share turnover, which measures the fraction of shares outstanding that trade over a given window. Let $V_{n\tau}$ be the number of shares (PX_VOLUME) of stock n that trade on day τ. Let $S_{n\tau}$ denote the total number of shares outstanding (EQY_SH_OUT) for stock n on day τ. Share turnover (STO) is defined as the EWMA mean fraction of shares that trade over an expanding window

$$STO_{nt} = \sum_{\tau=t_n}^{t} w_{n\tau} \left(\frac{V_{n\tau}}{S_{n\tau}} \right), \tag{2D.6}$$

where t_n is the inception date of the stock (first trade) and $w_{n\tau}$ is the EWMA weight assigned to day τ (defined by the HL parameter). If there is no trading volume on a given day, then $V_{n\tau} = 0$. All else equal, higher share turnover is associated with higher liquidity.

Bid-ask spread (BAS) represents a measure of *illiquidity*. Amihud and Mendelson (1986) wrote a classic paper describing the relationship between BAS and asset pricing. Let $P_{n\tau}^{A}$ be the closing ask-price (PX_ASK) for stock n on day τ, let $P_{n\tau}^{B}$ be the corresponding bid-price (PX_BID), and let $P_{n\tau}$ denote the closing price (PX_LAST). The BAS is computed as

$$BAS_{nt} = (-1) \sum_{\tau=t_n}^{t} w_{n\tau} \left(\frac{P_{n\tau}^{A} - P_{n\tau}^{B}}{P_{n\tau}} \right), \tag{2D.7}$$

where $w_{n\tau}$ is the EWMA weight assigned to day τ. Note that we have flipped the sign of the descriptor, so that stocks with low BAS will have positive standardized exposure and thus be associated with higher liquidity.

Another widely used measure of *illiquidity* was provided by Amihud (2002). This measure uses the ratio of daily absolute stock return to daily dollar trading volume. However, since large-cap stocks tend to have much higher daily dollar trading volume, to a large extent this definition is a proxy for inverse size. To mitigate collinearity with the size factor, we modify the original definition used by Amihud, as described below.

Let $r_{n\tau}$ be the local return of stock n on day τ, let $V_{n\tau}$ be the number of shares (PX_VOLUME) that trade on day τ, and let $S_{n\tau}$ denote the total number of shares outstanding (EQY_SH_OUT) on day τ. The modified Amihud descriptor is defined as

$$A_{nt} = (-1)\sum_{\tau=t_n}^{t} w_{n\tau}\left(\frac{|r_{n\tau}|}{(V_{n\tau}/S_{n\tau})}\right), \tag{2D.8}$$

where $w_{n\tau}$ is the EWMA weight assigned to day τ. Once again, we have flipped the sign of the descriptor so that stocks with small absolute returns and large share turnover have positive standardized exposure and are associated with higher liquidity.

Growth. The growth factor consists of four descriptors: historical sales growth (HSG), historical earnings growth (HEG), analyst-predicted medium-term growth (MTG), and analyst-predicted long-term growth (LTG).

We compute HSG using sales (SALES_REV_TURN) and total assets (BS_TOT_ASSET). More specifically, we do a time-series regression (including intercept) of annual sales versus time over the trailing five years. The slope coefficient represents the average annual growth in sales. This quantity is normalized by dividing by the mean total assets of the previous five years. HEG is computed in the exact same fashion, except that we use earnings (NET_INCOME) in place of sales (SALES_REV_TURN).

The raw descriptor for MTG is computed by taking the ratio of the Bloomberg forecast EPS (BEST_EPS) for Fiscal Year 3 (FY3) and Fiscal Year 2 (FY2). The raw descriptor for LTG is given by the mean analyst-predicted earnings growth over the next 3-5 years.

Variability. This factor is often regarded as a "quality" signal. In particular, stocks with low variability (i.e., greater stability) are typically associated with higher quality. The variability factor is composed of three descriptors: variability in net income (VNI), variability in sales (VSA), and variability of cash flow (VCF).

The first step in constructing the variability descriptors is to compute the standard deviation of the appropriate annual measure over the trailing five years. For instance, VNI uses the standard deviation of annual net income (NET_INCOME) over the last five years. Similarly, VSA uses annual sales (SALES_REV_TURN), while VCF uses annual cash flow (CF_CASH_FROM_OPER).

Clearly, large companies tend to have higher standard deviations of net income, sales, and cash flow. Hence, the final step in constructing the raw descriptor is to normalize the standard deviation by dividing by mean total assets (BS_TOT_ASSET) over the last five years.

Profit. This factor is also often regarded as an indicator of investment "quality." That is, highly profitable companies are typically considered as quality stocks.

The profit factor is composed of three descriptors: return on assets (ROA), return on equity (ROE), and profit margin (PRM). The ROA descriptor is computed as net income (NET_INCOME) divided by total assets (BS_TOTAL_ASSETS). The ROE descriptor is given by net income divided by book value (TOT_COMMON_EQY). Finally, the PRM descriptor is computed as net income divided by sales (SALES_REV_TURN).

Leverage. Many investors also view leverage as a component of investment "quality." More specifically, highly levered firms are usually regarded as being of lower quality.

The leverage factor is based on three descriptors: asset leverage (D2A), book leverage (D2B), and market leverage (D2M). All three ratios employ the same numerator, which is given by the sum of long-term debt (BS_LT_BORROW) and short-term debt (BS_ST_BORROW). For D2A, the denominator is total assets (BS_TOT_ASSET), while for D2B the denominator is the book value (TOT_COMMON_EQY), and for D2M the denominator is the total market capitalization (CUR_MKT_CAP).

(E) Industry Beta Factors

Industry factors are based on the Bloomberg Industry Classification System (BICS). The BICS structure contains 11 sectors at the top level, which are subdivided into 20 industry groups, 59 industries, and 214 sub-industries.

The twin objectives of industry factor identification are: (a) to maximize the explanatory power of the model, and (b) to maximize the statistical significance of the factor returns. Unfortunately, these two objectives tend to be in conflict with each other.

At one extreme, we could include every sub-industry as a separate factor. Using the most granular set of industry factors maximizes the in-sample explanatory power (R^2) of the model. One problem with this approach, however, is that in most markets at least some of the sub-industries would be empty. Even if all sub-industries were populated, many would be very thin (i.e., few stocks), leading to industry factor returns dominated by noise and thus having very low statistical significance.

At the other extreme, we could employ the 11 BICS sectors as our industry factors. The advantages of this approach are that it mitigates thin industries and the factor returns generally have high statistical significance. The disadvantage of this approach is that the sectors are not sufficiently granular to capture high explanatory power.

Our objective in identifying industry factor structure is to capture the vast majority of explanatory power that can be obtained from the sub-industries, while avoiding thin industries and retaining a high level of statistical significance. This exercise requires extensive empirical analysis, testing which combinations of industries and sub-industries can be effectively grouped together to achieve our twin objectives of high explanatory power and high statistical significance.

With the industry factors thus defined, the traditional approach assigns binary values for the industry exposures. That is, a stock exposure to an industry is equal to 1 if the stock is a member of the industry, and is equal to 0 otherwise. In this approach, if the industry factor return is say 100 bps, it contributes 100 bps to the return of *every* stock in the industry (i.e., homogeneous impact).

It is unrealistic, however, to assume that all stocks within an industry will respond in an identical fashion to movements in the underlying industry factor. Naturally, as with market betas, we expect some stocks to be more sensitive to industry movements, while others to be less sensitive. This effect can be captured through *industry betas*.

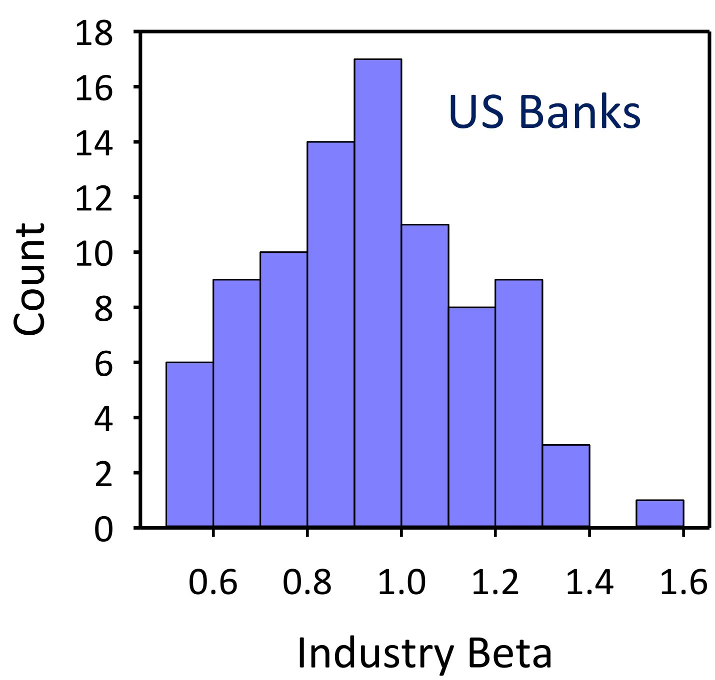
Industry betas are estimated in a similar fashion as market betas, but they are standardized differently. The first step in estimating industry betas is to regress the local excess stock returns (r_{nt}) against the cap-weighted local excess return (R_{it}) of the industry to which it belongs,

$$r_{nt} = \beta_{ni} R_{it} + e_{nt} \ . \tag{2E.1}$$

Here, β_{ni} is the beta of stock n to industry i. When doing the time-series regression, we employ the same estimation techniques as used for the market beta factor.

As with raw market betas, we expect the mean industry beta to be close to 1, with some dispersion around the mean. Estimated betas may be unreliable if the stock has a high frequency of missing/zero returns. For such stocks, we shrink the industry betas towards 1, with the shrinkage intensity depending on the frequency of missing/zero returns. For instance, recently issued IPOs have a shrinkage intensity of 100%, whereas stocks with no missing returns would have shrinkage intensity of zero.

Figure 2.4. Distribution of industry beta exposure for US banks in the MAC3 US Equity Model. The analysis date was 30-Apr-2020.



Next, we trim extreme betas at 0.6 below and 1.0 above the cap-weighted mean industry beta. That is, the raw industry betas will be roughly contained within the interval [0.4, 2.0], with a cap-weighted mean close to 1.0. The final step is to rigidly shift the distribution slightly so that the cap-weighted mean industry beta is *exactly* 1.

This standardization convention preserves two intuitive features of the model. First, it ensures that the market portfolio (ESTU) will have an exposure to each industry exactly given by its cap weight in the industry. The second intuitive result is that it preserves the interpretation of the market factor as the cap-weighted market portfolio, as described below in Section (3C).

Example. In Figure 2.4, we present a histogram of industry beta exposures for US banks in the MAC3 US Equity Model on analysis date 30-Apr-2020. The industry betas are centered about 1, and fall within the range from 0.5 to 1.6.

We find several important advantages to using industry betas rather than (0,1) exposures. First, we find a significant increase in the explanatory power of the model. It should be stressed that this increase in explanatory power is achieved solely through improvement in factor definition, not by simply adding more factors to the model.

Second, we find that industry betas increase the statistical significance of the industry factor returns. This is not surprising, since we increase the explanatory power of the model while keeping the number of factors fixed.

The third and most important benefit of using industry betas is that it improves the model specification. A basic premise of factor models is that the idiosyncratic returns are uncorrelated with the factor returns. However, if we used the traditional (0,1) exposures, we would find that stocks with high industry betas would tend to have specific returns positively correlated with the industry factors, while low industry beta stocks would tend to be negatively correlated. The use of industry betas helps mitigate spurious correlations between factor returns and specific returns.

(F) Country Beta Factors

In multi-country models, country factors represent an important component in the cross-sectional variation of stock returns. As with industry factors, the traditional approach is to assign (0,1) exposures, depending on whether or not the company is classified within the country.

Just as we find significant benefits to using industry beta factors, the same holds true for country beta factors. That is, employing country beta factors, we find: (a) an increase in the explanatory power of the model, (b) an increase in the statistical significance of the factor returns, and (c) an improvement in model specification by mitigating spurious correlations between factor returns and non-factor (idiosyncratic) returns.

Country betas are estimated and standardized in the same way as industry betas. Namely, to obtain the raw country betas, we regress the local excess returns of each stock against the local excess returns of the cap-weighted country portfolio to which it belongs. The raw betas are then trimmed for outliers at 0.6 below and 1.0 above the cap-weighted mean country beta. The final step is to standardize the country betas to be cap-weighted mean 1.

Note that single-country models do not contain country-beta factors, since that source of return variation is already captured by the market beta (style) factor. An exception to this rule occurs for *satellite countries* within a single-country model. For instance, Canada is a satellite factor within the US risk model. The exposures of Canadian stocks to the Canada satellite factor are given by the country betas.

(G) Coverage Universe and Treatment of Linked Securities

Before describing the coverage universe for each model, we quickly review the three model choices available within the MAC3 suite of equity models. The first set of models are the local models, which employ a detailed set of local factors to describe the risk characteristics of individual markets. The MAC3 equity suite contains 13 such local models. The second model choice is the MAC3 Global Equity Model, which contains a parsimonious set of global factors. The third model choice is the MAC3 Integrated Model, which represents the collection of all local models.

Coverage Universe. The MAC3 Global Equity Model and the MAC3 Integrated Model span the same coverage universe. They aim to cover the broadest range of security types, including common stocks, REIT's, Unit Trusts, stapled securities, depository receipts, and preferred stocks.

Local models cover only stocks that are associated with the local market through one of the following four fields: (1) Country of Risk, (2) Country of Domicile, (3) Country of Incorporation, or (4) Country of Issue. For example, Sony ADR trading in New York has US Country of Issue, hence it is covered by the US equity model. By contrast, the parent security (Sony trading in Tokyo), does not meet any of the four criteria, hence it is not covered by the US model. Microsoft trading in Frankfurt is covered by the Europe model since the security has German Country of Issue, but also by the US model through the US Country of Risk.

Exposures for Linked Securities. We now describe rules for assigning exposures to linked securities in the MAC3 suite of equity models. In the Global Equity Model and the Integrated Model, all factor exposures from the Fundamentals Ticker (the "parent") are copied over to the linked securities (the "children"). For example, Sony ADR (SNE US) trades in New York, while the Fundamentals Ticker is Sony Japan (6758 JP), which trades in Tokyo. Hence, in the Global Equity Model or the Integrated Model, the factor exposures of Sony ADR and Sony Japan are identical, meaning that Sony ADR is exposed to Japanese equity factors.

The local models cover only a subset of the securities. Hence, in a local model, the Fundamentals Ticker may or may not be covered by the model. If the Fundamentals Ticker *is* covered by the local model, we again simply copy the exposures of the Fundamentals Ticker over to the linked security.

For example, Google (Alphabet) contains two share classes: A and C. The Fundamentals Ticker is Google A (GOOGL US). Since the Fundamentals Ticker is covered in the US equity model, the exposures of Google A are simply copied over to Google C (GOOG US).

The second possibility is that the Fundamentals Ticker is *not* covered by the model. In this case, we compute all technical exposures using local excess returns of the security in question. The fundamental exposures are obtained by importing the raw descriptors of the Fundamentals Ticker and standardizing them with respect to the ESTU of the local model.

For example, consider Sony ADR from the perspective of the MAC3 US Equity Model. In this case, we take the local excess returns of Sony ADR (in USD) and use them to compute all technical exposures (e.g., market beta, industry beta, momentum, etc.). For the fundamental factors (e.g., earnings yield, leverage, etc.), we import the raw descriptors from the Fundamentals Ticker (i.e., Sony Japan) and standardize them relative to the US ESTU.

Section (3): Factor Returns

(A) Cross-sectional Regression

Factor returns are estimated on a daily basis. Let $\mathbf{X}$ denote the $N\text{x}K$ factor exposure matrix, where N is the number of stocks in the estimation universe, and K is the number of factors in the model. The $K\text{x}1$ vector of factor returns ($\mathbf{f}$) is estimated by regressing the $N\text{x}1$ vector of *local excess* returns ($\mathbf{r}$) against the factor exposures,

$$\mathbf{r} = \mathbf{X}\mathbf{f} + \mathbf{u} \ , \tag{3A.1}$$

where $\mathbf{u}$ is the $N\text{x}1$ vector of specific returns (i.e., unexplained residuals).

In single-country models, to retain an intuitive interpretation for the market factor, we impose the constraint that the cap-weighted industry factor returns add to zero,

$$\sum_{i=1}^{I} w_i f_i = 0 \ , \tag{3A.2}$$

where w_i is the cap weight of industry i, f_i is the industry factor return, and I is the total number of industry factors. As described by Menchero (2010), the regression solution can be written as

$$\mathbf{f} = \mathbf{\Omega}\mathbf{r} \ , \tag{3A.3}$$

where $\mathbf{\Omega}$ is the $K\text{x}N$ matrix of pure factor portfolio weights, given by

$$\mathbf{\Omega} = \mathbf{R}\left(\mathbf{R}'\mathbf{X}'\mathbf{V}\mathbf{X}\mathbf{R}\right)^{-1}\mathbf{R}'\mathbf{X}'\mathbf{V} \ . \tag{3A.4}$$

Here, $\mathbf{R}$ is the $K\text{x}(K-1)$ constraint matrix, and $\mathbf{V}$ is the $N\text{x}N$ diagonal matrix of regression weights (described in the next section). Note that the element in row k and column n of matrix $\mathbf{\Omega}$ (i.e., Ω_{kn}) gives the weight of stock n in pure factor portfolio k.

The constraint equation can be expressed in matrix form as

$$\mathbf{f} = \mathbf{R}\mathbf{g} \ , \tag{3A.5}$$

where $\mathbf{g}$ is a $(K-1)$ vector of auxiliary variables. The constraint equation is best illustrated by example. Suppose that there are three factors and the weighted sum is set equal to zero, i.e., $w_1 f_1 + w_2 f_2 + w_3 f_3 = 0$. In this case, the constraint equation is

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -(w_1/w_3) & -(w_2/w_3) \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} . \tag{3A.6}$$

For any set of values g_1 and g_2, it is easily verified that $w_1 f_1 + w_2 f_2 + w_3 f_3 = 0$ is always satisfied.

Multi-country models also contain country factors. In this case, an additional constraint is imposed such that the cap-weighted country factor returns sum to zero,

$$\sum_{c=1}^{C} w_c f_c = 0 \ , \tag{3A.7}$$

where w_c is the cap weight of country c, f_c is the country factor return, and C is the total number of country factors[5].

Since multi-country models impose two constraints, the constraint matrix now has K rows and $K-2$ columns. The general form of the constraint matrix is given by

$$\mathbf{R} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_C \end{bmatrix} , \tag{3A.8}$$

where $\mathbf{R}_I$ is an $I\text{x}(I-1)$ industry constraint matrix of the form given in Equation (3A.6), $\mathbf{R}_C$ is a $C\text{x}(C-1)$ country constraint matrix of the same form, and $\mathbf{I}$ is the identity matrix of dimension $(K-I-C)$. Note that the factors that map onto block $\mathbf{I}$ are the market factor and the style factors.

(B) Regression Weights and Return Winsorization

Regression Weights. The traditional regression weighting scheme for equity factor models uses the square root of market capitalization. The rationale for this choice is that since small-cap stocks are "noisier" than their large-cap counterparts, they should be down-weighted in the regression.

While this is the right intuition, square root of market capitalization is not the optimal set of regression weights. From an econometric view point, the optimal weights to use are inverse specific variance,

[5] Satellite countries are excluded from the constraint equation since they do not enter the main regression.

$$\mathbf{V} = \text{var}^{-1}(\mathbf{u}) \ , \tag{3B.1}$$

as this minimizes the noise in the factor return estimates. In the MAC3 equity risk models, we use inverse *residual* variance for the regression weights, where the residual variance is taken directly from the residual volatility style factor, as described in Section (2D). Empirically, we find that inverse residual variance is an excellent proxy for inverse specific variance, and these regression weights produce significantly more precise factor return estimates (lower noise) than square-root-of-market-cap weights.

Stock Return Winsorization. From Equation (3A.3), we see that factor returns depend on the stock return vector $\mathbf{r}$. Extreme stock returns—whether legitimate or spurious—can thus have an undue impact on factor returns. To ensure model robustness, we must impose winsorization limits on stock returns.

While some kind of trimming of stock returns is necessary, it should be approached with caution, since excessive trimming may induce spurious drifts in the long-term factor returns. For example, consider the size pure factor portfolio, which takes a net long position in large-cap stocks and a net short position in small-cap stocks. Since small-cap stocks with large positive returns are the observations most likely to be winsorized, this has the effect of inducing a spurious upward drift in the size factor return. Note that in most markets, the long-term drift of the size factor is negative, consistent with the small-cap premium. However, if the trimming is sufficiently aggressive, this may be enough to flip the sign on the long-term return. Clearly, this would be quite problematic for any type of long-term performance attribution analysis. The trimming limits are therefore set to ensure a robust production environment, while not leading to a material change in the long-term drift of the pure factor portfolios.

(C) Interpretation of Pure Factor Portfolios

Using a model with an intercept term representing the market factor (f_M), a set of country and industry betas, and a set of style factors, the local excess return of stock n may be written

$$r_n = f_M + \sum_c \beta_{nc} f_c + \sum_i \beta_{ni} f_i + \sum_s X_{ns} f_s + u_n \ , \tag{3C.1}$$

where β_{nc} is the beta of stock n to country c, β_{ni} is the beta of stock n to industry i, and X_{ns} is the stock exposure to style factor s.

We wish to preserve the intuitive interpretation of the market factor as representing the cap-weighted market portfolio. Let R_M be the local excess return of the market portfolio (ESTU), which has weights w_n^M. The market portfolio return can be written as

$$R_M = \sum_n w_n^M r_n = f_M + \sum_n \sum_c w_n^M \beta_{nc} f_c + \sum_n \sum_i w_n^M \beta_{ni} f_i + \sum_n \sum_s w_n^M X_{ns} f_s + \sum_n w_n^M u_n \ . \tag{3C.2}$$

Using the fact that the industry/country exposures of the market portfolio are given by the cap weights of the market portfolio in these segments, we find

$$R_M = f_M + \sum_c W_c^M f_c + \sum_i W_i^M f_i + \sum_n \sum_s w_n^M X_{ns} f_s + \sum_n w_n^M u_n \ , \tag{3C.3}$$

where W_c^M is the cap weight of the market portfolio in country c, and W_i^M is the corresponding weight in industry i. As described in Section (3A), we impose two constraints on the factor returns. First, the cap-weighted country factor returns sum to zero,

$$\sum_c W_c^M f_c = 0 \ . \tag{3C.4}$$

Second, the cap-weighted industry factor returns also sum to zero, i.e.,

$$\sum_i W_i^M f_i = 0 \ . \tag{3C.5}$$

Finally, note that the style factors are standardized to be cap-weighted mean zero,

$$\sum_n w_n^M X_{ns} = 0 \ . \tag{3C.6}$$

Hence, Equation (3C.3) simplifies to

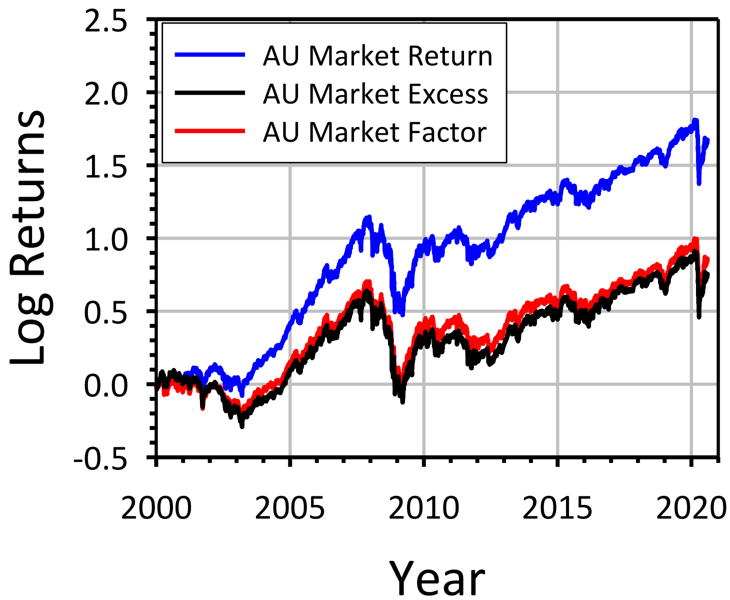
$$R_M = f_M + \sum_n w_n^M u_n \ . \tag{3C.7}$$

The idiosyncratic component is diversifiable and typically insignificant in magnitude compared to the market factor. This result shows that under this model specification, the market factor return can be intuitively interpreted as the local excess return of the cap-weighted market portfolio.

Example. In Figure 3.1, we plot cumulative log returns for: (a) the Australia cap-weighted market portfolio, (b) the market portfolio *excess* return (above the risk-free rate), and (c) the market factor from the MAC3 Australia Equity Model. Note that the market excess return and the market factor return track each other very closely (above 99% correlation) over the sample period, with each earning log returns of about 80%. By contrast, the cumulative log return of the market portfolio was much higher (above 160%). This large return difference is attributable to the relatively high risk-free rate in Australia over this period.

The market factor is the only factor that is 100% net long. All other factor portfolios are strictly dollar neutral. Style pure factor portfolios have unit exposure to their own style, zero exposures to all other styles, and have zero return contribution from industries and countries. Country pure factor portfolios are long a portfolio with unit exposure to the country, are short an equal dollar amount of the market, and have zero exposure to every style factor and zero return contribution from industries. Similarly, industry pure factor portfolios are long a portfolio with unit exposure to the industry, are short the market, and have zero exposure to every style and zero return contribution from countries.

Figure 3.1. Cumulative log returns for the Australia market. The blue line is the local return of the ESTU, the black line is the local excess return (above the risk-free rate) of the ESTU, and the red line is the cumulative log return of the market factor in the MAC3 Australia Equity Model.



Just as factor returns represent the returns of well-defined pure factor portfolios, specific returns also represent the returns of well-defined *pure specific portfolios*. These portfolios can be easily interpreted by writing the specific return as the difference between the local excess return of the stock and the factor return contribution,

$$u_n = r_n - \sum_k X_{nk} f_k \ . \tag{3C.8}$$

Hence, pure specific portfolios are strictly dollar neutral. In particular, they go 100% long the stock, go 100% short the market factor, and take short positions X_{nk} in each of the pure factor portfolios k. As a result, pure specific portfolios are void of factor risk.

(D) Estimation of Satellite Factor Returns

The primary ESTU of a model represents the set of stocks that are used to estimate the returns for the market, industry, and style factors. In a multi-country model, it is also used to estimate the country factor returns. As described in Section (2A), special care must be taken to ensure that the ESTU reflects the broad market (i.e., the investable universe) and does not include extremely small or illiquid stocks.

By contrast, the coverage universe should be as broad as possible, ideally consisting of all stocks—no matter how small or illiquid. For instance, in the MAC3 Global Equity Model, we aim to cover all frontier markets, even though frontier markets do not form part of the investable universe for a majority of portfolio managers. Hence, our objective is to cover frontier markets and explain their unique characteristics with dedicated factors, without having frontier markets distort the risk and return of the other pure factor portfolios. This objective is accomplished by means of *satellite* factors, which capture the risk associated with countries that are not part of the primary ESTU, but without impacting the other factor returns in the model.

As an example, Canada is treated as a satellite factor in the MAC3 US Equity Model. This means that Canadian stocks are covered by the model, but they are *not* in the pure factor portfolios of the US market factor, industry factors, or style factors.

While US factors will surely explain a significant portion of Canadian stock returns, it is naïve to think that Canadian stocks are "just like" US stocks. The unique behavior of Canadian stocks is captured by means of the satellite country factor.

Note that stocks in satellite countries have exposures to all factors in the model. That is, in the MAC3 US Equity Model, Canadian stocks have exposure to the US market factor, industry factors, and style factors.

Satellite factor returns are estimated by subtracting the return contributions of the factors (market, industry, and styles) from the local excess return of each stock, and then computing the univariate slope coefficient of the resulting residuals. More specifically, let r_n be the (trimmed) local excess return of stock n in satellite country s, and let X_{nk} be the stock exposure to factor k. We compute the residual returns by subtracting the return contributions of the market, industry, and style factors,

$$e_n \equiv r_n - \sum_k X_{nk} f_k \quad . \tag{3D.1}$$

Finally, we compute the satellite factor return f_s by univariate cross-sectional regression,

$$e_n \equiv \beta_{ns} f_s + u_n \quad , \tag{3D.2}$$

where β_{ns} is the country beta of stock n relative to the satellite country s. The regression solution is given by

$$f_s = \frac{\sum_n v_n \beta_{ns} e_n}{\sum_n v_n \beta_{ns}^2} \quad , \tag{3D.3}$$

where v_n is the regression weight of the stock. Note that Equation (3D.3) produces *identical* factor returns as would be obtained if the satellite country were included in the main regression, but with the regression weights for stocks in the satellite country multiplied by an infinitesimal constant.

Section (4): Factor Volatilities

(A) Term Structure of Risk

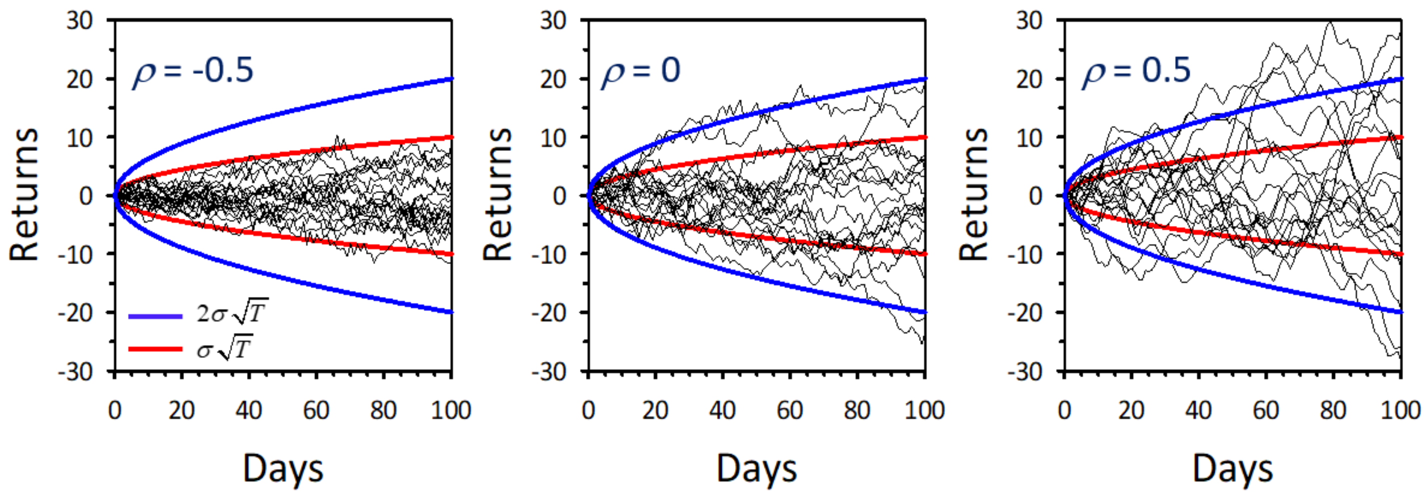
The MAC3 suite of equity models employs high-frequency observations (daily) to forecast risk at multiple horizons, including daily, weekly, monthly, quarterly, annual, and long-term. Utilizing high-frequency observations to forecast risk brings two significant advantages to the task of volatility forecasting: (1) It increases the number of observations, which can be used to significantly reduce the sampling error ("noise") in the volatility forecasts, and (2) high-frequency observations contain important information about the serial correlation of daily factor returns, which can in turn be utilized to refine and improve risk forecasts at multiple horizons.

When considering how risk forecasts scale across time, it is important to distinguish between the *reporting horizon* and the *prediction horizon*. The reporting horizon simply specifies the *units* of volatility, whereas the prediction horizon refers to the length of time over which the volatility is being forecast. Both horizons must be specified in order to uniquely define the risk forecast.

The most common reporting horizon is one year, referred to as *annualized* risk forecasts. For instance, suppose our best estimate of one-day volatility is 20 bps per day. In this case, both the reporting horizon and the prediction horizon correspond to one day. The reporting horizon can be easily changed by applying the *square-root-of-time scaling rule*. In this example, given 252 trading days per year, the annualized volatility of the one-day prediction horizon would be $20\sqrt{252}$, or 3.17%. The same one-day volatility forecast, instead stated with a one-week reporting horizon, would be $20\sqrt{5}$, or 44.7 bps.

The origin of the square-root-of-time scaling rule rests on the assumption that returns from one period to the next are independent (i.e., no serial correlation) and stationary. When this holds, if the one-period volatility is σ, the *T*-period volatility is $\sigma\sqrt{T}$. Even if independence does not hold, the rule is still applied in practice to translate volatility forecasts to any desired reporting horizon. Note that changing the reporting horizon is a trivial transformation that does not in any way reflect new information regarding the volatility forecast.

Figure 4.1. Simulated results (for 20 paths) showing the effect of serial correlation. The left panel is simulated using serial correlation of -0.50. The center panel is for zero serial correlation, and the right panel is for a serial correlation of 0.50. If the square-root-of-time scaling rule holds, then approximately 68% of the paths should fall within the red envelope (1σ), while 95% of the paths should fall within the blue envelope (2σ). This result indeed roughly holds for the center panel. However, the left panel has far too few exceedances, while the right panel shows far too many.



The *term structure of risk*, by contrast, refers to how volatility forecasts scale with *prediction* horizons. In this case, changing the prediction horizon is a non-trivial exercise that requires advanced econometric techniques to properly estimate volatility from the high-frequency observations.

The term structure of risk arises from non-stationarity and serial correlation in daily factor returns. Non-stationarity is handled through the use of EWMA, which assigns more weight to recent observations and implicitly recognizes that volatilities change across time. Our focus here, however, is to understand the effect of serial correlation on volatility forecasts at different horizons.

If the one-day volatility forecast is σ and daily factor returns exhibit positive serial correlation, then the *T*-period volatility will be greater than $\sigma\sqrt{T}$. On the other hand, if daily factor returns have negative serial correlation, the volatility will be less than $\sigma\sqrt{T}$. This is intuitive, since an asset that tends to move in opposite directions from one day to the next will have a high one-day volatility, but the returns will diffuse slowly over time (i.e., lower volatility over longer horizons).

Example. This behavior is clearly illustrated in Figure 4.1, which simulates results for 20 random paths over 100 days under three different serial-correlation assumptions. The volatility of one-day returns (normally distributed) in each case was 1%. The center panel is for zero serial correlation, in which case the square-root-of-time scaling holds. Hence, the volatility over a 100-day window is $\sqrt{100}$, or 10%. We expect approximately 68% of the paths to fall within 1σ of the mean (as given by the red line), while roughly 95% of the paths should fall within 2σ of the mean (as given by the blue line). The center panel is clearly consistent with these expectations.

The left panel in Figure 4.1 is for a serial correlation of -0.50. In this case, it is clear that the square-root-of-time scaling breaks down, since almost no paths fall outside of the 1σ interval. We also note that in the left panel, positive returns one day are likely to be followed by negative returns the next day, which serves to dampen the volatility over longer horizons. Similarly, in the right panel, which corresponds to a serial correlation of 0.50, far too many observations fall outside of the 1σ and 2σ envelopes. We also observe in the right panel that assets are more likely to have long strings of consecutive positive or negative returns.

Global factors often exhibit positive serial correlation in daily factor returns. One basic mechanism that drives this effect is asynchronous trading. For instance, on any given calendar date, the 24-hour trading day (market close to market close) in the US market overlaps significantly with the next calendar day of trading in Japan. Hence, a big move in the US market on a given day is reflected in the *next-day* returns in Japan.

Serial correlation in daily factor returns may exist even in the absence of asynchronous trading. For instance, the serial correlation of daily factor returns for the momentum factor in the MAC3 US Equity Model was 0.25 over the 21-year period running from 31-Dec-1998 to 30-Apr-2020. The US market factor had a serial correlation of -0.08 over the same period. The standard error of the correlation estimate is about 0.014, so both correlations are of high statistical significance.

Table 4.1. Annualized volatilities of the market factor and the momentum factor from the MAC3 US Equity Model over the period 31-Dec-1998 to 30-Apr-2020.

Aggregation Window Size	US Market Volatility	US Momentum Volatility
Day	19.56%	3.43%
Week	18.11%	4.27%
Month	17.04%	4.99%
Quarter	15.89%	5.17%

Example. To illustrate the effect of serial correlation on volatility over different horizons, we aggregated market and momentum factor returns over window sizes ranging from one day to one quarter. We then rolled the window forward one day at a time and computed the realized volatility (annualized) over the full sample period.

In Table 4.1, we report the resulting volatilities for various aggregation windows. The quarterly volatility of the market factor was considerably lower than the daily volatility, consistent with negative serial correlation. By contrast, momentum volatility increased sharply with increasing aggregation window size, consistent with positive serial correlation.

Evaluating the Accuracy of Risk Forecasts. Let $f_{k\tau}$ be the return to factor k for period τ, where the period may correspond to a single day or multiple days. For example, weekly and monthly factor returns are obtained by aggregating daily factor returns over 5-day or 21-day windows, respectively. The standardized return is given by

$$z_{k\tau} = \frac{f_{k\tau}}{\sigma_{k\tau}} \ , \tag{4A.1}$$

where $\sigma_{k\tau}$ denotes the beginning-of-period volatility forecast. The bias statistic is defined as the standard deviation of the z-scores,

$$B_k = \sqrt{\frac{1}{T}\sum_{\tau=1}^{T} z_{k\tau}^2} \ , \tag{4A.2}$$

where T is the total number of periods. Equation (4A.2) assumes that the z-scores are mean zero, which is a good approximation for short horizons such as one month. If risk forecasts are accurate, then the expected value of the bias statistic will be close to 1. On the other hand, bias statistics significantly above or below 1 indicate underprediction or overprediction of risk, respectively.

While the bias statistic is useful for identifying biases, it suffers from one significant shortcoming: It is subject to *error cancelation*. For instance, it is typical to underforecast risk entering a crisis, and to overforecast risk following a crisis. Over the entire crisis period, the bias statistic may be very close to 1, even though the risk forecasts may have been significantly biased throughout the crisis.

As discussed by Patton (2011), a more rigorous measure of forecasting accuracy is provided by the Q-statistic, defined as

$$Q_{k\tau} = z_{k\tau}^2 - \ln\left(z_{k\tau}^2\right). \tag{4A.3}$$

As shown in Technical Appendix A, the Q-statistic has several attractive properties. First, the Q-statistic is minimized in expectation when the true volatility is used to make every forecast. This property makes the Q-statistic a valuable tool for model calibration.

Second, the Q-statistic isn't subject to error cancelation. More specifically, overforecasting risk for one observation leads to a large expected value for $-\ln(z_{k\tau}^2)$, whereas underforecasting risk for another observation leads to a large expected value for $z_{k\tau}^2$. In either case, forecasting error leads to an increase in the expected value of the Q-statistic, and these errors cannot cancel.

A third attractive feature of the Q-statistic is that there exists a simple analytic formula that relates the expected increase in the Q-statistic to forecasting errors. Consider a portfolio with true volatility σ_T and predicted volatility σ_P. Let Q_T denote the Q-statistic obtained using σ_T as the forecast, let Q_P be the corresponding quantity using σ_P, and let $\Delta Q \equiv Q_P - Q_T$ be the difference. In Technical Appendix A, we derive the following result for the expected increase in the Q-statistic:

$$E\left[\Delta Q\right] = \frac{\sigma_T^2}{\sigma_P^2} + 2\ln\left(\frac{\sigma_P}{\sigma_T}\right) - 1. \tag{4A.4}$$

Remarkably, this result holds *independent of the return distribution*. As an example, suppose $\sigma_P = 1.10 \cdot \sigma_T$, so that we overforecast risk by 10 percent. Plugging this into Equation (4A.4), we find $E[\Delta Q] = 0.017$. Now suppose that $\sigma_P = 0.90 \cdot \sigma_T$, so that we underforecast risk by 10 percent. In this case, it is easy to verify that $E[\Delta Q] = 0.024$.

Most practitioners would agree that a 10% reduction in volatility forecast error represents a material improvement in the accuracy of the risk forecast. Hence, a useful rule of thumb when comparing Q-statistics of competing models is that a 0.02 reduction in the Q-statistic represents a material improvement in the risk forecasts. Arguably, however, even a 0.01 reduction (corresponding to a 7% error reduction) may be considered material.

A fourth attractive feature of the Q-statistic is that it places a higher penalty on underforecasting risk than on overforecasting. For instance, we saw from the previous example that a 10% overforecasting leads to an increase of 0.017, whereas a 10% underforecasting produces an increase of 0.024. The asymmetry becomes even more pronounced as the forecasting errors increase. For instance, overforecasting by 100% $\left(\sigma_P = 2 \cdot \sigma_T\right)$ leads to an increase of 0.636 in the expected Q-statistic, whereas underforecasting by 50% $\left(\sigma_P = 0.50 \cdot \sigma_T\right)$ leads to an increase of 1.614.

Model Calibration for Different Horizons. Different types of investors naturally have different objectives when using a risk model. Short-term investors—such as traders, market makers, and risk managers—typically combine very short investment horizons with a disciplined risk-management process. These investors therefore demand a model that will be sufficiently responsive to provide the most accurate risk forecasts over a short horizon, such as one day or one week.

Other investors, such as hedge fund managers and traditional asset managers with fairly short investment horizons, often embrace disciplined risk management as an integral part of their investment process. Such investors demand a responsive model that provides the most accurate forecasts over fairly short horizons, such as 1-3 months.

Traditional asset managers with longer investment horizons typically desire a risk model that provides accurate forecasts over prediction horizons ranging from 3-12 months. These investors usually would not make major adjustments to their portfolio based on short-term changes in volatility. Hence, such managers typically value a model with greater stability.

Finally, some investors—such as pension funds, foundations, and endowments—prefer models with very long-term horizons. For example, asset owners often make asset allocation decisions based on long-run historical averages of factor volatilities and correlations. Such investors naturally place a high premium on model stability.

With these different objectives in mind, the MAC3 equity models are designed for six distinct prediction horizons: daily, weekly, monthly, quarterly, annual, and long-term. All models share the same factor structure and have identical daily factor returns. What differentiates the models are the factor covariance matrices and the specific risk estimates that are used to forecast volatility.

The MAC3 parameters are calibrated with three distinct objectives in mind: (1) forecasting accuracy, as measured by bias statistics and Q-statistics, (2) model stability, and (3) smoothness in the term structure of risk. For the daily, weekly, and monthly models, our primary aim is to provide the most accurate risk forecasts at their respective horizons, attaching little weight to the stability of the forecasts. Hence, these models tend to be quite responsive. The quarterly model is calibrated giving most weight to forecasting accuracy, although model stability also plays a significant role. Model stability is a primary consideration in the annual model, which is calibrated to provide the most accurate risk forecasts for a one-year prediction horizon, subject to a constraint on model responsiveness. The long-term model places a high premium on model stability, using very long HL parameters that attach significant weight to the distant past.

Finally, we discuss the role of our third design objective, i.e., smoothness in the term structure of risk. More specifically, we aim for a term structure of risk that varies smoothly with prediction horizon. That is, we wish to avoid a lot of "bumps and wiggles" in the volatility forecast as the horizon is varied from short term to long term. This is achieved by *joint calibration* of all model parameters. Typically, we find that in times of extreme market turmoil, volatility forecasts are highest for the shortest prediction horizons. By contrast, in periods of historically low volatility, the most responsive models (e.g., daily or weekly) usually have the lowest risk forecasts, while the long-term model has the highest.

(B) Factor Volatility Forecasts

Whenever high-frequency data (e.g., daily) are used to forecast risk over longer horizons (e.g., weekly or monthly), it is important that serial correlation be properly taken into account. By contrast, if the prediction horizon matches the observation frequency (as in the daily model), then serial correlation adjustments are inappropriate.

Daily Forecasts. Let f_τ denote the outlier-adjusted factor return on day τ. The one-day variance forecast σ_t^2 at the start of day t is found using simple EWMA,

$$\sigma_t^2 = \sum_{\tau=1}^{t-1} w_\tau f_\tau^2 \ , \tag{4B.1}$$

where w_τ is the EWMA weight assigned to day τ. The HL parameter is calibrated to provide the most accurate forecast for a one-day horizon.

Other Forecast Horizons. In this case, the prediction horizon exceeds one day, so serial correlation must explicitly be taken into account. There are two main approaches for modeling serial correlation. One approach is due to Newey and West (1987), who explicitly estimate the serial correlation across different days. The other approach implicitly takes serial correlation into account by aggregating high-frequency returns over longer horizons. We refer to this as the low-frequency method, as described next.

Low-Frequency Volatilities. The first step in the low-frequency method is to aggregate daily factor returns over longer horizons. Let f_τ denote the return to a factor on day τ. The trailing T-day return[6] at the start of day t is given by

$$f_t^T = \sum_{\tau=t-T}^{t-1} f_\tau \ . \tag{4B.2}$$

For the weekly and monthly models, the aggregation window matches the prediction horizon, i.e., T=5 days and T=21 days, respectively. For the other prediction horizons, the aggregation windows are shorter than the prediction horizon. The rationale for this approach is that most of the effect of serial correlation can be captured using shorter windows, and shorter windows reduce the sampling error relative to longer windows.

The predicted variance for the low-frequency method is

$$\sigma_{LF}^2 = \mathrm{var}\left(f_t^T \right) \ , \tag{4B.3}$$

where the variance is computed using EWMA. Note that all variance forecasts are updated on a *daily* basis using T-day rolling windows.

Newey-West Volatilities. The second method for estimating the variance of serially correlated returns is due

[6] Note that for purposes of forecasting portfolio risk, it makes no material difference whether we aggregate arithmetic returns or log returns.

to Newey and West (1987), in which the serial correlations are explicitly estimated. The Newey-West formula for the *T*-period variance is given by

$$\sigma_{NW}^2 = T\sigma_0^2\left[1+2\sum_{l=1}^{L}\left(1-\frac{l}{L+1}\right)\rho_l\right], \tag{4B.4}$$

where $\sigma_0^2 = \mathrm{var}(f_\tau)$ is the variance of the daily factor returns, ρ_l is the correlation between f_τ and the lagged factor returns $f_{\tau-l}$, l is the lag number, and L is the total number of lags. The Newey-West variance contains three parameters: (1) the volatility HL used to compute σ_0^2, (2) the correlation HL used to estimate ρ_l, and (3) the total number of lags L.

The quantity in square brackets of Equation (4B.4) represents the variance adjustment due to serial correlation. If serial correlations are positive, then the variance adjustment is greater than 1, whereas it is less than 1 if serial correlations are negative. Finally, note that if all serial correlations are zero, the Newey-West formula reduces to the standard result of square-root-of-time scaling, i.e., $\sigma_{NW} = \sigma_0\sqrt{T}$.

In statistics, it is well known that blending forecasts can be an effective method for reducing estimation error. In particular, we show in Technical Appendix B that by properly blending two biased and noisy forecasts, we can obtain an estimate that has lower mean-squared error (MSE) than either estimate taken separately. We apply blending techniques broadly throughout the MAC3 Model.

Example. As a simple illustration, consider a portfolio whose true variance is 1. Suppose we have a biased and noisy forecast n,

$$\hat{\sigma}_n^2 = 1+\mu_n+e_n, \tag{4B.5}$$

where μ_n is the bias and e_n is a random error term (noise) which is drawn from a mean-zero distribution with standard deviation ε_n. Consider two such variance forecasts, both of which are biased and noisy. Estimate 1 has a relatively large bias $(\mu_1 = 0.2)$ and small noise $(\varepsilon_1 = 0.1)$, whereas Estimate 2 has relatively small bias $(\mu_2 = -0.05)$ and large noise $(\varepsilon_2 = 0.2)$. Suppose that the correlation between the two random error terms $(e_1$ and $e_2)$ is ρ. Now consider the blended estimate

$$\hat{\sigma}_B^2 = w_1\hat{\sigma}_1^2 + (1-w_1)\hat{\sigma}_2^2, \tag{4B.6}$$

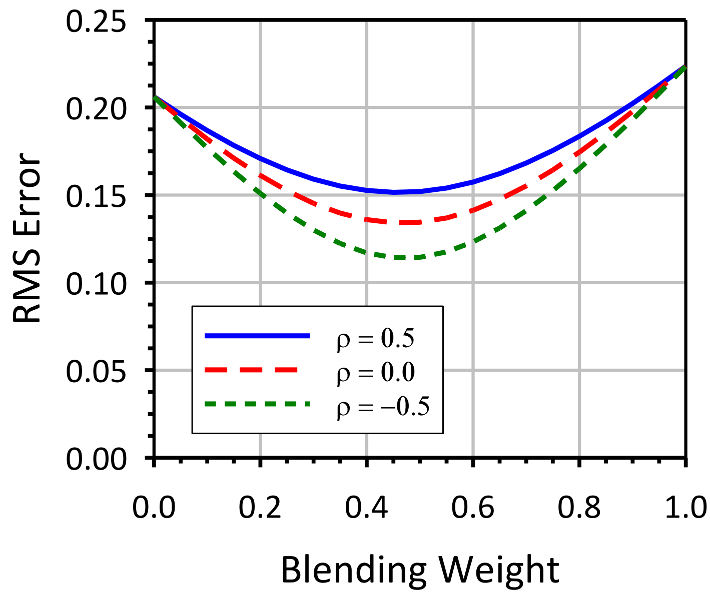
where w_1 is the weight assigned to Estimate 1. The root-mean-square (RMS) forecast error is defined as

$$RMS = \sqrt{E\left[\left(\hat{\sigma}_B^2-1\right)^2\right]}. \tag{4B.7}$$

In Technical Appendix B, we provide an analytic expression for the RMS error as a function of the six parameters: μ_1, ε_1, μ_2, ε_2, w_1, and ρ.

In Figure 4.2, we plot the RMS error as a function of the blending weight w_1, for three values of correlation ρ (0.5, 0.0, and -0.5). We see that RMS error can be reduced considerably by blending the two forecasts. The benefits are more modest if the errors are positively correlated, whereas they are more dramatic if the errors are negatively correlated.

Figure 4.2. Plot of RMS error versus blending weight, for three values of correlation.



Blended Volatilities. We apply this blending technique for estimating factor variance. In particular, the blended estimate is given by a weighted average of the Newey-West and the low-frequency volatility forecasts,

$$\sigma_B^2 = w_{NW}\sigma_{NW}^2 + (1 - w_{NW})\sigma_{LF}^2 \ , \tag{4B.8}$$

where w_{NW} is the weight assigned to the Newey-West variance. Empirically, we find that the blended forecast generally produces more accurate risk forecasts than either the Newey-West forecast or the low-frequency forecast.

(C) Cross-Sectional Volatility (CSV) Scaling

Accurate risk forecasting requires finding the right balance between two opposing forces. On the one hand, we have *sampling error*, which arises whenever we estimate volatility using a finite number of observations. On the other hand, we have *non-stationarity* of the return distribution, which implies that volatility is a dynamic variable.

The HL parameter of the model determines just how much weight we assign to recent observations. Sampling error considerations by themselves would lead us to select the *longest* possible HL, since this minimizes sampling error. By contrast, non-stationarity, considered in isolation, would lead us to select the *shortest* possible HL, since the recent past is the best indicator of volatility for the immediate future. Minimizing forecasting error thus requires finding the optimal balance between these two opposing effects.

As described by Menchero and Morozov (2015), cross-sectional observations can be utilized to improve the accuracy of volatility forecasts. More specifically, we use the cross-sectional bias statistic to assign more weight to recent observations, without paying a high penalty in terms of sampling error.

Bias statistics are typically computed for a single factor from a time series of z-scores, as in Equation (4A.2). However, bias statistics may also be computed for a *single* time period from a cross section of z-scores. In particular, let $z_{k\tau}^{T}$ be the trailing *T*-day return of factor k at the end of day τ, divided by the volatility forecast *T* days ago. The squared cross-sectional bias statistic for day τ is computed as

$$B_{\tau}^{2} = \frac{1}{K}\sum_{k}\left(z_{k\tau}^{T}\right)^{2}, \tag{4C.1}$$

where K is the total number of factors in the local model.

The traditional (time-series) bias statistic indicates whether risk forecasts for a single factor k were biased upward or downward over an interval of time. By contrast, the cross-sectional bias statistic indicates whether risk forecasts were biased upward or downward across a set of factors for an individual *T*-day window ending on day τ.

The CSV multiplier is computed by taking a weighted average of trailing squared cross-sectional bias statistics,

$$B_{t}^{2} = \sum_{\tau=1}^{t-1} w_{\tau} B_{\tau}^{2}, \tag{4C.2}$$

where w_{τ} is the EWMA weight assigned to period τ. The CSV-adjusted volatility forecasts are given by

$$\tilde{\sigma}_{kt} = B_{t}\sigma_{kt}, \tag{4C.3}$$

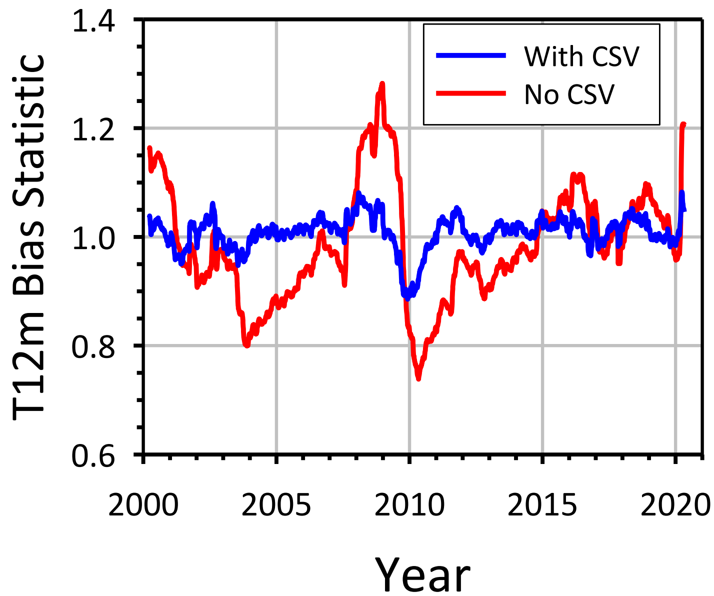
where σ_{kt} is the unadjusted forecast used to compute the z-scores $z_{k\tau}^{T}$.

The MAC3 Model applies a MXF approach to CSV scaling. That is, separate CSV multipliers are applied to both low-frequency (σ_{LF}^{2}) and high-frequency (σ_{NW}^{2}) components. We find that the MXF implementation of CSV scaling leads to more accurate risk forecasts and a more intuitive term structure of risk.

The idea behind the CSV method is that cross-sectional bias statistics contain important information on "instantaneous" biases in risk forecasts, and this information can be utilized to adjust and refine the original risk forecasts. For instance, entering a financial crisis, risk models tend to underforecast risk since they use the low-volatility past to predict the high-volatility future. In this case, the cross-sectional bias statistics tend to be greater than 1 nearly every single day. The CSV multiplier B_{t} quantifies the volatility adjustment required to remove the bias. Similarly, upon exiting a financial crisis, the cross-sectional bias statistics tend to be less than 1, and the CSV adjustment serves to lower the risk forecasts.

Example. As an illustration, in Figure 4.3, we show the effect of CSV scaling in the MAC3 US Equity Model (weekly horizon). We plot the mean trailing 252-day bias statistic of rolling weekly factor returns, averaged across all US factors. The bias statistic essentially represents the ratio of realized volatility to predicted volatility. The blue line is the bias statistic obtained using the CSV method, whereas the red line is without CSV. To isolate the impact of CSV, we kept all other model parameters fixed for the comparison. Figure 4.3 clearly shows that the CSV technique is very effective at reducing biases in risk forecasts. For example, during the 2008 Financial Crisis, the maximum bias statistic reached 1.08 in the model with CSV, versus 1.28 in the model without CSV. Similarly, in the aftermath of the Financial Crisis, the minimum bias statistic was 0.89 with the CSV method, versus 0.74 without CSV.

Figure 4.3. Plot of trailing 252-day bias statistic, averaged across all factors, for the MAC3 US Equity Model (weekly horizon), with and without the CSV adjustment. All other model parameters were left unchanged.



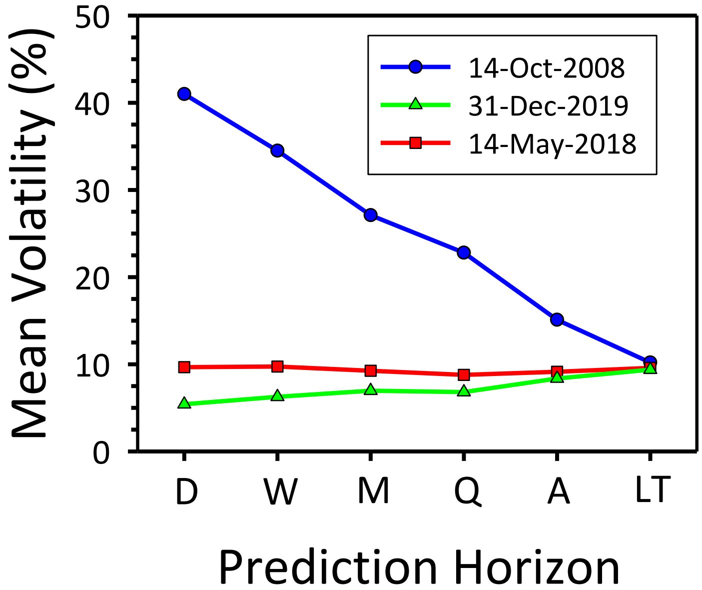
Not only is the CSV method effective at mitigating biases, but it also recovers much more quickly following a financial crisis. For instance, by early 2011, the trailing 12-month bias statistic with CSV adjustment had already reverted to 1 (meaning that the risk forecasts on average were unbiased for the 2010 calendar year), whereas without CSV it took until nearly 2015 for the bias statistics to return to 1. A similar effect is seen after the Internet Bubble period, in which the model without CSV overforecasts for about five years.

The CSV approach allows risk forecasts to be more responsive, without adding undue noise to the forecasts. To illustrate, suppose that the CSV HL parameter were four days. If this were applied to a single factor, the effective number of observations would be roughly three times the HL parameter, or 12 days. The sampling error associated with 12 observations is very high. However, if there are 50 factors in the cross section, then the effective number of observations used to compute the CSV multiplier is roughly 600 (50x12), thus greatly mitigating the adverse effects of sampling error.

The previous arguments, though valid, are heuristic in nature. However, CSV scaling also rests upon a theoretical foundation. In particular, we show in Technical Appendix A that the CSV adjustment given by Equation (4C.3) is precisely the multiplicative adjustment required to minimize the realized Q-statistic over a trailing window.

Example. With the factor volatilities thus estimated, we now illustrate the term structure of risk with a specific example. In particular, we compute the mean factor volatilities for the MAC3 US Equity Model, for each of the six forecasting horizons. We consider three select dates: (a) 14-Oct-2008, at the height of the Financial Crisis, (b) 31-Dec-2019, which was a period of historically low volatility, and (c) 14-May-2018, which was a stable period of relatively low volatility.

Figure 4.4. Mean factor volatility for the MAC3 US Equity Model on three select dates for six prediction horizons: daily (D), weekly (W), monthly (M), quarterly (Q), annual (A) and long term (LT).



In Figure 4.4, we plot the resulting term structure of risk. At the height of the Financial Crisis, we see a strong downward slope in the term structure, with the daily model producing an average volatility of more than 40%, versus about 10% for the long-term model. At the end of 2019, which was a period of extremely low volatility, we see that term structure of risk is upward sloping, with a mean volatility of roughly 5% for the daily model versus about 10% for the long-term model. The third date (14-May-2018) represents a fairly flat term structure of risk, where the mean factor volatilities did not vary much for the different horizons.

Section (5): Factor Correlations

(A) Term Structure of Correlations

Just as factor volatilities exhibit a term structure, factor correlations do as well. We apply the same two basic methods that were used to estimate the term structure of factor volatilities. The first approach is based on the low-frequency method, which involves aggregating daily factor returns over rolling T-day windows. The second method is Newey-West, which explicitly estimates factor covariances across different trading days.

Low-Frequency Correlations. We begin by describing the low-frequency method for estimating factor correlations. Let K be the total number of factors and let $\mathbf{f}_t^T$ denote the Kx1 vector of trailing *T*-day factor returns at the start of day t, as defined in Equation (4B.2). The low-frequency KxK factor covariance matrix is given by

$$\mathbf{F}_{LF} = \operatorname{cov}\left(\mathbf{f}_t^T, \mathbf{f}_t^T\right) , \tag{5A.1}$$

where the covariances are estimated using EWMA. Note that $\mathbf{F}_{LF}$ is updated for all models on a daily basis using rolling windows. Also note that the aggregation windows used for the term structure of correlations are identical to those used for the term structure of volatilities, as described in Section (4B). For instance, the weekly model uses $T = 5$ for the aggregation window, while the monthly model uses $T = 21$.

The main objective here is to extract the factor *correlation* matrix, which is done by scaling the rows and columns of $\mathbf{F}_{LF}$ by the inverse volatility,

$$\mathbf{C}_{LF} = \mathbf{d}_{LF}^{-1}\mathbf{F}_{LF}\mathbf{d}_{LF}^{-1} , \tag{5A.2}$$

where $\mathbf{d}_{LF}$ is a diagonal KxK matrix whose elements are given by the square root of the diagonal elements of $\mathbf{F}_{LF}$. Once we have the correlation matrix, the factor *covariance* matrix is produced by scaling the rows and columns using our best estimates of factor volatilities, as described in Section (4).

Newey-West Correlations. We now discuss the second approach for estimating factor correlations, which is based on the Newey-West factor covariance matrix,

$$\mathbf{F}_{NW} = \mathbf{F}_0 + \sum_{l=1}^{L}\left(1 - \frac{l}{L+1}\right)\left(\mathbf{F}_l + \mathbf{F}_l'\right) , \tag{5A.3}$$

where $\mathbf{F}_0$ is the contemporaneous (i.e., zero lags) factor covariance matrix, $\mathbf{F}_l$ is the factor covariance matrix with l lags, and L is the total number of lags. The matrix elements of $\mathbf{F}_l$ are given by

$$\mathbf{F}_l^{jk} = \operatorname{cov}\left(f_t^j, f_{t-l}^k\right) , \tag{5A.4}$$

where f_t^j is the return to factor j on day t, and f_{t-l}^k is the return to factor k on day $t-l$. The Newey-West factor covariance matrix contains two parameters: (1) the HL used to estimate $\mathbf{F}_l$, and (2) the number lags L. Note that the contemporaneous term $\mathbf{F}_0$ uses the same HL as $\mathbf{F}_l$.

With the Newey-West factor covariance matrix thus defined, we extract the factor correlation matrix in the usual manner,

$$\mathbf{C}_{NW} = \mathbf{d}_{NW}^{-1}\mathbf{F}_{NW}\mathbf{d}_{NW}^{-1} , \tag{5A.5}$$

where $\mathbf{d}_{NW}$ is a diagonal KxK matrix whose elements are given by the square root of the diagonal elements of $\mathbf{F}_{NW}$.

Blended Correlations. As done with volatilities, where we blend the low-frequency estimate with the Newey-West estimate, we apply the same approach for estimating factor correlations. More specifically, the blended correlation matrix is given by

$$\mathbf{C}_B = w_{NW}\mathbf{C}_{NW} + \left(1 - w_{NW}\right)\mathbf{C}_{LF} \ , \tag{5A.6}$$

where w_{NW} is the weight assigned to the Newey-West estimate. The blended correlation matrix $\mathbf{C}_B$ represents the MAC3 estimate of the sample correlation.

Correlation Matrices and Portfolio Optimization. Before proceeding further, we first pause to stress that risk models are used for two basic purposes. The leading use case is to estimate the volatility of non-optimized portfolios. The second important use case is portfolio optimization. It is worth pointing out that most techniques used in the industry today for estimating covariance matrices work well for one use case *or* the other, but typically not for *both*. Our aim is to produce a model that works equally well for *both* applications.

Matrix $\mathbf{C}_B$ represents our best estimate of each pairwise factor correlation. As such, this matrix is reliable for predicting the volatility of non-optimized portfolios. However, this does not imply that $\mathbf{C}_B$ (in its present form) is reliable for portfolio optimization. In particular, if the number of factors is large (as is the case for multi-asset-class risk models), the matrix $\mathbf{C}_B$ may be *ill-conditioned* or even *rank deficient*.

To better understand these important concepts, we make a brief digression. As discussed by Menchero, Wang, and Orr (2012), an ill-conditioned matrix is one in which the smallest *eigenvalues* are too small. The eigenvalues of a covariance matrix represent the predicted variances of the *eigenportfolios*, which are simply linear combinations of the original factors (i.e., portfolios of factors). Hence, portfolio optimizers naturally gravitate toward these "low-risk" eigenportfolios, which are in reality far riskier than the covariance matrix implies.

A rank-deficient matrix is one in which the smallest eigenvalues are equal to zero. In this case, the risk model actually predicts the existence of "riskless" combinations of factors. A portfolio manager who constructs an optimized portfolio using such a matrix would surely be unpleasantly surprised when the supposedly "riskless" portfolio goes live and is found in fact to be highly volatile.

As discussed by Menchero and Ji (2019), doing portfolio optimization using ill-conditioned covariance matrices leads to several problems in portfolio construction. In particular, these problems include: (a) underestimation of risk of optimized portfolios, (b) reduced out-of-sample risk-adjusted performance, and (c) increased portfolio leverage and turnover.

PCA Shrinkage. The Bloomberg PORT research team has developed an innovative and robust solution to this problem. The aim is to produce a well-conditioned correlation matrix that deviates minimally from $\mathbf{C}_B$. The well-conditioned nature of the matrix means that it can be reliably used for portfolio optimization, while the fact that the matrix deviates minimally from $\mathbf{C}_B$ implies that the matrix can be reliably used for risk forecasting.

The Bloomberg solution is to blend matrix $\mathbf{C}_B$ with a correlation matrix that is derived from $\mathbf{C}_B$ using *Principal Component Analysis* (PCA). This blending method is known as *PCA Shrinkage*. PCA is a widely used statistical technique for extracting factors from a time series of multiple correlated variables. The basic idea behind

PCA is that a few factors (the principal components) may be sufficient to explain most of the variability and comovement of the original factors. The final correlation matrix is given by a weighted average between the blended matrix and the PCA matrix,

$$\tilde{\mathbf{C}}_B = w_B \mathbf{C}_B + \left(1 - w_B\right) \mathbf{C}_{PC} \,, \tag{5A.7}$$

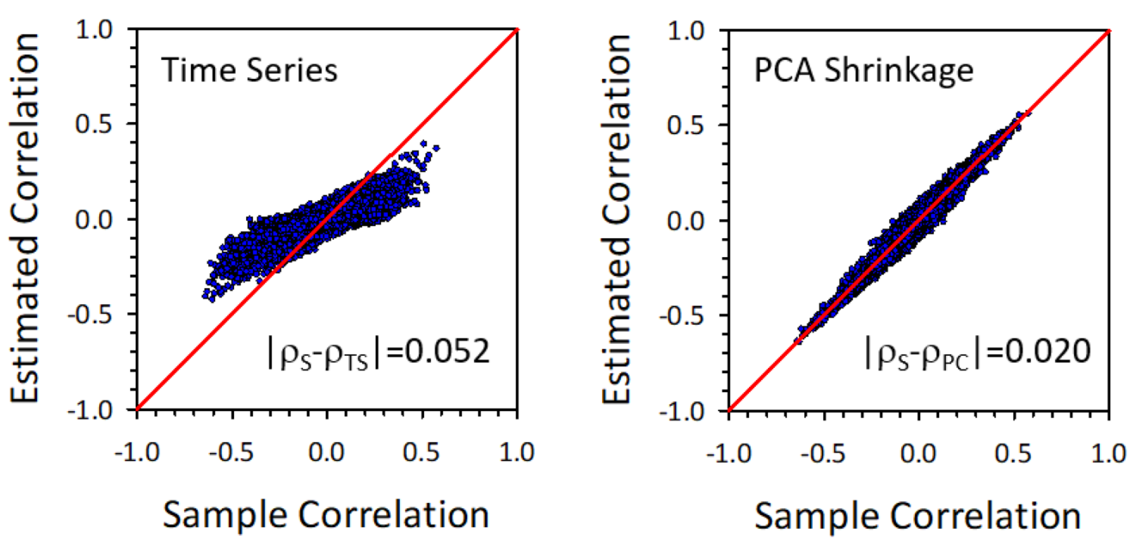
where w_B is the weight assigned to the blended matrix, and $\mathbf{C}_{PC}$ is the PCA correlation matrix derived from the blended correlation matrix $\mathbf{C}_B$.

By selecting the appropriate blending weight and number of principal components, we obtain a correlation matrix that can be reliably used both for risk forecasting and for portfolio optimization. That is, the correlation matrix in Equation (5A.7) is well-conditioned and its elements deviate minimally from those of matrix $\mathbf{C}_B$.

Bloomberg pioneered the use of PCA Shrinkage with the launch of the MAC2 Model in 2016. This estimation technique is carried over to the MAC3 Model. The PCA Shrinkage technique is described in detail by Menchero and Ji (2020).

Example. The most common alternative technique for estimating MAC factor correlations was developed by Shepard (2007/2008), and is known as the *time-series* method. In this approach, a time series of global factor returns are identified with the aim of explaining the correlations of local factors.

Figure 5.1. Scatterplot of estimated correlation versus sample correlation for a large number of equity and fixed income factor pairs. The time-series method systematically underforecasts correlations, while PCA Shrinkage provides an excellent fit to the sample correlation.

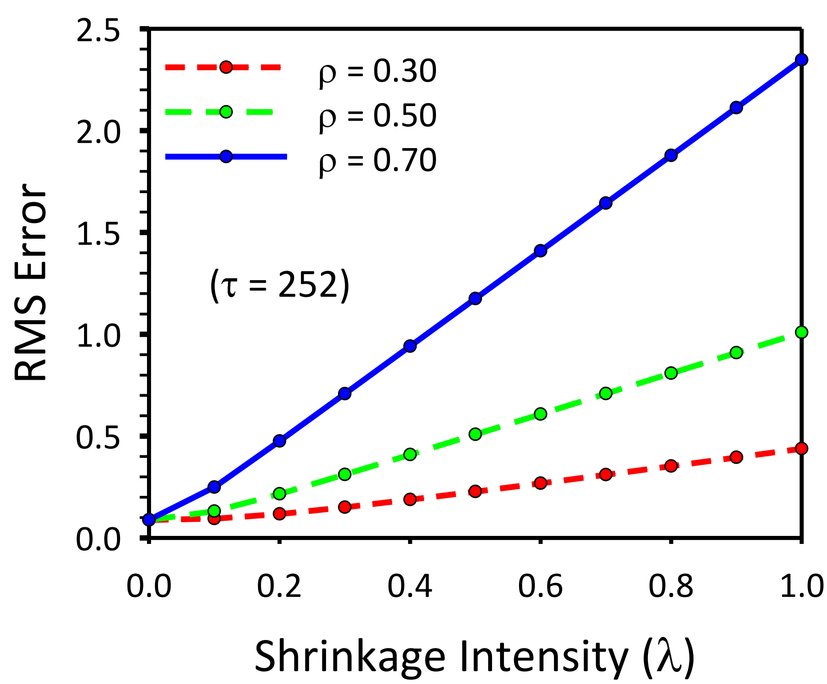


Menchero and Ji (2020) showed that the time-series method tends to systematically underforecast the magnitude of factor correlations, effectively "shrinking" them towards zero. To illustrate this effect, they estimated correlations using both the time-series method and PCA Shrinkage between a set of 387 equity factors and 431 fixed income factors (more than 166,000 factor pairs). Next, they computed the sample correlation for these factor pairs and created scatterplots of the estimated correlation versus the sample correlation.

The resulting scatterplots are shown in Figure 5.1 for both the time-series method (left panel) and PCA Shrinkage (right panel). The left panel shows that the time-series method systematically underforecasts the correlations by a large margin. By contrast, the right panel shows that PCA Shrinkage provides an excellent fit to the sample correlation.

Menchero and Ji (2020) also estimated a slope coefficient by fitting a regression line through the cloud of points. They found that the slope coefficient for the time-series method was approximately 0.30 (meaning correlations were "shrunk" by 70 percent towards zero). Moreover, this underforecasting of correlations was found to be highly persistent across time. By contrast, they found that slope coefficients using PCA Shrinkage were very close to 1 over all sample periods.

Figure 5.2. Plot of RMS error (in variance forecast) versus shrinkage intensity λ for asset-pair portfolios. We consider three values for the true correlation ρ, and estimate risk using $\tau = 252$ days. The RMS error is effectively minimized at a shrinkage intensity of zero, which corresponds to the sample correlation.



Example. In another paper, Menchero and Li (2020) studied the errors in variance forecasts caused by underestimation of correlations. In their study, they constructed asset-pair portfolios that went long one asset and short the other asset. The two assets were assumed to have equal true volatility and true correlation ρ. Under these assumptions, the true asset covariance matrix is known. They used the true asset covariance matrix to simulate τ periods of correlated asset returns, which in turn were used to *estimate* asset volatilities and their correlation. The estimated correlation $\hat{\rho}$ was then shrunk by a parameter λ, which represents the shrinkage intensity. That is, while the sample correlation was $\hat{\rho}$, the correlation actually used for making the risk forecast was $(1-\lambda)\hat{\rho}$. Note that $\lambda = 0$ corresponds to the sample correlation, while $\lambda = 1$ corresponds to uncorrelated assets. Menchero and Li (2020) derived an analytic expression for the root-mean-squared (RMS) error in variance forecasts of these portfolios as a function of three variables: (1) the true correlation ρ, (2) the number of periods τ used to estimate the volatilities and correlations, and (3) the shrinkage intensity λ.

In Figure 5.2, we plot the RMS error in variance forecasts of these asset-pair portfolios as a function of the shrinkage intensity λ, for three values of true correlation (0.30, 0.50, and 0.70), using a window length of $\tau = 252$ days to estimate the volatilities and correlations. From Figure 5.2, we see that RMS error is effectively minimized at a shrinkage intensity of zero. In other words, the sample correlation essentially represents the *optimal* estimate for purposes of predicting risk[7].

Note that RMS error rises dramatically as both the shrinkage intensity and the true correlation increase. For instance, using a shrinkage intensity of $\lambda = 0.70$ (as found in the time-series method) and a true correlation of $\rho = 0.50$, the RMS error is roughly 0.70, or 70%, which increases to roughly 160% if the true correlation is 0.70. By contrast, without any shrinkage ($\lambda = 0$), the RMS error is less than 0.10 (10%) in all cases.

This important result shows that for purposes of accurately forecasting portfolio risk, the estimated correlations should never deviate far from the sample correlation. As shown in Figure 5.1, this clearly holds for PCA Shrinkage, but not for the time-series method. Hence, PCA Shrinkage provides an important advantage over the time-series method when it comes to predicting portfolio volatility.

(B) Imputation Algorithm for Missing Factors

Factor covariance matrix estimation requires balanced panels of data, meaning that there are no missing factor returns. In reality, however, factor returns may be missing for a variety of reasons. National holidays represent one obvious source of missing factor returns. For instance, we cannot reliably estimate the return of the France factor on Bastille Day, since few (if any) French stocks actually trade. Another common source of missing factor returns comes from factors with short histories, such as countries that join the model at a later start date.

Each day, some factors may be missing, while other factors will be non-missing. The basic idea behind the imputation algorithm is to use the non-missing factor returns as explanatory variables to estimate the missing factor returns. This is essentially the same exercise one encounters in scenario analysis, in which several factors are "shocked" and these shocks are propagated to the other factors. In this case, the non-missing factor returns represent the shocked variables, while the missing factor returns are imputed using the propagated values. The MAC3 Model uses the factors from the MAC3 Global Equity Model for imputing the returns of all missing equity factors.

It is also important to recognize that the non-missing factors can never explain the full variability of the missing factor returns. In other words, the missing factor returns always contain a component that is uncorrelated with

---

[7] This statement must be qualified. Menchero and Li showed that in reality, RMS error is always minimized at a non-zero shrinkage intensity. It just so happens that for any reasonable window length τ, the optimal shrinkage intensity is so close to zero that the minimum is not visible to the naked eye.

the non-missing factors. This component is known as the residual variance of the factor.

It is important to include the residual variance of the factor, especially for factors with short return history. If the residual variance component is ignored, then each missing factor return would be perfectly correlated with a different linear combination of global equity factor returns prior to the factor start date. This would lead to spurious factor correlations that could persist for long periods, especially for models with long HL parameters (e.g., annual or long-term models). Hence, the MAC3 Model includes the residual component for factors with short history.

(C) Adjusting the Diagonal Blocks

The MAC3 Model is constructed in a hierarchical fashion, applying a technique known as *model integration*. At the bottom level of the hierarchy, we find the local models. Changing notation slightly, we let $\tilde{\mathbf{C}}_{jj}^{L}$ denote our best estimate of the local factor correlation matrix for model j, as given in Equation (5A.7). Similarly, we let $\tilde{\mathbf{C}}_{kk}^{L}$ denote our best estimate of the local factor correlation matrix for model k. As a concrete example, j might represent the MAC3 US Equity Model, while k might denote the MAC3 Europe Equity Model.

The collection of all MAC3 local equity models is known as the MAC3 Integrated Model. This model, which covers stocks around the globe, consists of 13 local models, each of which may contain several dozen factors. Hence, the MAC3 Integrated Model contains several hundred local equity factors.

To construct the MAC3 Integrated Model, we must estimate the correlations among all local factors. Matrices $\tilde{\mathbf{C}}_{jj}^{L}$ and $\tilde{\mathbf{C}}_{kk}^{L}$ provide our best factor correlation estimates for local models j and k, respectively, but tell us nothing about the correlations *between* factors in different models (e.g., US growth factor and Europe momentum factor).

To estimate correlations between factors in different local models, we again apply Equation (5A.7), except now we include all factors across all local equity models. In other words, we take a weighted average of the blended correlation matrix and the PCA correlation matrix, where now each of these matrices contain hundreds of local equity factors. While the MAC3 Integrated Model contains more than a dozen local equity models, the model integration technique can be illustrated in its full generality by considering only two such local models, j and k.

Let $\tilde{\mathbf{C}}_{G}$ denote the "global" (i.e., spanning models j and k) factor correlation matrix, computed using Equation (5A.7). This matrix can be written in block form,

$$\tilde{\mathbf{C}}_{G} = \begin{bmatrix} \tilde{\mathbf{C}}_{jj}^{G} & \tilde{\mathbf{C}}_{jk}^{G} \\ \tilde{\mathbf{C}}_{kj}^{G} & \tilde{\mathbf{C}}_{kk}^{G} \end{bmatrix} . \tag{5C.1}$$

Matrix $\tilde{\mathbf{C}}_G$ provides estimates for the off-diagonal blocks. Unfortunately, the diagonal blocks of $\tilde{\mathbf{C}}_G$ do not match our best estimates obtained from the local model. Hence, if we were to use matrix $\tilde{\mathbf{C}}_G$ as the official correlation matrix, we would be utilizing sub-optimal estimates for the diagonal blocks.

The objective of model integration is to obtain a final correlation matrix $\mathbf{C}_F$ in which the diagonal blocks exactly replicate the diagonal blocks of the local models, One way to achieve this is to simply "insert" the local correlation matrices into the diagonal blocks of $\tilde{\mathbf{C}}_G$. Unfortunately, there is no guarantee that the resulting correlation matrix would be positive definite. In order to preserve positive definiteness, we must adjust the off-diagonal blocks, but these adjustments should be minimal.

This task is accomplished by means of the integration matrix $\mathbf{M}$, defined as

$$\mathbf{M} = \begin{bmatrix} \mathbf{Q}_{jj} & 0 \\ 0 & \mathbf{Q}_{kk} \end{bmatrix}, \tag{5C.2}$$

where

$$\mathbf{Q}_{jj} \equiv \left(\tilde{\mathbf{C}}_{jj}^{L}\right)^{1/2} \left(\tilde{\mathbf{C}}_{jj}^{G}\right)^{-1/2}, \tag{5C.3}$$

and

$$\mathbf{Q}_{kk} \equiv \left(\tilde{\mathbf{C}}_{kk}^{L}\right)^{1/2} \left(\tilde{\mathbf{C}}_{kk}^{G}\right)^{-1/2}. \tag{5C.4}$$

Note that if the diagonal blocks of the global model match the diagonal blocks of the local models, then the integration matrix ($\mathbf{M}$) becomes simply the identity matrix. The final correlation matrix is found by applying the transformation,

$$\mathbf{C}_F = \mathbf{M}\tilde{\mathbf{C}}_G\mathbf{M}'. \tag{5C.5}$$

Carrying out the matrix multiplication, we obtain

$$\mathbf{C}_F = \begin{bmatrix} \tilde{\mathbf{C}}_{jj}^{L} & \mathbf{Q}_{jj}\tilde{\mathbf{C}}_{jk}^{G}\mathbf{Q}'_{kk} \\ \mathbf{Q}_{kk}\tilde{\mathbf{C}}_{kj}^{G}\mathbf{Q}'_{jj} & \tilde{\mathbf{C}}_{kk}^{L} \end{bmatrix}. \tag{5C.6}$$

Note that the diagonal blocks of the reconstituted matrix $\mathbf{C}_F$ equal the diagonal blocks of the local models. Moreover, in practice we find that the off-diagonal blocks of $\mathbf{C}_F$ are nearly equal to the off-diagonal blocks of $\tilde{\mathbf{C}}_G$. Hence, matrix $\mathbf{C}_F$ accomplishes our twin objectives: (a) it replicates the diagonal blocks from the local models, and (b) it deviates minimally from the off-diagonal blocks of the global correlation matrix $\tilde{\mathbf{C}}_G$.

Equation (5C.6) represents the final correlation matrix for the MAC3 Integrated Model. We apply the same technique to the other asset classes, to produce integrated global correlation matrices for each of the asset classes (e.g., fixed income, commodities, etc.). The integrated global models for each asset class represent the second level of the hierarchy.

We still face the task of estimating the factor correlations between local factors belonging to different asset classes. This is accomplished using the same technique as described above and represents the third and final level of the hierarchy. In this case, we estimate the "universal" correlation matrix $\tilde{\mathbf{C}}_U$, which combines all factors across all asset classes. In total, there could be more than 2000 factors contained in $\tilde{\mathbf{C}}_U$. We then apply matrix integration so that we replicate global correlation matrices $\mathbf{C}_F$ along the diagonal blocks (corresponding now to individual asset classes). Note that the diagonal blocks of $\mathbf{C}_F$ already replicate the diagonal blocks of the local models (e.g., $\tilde{\mathbf{C}}_{jj}^L$).

Section (6): Specific Risk

(A) Time-Series Forecasts

Specific returns represent the portion of stock returns that cannot be explained by the factors. As such, they are also known as non-factor returns or idiosyncratic returns. Specific returns represent an important component of risk, especially for portfolios that are not well diversified.

Specific returns are estimated using the same cross-sectional regression employed to estimate factor returns. That is, the specific return u_{nt} of stock n on day t is given by

$$u_{nt} = r_{nt} - \sum_k X_{nk}^t f_{kt} \quad , \tag{6A.1}$$

where r_{nt} is the local excess return of the stock, X_{nk}^t is the stock exposure to factor k at the start of day t, and f_{kt} is the factor return.

Given the time series of specific returns u_{nt}, we estimate the term structure of specific risk using the same techniques as those used in Section (4) to estimate factor volatilities. For instance, one-day specific-volatility forecasts are found by simply substituting the outlier-adjusted specific returns in place of the outlier-adjusted factor returns in Equation (4B.1).

Similarly, to estimate the low-frequency volatility $\sigma_{LF}(n)$ for stock n, we first aggregate specific returns over *T*-day rolling windows, as in Equation (4B.2). We then substitute these aggregated returns into Equation (4B.3) to estimate volatility.

Likewise, to estimate the Newey-West specific volatility $\sigma_{NW}(n)$, we apply Equation (4B.4) to the outlier-adjusted specific returns. As with factor volatilities, the final time-series specific risk forecast $\sigma_{TS}(n)$ is given by a weighted average between the Newey-West forecasts and the low-frequency forecasts,

$$\sigma_{TS}^2(n) = w_{NW}\sigma_{NW}^2(n) + \left(1 - w_{NW}\right)\sigma_{LF}^2(n) \quad , \tag{6A.2}$$

where w_{NW} is the weight assigned to the Newey-West forecast.

(B) Structural Forecasts

Let $\boldsymbol{\sigma}_{TS}$ denote the Nx1 vector of specific risk forecasts whose elements are given by the time-series estimates $\sigma_{TS}(n)$, where N is the number of stocks in the estimation universe. Since not all stocks (e.g., IPOs) have a full history of specific returns, we must also estimate a structural model, which assigns specific risk forecasts based on the stock's factor exposures.

In order to ensure that all specific risk forecasts are positive, we regress the log of the specific volatility against the factor exposures

$$\ln\left(\boldsymbol{\sigma}_{TS}\right) = \mathbf{X}\mathbf{b} + \boldsymbol{\varepsilon} \ , \tag{6B.1}$$

where $\mathbf{X}$ is the NxK factor exposure matrix, $\mathbf{b}$ is a Kx1 vector of slope coefficients, and $\boldsymbol{\varepsilon}$ is an Nx1 vector of unexplained residuals. We impose the usual constraint, which is to set the cap-weighted industry slope coefficients to zero each period. For multi-country models, we also impose the constraint that the cap-weighted country factor slope coefficients add to zero.

The structural estimate is found by eliminating the residuals and taking the exponential

$$\boldsymbol{\sigma}_{STR} = c \cdot \exp\left(\mathbf{X}\mathbf{b}\right) \ , \tag{6B.2}$$

where c is a multiplicative constant designed to remove the small bias created by the exponential function. The bias arises because although the residuals ε_n are mean zero, the exponential $\exp\left(\varepsilon_n\right)$ will not be mean 1 due to Jensen's inequality. Hence, the multiplicative constant is given by

$$c = \sum_n v_n \exp\left(\varepsilon_n\right) \ , \tag{6B.3}$$

where v_n is the regression weight.

Since the exponential of a sum is the product of the exponentials, the structural estimate for stock n may be written as

$$\sigma_{STR}(n) = c \prod_k \exp\left(X_{nk} b_k\right) \ , \tag{6B.4}$$

where X_{nk} is the stock exposure to factor k, and b_k is the slope coefficient. Since all stocks have unit exposure to the market factor M, the term $\exp\left(b_M\right)$ can be interpreted as the *average* specific risk of the ESTU. The other terms in Equation (6B.4) can be seen as providing a multiplier to scale the average specific risk forecast up or down, depending on the stock's exposure to each factor.

For example, we expect that the slope coefficient associated with the residual volatility factor will be strongly positive, since stocks with positive exposure to this factor tend to have high specific risk. Hence, if a stock has negative exposure to residual volatility, the multiplier will be a positive number less than 1 (i.e., lowering specific risk), whereas it will be greater than 1 if the stock has positive exposure to the factor.

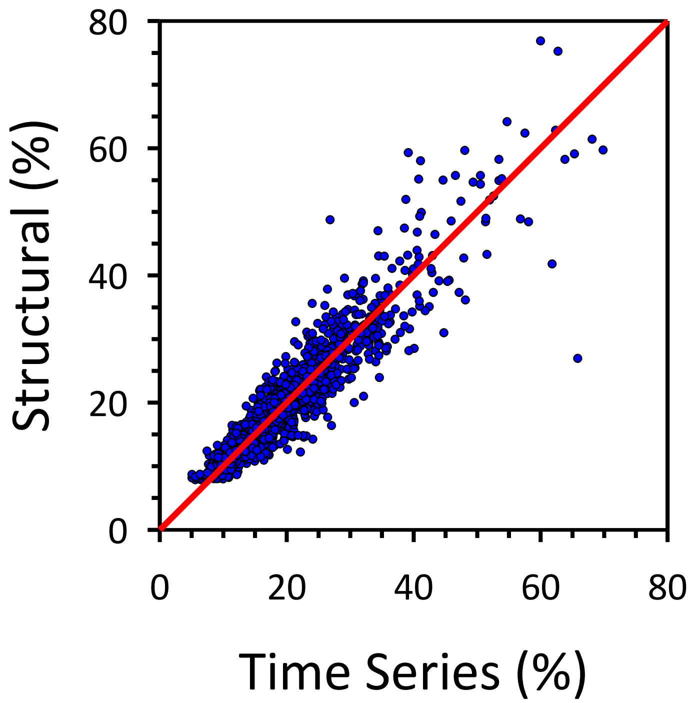
(C) Blended Forecasts

The next step in obtaining the specific risk forecast is to blend the time-series estimate with the structural estimate. In particular, the blended estimate is given by

$$\sigma_{BL}^2(n) = w_{TS}\sigma_{TS}^2(n) + \left(1 - w_{TS}\right)\sigma_{STR}^2(n) \ , \qquad (6C.1)$$

where w_{TS} is the weight assigned to the time-series estimate.

Figure 6.1. Scatterplot of annualized structural specific risk estimates versus time-series estimates for the MAC3 US Equity Model (monthly horizon) on analysis date 31-Dec-2019.



Example. It is instructive to compare the structural forecasts with the time-series forecasts. In Figure 6.1, we present a scatterplot of the structural estimate (y-axis) versus the time-series estimate (x-axis) for the MAC3 US Equity Model (monthly horizon) on 31-Dec-2019. We see that structural forecasts fit the time-series estimates quite well. Moreover, as discussed in Section (4), proper blending of the two forecasts, as in Equation (6C.1), generally produces more accurate risk forecasts.

(D) Stocks with Missing Data

Special treatment is required for IPOs and stocks with a high frequency of missing/zero returns. For such stocks, the time-series estimate becomes unreliable and we must assign more weight to the structural forecast. In this case, we "shrink" the blended estimate, given by Equation (6C.1), toward the structural forecast, given by Equation (6B.4). More specifically, the "shrunk" estimate of specific risk is given by

$$\sigma_{\gamma}^{2}(n)=\left(1-\gamma_{n}\right)\sigma_{BL}^{2}(n)+\gamma_{n}\sigma_{STR}^{2}(n) \ , \tag{6D.1}$$

where γ_n is the *shrinkage intensity*. For recently issued IPOs, $\gamma_n = 1$, so that the specific risk forecast is based entirely on the structural model. Once the IPO has acquired sufficient history, we smoothly transition the shrinkage intensity and assign more weight to the blended estimate.

Thinly traded stocks are treated similarly to IPOs. More specifically, if their frequency of missing/zero returns is above some threshold, the shrinkage intensity is set to $\gamma_n = 1$. Stocks with missing/zero returns below the threshold will have $0 < \gamma_n \leq 1$. In the limiting case that the stock has a full history of non-zero returns, $\gamma_n = 0$, which means we give 100% weight to the blended forecasts.

(E) Cross-Sectional Volatility (CSV) Scaling

The final step in estimating specific variance is to apply CSV scaling. The method is identical to the CSV scaling applied to factors in Section (4C), except that now specific returns and specific risk forecasts are used to construct the z-scores that feed into the algorithm.

Let σ_{nt} be the specific risk forecast of stock n for period t. The CSV-adjusted specific risk forecasts $\tilde{\sigma}_{nt}$ are given by

$$\tilde{\sigma}_{nt} = B_t \sigma_{nt} \ , \tag{6E.1}$$

where B_t is the CSV multiplier. Note that B_t is computed by taking an EWMA average of squared cross-sectional bias statistics, as in Equation (4C.2). The CSV technique utilizes recent observations to create a feedback loop that rapidly adjusts volatility forecasts to mitigate biases.

(F) Linked Specific Risk

For most securities, as long as the factor structure is properly specified, it is fair to assume that the specific returns are mutually uncorrelated. This assumption breaks down in the case of linked securities, which represent different *issues* from the same *issuer*. Since linked securities represent the same fundamental claim on the assets of the firm, the no-arbitrage argument ensures that the returns of these securities closely track each other across time. This is not to say that the linked securities move in lock-step, since short-run return differences may arise from liquidity considerations or asynchronous trading.

In the MAC3 Model, we apply non-zero correlations to the specific returns of linked assets. Intuitively, we expect to find a pronounced term structure of linked correlations, with the smallest correlations observed in the daily model, but rising sharply as the prediction horizon increases. We estimate linked correlations for different prediction horizons by aggregating the specific returns over the appropriate window length. We then average these correlations across a large number of linked securities, assigning more weight to large-cap stocks. For a given horizon, all linked securities will be assigned the same average correlation, but the correlation will vary with the prediction horizon.

We estimate the specific risk of linked securities in an analogous fashion to that used for assigning factor exposures to linked securities, as described in Section (2). More specifically, if the Fundamentals Ticker (the "parent") is covered by the model in question, then the specific risk forecast of the "child" (i.e., the linked security) is copied over from the parent. For instance, this treatment assures that Google A and Google C have the same specific risk forecast in the US model.

The second possibility is that the Fundamentals Ticker is *not* covered by the model in question. An example of this is Sony ADR (trading in New York) from the perspective of the MAC3 US Equity Model. In this case, the specific risk is estimated in the same way as any other US stock. That is, specific returns of the security are computed using the US Equity Model, and these returns are used to obtain a time-series forecast. Similarly, the structural forecast is obtained from the factor exposures of the security.

Section (7): Finite-Sample Adjustment

(A) Misspecification of Traditional Factor Models

Factor models have been used in finance since the mid-1970s. Remarkably, the standard approach for estimating factor models has been misspecified since inception. The Bloomberg MAC3 suite of equity models provides an innovative solution to this outstanding problem, as we now illustrate.

To understand the origin of the misspecification, we first review the basics of factor risk models. These models posit that asset returns ($\mathbf{r}$) are driven by a set of underlying factors,

$$\mathbf{r} = \mathbf{X}\mathbf{f} + \mathbf{u}, \tag{7A.1}$$

where $\mathbf{X}$ denotes the true NxK factor exposure matrix, $\mathbf{f}$ denotes the true Kx1 vector of factor returns, and $\mathbf{u}$ denotes the true Nx1 vector of idiosyncratic returns.

One of the basic assumptions behind factor models is that the specific returns $\mathbf{u}$ are mutually uncorrelated and also uncorrelated with the factors. Under these assumptions, the true NxN asset covariance matrix is given by

$$\mathbf{\Omega} = \mathbf{X}\mathbf{F}\mathbf{X}' + \mathbf{\Delta}, \tag{7A.2}$$

where $\mathbf{F} = \mathrm{cov}(\mathbf{f}, \mathbf{f})$ is the true KxK factor covariance matrix, and $\mathbf{\Delta} = \mathrm{cov}(\mathbf{u}, \mathbf{u})$ is the true NxN diagonal matrix of specific variance.

In reality, the only quantity in Equation (7A.1) that can be directly observed is the vector of stock returns $(\mathbf{r})$. Nevertheless, for illustrative purposes, we will make the simplifying assumption that the true factor exposure matrix $\mathbf{X}$ is also known.

In practice, factor returns are estimated by cross-sectional regression,

$$\mathbf{r} = \mathbf{X}\hat{\mathbf{f}} + \hat{\mathbf{u}}, \tag{7A.3}$$

where $\hat{\mathbf{f}}$ and $\hat{\mathbf{u}}$ are the *estimated* vectors of factor returns and specific returns, respectively. From these estimated returns, we obtain the *estimated* factor covariance matrix, $\hat{\mathbf{F}} = \mathrm{cov}(\hat{\mathbf{f}}, \hat{\mathbf{f}})$, and the *estimated* specific variance matrix, $\hat{\mathbf{\Delta}} = \mathrm{cov}(\hat{\mathbf{u}}, \hat{\mathbf{u}})$, which is assumed diagonal. Hence, the traditional estimate of the asset covariance matrix is given by

$$\hat{\mathbf{\Omega}} = \mathbf{X}\hat{\mathbf{F}}\mathbf{X}' + \hat{\mathbf{\Delta}}. \tag{7A.4}$$

The problem with the traditional approach is that the estimated quantities in Equation (7A.4) are biased estimators for the *true* quantities in Equation (7A.2). More specifically, the estimated factor variance systematically overpredicts the true factor variance, i.e., $E[\hat{\mathbf{F}}_{kk}] > \mathbf{F}_{kk}$, while the estimated specific variance systematically underpredicts the true specific variance, i.e., $E[\hat{\mathbf{\Delta}}_{nn}] < \mathbf{\Delta}_{nn}$.

(B) Illustrative Examples

We illustrate these concepts first with a simple toy model, and then follow with a more realistic example. Consider a universe of two assets, each of which has unit exposure to a single factor (representing the market). The factor model in this case is

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} f + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tag{7B.1}$$

where f is the return of the market factor, and $\mathbf{X}' = [1 \quad 1]$ is the factor exposure matrix. Note that the estimated factor return is easily solved by OLS regression, $\hat{f} = (r_1 + r_2)/2$. Similarly, the estimated specific returns are $\hat{u}_1 = (r_1 - r_2)/2$, and $\hat{u}_2 = -\hat{u}_1$.

For simplicity, assume that the true factor has unit variance $(\mathbf{F} = 1)$ and that both assets have the same true specific variance $(\Delta_1 = \Delta_2 = \varepsilon)$. In this case, the true asset covariance matrix is easily computed,

$$\mathbf{\Omega} = \begin{bmatrix} 1+\varepsilon & 1 \\ 1 & 1+\varepsilon \end{bmatrix}. \tag{7B.2}$$

In the traditional approach, the factor variance is estimated by the realized variance of the pure factor portfolio, with holdings vector $\mathbf{w}_f' = [0.5 \quad 0.5]$. The true variance of the factor portfolio is easily computed

$$\mathbf{w}_f'\mathbf{\Omega}\mathbf{w}_f = 1 + \frac{\varepsilon}{2}. \tag{7B.3}$$

We see that this overshoots the true factor variance $(\mathbf{F}=1)$ by $\varepsilon/2$. Similarly, it is easy to show that the true variance of the estimated specific returns is $\operatorname{var}(\hat{u}_1)=\varepsilon/2$, which underestimates the true specific variance (ε) by a factor of 2.

We now consider a more realistic example and show that the traditional asset covariance matrix, given by Equation (7A.4), leads to biased forecasts. Consider a multi-factor model with K factors. Let $\mathbf{w}_k$ be the Nx1 holdings vector of style pure factor portfolio k. Recall that style pure factor portfolios have unit exposure to their own factor, and zero exposure to all other factors. The predicted variance of the pure factor portfolio using $\hat{\mathbf{\Omega}}$ is therefore

$$\mathbf{w}_k'\hat{\mathbf{\Omega}}\mathbf{w}_k=\hat{\mathbf{F}}_{kk}+\mathbf{w}_k'\hat{\mathbf{\Delta}}\mathbf{w}_k, \tag{7B.4}$$

where $\hat{\mathbf{F}}_{kk}$ is the diagonal element of the estimated KxK factor covariance matrix $\hat{\mathbf{F}}$. However, note that $\hat{\mathbf{F}}_{kk}$ is *already* an unbiased estimate of the variance of the pure factor portfolio $\mathbf{w}_k$, whose returns we directly observe. Hence, the additional specific risk term $\mathbf{w}_k'\hat{\mathbf{\Delta}}\mathbf{w}_k$ in Equation (7B.4) leads to an *upward bias* in the risk forecast.

This effect can be understood by noting that the pure factor portfolio (like any portfolio) already contains an idiosyncratic component, which contributes part of the variability of the estimated factor returns. That is, the pure factor portfolio return is driven by the true (unobservable) factor and the true (unobservable) specific component. The variance estimate $\hat{\mathbf{F}}_{kk}$ already reflects both sources of risk. However, when we use the asset covariance matrix $\hat{\mathbf{\Omega}}$ to forecast the risk, it adds on another layer of specific risk. As a result, the traditional approach overestimates the risk of pure factor portfolios by essentially "double counting" the specific risk component.

(C) The MAC3 Solution

The Bloomberg PORT research team has pioneered an innovative solution to this problem. Our solution is called the finite-sample adjustment (FSA) due to the fact that the effect is entirely driven by having a finite number of stocks in the ESTU, which in turn leads to noise in the factor return estimates. Note that if there were an infinite number of stocks in the ESTU, then the specific risk of the pure factor portfolios would completely vanish, and no adjustment would be necessary.

The first step in the FSA method is to find unbiased estimates $\tilde{\mathbf{\Delta}}$ and $\tilde{\mathbf{F}}$ for the true specific variance $\mathbf{\Delta}$ and the true factor covariance matrix $\mathbf{F}$, respectively. In other words, $E[\tilde{\mathbf{\Delta}}]=\mathbf{\Delta}$, and $E[\tilde{\mathbf{F}}]=\mathbf{F}$. Given the FSA estimates for $\tilde{\mathbf{\Delta}}$ and $\tilde{\mathbf{F}}$, the NxN FSA asset covariance matrix becomes

$$\tilde{\mathbf{\Omega}}=\mathbf{X}\tilde{\mathbf{F}}\mathbf{X}'+\tilde{\mathbf{\Delta}}. \tag{7C.1}$$

Recall that $\hat{\mathbf{F}}$ represents our best estimate of the covariance matrix of pure factor portfolios, which is different from the covariance matrix $\mathbf{F}$ of the unobservable true factors. Hence, when constructing the FSA factor covariance matrix $\tilde{\mathbf{F}}$, our objective is to replicate $\hat{\mathbf{F}}$ using the FSA asset covariance matrix $\tilde{\mathbf{\Omega}}$ applied to the pure factor portfolios. In other words, our aim is to construct $\tilde{\mathbf{\Omega}}$ such that

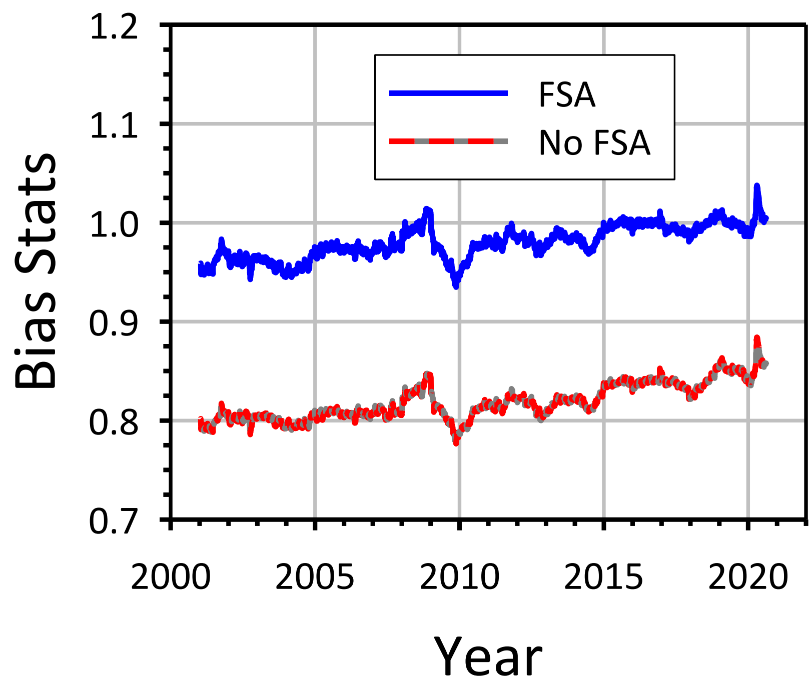
$$\mathbf{w}_j'\tilde{\mathbf{\Omega}}\mathbf{w}_k = \hat{\mathbf{F}}_{jk}, \qquad (7C.2)$$

where $\mathbf{w}_j$ and $\mathbf{w}_k$ are Nx1 holdings vectors for pure factor portfolios j and k, respectively.

To illustrate the benefits of the FSA method in practice, we compute the mean trailing 252-day bias statistics averaged across all factors in the MAC3 Australia Equity Model (daily horizon). The out-of-sample test period runs from 31-Dec-1999 to 17-Jul-2020.

In Figure 7.1, we plot the mean bias statistics for both the FSA model (blue line) and for the model without FSA (red/grey line). The two models are identical in every other respect. Note that the FSA model produces bias statistics very close to 1 over the entire sample period, indicating nearly unbiased estimates of portfolio risk. By contrast, the model without FSA systematically overforecasts portfolio risk across the entire sample period. The mean bias statistics, averaged across time, were 0.98 for the FSA model, versus 0.82 for the model without FSA.

Figure 7.1. Plot of trailing 252-day bias statistic, averaged across all factors, for the MAC3 Australia Equity Model (daily horizon). The blue line is for the FSA model, whereas the red/grey line is for the model without FSA. Note that the two models were identical in every other respect.



Although not plotted here, we find similar biases when it comes to specific risk forecasts. In this case, however, the model without FSA *underforecasts* risk, as expected by theory. We computed bias statistics for the pure specific portfolios of the MAC3 Australia Equity Model (daily horizon). Over the full sample period, the FSA model produced mean bias statistics of 1.03, versus 1.09 for the model without FSA.

Technical Appendix A: Review of Bias Statistics and Q-Statistics

Let $f_{k\tau}$ be the return to factor k for period τ, and let $\sigma_{k\tau}$ denote the beginning-of-period volatility forecast. The standardized return, or z-score, is defined as

$$z_{k\tau} = \frac{f_{k\tau}}{\sigma_{k\tau}} \ . \tag{A.1}$$

We assume that the z-scores are mean zero, which is a sound approximation for the relatively short horizons considered here. The bias statistic for a given factor k is defined as the standard deviation of the z-scores,

$$B_k = \sqrt{\frac{1}{T}\sum_{\tau=1}^{T} z_{k\tau}^2} \ , \tag{A.2}$$

where T is the total number of periods. If the risk forecasts are accurate, then the expected value of the bias statistic should be close to 1.

Of course, even if the risk forecasts are *perfect*, the realized bias statistic will never be *exactly* 1 due to sampling error. Nonetheless, it can be shown that if risk forecasts are perfect and returns are normally distributed, roughly 95% of the realized bias statistics should fall within the confidence interval,

$$B_k \in \left[1-\sqrt{\frac{2}{T}}, 1+\sqrt{\frac{2}{T}}\right] \ . \tag{A.3}$$

If $B_k < \left(1-\sqrt{2/T}\right)$, we conclude that our risk forecasts were too high. However, if we find $B_k > \left(1+\sqrt{2/T}\right)$, we conclude that our risk forecasts were too low.

As described in the main text, one shortcoming of the bias statistic is that it is prone to *error cancelation*. For instance, it is possible to underforecast risk one year and to overforecast risk the following year, while maintaining a bias statistic close to 1 over the two-year period. Hence, a bias statistic close to 1 does not necessarily imply that the risk forecasts were accurate.

A more rigorous measure of forecasting accuracy is the Q-statistic, given by

$$Q_{k\tau} = z_{k\tau}^2 - \ln\left(z_{k\tau}^2\right) \ . \tag{A.4}$$

The first term ($z_{k\tau}^2$) can be regarded as an underforecasting penalty, which becomes large if the risk forecasts are too low. Similarly, the second term, $-\ln(z_{k\tau}^2)$, can be regarded as an overforecasting penalty, which becomes large when the risk forecasts are too high.

Proof that True Volatility Minimizes the Expected Q-Statistic. We now show that the Q-statistic is minimized in expectation when the true volatility is used to make every forecast. This property makes the Q-statistic a useful tool for model calibration.

The expected value of the Q-statistic is given by

$$E\left[Q_{k\tau}\right]=E\left[\frac{f_{k\tau}^2}{\sigma_{k\tau}^2}\right]-E\left[\ln\left(\frac{f_{k\tau}^2}{\sigma_{k\tau}^2}\right)\right], \tag{A.5}$$

which can be rewritten as

$$E\left[Q_{k\tau}\right]=\frac{1}{\sigma_{k\tau}^2}E\left[f_{k\tau}^2\right]-2E\left[\ln\left|f_{k\tau}\right|\right]+2\ln\left(\sigma_{k\tau}\right) . \tag{A.6}$$

Note that the volatility forecast is not considered to be a random variable in this context. Let $a\equiv E[f_{k\tau}^2]$ and $b\equiv E[\ln|f_{k\tau}|]$. These are just numbers whose precise values depend on the return distribution. Hence, Equation (A.6) can be rewritten

$$E\left[Q_{k\tau}\right]=\frac{a}{\sigma_{k\tau}^2}-2b+2\ln\left(\sigma_{k\tau}\right) . \tag{A.7}$$

Setting the derivative to zero, we obtain,

$$\frac{dE\left[Q_{k\tau}\right]}{d\sigma_{k\tau}}=-\frac{2a}{\sigma_{k\tau}^3}+\frac{2}{\sigma_{k\tau}}=0 . \tag{A.8}$$

The solution is given by

$$\sigma_{k\tau}^2=a=E\left[f_{k\tau}^2\right] . \tag{A.9}$$

To determine whether this is a maximum or a minimum, we take the second derivative,

$$\frac{d^2E\left[Q_{k\tau}\right]}{d\sigma_{k\tau}^2}=\frac{4}{\sigma_{k\tau}^2} . \tag{A.10}$$

Since the second derivative is positive, it proves that the solution is a minimum. Hence, $E[Q_{k\tau}]$ is minimized in expectation when we use the true volatility.

Expected Increase in Q-Statistic due to Forecast Error. Next, we derive an expression for the expected increase in the Q-statistic due to noisy and biased forecasts. Let σ_T denote the true volatility of a portfolio, and let σ_P be predicted volatility, which contains estimation error. Let R denote the portfolio return. The Q-statistic using the true forecast is given by

$$Q_T=\frac{R^2}{\sigma_T^2}-\ln\left(\frac{R^2}{\sigma_T^2}\right), \tag{A.11}$$

whereas the Q-statistic using the estimated forecast is given by

$$Q_P = \frac{R^2}{\sigma_T^2}\frac{\sigma_T^2}{\sigma_P^2} - \ln\left(\frac{R^2}{\sigma_T^2}\frac{\sigma_T^2}{\sigma_P^2}\right) . \tag{A.12}$$

Using the fact that $E[R^2] = \sigma_T^2$, after a couple of lines of algebra, we find

$$E\left[Q_P - Q_T\right] = \frac{\sigma_T^2}{\sigma_P^2} + 2\ln\left(\frac{\sigma_P}{\sigma_T}\right) - 1 , \tag{A.13}$$

which is Equation (4A.4) of the main text. Note that the derivation does not make any distributional assumptions. Hence, Equation (A.13) holds independent of distribution.

Proof that the CSV Adjustment Minimizes the Realized Q-Statistic. Suppose that we multiply all volatility forecasts $\sigma_{k\tau}$ by a single scaling parameter γ, so that the scaled forecasts are given by $\tilde{\sigma}_{k\tau} = \gamma\sigma_{k\tau}$. The Q-statistic associated with the scaled forecasts are

$$Q_{k\tau}^{\gamma} = \frac{z_{k\tau}^2}{\gamma^2} - \ln\left(\frac{z_{k\tau}^2}{\gamma^2}\right), \tag{A.14}$$

where $z_{k\tau}$ is the z-score computed using volatility forecast $\sigma_{k\tau}$. The mean Q-statistic, averaged across factors and time, is given by

$$Q_{\gamma} = \frac{1}{K}\sum_{k\tau} w_{\tau} Q_{k\tau}^{\gamma} , \tag{A.15}$$

where w_{τ} is the EWMA weight assigned to period τ. We then solve for the value of the scaling parameter γ that minimizes the average Q-statistic. Setting the derivative of Q_{γ} equal to zero,

$$\frac{dQ_{\gamma}}{d\gamma} = 0 , \tag{A.16}$$

and solving for γ, we find

$$\gamma^2 = \frac{1}{K}\sum_{k\tau} w_t z_{k\tau}^2 , \tag{A.17}$$

which is exactly the CSV multiplier in Equation (4C.2). Hence, the CSV multiplier represents the volatility adjustment that minimizes the average realized Q-statistic.

Technical Appendix B: Error Reduction through Forecast Blending

In this Appendix, we analyze the reduction in forecast error that can be achieved by blending risk forecasts. We also derive an analytic formula for the optimal blending weight.

Consider a portfolio whose true variance is equal to 1. Now consider a noisy and biased estimate of that variance,

$$\hat{\sigma}_1^2 = 1 + \mu_1 + e_1 \ , \tag{B.1}$$

where μ_1 is the bias of the variance estimate and e_1 is a random error term with mean zero and variance ε_1^2. The mean-squared error (MSE) of the forecast is

$$\delta_1^2 = E\left[\left(\hat{\sigma}_1^2 - 1\right)^2\right] . \tag{B.2}$$

Substituting Equation (B.1) into Equation (B.2), and taking expectations, we find

$$\delta_1^2 = \mu_1^2 + \varepsilon_1^2 \ . \tag{B.3}$$

Now consider a second noisy and biased forecast

$$\hat{\sigma}_2^2 = 1 + \mu_2 + e_2 \ , \tag{B.4}$$

where μ_2 is the bias, and e_2 is a random error term with $E[e_2] = 0$ and $\mathrm{var}[e_2] = \varepsilon_2^2$. The MSE for forecast $\hat{\sigma}_2^2$ is computed in the same way as Equation (B.3), giving $\delta_2^2 = \mu_2^2 + \varepsilon_2^2$.

Next, we blend the two biased and noisy forecasts,

$$\hat{\sigma}_B^2 = w_1 \hat{\sigma}_1^2 + w_2 \hat{\sigma}_2^2 \ , \tag{B.5}$$

where w_1 is the weight assigned to the first forecast, and $w_2 = 1 - w_1$. The MSE of the blended forecast is

$$\delta_B^2 = E\left[\left(w_1\left(1 + \mu_1 + e_1\right) + w_2\left(1 + \mu_2 + e_2\right) - 1\right)^2\right] , \tag{B.6}$$

which can be simplified to

$$\delta_B^2 = E\left[\left(w_1\mu_1 + w_1 e_1 + w_2\mu_2 + w_2 e_2\right)^2\right] . \tag{B.7}$$

In general, the forecast errors e_1 and e_2 will have a correlation, denoted ρ_{12}, which implies that $E[e_1 e_2] = \varepsilon_1 \varepsilon_2 \rho_{12}$. Completing the square in Equation (B.7), and taking expectations, we obtain

$$\delta_B^2 = w_1^2\mu_1^2 + w_1^2\varepsilon_1^2 + w_2^2\mu_2^2 + w_2^2\varepsilon_2^2 + 2w_1 w_2 \mu_1 \mu_2 + 2 w_1 w_2 \varepsilon_1 \varepsilon_2 \rho_{12} \ . \tag{B.8}$$

To find the optimal blending weight, we take the first derivative of δ_B^2 with respect to w_1 and set it equal to zero. The optimal blending weight is easily found,

$$\tilde{w}_1 = \frac{\mu_2^2 + \varepsilon_2^2 - \mu_1\mu_2 - \varepsilon_1\varepsilon_2\rho_{12}}{\mu_1^2 + \varepsilon_1^2 + \mu_2^2 + \varepsilon_2^2 - 2\mu_1\mu_2 - 2\varepsilon_1\varepsilon_2\rho_{12}} \ . \tag{B.9}$$

For a general set of parameters, there is no guarantee that the optimal weight $\tilde{w}_1$ will fall within the interval $[0,1]$. Nonetheless, we find that for realistic sets of parameters, this is typically the case.

References

Amihud, Yakov., and Haim Mendelson. 1986. “Asset Pricing and the Bid-Ask Spread.” *Journal of Financial Economics*, vol. 17, no. 2 (December): 223-249.

Amihud, Yakov. 2002. “Illiquidity and Stock Returns: Cross-Section and Time-Series Effects.” *Journal of Financial Markets*, vol. 5, no. 1 (January): 31-56.

Ang, Andrew., Robert Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2009. "High Idiosyncratic Volatility and Low Returns: International and Further US Evidence." *Journal of Financial Economics*, vol. 91, no. 1 (January): 1-23.

Banz, R. 1981. "The Relationship Between Return and Market Value of Common Stock." *Journal of Financial Economics*, vol. 9, no. 1 (March): 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, no. 3 (June): 663-682.

De Bondt, Werner., and Richard Thaler. 1985. “Does the Stock Market Overreact?” *Journal of Finance*, vol. 40, no. 3 (July): 793-805.

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

Jegadeesh, Narasimhan. 1990. "Evidence of Predictable Behavior of Security Returns." *Journal of Finance*, vol. 45, no. 3 (July): 881-898.

Jegadeesh, Narasimhan., and Sheridan Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance*, vol. 48, no. 1 (March): 65-91.

Markowitz, Harry. 1952. "Portfolio Selection." *Journal of Finance,* vol. 7, no.1 (March): 77-91.

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

Menchero, Jose., and Jyh-Huei Lee. 2015. "Efficiently Combining Multiple Sources of Alpha." *Journal of Investment Management*, vol. 13, no. 4 (Q4): 71-86.

Menchero, Jose., and Lei Ji. 2019. "Portfolio Optimization with Noisy Covariance Matrices." *Journal of Investment Management*, vol. 17, no. 1 (Q1): 77-91.

Menchero, Jose., and Lei Ji. 2020. "Advances in Estimating Covariance Matrices." *Bloomberg Whitepaper* (to appear in the Journal of Investment Management).

Menchero, Jose., and Peng Li. 2020. "Correlation Shrinkage: Implications for Risk Forecasting." *Journal of Investment Management*, vol. 18, no. 3 (Q3): 92-108.

Menchero, Jose., and Andrei Morozov. 2015. "Improving Risk Forecasts Through Cross-Sectional Observations." *Journal of Portfolio Management*, vol. 41, no. 3 (Spring): 84-96.

Menchero, Jose., Jun Wang, and D.J. Orr. 2012. “Improving Risk Forecasts for Optimized Portfolios.” *Financial Analysts Journal*, vol. 68, no. 3 (May/June): 40-50.

Michaud, R. 1989. "The Markowitz Optimization Enigma: Is ‘Optimized’ Optimal?" *Financial Analysts Journal*, vol. 45, no. 1 (January/February): 31-42.

Naranjo, Andy., M. Nimalendran, and Mike Ryngaert. 1998. "Stock Returns, Dividend Yields, and Taxes." *Journal of Finance*, vol. 53, no. 6 (December): 2029-2057.

Newey, Whitney., and Kenneth West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Matrix." *Econometrica*, vol. 55, no. 3: 703-708.

Patton, Andrew. 2011. "Volatility Forecast Comparison using Imperfect Volatility Proxies." *Journal of Econometrics*, vol. 160, no. 1: 246-256.

Rosenberg, Barr. 1974. "Extra-Market Component of Covariance in Security Returns." *Journal of Financial and Quantitative Analysis*, vol. 9, no. 2 (March): 263-274.

Sharpe, William. 1964. "A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance*, vol. 19, no. 3 (September): 425-442.

Shepard, Peter. 2007/2008. "Integrating Multi-Market Risk Models." *Journal of Risk*, vol. 10, no. 2 (Winter): 25-45.

Tobin, J. 1958. “Liquidity Preference as Behavior Towards Risk.” *Review of Economic Studies*, vol. 25, no. 2 (February): 65-86.

Disclaimer:

Notwithstanding anything in the **MAC3 Global Equity Risk Model** (this “White Paper”) to the contrary, the information included in this White Paper is for informational purposes only. Bloomberg Finance L.P. and/or its affiliates (as applicable, "Bloomberg") makes no guarantee as to the adequacy, correctness or completeness of, or make any representation or warranty (whether express or implied) with respect to this White Paper. To the maximum extent permitted by law, Bloomberg and its affiliates shall not be responsible for or have any liability for any injuries or damages arising out of or in connection with this White Paper. All the information provided in this White Paper is confidential and proprietary to Bloomberg. Customer may not share, reproduce, publish, distribute or communicate this White Paper or information of any kind relating to these responses to any company, third parties or persons other than within Customer and only to such persons on a need-to-know basis in connection with this review.

Take the next step.

For additional information, press the <HELP> key twice on the Bloomberg Terminal®.

Beijing +86 10 6649 7500
Dubai +971 4 364 1000
Frankfurt +49 69 9204 1210
Hong Kong +852 2977 6000
London +44 20 7330 7500
Mumbai +91 22 6120 3600
New York +1 212 318 2000
San Francisco +1 415 912 2960
São Paulo +55 11 2395 9000
Singapore +65 6212 1000
Sydney +61 2 9777 8600
Tokyo +81 3 3201 8900

bloomberg.com/professional