# The Barra Global Equity Model (GEM2)

*Research Notes*

**September 2008**

Jose Menchero
Andrei Morozov
Peter Shepard

## Table of Contents

## 1. Introduction

This document describes the new Barra Global Equity Model, GEM2, and provides an in-depth comparison with the model it replaces, the Global Equity Model (GEM). With GEM, originally released in 1989, Barra pioneered the use of factor models for the forecasting of global equity portfolio risk. In the nearly 20 years since the launch of GEM, changes have occurred that warrant the introduction of a completely new model. These changes fall into two categories. First, additional data sources are available. These new data permit risk to be captured and quantified along dimensions that were previously unattainable. Second, there have been significant methodological advances in global equity risk modeling. GEM2 exploits both of these changes to provide users with more accurate risk forecasts within a highly intuitive model structure.

Two key elements are required to construct an accurate equity risk model. First, the model must employ a set of factors that explain as fully as possible the cross section of stock returns. The factors should capture all sources of return commonality, and thereby cleanly separate the systematic component of return from the idiosyncratic, or stock-specific, component. The second element pertains to the temporal resolution of the factor structure. In other words, the factor returns must be observed at a sufficiently high frequency to accurately estimate their volatilities and correlations.

The main advances of GEM2 are:

- a revised and updated industry factor structure based on the *Global Industry Classification Standard* (GICS®), better reflecting the global economy's current industrial structure
- a broader estimation universe based on the MSCI *Global Investable Market Indices* (GIMI) for greater precision in factor return estimates
- an improved and expanded set of style factors that captures additional sources of style risk
- the introduction of a World factor, giving a cleaner separation of country and industry effects
- higher frequency observations, allowing for more responsive risk forecasts
- a structural specific risk model that generates more accurate specific risk estimates
- greater stability in the proportion of risk coming from equity factors, currencies, and stock-specific sources

- more stable and accurate currency volatility forecasts
- a significant increase in the explanatory power of the model
- consistency between MSCI and risk model country classifications

GEM2 is being offered in short-term (GEM2S) and long-term (GEM2L) versions. Both versions share the same factor structure, and have identical factor returns. They differ in their responsiveness. GEM2S is designed to provide users with the most accurate and responsive risk forecasts. GEM2L, by contrast, is intended for longer-term investors who value more stable risk forecasts.

GEM also comes in two versions, one using MSCI industries, and the other based on FTSE classifications. Both versions of GEM employ the same methodology and produce very similar risk forecasts. This paper, however, considers only the MSCI version, which is the more widely used of the two.

## 2. Forecasting Global Equity Portfolio Risk

Global investors derive return from two basic sources: price appreciation in the local currency of the asset, and repatriation of the asset value back to the base currency (numeraire) of the investor.

For concreteness, consider an investment from the perspective of a US portfolio manager. Let $C(n)$ denote the local currency of stock $n$, and let $r_{FX}^{C(n),\$}$ be the return of currency $C(n)$ due to exchange-rate fluctuations against the dollar. The return of stock $n$ in US dollars, denoted $\hat{R}_n^{\$}$, is given by the usual expression

$$\hat{R}_n^{\$} = \hat{r}_n + r_{FX}^{C(n),\$} + \hat{r}_n r_{FX}^{C(n),\$}, \tag{2.1}$$

where $\hat{r}_n$ is the local return of the asset. The cross term, $\hat{r}_n r_{FX}^{C(n),\$}$, is typically minute, and can be safely ignored for risk purposes.[1]

Our primary interest is to forecast the volatility of *excess* returns (i.e., above the risk-free rate). If $r_f^{\$}$ denotes the risk-free rate of US dollars, then $R_n^{\$} = \hat{R}_n^{\$} - r_f^{\$}$ is the excess return of stock $n$ measured in US dollars.

[1] A notable exception to this is during times of extreme inflation, when the cross term can become significant.

Suppressing the cross term, and utilizing Equation 2.1, we decompose the excess return into an equity component $r_n$ and a currency component $q_n^{\$}$:

$$R_n^{\$} = r_n + q_n^{\$} . \tag{2.2}$$

The equity component is the *local excess return* of the stock,

$$r_n = \hat{r}_n - r_f^{C(n)} , \tag{2.3}$$

where $r_f^{C(n)}$ denotes the local risk-free rate for stock $n$. Note that $r_n$ is independent of numeraire, so that it is the same for all global investors.

The currency component in Equation 2.2 is given by

$$q_n^{\$} = r_f^{C(n)} + r_{FX}^{C(n),\$} - r_f^{\$} . \tag{2.4}$$

This term represents the excess return in US dollars due to holding cash denominated in the local currency of the stock.

We adopt a multi-factor framework to explain the local excess returns. This approach yields valuable insight into the underlying sources of portfolio return by separating systematic effects from the purely stock-specific component that can be diversified away. More specifically, we posit that the local excess returns are driven by a relatively small number, $K_E$, of global equity factors, plus an idiosyncratic component unique to the particular stock,

$$r_n = \sum_{k=1}^{K_E} X_{nk} f_k + u_n . \tag{2.5}$$

Here, $X_{nk}$ $(k \le K_E)$ is the exposure of stock $n$ to equity factor $k$, $f_k$ is the factor return, and $u_n$ is the specific return of the stock. The specific returns $u_n$ are assumed to be uncorrelated with the factor returns. The factor exposures are known at the start of each period, and the factor returns are estimated via cross-sectional regression.

Suppose that there are $K_C$ currencies in the model. Ordering the currencies after the equity factors, we can express the excess currency returns as

$$q_n^{\$} = \sum_{k=K_E+1}^{K_E+K_C} X_{nk} f_k^{\$} , \tag{2.6}$$

where $X_{nk}$ $(k > K_E)$ is the exposure of stock $n$ to currency $k$, and $f_k^{\$}$ is the excess return of the currency with respect to US dollars (which we calculate from risk-free rates and exchange rates, as in Equation 2.4). We take the currency exposures $X_{nk}$ to be equal to 1 if $k$ corresponds to the local currency of stock $n$, and 0 otherwise.

Equations 2.2-2.6 can be combined to obtain

$$R_n^{\$} = \sum_{k=1}^{K} X_{nk} f_k^{\$} + u_n , \tag{2.7}$$

where $K = K_E + K_C$ is the total number of combined equity and currency factors. The elements of $X_{nk}$ define the $N \times K$ factor exposure matrix, where $N$ is the total number of stocks. Note that the factor exposure matrix is independent of numeraire.

Our treatment thus far has considered only a single asset. Investors, however, are more concerned with portfolio risk. The portfolio excess return (in US dollars) is given by

$$R_P^{\$} = \sum_{n=1}^{N} h_n^P R_n^{\$} , \tag{2.8}$$

where $h_n^P$ is the portfolio weight of asset $n$. Note that, in general, the assets include both stocks and cash. The portfolio exposure to factor $k$ is given by the weighted average of the asset exposures,

$$X_k^P = \sum_{n=1}^{N} h_n^P X_{nk} . \tag{2.9}$$

In order to estimate portfolio risk, we also require the factor covariance matrix and the specific risk forecasts. The elements of the factor covariance matrix are

$$F_{kl}^{\$} = \mathrm{cov}\left(f_k^{\$}, f_l^{\$}\right), \tag{2.10}$$

where the dollar superscript indicates that at least some of the matrix elements depend on the base currency of the investor. Note, however, that the $K_E \times K_E$ block of the factor covariance matrix corresponding to equity factors is independent of numeraire.

The covariance of specific returns is

$$\Delta_{mn} = \mathrm{cov}\left(u_m, u_n\right). \tag{2.11}$$

For most stocks, we assume that the specific returns are uncorrelated, so that the off-diagonal elements of $\Delta_{mn}$ are zero. However, Equation 2.11 represents the generalized case in which the specific returns of some securities are linked.[2]

The portfolio risk in US dollars can now be obtained as the square root of the variance,

$$\sigma(R_P^{\$}) = \left[ \sum_{kl} X_k^P F_{kl}^{\$} X_l^P + \sum_{mn} h_m^P \Delta_{mn} h_n^P \right]^{1/2} . \tag{2.12}$$

[2] We relax the assumption of uncorrelated specific returns for different share classes of the same stock.

Although Equation 2.12 is written for portfolio risk, it is equally valid for tracking error by replacing portfolio weights and exposures with their active counterparts.

Many global investors, of course, are interested in risk forecasts from different numeraire perspectives. The only term on the right-hand-side of Equation 2.12 that depends on base currency is the factor covariance element, $F_{kl}^{\$}$. These elements, however, can be transformed to some other numeraire, $\gamma$. Substituting these transformed factor covariances $F_{kl}^{\gamma}$ into Equation 2.12 yields risk forecasts with respect to the new base currency $\gamma$.

# 3. Factor Exposures

## 3.1. Estimation Universe

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

A well-constructed equity index must address and overcome these very issues, and therefore serves as an excellent foundation for the estimation universe. The GEM2 estimation universe utilizes the MSCI *All Country World Investable Market Index* (ACWI IMI), part of the MSCI Global Investable Market Indices family which represents the latest in MSCI index-construction methodology. MSCI ACWI IMI aims to reflect the full breadth of global investment opportunities by targeting 99 percent of the float-adjusted market capitalization in 48 developed and emerging markets. The index-construction methodology applies innovative rules designed to achieve index stability, while reflecting the evolving equity markets in a timely fashion. Moreover, liquidity screening rules are applied to ensure that only investable stocks with reliable pricing are included for index membership.

The official MSCI ACWI IMI history begins in June 1994. To obtain a deeper model history, however, we backward-extend the estimation universe for two years by including select securities with existing GICS codes and clean daily returns. More specifically, for the period June 1992 to June 1994, we construct the estimation universe by including the constituents of the MSCI *All Country World Index* (ACWI), as well as all US and Japanese small-cap stocks with MSCI ACWI IMI membership as of June 1994.

The MSCI ACWI IMI aims to capture the full opportunity set open to global equity investors. As a result, if a country is deemed excessively restrictive to foreign investment, it is excluded. Nevertheless, many of these frontier markets represent important opportunities for certain global investors. For instance, since 2002, Qualified Foreign Institutional Investors (QFIIs) are permitted to invest directly in the Chinese domestic A-share market, which is excluded from MSCI ACWI IMI. Similarly, the six Gulf Cooperation Council (GCC) markets (Saudi Arabia, Oman, Qatar, UAE, Kuwait, and Bahrain) are important to many global investors, although they are also excluded from MSCI ACWI IMI. We therefore supplement the estimation universe with stocks from China Domestic and the six GCC countries, represented by the MSCI China A Index and the MSCI GCC Countries Index. We down-weight these markets in the regression in order to not distort the factor relationships within the investable universe. In the future, as frontier markets become more liquid and data becomes available, we plan to extend GEM2 coverage to many of these markets.

In Figure 3.1, we show the number of securities in MSCI ACWI IMI over time. Several trends are discernable. First, the number of assets rises significantly in the early history (1994-1997); this is due to generally rising markets and an increase in the number of companies meeting the minimum capitalization requirements for index membership. The second clear trend is the gradual decline in the number of constituents that occurs from 1997 to 2002. This is a consequence of a series of financial shocks, beginning with the Asia and Russia crises of 1997/1998, and ending with the collapse of the internet bubble in 2000 and the ensuing bear market. Since June 2002, with the general recovery of the financial markets, we again observe a secular increase in the number of MSCI ACWI IMI constituents.

The most prominent event in Figure 3.1 occurs in June 2002, which corresponds to the transition to float-adjusted market-cap weighting and the application of liquidity screening rules for MSCI ACWI IMI. In this event, the number of constituents drops sharply from nearly 12,000 to under 7,000. Although most of the excluded assets were small-cap stocks, this event must be treated carefully in the estimation universe to avoid potentially spurious jumps in factor exposures. We eliminate the discontinuity in the estimation universe by backward-excluding all stocks that were dropped from MSCI ACWI IMI in June 2002. We allow exceptions to this rule in cases where additional stocks are needed to populate countries with few assets.

In Figure 3.1, we also plot the number of securities in the GEM2 estimation universe over time. We report only assets that receive full regression weight (i.e., we exclude China domestic and GCC countries). While the large discontinuity in MSCI ACWI IMI (June 2002) has been removed from the estimation universe by our backward-exclusion policy, we observe another discontinuity in June 1994, coinciding with the start of MSCI ACWI IMI history. Although this would appear to violate our goal of estimation universe stability, this event does not present model difficulties for two reasons. First, the transition occurs well before the first available factor covariance matrix, so that it cannot lead to spurious jumps in risk forecasts. Second, the impact is mitigated by the fact that the affected assets are small-cap stocks so that, on a regression-weighted basis, the factor exposures are quite stable.

In summary, the GEM2 estimation universe is constructed as follows.

- June 2002 to present: MSCI ACWI IMI assets
- June 1994 to June 2002: MSCI ACWI IMI assets less those dropped in June 2002 (assets dropped are backward-excluded)
- June 1992 to June 1994: MSCI ACWI assets plus US and Japanese assets in MSCI ACWI IMI as of June 1994 (MSCI ACWI IMI members backward-included)
- MSCI China A Index and MSCI GCC Countries Index, with down-weighting in regression

The GEM estimation universe is based on the MSCI World Index, which includes developed markets. The number of constituents in this index is shown in Figure 3.1. Clearly, the MSCI World Index is far narrower than the GEM2 estimation universe. Since estimation error increases as the sample size is reduced, the broader estimation universe of GEM2 results in more efficient estimates of the factor returns.

## 3.2. GEM2 Factor Structure

The equity factor set in GEM2 includes a World factor $(w)$, countries $(c)$, industries $(i)$, and styles $(s)$. Every stock is assigned an exposure of 1 to the World factor. Hence, the local excess returns in Equation 2.5 can be rewritten as

$$r_n = f_w + \sum_c X_{nc} f_c + \sum_i X_{ni} f_i + \sum_s X_{ns} f_s + u_n . \tag{3.1}$$

Mathematically, the World factor represents the intercept term in the cross-sectional regression. Economically, it describes the aggregate up-and-down movement of the global equity market. Typically, the World factor is the dominant source of total risk for a diversified long-only portfolio.

For most institutional investors, however, the primary concern is the risk of *active* long/short portfolios. If both the portfolio and benchmark are fully invested – as is typically the case – then the active exposure to the World factor is zero. Similarly, if the long and short positions of a long/short portfolio are of equal absolute value, the net exposure to the World factor is zero. Thus, we must look beyond the World factor to other sources of risk.

Country factors play a critical role in global equity risk modeling. One reason is that they are powerful indicator variables for explaining the cross section of global equity returns. A second, related, reason is that the country allocation decision is central to many global investment strategies, and portfolio managers often must carefully monitor their exposures to these factors. We therefore include explicit country factors for all markets covered.

In Table 3.1, we present a list of the 55 countries covered by GEM2, together with their corresponding currencies. The country exposures $X_{nc}$ in GEM2 are set equal to 1 if stock $n$ is in country $c$, and set equal to 0 otherwise. We assign country exposures based on country membership within the MSCI ACWI IMI, MSCI China A Index and MSCI GCC Countries Index. Note that depository receipts and cross-listed assets are assigned factor exposures for the underlying or primary asset, as defined by the MSCI Equity Indices.

We also report the average and ending weights over the period January 1997 to January 2008. It is interesting to note the relative decline of the US and Japanese markets, and the rise of markets such as Brazil, Russia, India and China.

Industries are also important variables in explaining the sources of global equity return co-movement. One of the major strengths of GEM2 is to employ the Global Industry Classification Standard (GICS®) for the industry factor structure. The GICS scheme is hierarchical, with 10 top-level sectors, which are then divided into 24 industry groups, 68 industries, and 154 sub-industries. GICS applies a consistent global methodology to classify stocks based on careful evaluation of the firm's business model and economic operating environment. The GICS structure is reviewed annually by MSCI Barra and Standard & Poor's to ensure it remains timely and accurate.

Identifying which industry factors to include in the model involves a combination of judgment and empirical analysis. At one extreme, we could use the 10 GICS sectors as industry factors. Such broad groupings, however, would certainly fail to capture much of the cross-sectional variation in stock returns. At the other extreme, we could use all 154 sub-industries as the factor structure. Besides the obvious difficulties associated with the unwieldy numbers of factors (e.g., risk reporting, thin industries), such an approach would present a more serious problem for risk forecasting: although adding more factors always increases the in-sample $R^2$ of the cross-sectional regressions, many of the factor returns would not be statistically significant. Allowing noise-dominated "factors" into the model defeats the very purpose of a factor risk model.

In GEM2, selection of the industry factor structure begins at the second level of the GICS hierarchy, with each of the 24 industry groups automatically qualifying as a factor. This provides a reasonable level of granularity, without introducing an excessive number of factors. We then analyze each industry group, carefully examining the industries and sub-industries contained therein to determine if a more granular factor structure is warranted. The basic criteria we use to guide industry factor selection are: (a) the groupings of industries into factors must be economically intuitive, (b) the industry factors should have a strong degree of statistical significance, (c) incorporating an additional industry factor should significantly increase the explanatory power of the model, and (d) thin industries (those with few assets) should be avoided.

The result of this process is the set of 34 GEM2 industry factors, presented in Table 3.2. Industries that qualify as factors tend to exhibit volatile returns and have significant weight. We find that this relatively parsimonious set of factors captures most of the in-sample $R^2$ explained by the 154 sub-industries, but with a much higher degree of statistical significance. Also reported in Table 3.2 are the average and end-of-period industry weights from January 1997 to January 2008. The weights were computed using the entire GEM2 estimation universe (i.e., including China Domestic and GCC countries). Only five industries have end-of-period weights less than 100 bps, and these tend to be highly volatile, thus making them useful risk factors.

In Table 3.3, we report the underlying GICS codes that map to each of the GEM2 industry factors. In five of the ten GICS sectors, there is a direct mapping between industry groups and factors. In the other sectors, however, the factors are composed of collections of industries or sub-industries. For instance, in the energy sector, the Oil & Gas Exploration & Production sub-industry forms a separate factor, with the other four sub-industries constituting another factor. Extracting this particular sub-industry is warranted due to its high statistical significance and reasonable weight within the estimation universe. Other notable examples of industry factor refinement include (a) carving out the Airlines industry from Transportation, (b) including Biotechnology as a separate industry within the Health Care sector, and (c) using an Internet factor within the Information Technology sector. In all of these cases, the resulting factor groupings are financially intuitive and statistically significant.

Investment style represents another major source of systematic risk. Style factors, also known as *risk indices*, are designed to capture these sources of risk. They are constructed from financially intuitive stock attributes called *descriptors,* which serve as effective predictors of equity return covariance. Since the descriptors within a particular style factor are meant to capture the same underlying driver of returns, these descriptors tend to be significantly collinear. For instance, price-to-book ratio, dividend yield, and earnings yield are all attributes used to identify value stocks, and they tend to exhibit significant cross-sectional correlation. Although these descriptors have significant explanatory power on their own, naively including them as separate factors in the model may lead to serious multi-collinearity problems. Combining these descriptors into a single style factor overcomes this difficulty, and also leads to a more parsimonious factor structure.

Unlike country and industry factors, which are assigned exposures of either 0 or 1, style factor exposures are continuously distributed. To facilitate comparison across style factors, they are standardized to have a mean of 0 and a standard deviation of 1. In other words, if $d_{nl}^{Raw}$ is the raw value of stock $n$ for descriptor $l$, then the standardized descriptor value is given by

$$d_{nl} = \frac{d_{nl}^{Raw} - \mu_l}{\sigma_l}, \tag{3.2}$$

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

Some of the style factors are standardized on a *global-relative* basis, others on a *country-relative* basis. In the former case, the mean and standard deviation in Equation 3.2 are computed using the entire global cross section. In the latter case, the factors have mean 0 and standard deviation 1 within each country. When deciding which standardization convention to adopt, we consider both the intuitive meaning of the factor and its explanatory power.

Formally, descriptors are combined into risk indices as follows

$$X_{nk} = \sum_{l \in k} w_l \, d_{nl} \,, \tag{3.3}$$

where $w_l$ is the descriptor weight, and the sum takes place over all descriptors within a particular risk index. Descriptor weights are determined using an optimization algorithm to maximize the explanatory power of the model.

One of the major advances of GEM2 is its refined and expanded set of style factors. GEM2 uses eight style factors, compared to four in its predecessor. Below, we provide a qualitative description of each of the style factors:

- The *Volatility* factor is typically the most important style factor. In essence, it captures market risk that cannot be explained by the World factor. The most significant descriptor within the Volatility index is historical beta relative to the World portfolio (as proxied by the estimation universe). To better understand this factor, consider a fully invested long-only portfolio that is strongly tilted toward high-beta stocks. Intuitively, this portfolio has greater market risk than a portfolio with beta equal to one. This additional market risk is captured through positive exposure to the Volatility factor. Note that the time-series correlation between the World factor and the Volatility factor is typically very high, so that these two sources of risk reinforce one another in this example. If, by contrast, the portfolio is invested in low-beta stocks, then the risk from the Volatility and the World factors is partially cancelled, as intuitively expected. We standardize the Volatility factor on a global-relative basis. As a result, the mean exposure to Volatility within a country can deviate significantly from zero. This standardization convention is a natural one for a global model, as most investors regard stocks in highly volatile markets as having more exposure to the factor than those in low-volatility markets. This view is reflected in the data, as we find that the explanatory power of the factor is greater using the global-relative standardization.
- The *Momentum* factor often ranks second in importance after Volatility. Momentum differentiates stocks based on recent relative performance. Descriptors within Momentum include historical alpha from a 104-week regression and relative strength (over trailing six and 12 months) with a one-month lag. Similarly to Volatility, Momentum is standardized on a global-relative basis. This is also an intuitive convention for a global model. From the perspective of a global investor, a stock that strongly outperforms the World portfolio is likely to be considered a positive momentum stock, even if it slightly underperforms its country

peers. The empirical results support this view, as the Momentum factor standardized globally has greater explanatory power than one standardized on a country-relative basis.

- The *Size* factor represents another well-known source of return covariance. It captures the effect of large-cap stocks moving differently from small-cap stocks. We measure Size by a single descriptor: log of market capitalization. The explanatory power of the model is quite similar whether Size is standardized globally or on a country-by-country basis. We adopt the country-relative standardization, however, since it is more intuitive and consistent with investors' perception of the markets. For instance, major global equity indices, such as the MSCI Global Investable Market Indices, segment each country according to size, with the largest stocks inside each country always being classified as large-cap stocks. Moreover, standardizing the Size factor on a global-relative basis would serve as an unintended proxy for developed markets versus emerging markets, and increases collinearity with the country factors.

- The *Value* factor describes a major investment style which seeks to identify stocks that are priced low relative to fundamentals. We standardize Value on a country-relative basis. This again is consistent with the way major indices segment each market, with each country divided roughly equally into value and growth sub-indices. This convention also circumvents the difficulty of comparing fundamental data across countries with different accounting standards. GEM2 utilizes official MSCI data items for Value factor descriptors, as described in the *MSCI Barra Fundamental Data Methodology* handbook.[3]

- *Growth* differentiates stocks based on their prospects for sales or earnings growth. It is standardized on a country-relative basis, consistent with the construction of the MSCI Value and Growth Indices. Therefore, each country has approximately half the weight in stocks with positive Growth exposure, and half with negative exposure. The GEM2 Growth descriptors also utilize official MSCI data items, as described in the *MSCI Barra Fundamental Data Methodology* handbook.

- The *Non-Linear Size (NLS)* factor captures non-linearities in the payoff to the size factor across the market-cap spectrum. NLS is based on a single raw descriptor: the cube of the log of market capitalization. Since this raw descriptor is highly collinear with the Size factor, we orthogonalize it to the Size factor. This procedure does not affect the fit of the model, but does mitigate the confounding effects of collinearity, and thus preserves an intuitive meaning

[3] http://www.msci.com/methodology/meth_docs/MSCI_Sep08_Fundamental_Data.pdf

for the Size factor. The NLS factor is represented by a portfolio that goes long mid-cap stocks, and shorts large-cap and small-cap stocks.

- The *Liquidity* factor describes return patterns to stocks based on relative trading activity. Stocks with high turnover have positive exposure to Liquidity, whereas low-turnover stocks have negative exposure. Liquidity is standardized on a country-relative basis.
- *Leverage* captures the return difference between high-leverage and low-leverage stocks. The descriptors within Leverage include market leverage, book leverage, and debt-to-assets ratio. This factor is standardized on a country-relative basis.

In Appendix A, we provide additional detail on the individual descriptors comprising each style factor.

### 3.3. Comparison with GEM Factor Structure

GEM also contains factors for countries, industries, styles, and currencies. While there are many similarities with GEM2, there are also important differences in factor structure between the two models. In this section, we highlight some of these differences.

The overlap of country factors for GEM and GEM2 is very large. Both models cover the 48 developed and emerging markets contained within MSCI ACWI IMI. GEM also contains the following eight non-MSCI ACWI IMI countries: Sri Lanka, Venezuela, Nigeria, Slovakia, Zimbabwe, Saudi Arabia, Oman, and Bahrain. GEM2, on the other hand, supplements the 48 MSCI ACWI IMI markets with China Domestic (A-shares traded in mainland China) and the six GCC countries. Given the growing importance of the GCC region and the China domestic market, we included these in GEM2. By contrast, due to tight capital controls, illiquidity, or extreme economic instability (e.g., hyperinflation), Sri Lanka, Slovakia, Venezuela, Nigeria, and Zimbabwe are not presently included in GEM2.

A more fundamental difference between the models is that, whereas GEM2 uses country exposures of 0 or 1, GEM determines country exposure as historical beta estimated by a time-series regression. More specifically, the GEM country exposures are found by regressing 60 months of local excess stock returns against the local excess country returns within MSCI ACWI. While this has some intuitive appeal, it also presents difficulties. For instance, the country betas can only be estimated with the usual sampling error. Furthermore, since the regression takes place over the trailing 60 months, the estimates can be stale. As a result, the GEM country exposures – one of the key drivers of global equity risk – are sometimes unintuitive. This can lead to unexpected active exposures to GEM country factors, even when the active country weight is zero.

GEM uses 38 industry factors, with exposures given by 0 or 1, as in GEM2. The GEM structure was a good representation of the global industrial structure at the time of the model's release but is now significantly out of date. In Table 3.4, we report the mapping of GEM2 industries to GEM industries. In some sectors, such as Energy, we see that more than one GEM2 factor maps to a single GEM factor. In other sectors, such as Information Technology, we find the reverse (i.e., single GEM2 industries mapping to multiple GEM industries). In yet other sectors, such as Financials, there is a near a one-to-one mapping between GEM2 and GEM industries.

In Table 3.5, we report the mapping of GEM industries to GEM2 industries. As of January 2008, 13 of the 38 GEM industries have weights less than 100 bps. Many of these thinly populated GEM industries are not particularly volatile, such as Textiles & Apparel, or Forestry and Paper Products, so that they no longer constitute ideal risk factors.

GEM contains four style factors. Some are very similar to their GEM2 counterparts, while others are significantly different. Size is an example of a style factor that is essentially the same in GEM and GEM2. Both models define the Size factor by a single descriptor: log of total market capitalization.

The GEM Value factor is similar to the GEM2 Value factor. Both models use forward and trailing earnings-to-price ratios, dividend yield, and book-to-price ratios as descriptors. GEM2 also includes a fifth descriptor, cash earning-to-price ratio, in the Value factor. Descriptor weights, of course, are different between the two models.

The *Success* factor in GEM roughly corresponds to the GEM2 Momentum factor. Whereas GEM2 uses three descriptors in the factor, GEM uses only 12-month relative strength. Another difference is that the GEM2 factor applies a one-month lag to relative strength, whereas GEM does not. The GEM2 treatment is more consistent with the way most practitioners define momentum.

The biggest difference is with the GEM *Variability in Markets* (VIM) factor, which roughly corresponds to the GEM2 Volatility factor. The GEM factor is a single-descriptor risk index which uses the residual volatility (historical sigma) from a CAPM country regression. Since this descriptor does not directly capture covariance, however, it is weaker than the Volatility factor of GEM2.

A final comment is in order regarding country factors and volatility factors. GEM estimates country exposure by historical beta, and includes VIM as a separate factor. GEM2, by contrast, uses indicator variables for country exposure and includes Volatility (with historical beta as a descriptor) as a separate factor. The most meaningful direct comparison between the GEM and GEM2 country factors, therefore, should include not only the country factors, but also the Volatility and VIM factors as well.

## 4. Factor Returns

### 4.1. Data Quality and Outlier Treatment

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

In order to obtain the highest quality inputs, GEM2 leverages the same data infrastructure used for the construction of MSCI Global Investable Market Indices. Data items such as raw descriptors, GICS codes, country classifications, and clean daily stock returns are obtained from in-house sources that have already undergone extensive quality control.

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

In GEM2, we employ a multi-step algorithm to identify and treat outliers. The algorithm assigns each observation into one of three groups. The first group represents values so extreme that they are treated as potential data errors and removed from the estimation process. The second group represents values that are regarded as legitimate, but nonetheless so large that their impact on the model must be limited. We *trim* these observations to three standard deviations from the mean. The third group of observations, which forms the bulk of the distribution, consists of values less than three standard deviations from the mean. These observations are left unadjusted.

The thorough data quality assurance process, coupled with the robust outlier algorithm, ensures that the input data are clean and reliable. There is still the issue, however, of dealing with *missing* data. Data could be missing either due to lack of availability or due to removal by the outlier algorithm. The data items most likely to be missing include descriptors for the Value, Growth, Leverage, and Liquidity style factors. More rarely, stocks may be missing exposure to Volatility and Momentum factors; this occurs for securities which lack return history, such as recent spin-offs or IPOs. Other data items, such as Size, industry, and country exposures, are never missing, since these items are required for coverage universe membership.

If the underlying data required to compute factor exposures are missing, then the factor exposures are generated using a data-replacement algorithm. A simple approach would be to assign zero exposure to all missing values. A better approach, however, is to exploit known relationships between factor exposures. For instance, if we know that a stock is in the Semiconductor industry and has a positive exposure to Volatility, then it is likely that the stock also has a positive exposure to Growth.

We apply the data-replacement algorithm to the Value, Growth, Leverage, and Liquidity factors. The algorithm works by regressing these four factors (using only non-missing exposures) against Size, Volatility and industries. The slope coefficients of the regression are then used to estimate the factor exposures for the stocks with missing data. Countries are not used as explanatory variables in the regression since the four style factors (Value, Growth, Leverage and Liquidity) are standardized to be mean zero with respect to countries. The algorithm is not applied to the other four style factors (Size, Non-Linear Size, Volatility and Momentum) since these factor exposures are rarely, if ever, missing. Note also that the algorithm is applied at the factor level as opposed to the descriptor level. In other words, if a stock has data for some descriptors within a risk index, but not others, then the non-missing data will be used to compute the factor exposures. Only if *all* descriptors are absent does the replacement algorithm become active.

## 4.2. GEM2 Regression Methodology

The equity factor returns $f_k$ in GEM2 are estimated by regressing the local excess returns $r_n$ against the factor exposures $X_{nk}$,

$$r_n = \sum_{k=1}^{K_E} X_{nk} f_k + u_n . \tag{4.1}$$

GEM2 uses weighted least squares, assuming that the variance of specific returns is inversely proportional to the square root of total market capitalization.

As described in Section 3, the GEM2 equity factors include the World factor, countries, industries, and styles. Every stock in GEM2 has unit exposure to the World factor, and indicator variable exposures of 0 or 1 to countries and industries. As a result, the sum of all country factors equals the World factor, and similarly for industries, i.e.,

$$\sum_c X_{nc} = 1, \quad \text{and} \quad \sum_i X_{ni} = 1, \tag{4.2}$$

for all stocks $n$. In other words, the sum of all country columns in the factor exposure matrix gives a column with 1 in every entry, which corresponds to the World factor. The same holds for industry factors. The GEM2 factor structure, therefore, exhibits exact two-fold collinearity. Constraints must be applied to obtain a unique solution.

In GEM2 we adopt an intuitive set of constraints that require the cap-weighted country and industry factor returns to sum to zero,

$$\sum_c w_c f_c = 0, \quad \text{and} \quad \sum_i w_i f_i = 0, \tag{4.3}$$

where $w_c$ is the weight of the estimation universe in country $c$, and $w_i$ is the corresponding weight in industry $i$. These constraints remove the exact collinearities from the factor exposure matrix, without reducing the explanatory power of the model.

We can now give a more precise interpretation to the factors. Consider the cap-weighted estimation universe, with holdings $h_n^E$. The return of this portfolio $R_E$ can be attributed using the GEM2 factors,

$$R_E = f_w + \sum_c w_c f_c + \sum_i w_i f_i + \sum_s X_s^E f_s + \sum_n h_n^E u_n . \tag{4.4}$$

The constraints imply that the first two sums in Equation 4.4 are equal to zero. The third sum is also zero since the style factors are standardized to be cap-weighted mean zero; i.e., $X_s^E = 0$, for all styles $s$. The final sum in Equation 4.4 corresponds to the specific return of a broadly diversified portfolio, and is *approximately* zero (note it would be *exactly* zero if we used regression weights instead of capitalization weights). Thus, to an excellent approximation, Equation 4.4 reduces to

$$R_E \approx f_w . \tag{4.5}$$

In other words, the return of the World factor is essentially the cap-weighted return of the estimation universe.

To better understand the meaning of the pure factor portfolios, we report in Table 4.1 the long and short weights (January 2008) of the World portfolio and several pure factor portfolios in various market segments. The World portfolio is represented by the cap-weighted GEM2 estimation universe. The pure World factor is 100 percent net long and the net weights closely match those of the World portfolio in each segment. The other pure factors all have net weight of zero, and therefore represent long/short portfolios.

As a first approximation, the pure country factors can be regarded as going long 100 percent the particular country, and going short 100 percent the World portfolio. For instance, going long 100 percent Japan and short 100 percent the World results in a portfolio with roughly 91 percent weight in Japan, and -91 percent in all other countries. The pure country factors, however, have zero exposure to industry factors. This is accomplished by taking appropriate long/short combinations in other countries. For instance, the Japanese market is over-represented in the segment corresponding to the Automobile factor. To partially hedge this exposure, the pure Japan factor takes a net short position of -1.08 percent in the US Automobile segment. A similar short position would be found in the German Automobile segment.

The pure Automobile factor can be thought, as a first approximation, to be formed by going 100 percent long the Automobile industry and 100 percent short the World portfolio. A more refined view of the factor takes into account that the net weight in each country is zero. The pure Automobile factor naturally takes a large long position in Japanese automobiles, but hedges the Japan exposure by taking short positions in other Japanese segments.

The pure Volatility factor is perhaps the easiest to understand, as it takes offsetting long and short positions within all segments corresponding to GEM2 factors (e.g., Japan, US, and Automobiles). Note that the weights are not equal to zero for segments that do not correspond to GEM2 factors, such as Japanese automobiles.

## 4.3. Characteristics of GEM2 Factors

In this section, we present and discuss some of the quantitative characteristics of the GEM2 factors. In particular, we investigate the degree of collinearity among the factor exposures, and report on the statistical significance, performance and volatility of the factor returns.

A feature of the multi-factor framework is that it disentangles the effects of many variables acting simultaneously. Multi-collinearity among factors, however, can confound this clean separation of effects.

One measure of collinearity is the pair-wise cross-sectional correlation between factor exposures. In Table 4.2, we report regression-weighted correlations among style factors and industries, averaged over the period January 1997 to June 2008. In general, the correlations are intuitive in sign. For instance, Volatility is positively correlated with Growth, Liquidity, and Semiconductors, and negatively with Value and Utilities. Value, as expected, is positively correlated with Banks and Utilities, and negatively with Growth and Biotechnology. Although none of the correlations are particularly large, the correlations between pairs of style factors are typically larger, on average, than those between styles and industries.

The statistical significance of factor returns plays a key role in the construction of a risk model. Let $f_k$ be the factor return for a particular period, and $\mathrm{se}(k)$ be the standard error of this estimate. The *t*-statistic is given by

$$t_k = \frac{f_k}{\mathrm{se}(k)} . \tag{4.6}$$

Generally, a *t*-statistic with absolute value greater than 2 is considered significant at the 95 percent confidence level. Ideally, a risk factor should have high *t*-statistics that are persistent across time. Useful measures of this are the average squared *t*-statistic over time, and the percentage of observations with significant *t*-statistics.

In Table 4.3, we report the statistical significance of the GEM2 World and country factor returns over the sample period January 1997 to June 2008. Not surprisingly, the World factor scores the highest average squared *t*-statistic. Countries that have highly significant factor returns are Japan, US, and China Domestic. Generally, countries with large weights and volatile returns tend to be highly significant. Also note that most countries have a high proportion of statistically significant factor returns, far higher than the five percent one would expect if the factors were driven by pure noise.

In Table 4.3, we also report the sample kurtosis for weekly factor returns, average factor returns, volatilities, and correlations with the estimation universe over the January 1997 to June 2008 sample period. All of the country factor returns have kurtosis in excess of 3 (which is the level of kurtosis of a normal distribution), indicating that the returns are not normally distributed. Four countries – Hong Kong, Malaysia, Russia, and Singapore – have kurtosis greater than 10.

Since the country factors are net of the World factor, positive factor returns indicate that the country portfolio has outperformed the World portfolio. We see that the GCC markets have performed particularly well over the sample period, while several Asian markets have underperformed. We also report the annualized factor volatility. Not surprisingly, the US factor has the lowest volatility, followed by Canada and the developed European markets. This is a reflection of the fact that these countries tend to track the World portfolio relatively closely. The Sharpe Ratio is the annualized factor return divided by the factor volatility. Again, the GCC markets have performed well on a risk-adjusted basis. The final column in Table 4.3 reports the time-series correlation with the estimation universe return. It is interesting to observe that this correlation for the World factor is 0.994, confirming that it represents the cap-weighted estimation universe.

In Table 4.4, we report statistical characteristics for the GEM2 industry factors. Clearly, the GEM2 industry factors are strongly significant, with roughly two-thirds of them exhibiting significant *t*-statistics over 50 percent of the time. Table 4.4 provides strong empirical support for our industry factor structure: in all instances where we added a refined industry structure (e.g., carving out Internet, Biotechnology or Airlines), the resulting factors are highly volatile and statistically significant. Note that industry factors, like country factors, all have kurtosis levels exceeding 3.

There are several other interesting observations in Table 4.4. First, the worst performing industry factor over this period was Airlines (-8.54%). This contrasts sharply with the Transportation Non-Airline factor, which is derived from the same GICS industry group and had a positive return over the sample period. The top-performing industry factor, both on an absolute basis and a risk-adjusted basis, was Biotechnology. Note that defensive sectors, such as Consumer Staples or Utilities, were negatively correlated with the estimation universe, whereas aggressive sectors, such as Information Technology, were positively correlated.

In Table 4.5, we report the summary statistics for the GEM2 style factors. In this case, we divide the sample period into two sub-periods. In the first sample period (January 1997 to June 2002), the style factors were generally more volatile and had a greater degree of statistical significance than in the second sub-period (July 2002 to June 2008). This is not too surprising, since the first sub-period contains the rise and fall of the internet bubble. Note that the kurtosis of the Value factor in the second sample period is unusually high. This is caused primarily by two six-sigma events in August 2007.

We can discern from Table 4.5 a clear risk hierarchy in the style factors. Alone at the top is the Volatility factor, which is far more volatile and significant than any of the other factors. In the second tier are Momentum and Size, which are clearly set apart from the third group, containing the other five styles. In this third group, the Value factor tends to be the most significant.

Performance, however, tells a different story. Volatility performed very poorly during the first sub-period, and was essentially flat during the second. Value and Momentum, on the other hand, were strongly positive over both sample periods.

There is substantial correlation between some of the style factors with the estimation universe. For instance, from Table 4.5, we see that Volatility is strongly correlated (about 80 percent) with the estimation universe. This makes sense since, during up-markets, high-beta stocks tend to outperform low-beta stocks, so that the return to Volatility – which includes beta – is also positive. A similar argument holds for down-markets. Liquidity and Size also tend to be positively correlated with the estimation universe over both sample periods.

### 4.4. GEM Regression Methodology

GEM estimates factor returns through a cross-sectional regression, as expressed in Equation 4.1. Letting a tilde denote the GEM quantities, the regression can be written as

$$r_n = \sum_{k=1}^{\tilde{K}_E} \tilde{X}_{nk} \tilde{f}_k + \tilde{u}_n . \tag{4.7}$$

GEM also uses square root of market-cap regression weighting, as in GEM2.

While the GEM regression is formally equivalent to that in GEM2, there are nonetheless significant differences between the two regressions. The first major difference is that the GEM regression imposes constraints on all of the country factor returns,

$$\tilde{f}_k = R_k^W , \tag{4.8}$$

where $R_k^W$ is the *cap-weighted* return of country $k$ in MSCI ACWI. In other words, GEM does not estimate the country factor returns by regression; rather, it takes them from the country indices. The effect of these multiple constraints is to reduce the explanatory power of the model.

Another difference is that GEM does not include a World factor. Reflecting the global economy of its time, GEM embeds the market effect in the country factor, giving them greater importance than industry factors, which are net of the market. In GEM2, by contrast, the market is captured by the World factor, and both industries and countries are treated on an equal footing net of the market.

### 4.5. Comparing the Explanatory Power of the Models

An important summary statistic used to measure the explanatory power of a model is the $R^2$ of the cross-sectional regressions,

$$R^2 = 1 - \frac{\sum_n w_n u_n^2}{\sum_n w_n r_n^2}, \tag{4.9}$$

where $w_n$ is the regression weight of stock $n$, $u_n$ is the specific return, and $r_n$ is the local excess return. The $R^2$ can be interpreted as the ratio of the average *explained* squared return to the average *total* squared return.

To investigate the explanatory power of GEM2 relative to GEM, we carry out monthly cross-sectional regressions for both models. Since differences in $R^2$ can be quite small, it is important to control all external variables that can impact the $R^2$ values.[4] Therefore, in our regressions, we use (a) identical set of stocks (namely, the GEM2 estimation universe excluding China domestic and GCC), (b) identical input stock returns, and (c) identical regression weighting (i.e., square root of market cap). The only quantity that was varied was the factor exposure matrix. For the GEM regressions, we also impose the constraint on country factor returns (that they be equal to the country index returns), given by Equation 4.8.

In Figure 4.1, we plot the trailing 12-month average $R^2$ for GEM2 and GEM. The GEM2 factors consistently dominate their GEM counterparts over the entire sample period. On average, GEM2 has an $R^2$ advantage of about 350 basis points (bps) over GEM, although this difference rises to well over 500 bps during the period 2000-2002. For equity factor models, where even a 50 bp boost in $R^2$ is considered a significant achievement, an increase of 350 bps is unusually large.

[4] When comparing the explanatory power of two sets of factors, it is usually important to use the *adjusted R-squared*, which applies a slight correction based on the number of explanatory variables. This adjustment is unnecessary in this case, since GEM and GEM2 have the same number of factors.

In Table 4.6, we investigate the sources of this increase in explanatory power. We examine seven special cases, carefully selecting combinations of factors and running side-by-side regressions to isolate the desired effect. The sample period is from January 1997 to June 2008. Regressions labeled "A" use only GEM2 factors, whereas those labeled "B" employ either GEM factors or a mix of GEM2 and GEM factors.

In Case 1 of Table 4.6, we compare GEM2 country factors against the GEM country factors. Recall that GEM country exposures are given by the country betas, whereas GEM2 uses simple indicator variables. One expects that the GEM country factors, in isolation, would explain more than the GEM2 countries. Surprisingly, we see that GEM2 country factors outperform GEM's by 94 bps.

This puzzling result can be explained in terms of the GEM constraints (i.e., Equation 4.8). In Case 2, we run identical regressions as in Case 1, except this time removing the GEM constraints. As expected, the GEM country factors now explain more than the GEM2 factors.

As discussed in Section 3.3, because the GEM country factors incorporate covariance with the market, a fair comparison between the GEM2 and GEM country factors must include the Volatility factor (for GEM2) and VIM (for GEM). In Case 3, we run this comparison and see that GEM2 country factors now outperform the GEM factors by 137 bps. This also reflects the strength of the GEM2 volatility factor, which boosts the $R^2$ by 361 bps. Case 3 therefore demonstrates that GEM2 country factors, in proper combination with Volatility, outperform the GEM country factors.

In Case 4 of Table 4.6, we compare the explanatory power of GEM2 style factors versus GEM's. In Regression A, we include the full set of GEM2 factors. In Regression B, we swap the GEM2 style factors with the GEM style factors. This comparison, therefore, isolates the effect of style factors. We see that the GEM2 style factors outperform their GEM counterparts by 71 bps over the sample period.

In Case 5 of Table 4.6, we compare the explanatory power of GEM2 industry factors versus GEM. Again, this was accomplished by pulling out the GEM2 industry factors and replacing them with the GEM industry factors in Regression B. The result is that GEM2 industries explain, on average, 109 bps per month more than the GEM industries.

In Case 6, we compare the explanatory power of the complete factor sets for both models, but *without* imposing the GEM country constraints. GEM2 outperforms GEM by 142 bps per month.

In the final case, we run the actual regressions for GEM, which *includes* the country constraints. Over this sample period, GEM2 explains, on average, 355 bps more each month than GEM. This represents an impressive increase in explanatory power.

While Table 4.6 provides insight into the sources of the *difference* in $R^2$ between GEM2 and GEM, it is also illuminating to consider the sources of *absolute* $R^2$ in GEM2. In Figure 4.2, we plot the trailing 12-month $R^2$ of GEM2 factor groups over time. The first case we examine is the World factor in isolation. This is the single most important factor and, on average, explains about 10 percent in $R^2$. Its explanatory power, however, fluctuates over time. For instance, in 2000 the World factor explained only a few percentage points, whereas this figure rose to nearly 20 percent by 2003.

From Figure 4.2, it is evident that most of the explanatory power of the model comes from the combined effect of country and industry factors. Interestingly, the *relative* explanatory power of these factor groups changes significantly over time. In the early period, industries explained relatively little compared to countries. Then, some time in 1999, the explanatory power of industries suddenly exploded. From 2000-2003, industries actually explained a greater proportion of trailing 12-month $R^2$ than countries. Since 2003, the situation has partially reversed, with country factors reasserting their dominance although industry factors remain very important. This example illustrates some the insights obtained by the GEM2 factor structure, which places countries and industries on an equal footing.

Figure 4.2 also demonstrates that style factors contribute about 250 bps in $R^2$ beyond that explained by countries and industries alone. This amount, however, also varies considerably over time. In the early period, style factors added only about 100 bps, whereas this widened to about 500 bps during the 2000-2004 period. More recently, style factors have added about 200 bps.

The black dots in Figure 4.2 represent the monthly $R^2$ values for the complete set of GEM2 factors. This quantity varies dramatically from month to month, and tends to be largest during big market moves. In extreme months, commonalities in returns swamp asset-specific returns. For example, the largest single $R^2$ observed during the sample period was 0.75. This occurred in August 1998, when the market dropped by more than 15 percent.

# 5. Factor Covariance Matrix

## 5.1. GEM2 Factor Covariance Matrix

The GEM2 factor covariance matrix is built from the time series of weekly factor returns, $f_{kt}$. Weekly observations reduce sampling error and provide more responsive risk forecasts than monthly observations. Daily factor returns are not used for computing global covariances due to asynchronous trading effects, but they are computed for performance attribution purposes.

We use exponentially weighted moving averages (EWMA) to estimate the factor covariance matrix for both equity and currency blocks. This approach gives more weight to recent observations and is a simple, yet robust, method for dealing with data non-stationarity. An alternative approach would be to use generalized auto-regressive conditional heteroskedasticity, or GARCH. We find that EWMA estimators are typically more robust to changing market dynamics and produce more accurate risk forecasts than their GARCH counterparts.

In GEM2, we must also account for the possibility of serial correlation in factor returns, which can affect risk forecasts over a longer horizon. Suppose, for instance, that high-frequency returns are negatively correlated. In this case, long-horizon risk forecasts estimated on high-frequency data will be lower than that implied using simple square-root-of-time scaling, since returns one period tend to be partially offset by opposing returns the following period.

The prediction horizon in Barra risk models is typically one month. Models that are estimated on daily or weekly returns, therefore, must adjust for serial correlation. Note that models estimated on monthly observations, such as GEM, need not adjust for serial correlation, since the observation frequency coincides with the prediction horizon.

Full treatment of serial correlation must account for not only the correlation of one factor with itself across time, but also for the correlation of two factors with each other across different time periods. In GEM2, we model serial correlation using the Newey-West methodology[5] with two lags. This assumes that the return of any factor may be correlated with the return of any other factor up to two weeks prior.

It is useful to think of the factor covariance matrix as being composed of a correlation matrix, which is scaled by the factor volatilities. Volatilities and correlations can then be estimated separately using EWMA with different half-life parameters.

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

The volatility half-life determines the overall responsiveness of the model. Selecting the proper volatility half-life involves analyzing the trade-off between accuracy and responsiveness on the one hand, and stability on the other. If the volatility half-life is too long, the model gives undue weight to distant observations that have little to do with current market conditions. This leads to stable risk forecasts, but at the cost of reduced accuracy. By contrast, a short volatility half-life makes the model more responsive, generally improving the bias statistics, but at the cost of jumpier risk forecasts. Of course, the volatility half-life cannot be lowered without bound; if the half-life is too short, then sampling error can become so large that the risk forecasts are not only less stable, but also less accurate.

We measure the accuracy of the risk forecasts using bias statistics, which essentially represent the ratio of realized risk to forecast risk. A bias statistic of 1 is considered ideal, but sampling error ensures that realized bias statistics will deviate from 1, even for perfect risk forecasts. In Appendix B, we provide a review of bias statistics.

When carrying out bias statistic testing, one must consider the number of months to include in the observation window. Sampling error suggests that one should use the largest interval possible, since that would minimize sampling error. This would be a compelling argument for stationary data. In reality, however, financial markets are non-stationary. It is possible to over-predict risk for several years and then under-predict for others, while obtaining a bias statistic close to 1. Getting the right *average* risk forecast over 10 or 20 years is small consolation to a portfolio manager who can be wiped out due to poor risk forecasts in any single year.

Rolling 12-month intervals provide a more relevant framework for evaluating the accuracy of risk forecasts. Although sampling error is larger, this approach penalizes the model for poor risk forecasts during any individual year. For a given portfolio, we compute the rolling absolute deviation (RAD) of the bias statistic as follows. First, we calculate the bias statistic over the first year of the observation window, using 12 monthly observations. We then take the absolute value of the deviation of the bias statistic from 1. We then roll the 12-month window forward one month at a time, computing absolute deviations each month, until we reach the end of the observation window. For instance, a 138-month observation window will accommodate 127 separate 12-month windows. The RAD for the portfolio is the average of the absolute deviations over these 127 sub-periods. The RAD measure captures in a single number the accuracy of risk forecasts over rolling 12-month intervals by penalizing all observations throughout the sample history that deviate from the ideal bias statistic of 1.

As shown in Appendix B, for perfect risk forecasts and normally distributed returns (kurtosis 3), the critical value of the RAD over a 138-month observation window is 0.22. In other words, if the observed RAD is larger than 0.22, then with 95-percent confidence we can reject the null hypothesis that the risk forecast was accurate. Returns, of course, are not normally distributed. For a kurtosis of 5, which corresponds to slightly fat tails, the critical value of RAD for a 138-month observation window is 0.29. For even fatter tails, say a kurtosis of 10, the critical value increases to 0.34.

To study the question of the stability of risk forecasts, we compute the monthly absolute change in volatility for each factor $k$,

$$v_k(t) = \frac{\left|\sigma_k(t+1) - \sigma_k(t)\right|}{\sigma_k(t)}. \tag{5.1}$$

If the volatility forecast changes over one month from 6.0 percent to 5.7 percent, for instance, then the monthly absolute change is 5 percent.

The *mean variability* provides a measure of the stability of the risk forecasts, and is given by the average of the monthly absolute changes,

$$\bar{v} = \frac{1}{KT}\sum_{k,t} v_k(t), \tag{5.2}$$

where $K$ is the number of factors and $T$ is the number of time periods.

In Figure 5.1, we show the mean RAD for all GEM2 equity factors, plotted as a function of volatility half-life. The bias statistics were computed over rolling 12-month windows using a 138-month observation window from January 1997 to June 2008. We also plot the mean monthly variability $\bar{v}$ of the equity factors. The volatility half-lives of the GEM2S and GEM2L models are indicated by vertical lines.

The solid lines in Figure 5.1 are the actual GEM2 results, which incorporate a two-lag autocorrelation adjustment. For comparison, we also show by dashed lines the corresponding results without serial correlation adjustments. Including autocorrelation adjustments makes the risk forecasts more accurate, but also slightly increases the monthly variability in the forecasts.

From Figure 5.1, we see that the volatility half-life of GEM2S (18 weeks) minimizes the mean RAD, and therefore optimizes the accuracy of the risk forecasts. The mean RAD for GEM2S is about 0.24, versus 0.27 for GEM2L, which has a volatility half-life of 52 weeks. While GEM2S is therefore more accurate than GEM2L, the price one pays is jumpier risk forecasts. For GEM2S, the monthly variability is about 6.7 percent, versus 2.5 percent for GEM2L.

Figure 5.1 also clearly illustrates the effect of sampling error. Namely, as the half-life is reduced below about 15 weeks, the RAD increases dramatically while the variability of the risk forecast also increases.

In Figure 5.2 we show the mean RAD and variability for the GEM2 currency factors. The results were weighted by market capitalization within the GEM2 estimation universe. Qualitatively and quantitatively, the results for GEM2 currency factors are very similar to those for equity factors. For GEM2S, the 18-week volatility half-life minimizes the mean RAD at about 0.26, versus about 0.30 for GEM2L. The GEM2L forecasts, however, are more stable, with mean monthly variability of 2.6 percent versus 6.6 percent for GEM2S. Figure 5.2 also demonstrates the increased accuracy of risk forecasts that comes with treatment of autocorrelation.

Special care must be exercised when selecting the correlation half-life. To ensure a well-conditioned correlation matrix, the number of observations $T$ must be significantly greater than the number of factors $K$. In the extreme case that $T < K$, the correlation matrix will have zero eigenvalues. This can present problems for portfolio optimization, since it would be possible to find spurious active portfolios with seemingly zero factor risk.

For the short model, we use a correlation half-life of 104 weeks. A useful rule of thumb is that the number of effective observations is roughly three times the half-life. Therefore, a correlation half-life of 104 weeks corresponds to roughly 300 effective observations. For the long model, we use a correlation half-life of 156 weeks. The volatility and correlation half-lives used in GEM2 are the same as those used in other Barra single-country models, such as USE3S and USE3L.

Another complication in estimating the factor covariance matrix arises for the case of missing factor returns. One possible reason for missing data is holidays. For instance, stocks do not trade in China during golden week, so it is not possible to directly estimate a return to the China Domestic factor. More commonly, however, missing factor returns arise from using time series of differing lengths. For instance, the GCC stocks appear in the estimation universe as of July 2000, so the factor returns prior to this date are missing.

We use the EM algorithm[6] to estimate the factor covariance matrix in the presence of missing factor returns. This method employs an iterative procedure to estimate the covariance matrix

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

conditional on all available data. The EM algorithm assures that the factor covariance matrix is positive semi-definite.

## 5.2. GEM Factor Covariance Matrix

GEM uses a multi-step procedure to estimate the factor covariance matrix. The starting point is to compute the factor *correlation* matrix using monthly factor returns for all equity and currency factors. A 48-month half-life is used for this step, and the EM algorithm is applied to factors with short return series.

The next step is to scale the correlation matrix with the factor volatilities to obtain the first-pass covariance matrix. Here, different methods are used to estimate the volatilities of different factors. For the US, Japan, and Sweden, the country factor volatilities are taken from a GARCH model. The volatilities of the other equity factors are estimated using EWMA with half-lives ranging from 36 months to 90 months. The currency volatilities, on other hand, are taken from a separate fixed income risk model that uses highly responsive GARCH estimators or EWMA based on daily data with a 24-day half-life.

At this stage of the GEM covariance matrix construction process, the currency volatilities match those of the fixed income risk model, but the correlations do not. The fixed income risk model uses an EWMA with a 17-week half-life to estimate the currency correlations. To ensure that the forecasts of pure currency portfolios are the same in GEM and the fixed income risk model, a transformation is applied to the factor covariance matrix.

## 5.3. Currency Risk Forecasts

In this section, we compare the accuracy of the one-month currency risk forecasts of GEM2S, GEM2L, and GEM over the 138-month period from January 1997 to June 2008. In Table 5.1, we report the 12-month RAD, the monthly variability, and the sample kurtosis for all currencies in GEM2.

The average kurtosis of monthly currency returns over the sample period – above 10 – is quite large compared to that of typical equity factors. The range of kurtosis values is also very large. On the one hand, many developed European countries have kurtosis values of 3 or less. By contrast, the kurtosis of emerging-market currencies is typically much higher. The extreme case is the Russian ruble, which exhibits a sample kurtosis of almost 60 over the sample period.

From Table 5.1, we see that on an equal-weighted basis, GEM2S had the most accurate currency risk forecasts, with a mean RAD of 0.36. GEM2L had a mean RAD of 0.40, followed by GEM, with a mean RAD of 0.43. On a cap-weighted basis, which is more relevant to most international investors, GEM2S produces a mean RAD of 0.26, versus 0.30 for GEM2L and 0.32 for GEM.

When comparing these values, the reader should keep in mind the mean RAD for perfect risk forecasts and normally distributed returns (kurtosis of 3) is about 0.17. If the returns have fat tails (kurtosis greater than 3), then the mean RAD will be greater than 0.17. For instance, our simulations indicate that for a kurtosis of 10, the mean RAD for perfect risk forecasts is about 0.23. Relative to the theoretical lower bound of perfect risk forecasts and stationary returns, a mean RAD of 0.26 is quite impressive.

In Table 5.1 we also report the mean variability of the currency risk forecasts for the three models. For GEM, which uses a highly responsive currency model, the mean monthly variability of currency risk forecasts is 11.2 percent on a cap-weighted basis. GEM2S and GEM2L provide more stable currency risk forecasts. The mean monthly variability for GEM2S is 6.6 percent and for GEM2L it is 2.6 percent.

## 6. Specific Risk

### 6.1. GEM2 Specific Risk Model

GEM2 uses a structural model to estimate specific risk. This methodology is similar to that used in other Barra risk models, but employs weekly rather than monthly observations. These higher-frequency observations allow us to make the forecasts more responsive while also reducing estimation error.

Mathematically, the specific risk forecast of a stock is given as the product of three components

$$\sigma_n = \hat{S}\left(1+\hat{V}_n\right)K_M , \tag{6.1}$$

where $\hat{S}$ is the forecast mean absolute specific return of stocks in the estimation universe, $\hat{V}_n$ is the forecast relative absolute specific return, and $K_M$ is a multiplicative scalar which depends on the market-cap segment $M$ of the stock. The overall responsiveness of the model is dominated by $\hat{S}$, whereas $\hat{V}_n$ determines the cross-sectional variation in specific risk. The role of $K_M$ is to convert from weekly absolute specific returns to monthly standard deviation. The motivation for modeling specific risk in terms of absolute specific returns is that the estimation process is less prone to outliers.

The *realized* mean absolute specific return for a given period $t$ is computed as the cap-weighted average,

$$S_t = \sum_n w_{nt} |u_{nt}|, \tag{6.2}$$

where $u_{nt}$ is the realized specific return for stock $n$. We use the cap-weighted average because large-cap assets are less prone to extreme outliers. The forecast mean absolute specific return is given by the exponentially weighted average of the realized values,

$$\hat{S} = \sum_t \gamma_t S_t . \tag{6.3}$$

For GEM2S, we use a half-life parameter of 9 weeks for $\gamma_t$, versus 24 weeks for GEM2L. To determine these half-life parameters, we constructed a group of active portfolios, and then computed their average forecast variability. We then selected the half-life parameters $\gamma_t$ so that the specific risk forecast variability of these portfolios roughly matched the variability in forecasts of the factors, while also taking into account model performance. This results in highly accurate specific risk forecasts and greater stability in the proportion of risk coming from factors and stock-specific components.

In Figure 6.1, we plot the realized mean weekly absolute specific return over the period June 1992 to June 2008. In the mid-1990s, the absolute specific returns were quite low (about two percent), but then began to rise, hitting a peak of nearly six percent in early 2000. After the collapse of the internet bubble, absolute specific return levels fell sharply over the next five years. Since 2007, however, specific risk levels have again begun to rise.

In Figure 6.1, we also plot the forecast values from GEM2S and GEM2L. Overall, both models do an excellent job forecasting the absolute specific return levels. In the early history, up to 1997, the forecasts are virtually identical. GEM2S, however, catches the upward trend more quickly than GEM2L, and likewise responds more quickly when the volatility levels decrease following the collapse of the internet bubble.

We use a factor model to forecast the relative absolute specific return, whose *realized* values are given by

$$\varepsilon_{nt} = \frac{|u_{nt}| - S_t}{S_t} . \tag{6.4}$$

GEM2 factors already provide a sound basis for explaining $\varepsilon_{nt}$, since the level of relative absolute specific risk depends on the country, industry, and style exposures of the individual stocks. For instance, large-cap utility stocks tend to have lower specific returns than, say, small-cap internet stocks.

While the model factors are well suited to forecast specific risk, the most powerful explanatory variable is simply the trailing realized specific volatility. We therefore augment the GEM2 factor exposure matrix with this additional factor for the purpose of forecasting the relative absolute specific returns.

To estimate the relationship between factor exposures and relative absolute specific returns, we perform a 104-week pooled cross-sectional regression,

$$\varepsilon_{nt} = \sum_k \tilde{X}^t_{nk} \, g_k + \lambda_{nt} , \tag{6.5}$$

where $\tilde{X}^t_{nk}$ represents the augmented factor exposure matrix element. We use capitalization weights in the regression to reduce the impact of high-kurtosis small-cap stocks, and apply exponential weights of 26 weeks for the short model and 52 weeks for the long model. The forecast relative absolute specific return is given by

$$\hat{V}_n = \sum_k \tilde{X}_{nk} \, g_k , \tag{6.6}$$

where $\tilde{X}_{nk}$ is the most recent factor exposure, and $g_k$ is the slope coefficient estimated over the trailing 104 weeks.

The product $\hat{S}(1+\hat{V}_n)$ is the weekly forecast absolute specific return for the stock. We need to convert this to a monthly standard deviation. If the specific returns were normally distributed (temporally), then the time-series standard deviation of the ratio

$$\frac{u_{nt}}{S_t(1+\hat{V}_{nt})} \tag{6.7}$$

would be $\sqrt{\pi / 2}$, where $\hat{V}_{nt}$ is the forecast relative absolute specific return at time $t$. However, since the specific returns exhibit kurtosis, the standard deviation is usually slightly greater than this theoretical value. For this reason, the multiplicative scalar $K_M$ is sometimes called the *kurtosis correction*.

There is one final adjustment that must be applied to the kurtosis correction. The prediction horizon of the risk model is one month. We observe, on average, a small but persistent negative serial correlation in weekly specific returns. As a result, the monthly specific volatility will be slightly less than that suggested by simple square root of time scaling. If we rank stocks by degree of serial correlation, however, we find no relationship between the rank of stocks over non-overlapping periods. We therefore define the standardized return with two-lag autocorrelation,

$$b_{nt} = \frac{u_{n,t} + u_{n,t-1} + u_{n,t-2}}{\sqrt{3}S_t(1+\hat{V}_{nt})} . \tag{6.8}$$

The kurtosis correction is given by the realized standard deviation of $b_{nt}$,

$$K_M = \sigma(b_{nt}) , \tag{6.9}$$

where the standard deviation is computed over all observations within a particular market-cap segment $M$ over the trailing 104 weeks.

## 6.2. GEM Specific Risk Model

GEM estimates specific risk as the standard deviation of realized residual returns (i.e., historical sigma) from a CAPM country regression over the trailing 60 months. Since specific risk is different (and generally lower) than residual risk, a multiplicative scalar is applied to remove this bias.

If there are fewer than 24 months of returns, then the specific risk is estimated by regressing historical sigma values against country, industry, and size factors. The slope coefficients produce fits given the country, industry and size exposures of the stock.

## 6.3. Specific Risk Bias Statistics

In this section, we evaluate the accuracy of the specific risk forecasts for the models. In Table 6.1, we report summary-level rolling 12-month mean bias statistics and mean RAD for the GEM2 estimation universe. The bias statistics are computed at the asset level using realized specific returns in the numerator and specific risk forecasts in the denominator. We consider both cap-weighted and equal-weighted averages. We also report the percentage of observations falling within the confidence interval, which for 12 months is taken as [0.59, 1.41].

In Panel A of Table 6.1, we consider the 66-month sample period from January 1997 to June 2002. This was a period of generally rising specific risk levels. Both GEM2S and GEM2L perform extremely well over this period. The mean bias statistics are close to 1, indicating that both models adapt quickly to rising volatility levels. For the cap-weighted case, over 85 percent of observations in GEM2S fall within the confidence interval, versus about 82 percent for GEM2L. On an equal-weighted basis, which over-represents small-cap stocks, the percentage falling within the confidence interval is slightly lower, but still in the neighborhood of 80 percent.

GEM does not fare so well over the first sample period. The mean bias statistic is 1.21 on a cap-weighted basis, and 1.32 on an equal-weighted basis, indicating that risk was significantly under-predicted. The mean RAD is also significantly higher, e.g., 0.37 for GEM versus 0.24 for GEM2S on a cap-weighted basis.

In Panel B of Table 6.1, we consider the 72-month sample period from July 2002 to June 2008. This was a time of generally falling specific risk levels. The GEM2 bias statistics are still very close to 1, indicating that GEM2 did an excellent job adapting to new specific risk levels. GEM, by contrast, exhibits a mean bias statistic of 0.80 for the cap-weighted case, indicating that specific risk was over-predicted during this period. The GEM2 models also have a higher percentage of observations falling within the confidence interval than GEM.

In Panel C of Table 6.1, we examine the entire 138-month sample period. The results are consistent with Panels A and B. Namely, GEM2S and GEM2L have mean bias statistics very close to 1, have small RAD compared to GEM, and have a greater percentage of observations falling within the confidence interval.

In Figure 6.2, we plot the frequency distribution of the rolling 12-month bias statistics for the GEM2 and GEM models. The results are cap-weighted, and the analysis period covers 138 months from January 1997 to June 2008. The confidence interval is indicated by the two vertical lines. The performance of the GEM2S and GEM2L specific risk models is evidently very similar. The GEM2S distribution is slightly more peaked in the center, and has slightly smaller tails than the GEM2L distribution.

## 7. Model Performance

In this section, we evaluate the accuracy of risk forecasts for a combination of country, industry and style portfolios. The country and industry portfolios were constructed by carving out the constituents of the GEM2 estimation universe from the appropriate segment, and then cap-weighting the stocks. The style portfolios were constructed by ranking all stocks according to their exposures to the particular style factor, and then cap-weighting the stocks that ranked in the top and bottom 20 percent. For each of the long-only country, industry, and style portfolios thus formed, we also construct long/short active portfolios by selecting MSCI ACWI IMI as the benchmark. The sample period in all cases was from January 1997 to June 2008.

In Table 7.1, we report the 12-month RAD and sample kurtosis for long-only country portfolios and active portfolios. The average kurtosis was about 4.5 for both sets of portfolios. Overall, the GEM2S model produces the most accurate risk forecasts. The 12-month mean RAD for the country portfolios was 0.23, versus 0.25 for GEM2L, and 0.29 for GEM. For the active portfolios, the differences were even more pronounced. The mean RAD for GEM2S was 0.23, versus 0.26 GEM2L and 0.33 for GEM.

The GEM mean RAD is more than 40 percent higher than the mean RAD of GEM2S. This difference is quite significant, especially when one considers the theoretical lower bound. Assuming normally distributed returns, the mean RAD is about 0.17 for perfect risk forecasts. For a kurtosis of 4.5, however, the mean RAD is about 0.20 for perfect risk forecasts. By this measure, the GEM2S model has a mean RAD only 0.03 above the theoretical lower bound, versus 0.13 for GEM.

In Table 7.2, we report the 12-month RAD and sample kurtosis values for the industry portfolios. Industry factor kurtosis levels are slightly below those of country factors. Similar to the case for the country portfolios, the GEM2S model produces the most accurate risk forecasts, followed by GEM2L, and then GEM. For instance, for active industry portfolios, the mean RAD for GEM2S is 0.24, versus 0.27 for GEM2L, and 0.38 for GEM. It is interesting to note that GEM fares particularly poorly in the Biotechnology and Internet industries, suggesting that the older GEM industry structure is not well suited to capture the risk of these newer industries.

In Table 7.3, we report the 12-month RAD and sample kurtosis values for the style portfolios. Note that the kurtosis levels for the active portfolios are considerably higher than those for the long-only portfolios. Once again, the GEM2S model produces the most accurate risk forecasts, followed by GEM2L and GEM. For example, the mean RAD of the active style portfolios is 0.28 for GEM2S, 0.31 for GEM2L, and 0.44 for GEM.

In Table 7.4, we report summary results of 12-month rolling bias statistics for the complete set of country, industry, and style portfolios from Tables 7.1, 7.2, and 7.3. The sample period is divided into two sub-periods. Sub-period A is from January 1997 to June 2002, and is a time of generally rising volatility. All risk models slightly underpredict risk during sub-period A, although GEM2S underpredicts by a smaller margin than either GEM2L or GEM. GEM2S and GEM2L have smaller mean RAD than GEM, and a higher percentage of observations falling within the confidence interval.

Sub-period B runs from July 2002 to June 2008, and generally corresponds to a falling volatility environment. From Table 7.4, we see that the mean bias statistics for GEM2S are very close to 1, indicating that the model is able to quickly adjust to the lower volatility levels. GEM2L also performed reasonably well, with mean bias statistics of 0.89 for both long-only and active portfolios. The mean bias statistics for GEM during this period were 0.78 for long-only portfolios and 0.70 for active portfolios, indicating that the relatively unresponsive GEM is slow to adjust to the reduced volatility environment.

The entire sample period, running from January 1997 to June 2008, is shown in Panel C. GEM2S has the greatest proportion of observations (above 86 percent) falling within the confidence interval. The GEM2L model also performs well, with about 81 percent of the observations fall within the confidence interval. For GEM, only about 66 percent of the observations for active portfolios fall within the confidence interval. Most of the cases falling outside correspond to over-forecasting of risk during sub-period B, and are the result of attaching too much weight to the highly volatile events surrounding the internet bubble period.

In Figure 7.1, we plot the frequency distribution of the 12-month rolling bias statistics for the complete set of country, industry, and style *active* portfolios. The sample period was January 1997 to June 2008. The vertical lines denote the approximate confidence interval assuming normally distributed returns. This graph clearly illustrates that the GEM2S bias statistic distribution is more centered about the ideal value of 1, and has thinner tails. The GEM2L distribution is shifted somewhat to the left, and has slightly broader tails. The GEM distribution is shifted significantly more to the left and has fatter tails on both sides of the distribution.

## 8. Conclusion

GEM was the pioneering global equity factor model. It has helped managers measure the risk of their portfolios, understand the sources of the risk and construct better portfolios. GEM2 applies the methodological advances and greater data availability to produce a new standard in global equity modeling.

GEM2 contains many improvements over GEM. In this document, we have presented many of the methodological innovations and enhancements in the new GEM2 model. The highlights are:

- Use of weekly factor returns, allowing us to build a more responsive and accurate model
- GICS-based industry factors
- A broader estimation universe, based on MSCI ACWI IMI
- An improved and expanded set of style factors
- A cleaner separation of country and industry effects through the introduction of the World factor
- The introduction of a structural specific risk model

These enhancements, combined with painstaking data and analytical quality checks, make GEM2 a great advance in global equity risk modeling.

**Figure 3.1**

Number of stocks versus time, for MSCI ACWI IMI, GEM2 ESTU, and MSCI World.

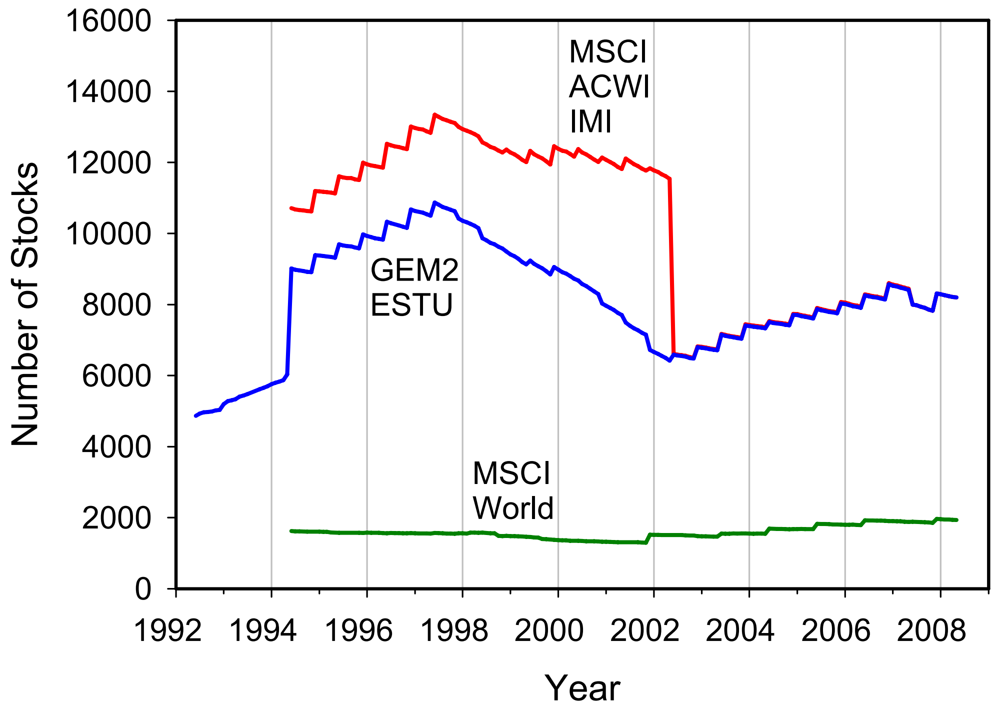


**Table 3.1**

GEM2 country factors and currencies. Weights are computed within the GEM2 estimation universe using total market capitalization. Average is taken over the period from January 1997 to January 2008.

| Country Code | Country Name | Currency Name | Average Weight | Jan-08 Weight |
|---|---|---|---|---|
| ARG | Argentina | Argentine Peso | 0.09 | 0.08 |
| AUS | Australia | Australian Dollar | 1.58 | 2.27 |
| AUT | Austria | Euro | 0.16 | 0.34 |
| BHR | Bahrain | Bahraini Dinar | 0.01 | 0.02 |
| BEL | Belgium | Euro | 0.62 | 0.63 |
| BRA | Brazil | Brazilian Real | 0.64 | 1.69 |
| CAN | Canada | Canadian Dollar | 2.53 | 3.14 |
| CHL | Chile | Chilean Peso | 0.17 | 0.24 |
| CHN | China Domestic | Chinese Yuan | 1.86 | 7.97 |
| CHX | China International | Hong Kong Dollar | 0.66 | 2.74 |
| COL | Colombia | Colombian Peso | 0.03 | 0.06 |
| CZE | Czech Republic | Czech Koruna | 0.06 | 0.14 |
| DNK | Denmark | Danish Krone | 0.38 | 0.42 |
| EGY | Egypt | Egyptian Pound | 0.05 | 0.15 |
| FIN | Finland | Euro | 0.58 | 0.64 |
| FRA | France | Euro | 4.08 | 4.80 |
| DEU | Germany | Euro | 3.23 | 3.41 |
| GRC | Greece | Euro | 0.27 | 0.41 |
| HKG | Hong Kong | Hong Kong Dollar | 1.05 | 1.47 |
| HUN | Hungary | Hungarian Forint | 0.06 | 0.08 |
| IND | India | Indian Rupee | 0.53 | 2.22 |
| IDN | Indonesia | Indonesian Rupiah | 0.13 | 0.31 |
| IRE | Ireland | Euro | 0.25 | 0.25 |
| ISR | Israel | Israeli Shekel | 0.19 | 0.29 |
| ITA | Italy | Euro | 2.00 | 1.91 |
| JPN | Japan | Japanese Yen | 11.24 | 8.23 |

**Table 3.1 (cont)**

| Country Code | Country Name | Currency Name | Average Weight | Jan-08 Weight |
|---|---|---|---|---|
| JOR | Jordan | Jordanian Dinar | 0.03 | 0.04 |
| KOR | Korea | Korean Won | 1.00 | 1.89 |
| KWT | Kuwait | Kuwaiti Dinar | 0.13 | 0.31 |
| MYS | Malaysia | Malaysian Ringgit | 0.42 | 0.48 |
| MEX | Mexico | Mexican Peso | 0.42 | 0.54 |
| MAR | Morocco | Moroccan Dirham | 0.03 | 0.08 |
| NLD | Netherland | Euro | 1.44 | 1.00 |
| NZL | New Zealand | New Zealand Dollar | 0.09 | 0.06 |
| NOR | Norway | Norwegian Krone | 0.33 | 0.65 |
| OMN | Oman | Omani Rial | 0.01 | 0.03 |
| PAK | Pakistan | Pakistan Rupee | 0.03 | 0.06 |
| PER | Peru | Peruvian Sol | 0.04 | 0.10 |
| PHL | Philippines | Philippine Peso | 0.06 | 0.13 |
| POL | Poland | Polish Zloty | 0.10 | 0.30 |
| PRT | Portugal | Euro | 0.19 | 0.22 |
| QAT | Qatar | Qatari Rial | 0.05 | 0.16 |
| RUS | Russia | Russian Ruble | 0.48 | 1.60 |
| SAU | Saudi Arabia | Saudi Rial | 0.31 | 0.82 |
| SGP | Singapore | Singapore Dollar | 0.47 | 0.66 |
| ZAF | South Africa | South African Rand | 0.62 | 0.79 |
| ESP | Spain | Euro | 1.38 | 1.76 |
| SWE | Sweden | Swedish Krone | 1.03 | 0.96 |
| CHE | Switzerland | Swiss Franc | 2.48 | 2.21 |
| TWN | Taiwan | Taiwan Dollar | 1.22 | 1.23 |
| THA | Thailand | Thailand Bhat | 0.18 | 0.33 |
| TUR | Turkey | New Turkish Lira | 0.17 | 0.41 |
| GBR | UK | U.K. Pound | 8.44 | 6.98 |
| ARE | UAE | Emirati Dirham | 0.06 | 0.28 |
| USA | US | US Dollar | 46.37 | 31.99 |

**Table 3.2**

GEM2 Industry Factors. Weights are computed within the GEM2 estimation universe using total market capitalization. Average is taken over the period from January 1997 to January 2008.

| GICS Sector | GEM2 Code | GEM2 Industry Factor Name | Average Weight | Jan-08 Weight |
|---|---|---|---|---|
| Energy | 1 | Energy Equipment & Services | 0.75 | 1.29 |
| | 2 | Oil, Gas & Consumable Fuels | 4.88 | 9.32 |
| | 3 | Oil & Gas Exploration & Production | 1.00 | 1.72 |
| Materials | 4 | Chemicals | 2.36 | 2.84 |
| | 5 | Construction, Containers, Paper | 1.38 | 1.24 |
| | 6 | Aluminum, Diversified Metals | 1.05 | 2.41 |
| | 7 | Gold, Precious Metals | 0.37 | 0.58 |
| | 8 | Steel | 0.79 | 1.83 |
| Industrials | 9 | Capital Goods | 7.33 | 8.60 |
| | 10 | Commercial & Professional Services | 1.43 | 0.77 |
| | 11 | Transportation Non-Airline | 1.82 | 2.32 |
| | 12 | Airlines | 0.37 | 0.45 |
| Consumer Discretionary | 13 | Automobiles & Components | 2.52 | 2.29 |
| | 14 | Consumer Durables & Apparel | 2.33 | 1.93 |
| | 15 | Consumer Services | 1.35 | 1.39 |
| | 16 | Media | 3.24 | 2.11 |
| | 17 | Retailing | 3.42 | 2.08 |
| Consumer Staples | 18 | Food & Staples Retailing | 1.82 | 1.76 |
| | 19 | Food, Beverage & Tobacco | 4.56 | 4.37 |
| | 20 | Household & Personal Products | 1.43 | 1.20 |
| Health Care | 21 | Health Care Equipment & Services | 2.13 | 1.93 |
| | 22 | Biotechnology | 0.78 | 0.68 |
| | 23 | Pharmaceuticals, Life Sciences | 6.17 | 3.82 |
| Financials | 24 | Banks | 10.52 | 10.83 |
| | 25 | Diversified Financials | 5.63 | 5.06 |
| | 26 | Insurance | 4.61 | 4.14 |
| | 27 | Real Estate | 2.08 | 3.07 |
| Information Technology | 28 | Internet Software & Services | 0.62 | 0.74 |
| | 29 | IT Services, Software | 3.24 | 2.56 |
| | 30 | Communications Equipment | 2.46 | 1.41 |
| | 31 | Computers, Electronics | 3.69 | 2.81 |
| | 32 | Semiconductors | 2.47 | 1.52 |
| Telecom | 33 | Telecommunication Services | 7.11 | 5.84 |
| Utilities | 34 | Utilities | 4.31 | 5.08 |

**Table 3.3**

Mapping of GEM2 industry factors to GICS codes.

| Code | GEM2 Industry Factor Name | GICS Codes |
|---|---|---|
| 1 | Energy Equipment & Services | 101010 |
| 2 | Oil, Gas & Consumable Fuels | 10102010, 10102030, 10102040, 10102050 |
| 3 | Oil & Gas Exploration & Production | 10102020 |
| 4 | Chemicals | 151010 |
| 5 | Construction, Containers, Paper | 151020, 151030, 151050 |
| 6 | Aluminum, Diversified Metals | 15104010, 15104020 |
| 7 | Gold, Precious Metals | 15104030, 15104040 |
| 8 | Steel | 15104050 |
| 9 | Capital Goods | 2010 |
| 10 | Commercial & Professional Services | 2020 |
| 11 | Transportation Non-Airline | 203010, 203030, 203040, 203050 |
| 12 | Airlines | 203020 |
| 13 | Automobiles & Components | 2510 |
| 14 | Consumer Durables & Apparel | 2520 |
| 15 | Consumer Services | 2530 |
| 16 | Media | 2540 |
| 17 | Retailing | 2550 |
| 18 | Food & Staples Retailing | 3010 |
| 19 | Food, Beverage & Tobacco | 3020 |
| 20 | Household & Personal Products | 3030 |
| 21 | Health Care Equipment & Services | 3510 |
| 22 | Biotechnology | 352010 |
| 23 | Pharmaceuticals, Life Sciences | 352020, 352030 |
| 24 | Banks | 4010 |
| 25 | Diversified Financials | 4020 |
| 26 | Insurance | 4030 |
| 27 | Real Estate | 4040 |
| 28 | Internet Software & Services | 451010 |
| 29 | IT Services, Software | 451020, 451030 |
| 30 | Communications Equipment | 452010 |
| 31 | Computers, Electronics | 452020, 452030, 452040 |
| 32 | Semiconductors | 4530 |
| 33 | Telecommunication Services | 5010 |
| 34 | Utilities | 5510 |

### Table 3.4

Mapping from GEM2 Industries to GEM Industries, as of January 2008. Mapping is based on capitalization weights using common assets between the GEM2 estimation universe and the GEM coverage universe. Mapped weights below 10-percent threshold are not shown.

| GEM2 Code | GEM2 Industry Factor Name | GEM Code | GEM Industry Factor Name | Mapped Weight |
|---|---|---|---|---|
| 1 | Energy Equipment & Services | 14 | Energy Equipment & Services | 97 |
| 2 | Oil, Gas & Consumable Fuels | 1 | Energy Sources | 98 |
| 3 | Oil & Gas Exploration & Production | 1 | Energy Sources | 100 |
| 4 | Chemicals | 4 | Chemicals | 98 |
| 5 | Construction, Containers, Paper | 3 | Building Materials & Components | 54 |
| 5 | | 5 | Forestry and Paper Products | 35 |
| 6 | Aluminum, Diversified Metals | 6 | Metals - Non-Ferrous | 55 |
| 6 | | 7 | Metals – Steel | 16 |
| 6 | | 38 | Gold Mines | 12 |
| 6 | | 1 | Energy Sources | 10 |
| 7 | Gold, Precious Metals | 38 | Gold Mines | 63 |
| 7 | | 6 | Metals - Non-Ferrous | 37 |
| 8 | Steel | 7 | Metals – Steel | 98 |
| 9 | Capital Goods | 37 | Multi-Industry | 24 |
| 9 | | 16 | Machinery & Engineering | 24 |
| 9 | | 9 | Aerospace & Military Technology | 13 |
| 9 | | 10 | Construction & Housing | 13 |
| 9 | | 12 | Electrical & Electronics | 10 |
| 10 | Commercial & Professional Services | 25 | Business & Public Services | 99 |
| 11 | Transportation Non-Airline | 30 | Transportation – Road & Rail | 50 |
| 11 | | 25 | Business & Public Services | 24 |
| 11 | | 31 | Transportation – Shipping | 23 |
| 12 | Airlines | 29 | Transportation - Airlines | 100 |
| 13 | Automobiles & Components | 18 | Automobiles | 69 |
| 13 | | 15 | Industrial Components | 27 |
| 14 | Consumer Durables & Apparel | 17 | Appliances & Household Durables | 27 |
| 14 | | 22 | Recreation & Other Consumer Goods | 24 |
| 14 | | 23 | Textiles & Apparel | 22 |
| 14 | | 10 | Construction & Housing | 13 |
| 15 | Consumer Services | 26 | Leisure & Tourism | 90 |
| 15 | | 25 | Business & Public Services | 10 |
| 16 | Media | 24 | Broadcasting & Publishing | 91 |
| 17 | Retailing | 27 | Merchandizing | 84 |
| 18 | Food & Staples Retailing | 27 | Merchandizing | 100 |
| 19 | Food, Beverage & Tobacco | 19 | Beverages & Tobacco | 56 |
| 19 | | 20 | Food & Household Products | 44 |
| 20 | Household & Personal Products | 20 | Food & Household Products | 70 |
| 20 | | 17 | Appliances & Household Durables | 18 |
| 20 | | 21 | Health and Personal Care | 10 |
| 21 | Health Care Equipment & Services | 21 | Health and Personal Care | 73 |
| 21 | | 25 | Business & Public Services | 27 |
| 22 | Biotechnology | 21 | Health and Personal Care | 99 |
| 23 | Pharmaceuticals, Life Sciences | 21 | Health and Personal Care | 100 |
| 24 | Banks | 33 | Banking | 100 |
| 25 | Diversified Financials | 34 | Financial Services | 98 |
| 26 | Insurance | 35 | Insurance | 100 |
| 27 | Real Estate | 36 | Real Estate | 95 |
| 28 | Internet Software & Services | 25 | Business & Public Services | 95 |
| 29 | IT Services, Software | 25 | Business & Public Services | 99 |
| 30 | Communications Equipment | 28 | Telecommunication | 98 |
| 31 | Computers, Electronics | 13 | Electrical Components & Instruments | 58 |
| 31 | | 11 | Data Processing & Reproduction | 27 |
| 31 | | 25 | Business & Public Services | 11 |
| 32 | Semiconductors | 13 | Electronic Components & Instruments | 100 |
| 33 | Telecommunication Services | 28 | Telecommunication | 100 |
| 34 | Utilities | 2 | Utilities-Electrical & Gas | 94 |

### Table 3.5

Mapping from GEM Industries to GEM2 Industries, as of January 2008. Mapping is based on capitalization weights using common assets between the GEM2 estimation universe and the GEM coverage universe. Mapped weights below 10-percent threshold are not shown.

| GEM Code | GEM Industry Name | ESTU Weight | GEM2 Code | GEM2 Industry Name | Mapped Weight |
|---|---|---|---|---|---|
| 1 | Energy Sources | 10.60 | 2 | Oil, Gas & Consumable Fuels | 82 |
| 1 | | | 3 | Oil & Gas Exploration & Production | 16 |
| 2 | Utilities - Electrical & Gas | 5.07 | 34 | Utilities | 99 |
| 3 | Building Materials & Components | 0.96 | 5 | Construction, Containers, Paper | 64 |
| 3 | | | 9 | Capital Goods | 34 |
| 4 | Chemicals | 2.78 | 4 | Chemicals | 98 |
| 5 | Forestry and Paper Products | 0.35 | 5 | Construction, Containers, Paper | 99 |
| 6 | Metals - Non Ferrous | 1.53 | 6 | Aluminum, Diversified Metals | 84 |
| 6 | | | 7 | Gold, Precious Metals | 13 |
| 7 | Metals - Steel | 2.02 | 8 | Steel | 81 |
| 7 | | | 6 | Aluminum, Diversified Metals | 18 |
| 8 | Misc. Materials & Commodities | 0.13 | 6 | Aluminum, Diversified Metals | 39 |
| 8 | | | 5 | Construction, Containers, Paper | 35 |
| 9 | Aerospace & Military Technology | 1.26 | 9 | Capital Goods | 100 |
| 10 | Construction & Housing | 1.28 | 9 | Capital Goods | 79 |
| 10 | | | 14 | Consumer Durables & Apparel | 17 |
| 11 | Data Processing & Reproduction | 0.99 | 31 | Computers, Electronics | 97 |
| 12 | Electrical & Electronics | 0.91 | 9 | Capital Goods | 91 |
| 13 | Electrical Components & Instruments | 3.31 | 31 | Computers, Electronics | 48 |
| 13 | | | 32 | Semiconductors | 45 |
| 14 | Energy Equipment & Services | 1.35 | 1 | Energy Equipment & Services | 91 |
| 15 | Industrial Components | 0.53 | 13 | Automobiles & Components | 88 |
| 16 | Machinery & Engineering | 1.80 | 9 | Capital Goods | 99 |
| 17 | Appliances & Household Durables | 0.76 | 14 | Consumer Durables & Apparel | 62 |
| 17 | | | 20 | Household & Personal Products | 27 |
| 18 | Automobiles | 1.89 | 13 | Automobiles & Components | 95 |
| 19 | Beverages & Tobacco | 2.66 | 19 | Food, Beverage and Tobacco | 100 |
| 20 | Food & Household Products | 2.94 | 19 | Food, Beverage and Tobacco | 69 |
| 20 | | | 20 | Household & Personal Products | 30 |
| 21 | Health & Personal Care | 6.30 | 23 | Pharmaceuticals & Life Sciences | 63 |
| 21 | | | 21 | Health Care Equipment & Services | 23 |
| 21 | | | 22 | Biotechnology | 11 |
| 22 | Recreation & Other Consumer Goods | 0.56 | 14 | Consumer Durables & Apparel | 86 |
| 22 | | | 13 | Automobiles & Components | 12 |
| 23 | Textiles & Apparel | 0.29 | 14 | Consumer Durables & Apparel | 90 |
| 24 | Broadcasting & Publishing | 2.08 | 16 | Media | 99 |
| 25 | Business & Public Services | 6.01 | 29 | IT Services, Software | 42 |
| 25 | | | 10 | Commercial & Professional Services | 13 |
| 25 | | | 28 | Internet Software & Services | 12 |
| 26 | Leisure & Tourism | 1.16 | 15 | Consumer Services | 98 |
| 27 | Merchandising | 3.67 | 18 | Food & Staples Retailing | 49 |
| 27 | | | 17 | Retailing | 49 |
| 28 | Telecommunications | 8.45 | 33 | Telecommunication Services | 80 |
| 28 | | | 30 | Communications Equipment | 20 |
| 29 | Transportation - Airlines | 0.36 | 12 | Airlines | 91 |
| 30 | Transportation - Road & Rail | 1.01 | 11 | Transportation Non-Airline | 99 |
| 31 | Transportation - Shipping | 0.40 | 11 | Transportation Non-Airline | 89 |
| 32 | Wholesale & International Trade | 0.62 | 9 | Capital Goods | 78 |
| 32 | | | 17 | Retailing | 19 |
| 33 | Banking | 11.19 | 24 | Banks | 100 |
| 34 | Financial Services | 5.23 | 25 | Diversified Financials | 100 |
| 35 | Insurance | 3.95 | 26 | Insurance | 100 |
| 36 | Real Estate | 2.40 | 27 | Real Estate | 99 |
| 37 | Multi-Industry | 2.46 | 9 | Capital Goods | 95 |
| 38 | Gold Mines | 0.75 | 7 | Gold, Precious Metals | 56 |
| 38 | | | 6 | Aluminum, Diversified Metals | 44 |

**Table 4.1**

Segment weights, as of January 2008, for the World portfolio and several pure factor portfolios. The World portfolio is represented by the cap-weighted GEM2 estimation universe.

| Segment | (ESTU) World Portfolio | Pure World Factor | Pure Japan Factor | Pure US Factor | Pure Auto Factor | Pure Volatility Factor |
|---|---|---|---|---|---|---|
| World (Net) | 100.00 | 100.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Long | 100.00 | 106.36 | 97.84 | 66.27 | 113.72 | 55.83 |
| Short | 0.00 | -6.36 | -97.84 | -66.27 | -113.72 | -55.83 |
| Japan (Net) | 9.36 | 9.33 | 90.67 | -9.33 | 0.00 | 0.00 |
| Long | 9.36 | 10.49 | 90.67 | 0.21 | 33.95 | 6.35 |
| Short | 0.00 | -1.16 | 0.00 | -9.54 | -33.95 | -6.35 |
| US (Net) | 35.68 | 35.18 | -35.18 | 64.82 | 0.00 | 0.00 |
| Long | 35.68 | 35.37 | 2.17 | 64.83 | 20.47 | 16.99 |
| Short | 0.00 | -0.19 | -37.35 | 0.00 | -20.47 | -16.99 |
| Auto (Net) | 2.35 | 2.40 | 0.00 | 0.00 | 97.60 | 0.00 |
| Long | 2.35 | 2.52 | 4.18 | 1.00 | 97.60 | 1.15 |
| Short | 0.00 | -0.12 | -4.18 | -1.00 | 0.00 | -1.15 |
| Japan Auto (Net) | 1.09 | 0.89 | 4.18 | -0.05 | 33.93 | 0.21 |
| Long | 1.09 | 0.93 | 4.18 | 0.06 | 33.93 | 0.39 |
| Short | 0.00 | -0.04 | 0.00 | -0.11 | 0.00 | -0.18 |
| US Auto (Net) | 0.22 | 0.29 | -1.08 | 0.92 | 15.69 | 0.26 |
| Long | 0.22 | 0.29 | 0.00 | 0.92 | 15.69 | 0.34 |
| Short | 0.00 | 0.00 | -1.08 | 0.00 | 0.00 | -0.08 |

### Table 4.2

Regression-weighted cross-sectional correlation of style and industry factor exposures. Results are averages over the period January 1997 to June 2008. Correlations above 0.10 in absolute value are shaded in gray.

| Factor | Volatility | Momentum | Size | Value | Growth | NL Size | Liquidity | Leverage |
|---|---|---|---|---|---|---|---|---|
| Volatility | 1.00 | -0.05 | -0.09 | -0.25 | 0.27 | -0.06 | 0.42 | -0.07 |
| Momentum | -0.05 | 1.00 | 0.12 | -0.17 | 0.08 | 0.15 | 0.09 | -0.08 |
| Size | -0.09 | 0.12 | 1.00 | -0.11 | -0.05 | 0.11 | 0.10 | -0.03 |
| Value | -0.25 | -0.17 | -0.11 | 1.00 | -0.24 | -0.03 | -0.15 | 0.22 |
| Growth | 0.27 | 0.08 | -0.05 | -0.24 | 1.00 | -0.02 | 0.18 | -0.17 |
| Non-linear Size | -0.06 | 0.15 | 0.11 | -0.03 | -0.02 | 1.00 | 0.11 | 0.01 |
| Liquidity | 0.42 | 0.09 | 0.10 | -0.15 | 0.18 | 0.11 | 1.00 | -0.03 |
| Leverage | -0.07 | -0.08 | -0.03 | 0.22 | -0.17 | 0.01 | -0.03 | 1.00 |
| Energy Equipment & Services | 0.06 | 0.02 | -0.03 | -0.02 | 0.06 | 0.02 | 0.07 | -0.03 |
| Oil, Gas & Consumable Fuels | -0.02 | 0.02 | 0.11 | 0.06 | -0.03 | -0.04 | 0.00 | -0.01 |
| Oil & Gas Exploration & Production | 0.01 | 0.02 | -0.01 | 0.01 | 0.00 | 0.00 | 0.02 | 0.01 |
| Chemicals | -0.03 | 0.00 | -0.01 | 0.04 | -0.04 | 0.03 | -0.01 | 0.00 |
| Construction, Containers, Paper | -0.04 | -0.01 | -0.02 | 0.07 | -0.06 | 0.04 | -0.04 | 0.06 |
| Aluminum, Diversified Metals | 0.03 | 0.02 | 0.02 | 0.03 | 0.00 | 0.01 | 0.03 | 0.00 |
| Gold, Precious Metals | 0.00 | -0.01 | 0.00 | -0.06 | 0.00 | 0.00 | -0.01 | -0.04 |
| Steel | 0.04 | 0.02 | -0.01 | 0.09 | -0.02 | 0.01 | 0.03 | 0.03 |
| Capital Goods | 0.01 | 0.01 | -0.07 | 0.04 | -0.02 | 0.01 | 0.00 | -0.01 |
| Commercial & Professional Services | 0.01 | 0.00 | -0.09 | -0.05 | 0.06 | -0.01 | -0.02 | -0.03 |
| Transportation Non-Airline | -0.06 | 0.01 | 0.00 | 0.02 | -0.02 | 0.02 | -0.04 | 0.08 |
| Airlines | 0.02 | -0.02 | 0.01 | 0.02 | -0.01 | 0.02 | 0.02 | 0.08 |
| Automobiles & Components | 0.01 | 0.00 | 0.01 | 0.10 | -0.02 | 0.00 | -0.01 | 0.00 |
| Consumer Durables & Apparel | 0.01 | -0.01 | -0.07 | 0.05 | 0.00 | 0.00 | 0.03 | -0.06 |
| Consumer Services | -0.02 | 0.00 | -0.06 | -0.03 | 0.04 | 0.02 | 0.01 | 0.02 |
| Media | -0.01 | -0.03 | 0.01 | -0.12 | 0.02 | 0.02 | -0.03 | 0.00 |
| Retailing | 0.03 | 0.01 | -0.05 | -0.01 | 0.05 | 0.01 | 0.04 | -0.06 |
| Food & Staples Retailing | -0.07 | -0.01 | 0.02 | -0.02 | 0.00 | 0.01 | -0.03 | -0.01 |
| Food, Beverage & Tobacco | -0.14 | 0.00 | 0.02 | -0.01 | -0.05 | 0.02 | -0.07 | -0.01 |
| Household & Personal Products | -0.05 | 0.00 | 0.03 | -0.04 | -0.02 | -0.01 | -0.03 | -0.03 |
| Health Care Equipment & Services | -0.02 | 0.02 | -0.06 | -0.08 | 0.09 | 0.01 | 0.04 | -0.05 |
| Biotechnology | 0.09 | 0.00 | -0.05 | -0.17 | 0.07 | -0.03 | 0.07 | -0.05 |
| Pharmaceuticals, Life Sciences | -0.04 | 0.00 | 0.06 | -0.10 | 0.02 | -0.04 | 0.00 | -0.10 |
| Banks | -0.08 | 0.00 | 0.12 | 0.12 | -0.09 | -0.05 | -0.10 | 0.10 |
| Diversified Financials | 0.05 | 0.00 | 0.02 | 0.04 | 0.01 | 0.01 | -0.03 | 0.04 |
| Insurance | -0.06 | 0.00 | 0.04 | 0.11 | -0.05 | 0.03 | -0.06 | -0.10 |
| Real Estate | -0.13 | 0.02 | -0.09 | 0.07 | -0.08 | 0.03 | -0.07 | 0.17 |
| Internet Software & Services | 0.15 | -0.02 | -0.04 | -0.12 | 0.12 | -0.03 | 0.07 | -0.06 |
| IT Services, Software | 0.15 | -0.01 | -0.06 | -0.14 | 0.12 | -0.02 | 0.09 | -0.12 |
| Communications Equipment | 0.16 | -0.02 | -0.01 | -0.09 | 0.06 | -0.03 | 0.08 | -0.07 |
| Computers, Electronics | 0.11 | -0.02 | -0.02 | -0.05 | 0.04 | -0.03 | 0.09 | -0.09 |
| Semiconductors | 0.21 | -0.02 | 0.00 | -0.10 | 0.09 | -0.02 | 0.16 | -0.08 |
| Telecommunication Services | 0.04 | -0.02 | 0.15 | -0.07 | 0.01 | -0.07 | 0.01 | 0.05 |
| Utilities | -0.17 | 0.01 | 0.06 | 0.11 | -0.13 | 0.04 | -0.07 | 0.20 |

**Table 4.3**

Country factor summary statistics based on weekly factor returns. Sample period is from January 1997 to June 2008. Averages computed excluding World, China Domestic, and GCC factors.

| Factor Name | Average Squared t-statistic | Percent Observations with \|t\|>2 | (weekly) Factor Kurtosis | Annualized Factor Return | Annualized Factor Volatility | Factor Sharpe Ratio | Correlation with ESTU |
|---|---|---|---|---|---|---|---|
| World Factor | 753.02 | 94.0 | 4.91 | 6.96 | 15.14 | 0.46 | 0.994 |
| Argentina | 9.29 | 44.6 | 6.31 | -0.32 | 24.70 | -0.01 | -0.109 |
| Australia | 18.61 | 64.6 | 3.52 | -1.30 | 9.11 | -0.14 | -0.303 |
| Austria | 3.84 | 27.3 | 3.95 | 2.27 | 11.33 | 0.20 | -0.199 |
| Bahrain | 4.25 | 26.0 | 5.03 | 10.79 | 15.54 | 0.69 | -0.454 |
| Belgium | 6.59 | 42.6 | 4.51 | 0.73 | 10.16 | 0.07 | -0.107 |
| Brazil | 33.38 | 72.5 | 5.63 | 5.96 | 20.06 | 0.30 | -0.029 |
| Canada | 13.32 | 56.7 | 5.07 | 0.61 | 6.56 | 0.09 | -0.219 |
| Chile | 6.96 | 42.4 | 3.88 | 0.09 | 13.41 | 0.01 | -0.184 |
| China Domestic | 748.35 | 95.0 | 4.76 | -1.10 | 27.56 | -0.04 | -0.350 |
| China International | 42.91 | 66.3 | 8.18 | 3.27 | 28.35 | 0.12 | -0.076 |
| Colombia | 4.40 | 25.8 | 4.29 | 6.27 | 21.71 | 0.29 | -0.208 |
| Czech Republic | 4.36 | 30.0 | 3.77 | 3.86 | 19.81 | 0.19 | -0.127 |
| Denmark | 7.20 | 45.0 | 3.98 | 0.54 | 11.44 | 0.05 | -0.159 |
| Egypt | 11.98 | 40.3 | 5.46 | 1.33 | 25.60 | 0.05 | -0.333 |
| Finland | 6.05 | 40.8 | 4.08 | 5.58 | 10.70 | 0.52 | 0.028 |
| France | 20.25 | 60.9 | 9.16 | 2.16 | 8.63 | 0.25 | 0.065 |
| Germany | 16.99 | 62.4 | 5.03 | -1.59 | 8.96 | -0.18 | 0.080 |
| Greece | 31.65 | 60.2 | 9.13 | 2.50 | 26.15 | 0.10 | -0.092 |
| Hong Kong | 37.62 | 66.4 | 10.07 | -1.65 | 18.66 | -0.09 | -0.079 |
| Hungary | 5.12 | 29.7 | 7.06 | -1.26 | 22.84 | -0.06 | -0.008 |
| India | 59.09 | 72.5 | 3.94 | 5.51 | 23.33 | 0.24 | -0.195 |
| Indonesia | 25.87 | 62.6 | 5.44 | -9.42 | 28.97 | -0.33 | -0.185 |
| Ireland | 4.59 | 32.4 | 3.84 | 2.94 | 12.02 | 0.24 | -0.069 |
| Israel | 14.00 | 56.2 | 4.46 | 2.41 | 17.36 | 0.14 | -0.144 |
| Italy | 21.86 | 62.4 | 8.97 | 2.45 | 12.06 | 0.20 | 0.052 |
| Japan | 281.56 | 90.3 | 3.64 | -3.62 | 14.90 | -0.24 | -0.200 |
| Jordan | 4.06 | 19.0 | 4.48 | 2.01 | 18.64 | 0.11 | -0.310 |
| Korea | 130.86 | 84.4 | 6.35 | 1.91 | 31.74 | 0.06 | -0.099 |
| Kuwait | 36.24 | 70.1 | 3.58 | 7.00 | 17.71 | 0.40 | -0.441 |
| Malaysia | 63.37 | 72.1 | 13.95 | -8.13 | 25.35 | -0.32 | -0.162 |
| Mexico | 11.11 | 51.3 | 5.92 | 4.50 | 14.54 | 0.31 | 0.001 |
| Morocco | 3.43 | 21.5 | 4.68 | 2.38 | 18.55 | 0.13 | -0.335 |
| Netherland | 9.55 | 46.5 | 6.53 | -0.52 | 9.33 | -0.06 | 0.078 |
| New Zealand | 3.22 | 26.2 | 3.58 | -3.03 | 11.65 | -0.26 | -0.289 |
| Norway | 8.05 | 45.3 | 4.32 | 0.00 | 12.69 | 0.00 | -0.084 |
| Oman | 4.26 | 24.7 | 3.98 | 14.98 | 17.61 | 0.85 | -0.459 |
| Pakistan | 9.41 | 41.1 | 4.85 | 9.63 | 32.66 | 0.29 | -0.211 |
| Peru | 2.61 | 19.6 | 4.32 | -0.96 | 15.83 | -0.06 | -0.204 |
| Philippines | 10.61 | 44.6 | 5.51 | -14.66 | 23.19 | -0.63 | -0.163 |
| Poland | 9.21 | 47.0 | 4.32 | -3.23 | 21.37 | -0.15 | -0.074 |
| Portugal | 5.14 | 33.9 | 5.62 | 0.40 | 13.90 | 0.03 | -0.160 |
| Qatar | 37.90 | 61.4 | 6.72 | 19.20 | 31.54 | 0.61 | -0.267 |
| Russia | 39.61 | 70.1 | 26.08 | 25.41 | 45.53 | 0.56 | -0.019 |
| Saudi Arabia | 140.00 | 80.8 | 8.65 | 12.41 | 35.83 | 0.35 | -0.251 |
| Singapore | 23.58 | 62.8 | 10.41 | -0.08 | 18.83 | 0.00 | -0.072 |
| South Africa | 20.34 | 63.1 | 3.97 | -1.06 | 12.80 | -0.08 | -0.139 |
| Spain | 9.87 | 48.5 | 4.80 | 3.67 | 9.31 | 0.39 | 0.007 |
| Sweden | 12.08 | 54.7 | 8.63 | 2.96 | 10.26 | 0.29 | 0.069 |
| Switzerland | 10.71 | 52.5 | 5.79 | 1.79 | 8.75 | 0.20 | 0.145 |
| Taiwan | 133.70 | 83.2 | 4.78 | -7.52 | 25.23 | -0.30 | -0.248 |
| Thailand | 35.40 | 69.8 | 4.44 | -5.65 | 27.05 | -0.21 | -0.163 |
| Turkey | 48.15 | 76.2 | 6.01 | 1.43 | 40.36 | 0.04 | -0.125 |
| UK | 35.56 | 74.0 | 3.62 | -3.23 | 7.18 | -0.45 | -0.052 |
| UAE | 27.04 | 54.2 | 4.44 | 24.07 | 26.82 | 0.90 | -0.347 |
| US | 195.52 | 88.4 | 5.30 | 2.33 | 5.98 | 0.39 | 0.200 |
| Average | 31.72 | 53.14 | 6.08 | | 18.07 | | |

## Table 4.4

Industry factor summary statistics based on weekly factor returns. Sample period is from January 1997 to June 2008.

| Factor Name | Average Squared t-statistic | Percent Observations with \|t\|>2 | (weekly) Factor Kurtosis | Annualized Factor Return | Annualized Factor Volatility | Factor Sharpe Ratio | Correlation with ESTU |
|---|---|---|---|---|---|---|---|
| World Factor | 753.02 | 94.0 | 4.91 | 6.96 | 15.14 | 0.46 | 0.994 |
| Energy Equipment & Services | 56.73 | 79.2 | 4.22 | 6.48 | 25.86 | 0.25 | -0.075 |
| Oil, Gas & Consumable Fuels | 22.45 | 64.1 | 4.38 | 2.34 | 10.56 | 0.22 | -0.125 |
| Oil & Gas Exploration & Production | 43.11 | 72.3 | 3.85 | 4.18 | 17.50 | 0.24 | -0.134 |
| Chemicals | 8.24 | 47.3 | 7.58 | 0.32 | 6.11 | 0.05 | 0.042 |
| Construction, Containers, Paper | 7.24 | 41.9 | 9.08 | -4.24 | 6.22 | -0.68 | -0.003 |
| Aluminum, Diversified Metals | 21.90 | 61.2 | 5.21 | 3.97 | 12.50 | 0.32 | 0.023 |
| Gold, Precious Metals | 37.46 | 72.5 | 5.66 | 6.13 | 24.29 | 0.25 | -0.213 |
| Steel | 13.17 | 51.3 | 4.46 | 0.95 | 10.17 | 0.09 | 0.080 |
| Capital Goods | 12.08 | 54.0 | 6.08 | -0.15 | 4.03 | -0.04 | 0.113 |
| Commercial & Professional Services | 4.83 | 34.9 | 5.58 | -0.98 | 4.86 | -0.20 | -0.110 |
| Transportation Non-Airline | 5.72 | 40.3 | 3.99 | 1.66 | 5.01 | 0.33 | -0.081 |
| Airlines | 9.44 | 51.0 | 6.05 | -8.54 | 13.68 | -0.62 | 0.148 |
| Automobiles & Components | 9.32 | 43.6 | 10.39 | -3.99 | 7.24 | -0.55 | 0.145 |
| Consumer Durables & Apparel | 7.09 | 44.5 | 5.76 | -4.11 | 4.88 | -0.84 | 0.041 |
| Consumer Services | 7.32 | 42.3 | 7.01 | -0.57 | 6.29 | -0.09 | -0.090 |
| Media | 9.60 | 48.8 | 8.32 | 0.57 | 5.96 | 0.10 | -0.045 |
| Retailing | 16.58 | 60.4 | 5.21 | -1.77 | 7.16 | -0.25 | -0.053 |
| Food & Staples Retailing | 4.61 | 33.6 | 4.98 | -1.10 | 6.06 | -0.18 | -0.327 |
| Food, Beverage & Tobacco | 7.23 | 48.3 | 5.62 | -0.64 | 4.93 | -0.13 | -0.483 |
| Household & Personal Products | 3.88 | 27.5 | 10.15 | 1.04 | 8.12 | 0.13 | -0.327 |
| Health Care Equipment & Services | 15.15 | 57.6 | 5.19 | 2.19 | 8.10 | 0.27 | -0.378 |
| Biotechnology | 25.67 | 61.4 | 9.95 | 12.50 | 19.04 | 0.66 | -0.065 |
| Pharmaceuticals, Life Sciences | 13.10 | 52.0 | 6.24 | 3.36 | 8.04 | 0.42 | -0.308 |
| Banks | 21.72 | 62.4 | 5.06 | -1.24 | 5.81 | -0.21 | 0.080 |
| Diversified Financials | 12.73 | 57.9 | 4.44 | -0.62 | 5.47 | -0.11 | 0.274 |
| Insurance | 11.97 | 52.7 | 8.20 | -3.09 | 6.86 | -0.45 | 0.036 |
| Real Estate | 13.31 | 47.0 | 4.76 | -2.10 | 5.53 | -0.38 | -0.087 |
| Internet Software & Services | 17.22 | 56.2 | 7.39 | 8.01 | 18.89 | 0.42 | 0.165 |
| IT Services, Software | 16.17 | 59.4 | 5.41 | 2.31 | 8.55 | 0.27 | 0.106 |
| Communications Equipment | 20.57 | 63.1 | 5.70 | 1.66 | 13.43 | 0.12 | 0.169 |
| Computers, Electronics | 20.99 | 63.6 | 3.54 | 0.92 | 7.93 | 0.12 | 0.166 |
| Semiconductors | 58.68 | 77.9 | 3.77 | 9.27 | 18.46 | 0.50 | 0.204 |
| Telecommunication Services | 15.42 | 52.5 | 13.53 | 0.68 | 8.12 | 0.08 | -0.055 |
| Utilities | 14.03 | 58.6 | 6.76 | -2.29 | 6.76 | -0.34 | -0.353 |
| Average (ex World Factor) | 17.20 | 54.16 | 6.28 | | 9.78 | | |

### Table 4.5

Style factor summary statistics based on weekly factor returns. The sample period, January 1997 to June 2008, is divided into two sub-periods, A and B.

A. January 1997 to June 2002 (66 months)

| Factor Name | Average Squared t-statistic | Percent Observations with \|t\|>2 | (weekly) Factor Kurtosis | Annualized Factor Return | Annualized Factor Volatility | Factor Sharpe Ratio | Correlation with ESTU |
|---|---|---|---|---|---|---|---|
| Volatility | 119.46 | 82.6 | 4.92 | -5.98 | 8.21 | -0.73 | 0.785 |
| Momentum | 48.12 | 78.5 | 6.34 | 7.62 | 4.46 | 1.71 | 0.109 |
| Size | 43.65 | 79.9 | 4.73 | 2.37 | 2.97 | 0.80 | 0.291 |
| Value | 11.76 | 57.6 | 4.55 | 7.65 | 1.96 | 3.89 | -0.156 |
| Growth | 9.75 | 50.7 | 4.49 | -0.09 | 1.61 | -0.05 | 0.407 |
| Non-linear Size | 8.86 | 49.7 | 4.31 | 1.61 | 1.89 | 0.85 | 0.151 |
| Liquidity | 10.18 | 53.8 | 4.00 | 4.99 | 1.63 | 3.06 | 0.506 |
| Leverage | 4.73 | 33.0 | 5.55 | -1.18 | 1.13 | -1.04 | -0.029 |
| Average | 32.06 | 60.72 | 4.86 | 2.12 | 2.98 | 1.06 | 0.26 |

B. July 2002 to June 2008 (72 months)

| Factor Name | Average Squared t-statistic | Percent Observations with \|t\|>2 | (weekly) Factor Kurtosis | Annualized Factor Return | Annualized Factor Volatility | Factor Sharpe Ratio | Correlation with ESTU |
|---|---|---|---|---|---|---|---|
| Volatility | 89.50 | 84.8 | 4.07 | 0.36 | 4.98 | 0.07 | 0.836 |
| Momentum | 35.37 | 76.4 | 5.71 | 4.62 | 2.65 | 1.75 | -0.113 |
| Size | 26.45 | 65.0 | 5.33 | 0.71 | 1.67 | 0.42 | 0.411 |
| Value | 11.89 | 53.1 | 10.74 | 3.86 | 1.37 | 2.83 | 0.082 |
| Growth | 4.45 | 30.4 | 4.86 | 1.00 | 0.80 | 1.24 | 0.039 |
| Non-linear Size | 6.85 | 44.0 | 5.34 | 1.25 | 1.11 | 1.13 | 0.171 |
| Liquidity | 8.46 | 46.3 | 4.34 | 0.35 | 1.25 | 0.28 | 0.462 |
| Leverage | 4.30 | 33.7 | 8.59 | -0.09 | 0.84 | -0.11 | 0.016 |
| Average | 23.41 | 54.21 | 6.12 | 1.51 | 1.83 | 0.95 | 0.24 |

**Table 4.6**

Explanatory power of GEM2 factors versus GEM. Results are averages of monthly $R^2$ values over the sample period January 1997 to June 2008. Notation: G2 indicates GEM2 factors, G indicates GEM factors with constraints (Equation 4.8 of text), and GU indicates GEM factors unconstrained. Furthermore, W, C, I, and S denote World, country, industry, and style factors respectively; Vol denotes the GEM2 Volatility factor, and VIM denotes the GEM Variability in Markets factor.

Case 1 compares the explanatory power of GEM2 country factors versus GEM with constraints. Case 2 shows the increase in $R^2$ explained by countries when the GEM constraints are removed. Case 3 compares the explanatory power of GEM2 and GEM country factors in conjunction with the Volatility and VIM factors. Cases 4 and 5 demonstrate the superiority of the GEM2 style factors and industry factors, respectively. Case 6 demonstrates the increased explanatory power of GEM2 versus an unconstrained GEM model. The final case shows the added explanatory power of the actual GEM2 model versus the actual GEM model.

| Case | Regression A | Regression B | $R^2$(A) | $R^2$(B) | $\Delta R^2$ |
|---|---|---|---|---|---|
| 1 | G2(W,C) | G(C ) | 0.2082 | 0.1988 | 0.0094 |
| 2 | G2(W,C) | GU(C ) | 0.2082 | 0.2229 | -0.0147 |
| 3 | G2(W,C,Vol) | GU(C, VIM) | 0.2443 | 0.2306 | 0.0137 |
| 4 | G2(W,C,I,S) | G2(W,C,I);G(S) | 0.3094 | 0.3023 | 0.0071 |
| 5 | G2(W,C,I,S) | G2(W,C,S);G(I) | 0.3094 | 0.2985 | 0.0109 |
| 6 | G2(W,C,I,S) | GU(C,I,S) | 0.3094 | 0.2952 | 0.0142 |
| Actual | G2(W,C,I,S) | G(C,I,S) | 0.3094 | 0.2739 | 0.0355 |

**Figure 4.1**

Trailing 12-month $R^2$ for GEM2 and GEM models.

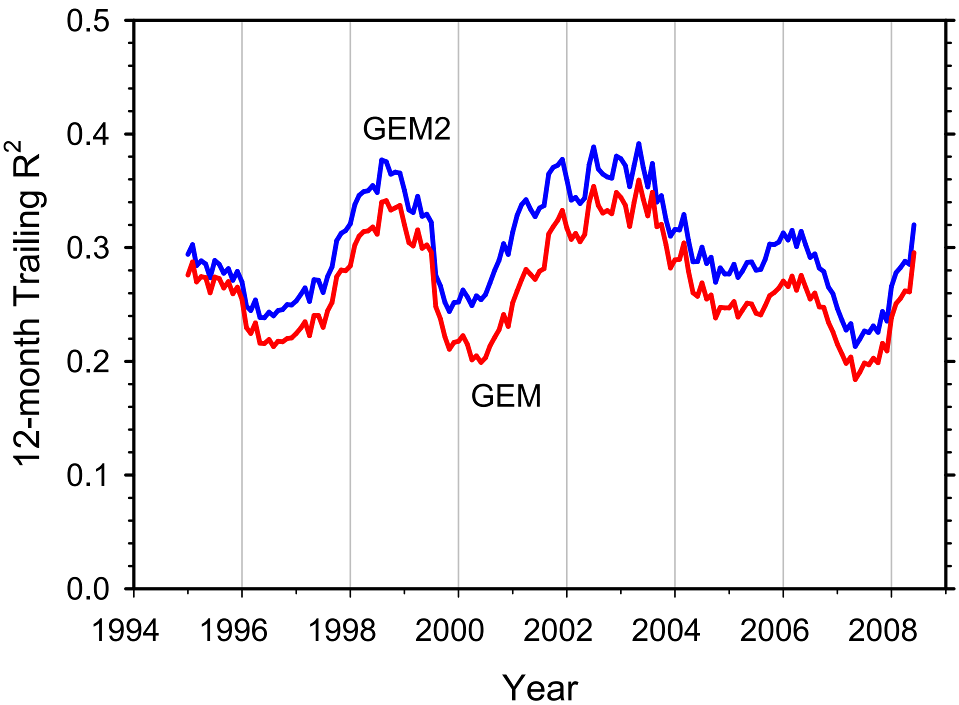


**Figure 4.2**

Explanatory power of GEM2 factor groups, expressed as trailing 12-month $R^2$. Here, W indicates the World factor in isolation, W/C denotes World factor plus countries, and W/C/I represents World, country, and industry factors. The black dots indicate the actual monthly $R^2$ values.

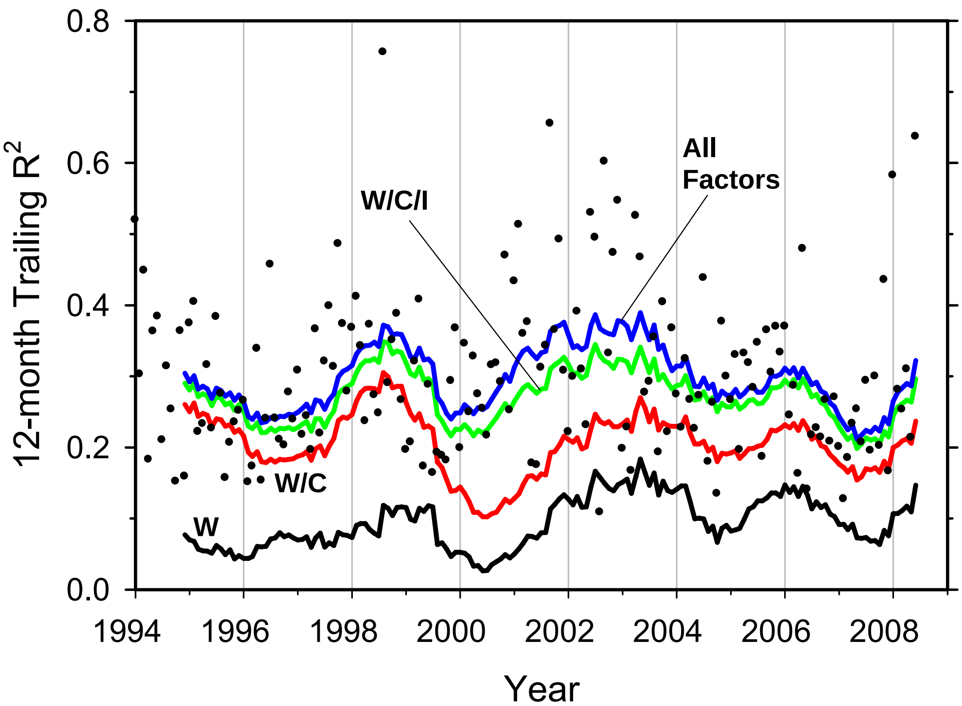


**Figure 5.1**

Mean 12-month rolling absolute deviation (RAD) and mean Variability $(\bar{v})$ of GEM2 equity factors, plotted as a function of volatility half-life. The sample period was 138 months from January 1997 to June 2008. The solid lines indicate results with the two-lag autocorrelation that is used GEM2, and the dashed lines are for no autocorrelation adjustment. The vertical lines indicate the actual half-life parameters for GEM2S and GEM2L.

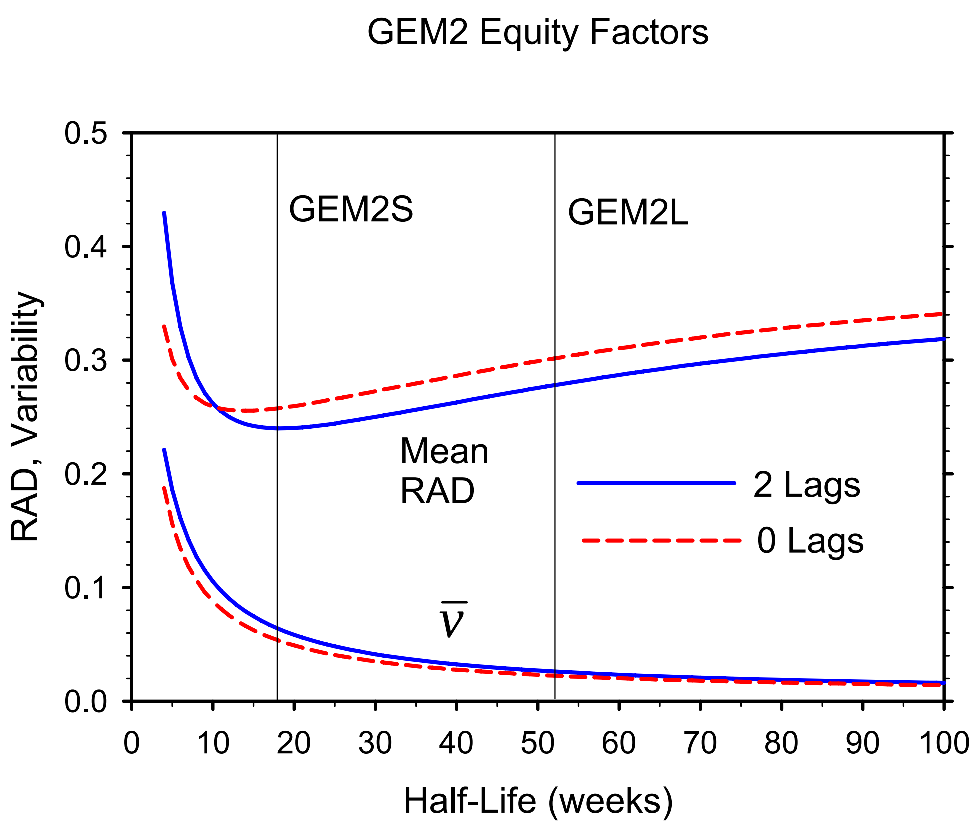
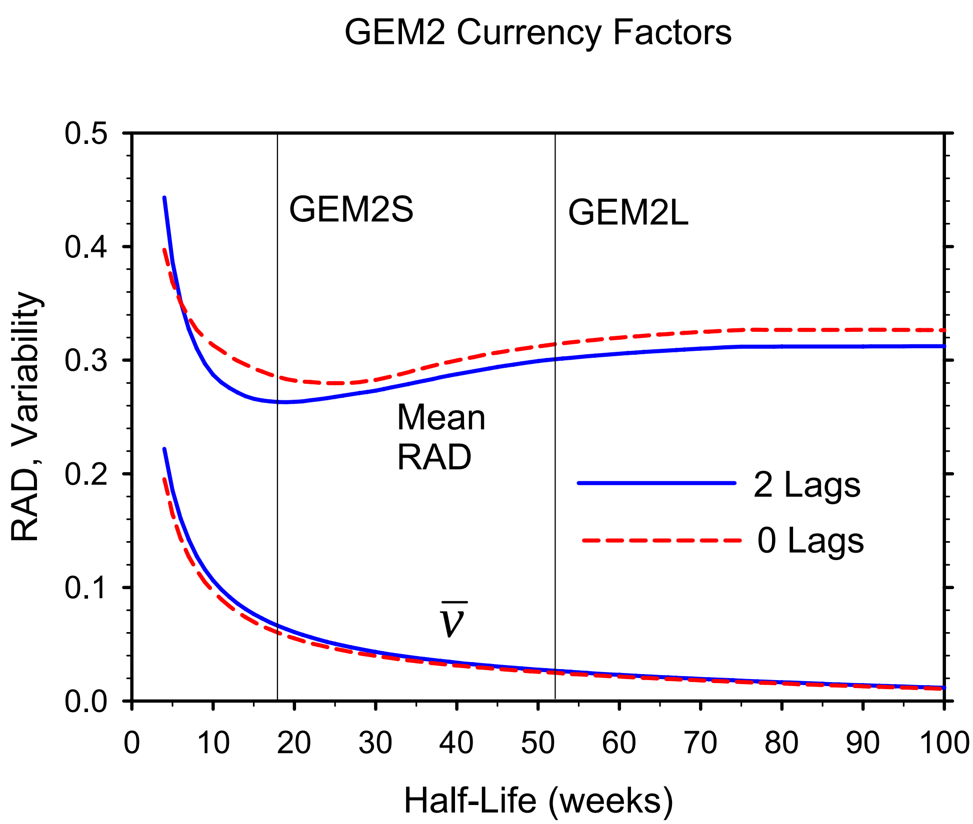


**Figure 5.2**

Mean 12-month rolling absolute deviation (RAD) and mean Variability $(\bar{v})$ of GEM2 currency factors, plotted as a function of volatility half-life. Results are cap-weighted averages as of June 2008. Sample period was 138 months from January 1997 to June 2008. The solid lines indicate results with the two-lag autocorrelation used in GEM2, and the dashed lines are for no autocorrelation adjustment. The vertical lines indicate the actual half-life parameters for GEM2S and GEM2L.



## Table 5.1

Currency 12-month RAD and forecast variability for GEM2S, GEM2L, and GEM models. The sample period contains 138 months from January 1997 to June 2008. Also reported is the sample monthly kurtosis. Monthly standardized returns were trimmed at $10\sigma$, and monthly variability was trimmed at 2. Trimming primarily affected only the GEM reported values. Results are quoted from a US Dollar numeraire perspective.

| Currency | (Monthly) | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|---|
| Name | Kurtosis | 12m RAD | Variability | 12m RAD | Variability | 12m RAD | Variability |
| Argentine Peso | 35.4 | 0.75 | 0.106 | 0.93 | 0.051 | | |
| Australian Dollar | 2.5 | 0.17 | 0.062 | 0.20 | 0.022 | 0.16 | 0.104 |
| Bahraini Dinar | 8.6 | 0.33 | 0.078 | 0.36 | 0.036 | 0.32 | 0.124 |
| Brazilian Real | 9.0 | 0.58 | 0.087 | 0.62 | 0.035 | | |
| Canadian Dollar | 3.0 | 0.21 | 0.059 | 0.20 | 0.023 | 0.46 | 0.069 |
| Chilean Peso | 3.2 | 0.21 | 0.071 | 0.21 | 0.026 | 0.56 | 0.253 |
| Chinese Yuan | 11.4 | 0.57 | 0.080 | 0.67 | 0.042 | 1.23 | 0.199 |
| Colombian Peso | 4.4 | 0.35 | 0.081 | 0.35 | 0.030 | 0.34 | 0.208 |
| Czech Koruna | 2.5 | 0.16 | 0.057 | 0.17 | 0.023 | 0.22 | 0.136 |
| Danish Krone | 2.9 | 0.13 | 0.054 | 0.16 | 0.019 | 0.15 | 0.061 |
| Egyptian Pound | 29.6 | 0.74 | 0.123 | 0.81 | 0.058 | 0.54 | 0.285 |
| Euro | 2.9 | 0.13 | 0.052 | 0.16 | 0.018 | 0.14 | 0.066 |
| Hong Kong Dollar | 7.2 | 0.55 | 0.091 | 0.68 | 0.050 | 0.64 | 0.058 |
| Hungarian Forint | 3.2 | 0.17 | 0.062 | 0.19 | 0.024 | 0.26 | 0.100 |
| Indian Rupee | 8.3 | 0.48 | 0.097 | 0.47 | 0.033 | 0.32 | 0.182 |
| Indonesian Rupiah | 14.0 | 0.39 | 0.107 | 0.59 | 0.058 | 0.79 | 0.292 |
| Israeli Shekel | 5.7 | 0.29 | 0.077 | 0.31 | 0.030 | 0.24 | 0.135 |
| Japanese Yen | 7.1 | 0.20 | 0.061 | 0.24 | 0.022 | 0.25 | 0.072 |
| Jordanian Dinar | 3.8 | 0.36 | 0.100 | 0.36 | 0.032 | 0.78 | 0.355 |
| Korean Won | 23.0 | 0.43 | 0.092 | 0.52 | 0.047 | 0.63 | 0.251 |
| Kuwaiti Dinar | 11.4 | 0.56 | 0.124 | 0.48 | 0.055 | | |
| Malaysian Ringgit | 22.8 | 0.83 | 0.148 | 0.97 | 0.098 | 1.83 | 0.485 |
| Mexican Peso | 6.9 | 0.24 | 0.063 | 0.24 | 0.019 | 0.33 | 0.153 |
| Moroccan Dirham | 2.9 | 0.14 | 0.060 | 0.16 | 0.022 | | |
| New Zealand Dollar | 3.0 | 0.17 | 0.068 | 0.20 | 0.027 | 0.23 | 0.100 |
| Norwegian Krone | 3.1 | 0.16 | 0.056 | 0.16 | 0.020 | 0.15 | 0.105 |
| Omani Rial | 1.9 | 0.40 | 0.072 | 0.39 | 0.024 | 0.19 | 0.153 |
| Pakistan Rupee | 11.3 | 0.48 | 0.117 | 0.43 | 0.042 | 0.43 | 0.118 |
| Peruvian Sol | 6.3 | 0.39 | 0.115 | 0.33 | 0.045 | 0.27 | 0.194 |
| Philippine Peso | 8.0 | 0.42 | 0.096 | 0.42 | 0.039 | 0.40 | 0.213 |
| Polish Zloty | 3.0 | 0.21 | 0.068 | 0.21 | 0.025 | 0.29 | 0.111 |
| Qatari Rial | 11.4 | 0.31 | 0.085 | 0.41 | 0.029 | | |
| Russian Ruble | 59.5 | 0.68 | 0.088 | 0.79 | 0.041 | 0.47 | 0.245 |
| Saudi Rial | 21.2 | 0.46 | 0.108 | 0.58 | 0.046 | 0.44 | 0.267 |
| Singapore Dollar | 5.3 | 0.26 | 0.074 | 0.30 | 0.030 | 0.24 | 0.216 |
| South African Rand | 3.4 | 0.38 | 0.082 | 0.30 | 0.030 | 0.37 | 0.298 |
| Swedish Krone | 2.7 | 0.11 | 0.052 | 0.13 | 0.018 | 0.14 | 0.064 |
| Swiss Franc | 2.4 | 0.13 | 0.055 | 0.15 | 0.019 | 0.16 | 0.053 |
| Taiwan Dollar | 7.8 | 0.36 | 0.092 | 0.36 | 0.033 | 0.31 | 0.233 |
| Thailand Bhat | 16.0 | 0.33 | 0.108 | 0.52 | 0.056 | 0.91 | 0.339 |
| New Turkish Lira | 6.7 | 0.46 | 0.088 | 0.48 | 0.041 | 0.60 | 0.139 |
| U.K. Pound | 3.1 | 0.15 | 0.056 | 0.16 | 0.019 | 0.14 | 0.071 |
| UAE Dirham | 29.3 | 0.61 | 0.119 | 0.67 | 0.060 | | |
| Average (Equal weight) | 10.2 | 0.36 | 0.084 | 0.40 | 0.035 | 0.43 | 0.176 |
| Average (Cap weight) | 8.3 | 0.26 | 0.066 | 0.30 | 0.026 | 0.32 | 0.112 |

**Figure 6.1**

Plot of weekly realized and forecast mean absolute specific returns for the GEM2 estimation universe.

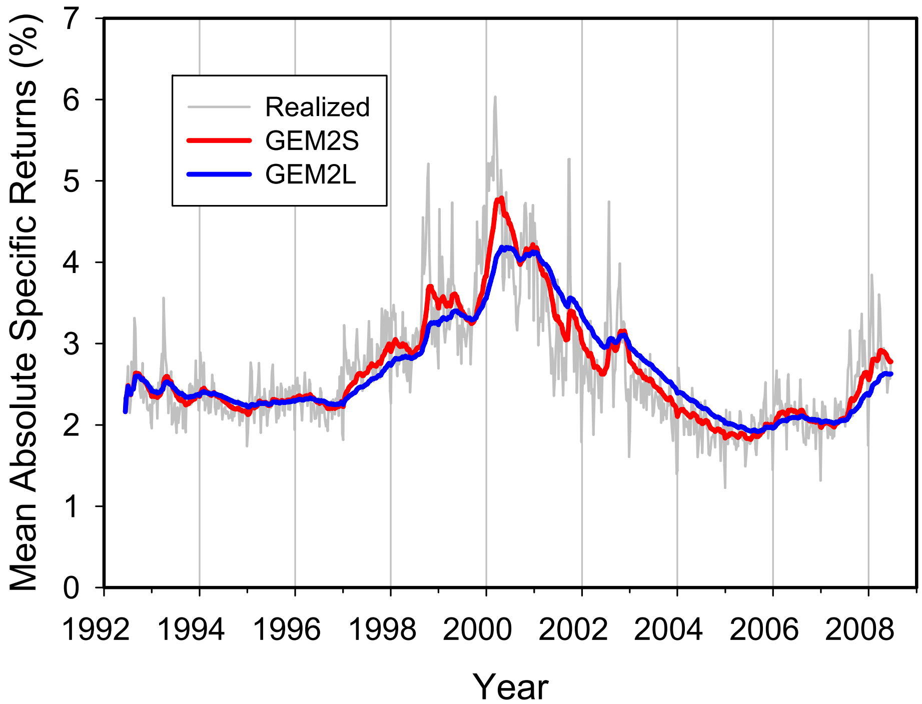


**Table 6.1**

Specific risk bias statistic results, based on 12-month rolling windows. Results were computed on the GEM2 estimation universe, and aggregated on cap-weighted and equal-weighted bases. The full sample period, January 1997 to June 2008, is divided into two sub-periods, A and B. The 12-month confidence interval is taken as [0.59, 1.41].

A. January 1997 to June 2002 (66 months)

| | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|
| Item | Cap Wt | Equal Wt | Cap Wt | Equal Wt | Cap Wt | Equal Wt |
| Mean Rolling Bias Statistic | 1.02 | 1.00 | 1.04 | 1.03 | 1.21 | 1.32 |
| Mean Rolling Absolute Deviation | 0.24 | 0.28 | 0.26 | 0.29 | 0.37 | 0.45 |
| Over-forecast | 4.8% | 8.6% | 5.2% | 8.6% | 4.8% | 3.5% |
| Within Confidence Interval | 85.4% | 80.2% | 82.1% | 78.0% | 67.8% | 63.1% |
| Under-forecast | 9.8% | 11.1% | 12.8% | 13.4% | 27.4% | 33.4% |

B. July 2002 to June 2008 (72 months)

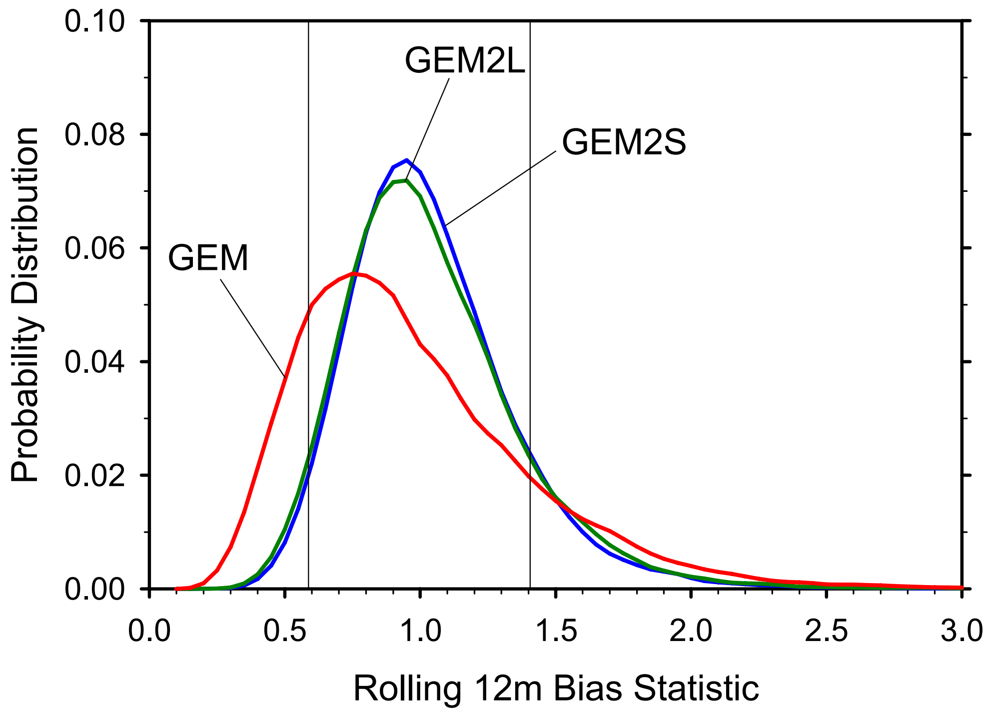
| | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|
| Item | Cap Wt | Equal Wt | Cap Wt | Equal Wt | Cap Wt | Equal Wt |
| Mean Rolling Bias Statistic | 1.02 | 1.02 | 1.01 | 1.00 | 0.80 | 0.88 |
| Mean Rolling Absolute Deviation | 0.23 | 0.25 | 0.24 | 0.26 | 0.34 | 0.32 |
| Over-forecast | 4.0% | 5.5% | 5.3% | 7.3% | 29.6% | 20.9% |
| Within Confidence Interval | 86.7% | 84.0% | 85.1% | 82.8% | 65.0% | 70.9% |
| Under-forecast | 9.3% | 10.5% | 9.6% | 9.9% | 5.4% | 8.2% |

C. January 1997 to June 2008 (138 months)

| | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|
| Item | Cap Wt | Equal Wt | Cap Wt | Equal Wt | Cap Wt | Equal Wt |
| Mean Rolling Bias Statistic | 1.02 | 1.01 | 1.02 | 1.01 | 0.96 | 1.12 |
| Mean Rolling Absolute Deviation | 0.23 | 0.26 | 0.24 | 0.28 | 0.35 | 0.39 |
| Over-forecast | 4.3% | 7.4% | 5.4% | 8.3% | 19.5% | 11.1% |
| Within Confidence Interval | 86.3% | 82.1% | 84.1% | 80.3% | 66.6% | 67.1% |
| Under-forecast | 9.4% | 10.6% | 10.5% | 11.5% | 13.9% | 21.8% |

**Figure 6.2**

Frequency distribution of 12-month rolling bias statistics for specific risk forecasts. Sample period is January 1997 to June 2008. Results are cap-weighted within the GEM2 estimation universe. Vertical lines indicate confidence interval [0.59, 1.41].



## Table 7.1

Country 12-month RAD and monthly kurtosis values for sample period January 1997 to June 2008. Portfolios are cap-weighted carve-outs of GEM2 estimation universe. Active portfolios are obtained by running portfolios against MSCI ACWI IMI as the benchmark.

| | Kurtosis | | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|---|---|
| Portfolio Name | Portfolio | Active | Portfolio | Active | Portfolio | Active | Portfolio | Active |
| Argentina | 4.55 | 4.77 | 0.32 | 0.31 | 0.30 | 0.30 | 0.38 | 0.38 |
| Australia | 3.04 | 3.48 | 0.16 | 0.19 | 0.20 | 0.22 | 0.20 | 0.31 |
| Austria | 2.96 | 3.39 | 0.20 | 0.23 | 0.22 | 0.26 | 0.24 | 0.31 |
| Bahrain | 3.78 | 2.58 | 0.16 | 0.13 | 0.19 | 0.14 | 0.12 | 0.16 |
| Belgium | 4.42 | 3.56 | 0.18 | 0.24 | 0.23 | 0.28 | 0.29 | 0.34 |
| Brazil | 4.26 | 4.60 | 0.38 | 0.33 | 0.35 | 0.32 | 0.39 | 0.48 |
| Canada | 4.61 | 2.85 | 0.24 | 0.22 | 0.26 | 0.21 | 0.29 | 0.24 |
| Chile | 4.73 | 3.71 | 0.22 | 0.21 | 0.25 | 0.25 | 0.25 | 0.30 |
| China International | 5.61 | 6.71 | 0.24 | 0.27 | 0.32 | 0.35 | 0.51 | 0.59 |
| Colombia | 3.18 | 3.04 | 0.23 | 0.25 | 0.23 | 0.25 | 0.27 | 0.27 |
| Czech Republic | 3.77 | 3.84 | 0.19 | 0.19 | 0.20 | 0.24 | 0.29 | 0.32 |
| Denmark | 3.07 | 2.59 | 0.16 | 0.19 | 0.21 | 0.23 | 0.20 | 0.25 |
| Egypt | 4.66 | 5.14 | 0.30 | 0.27 | 0.27 | 0.24 | 0.29 | 0.26 |
| Finland | 4.22 | 4.60 | 0.19 | 0.20 | 0.22 | 0.21 | 0.28 | 0.37 |
| France | 3.72 | 3.86 | 0.21 | 0.21 | 0.27 | 0.24 | 0.26 | 0.32 |
| Germany | 5.49 | 6.54 | 0.23 | 0.24 | 0.29 | 0.26 | 0.28 | 0.38 |
| Greece | 3.95 | 4.64 | 0.22 | 0.23 | 0.29 | 0.34 | 0.29 | 0.41 |
| Hong Kong | 5.67 | 6.94 | 0.18 | 0.26 | 0.27 | 0.31 | 0.37 | 0.45 |
| Hungary | 5.40 | 4.82 | 0.21 | 0.21 | 0.21 | 0.21 | 0.27 | 0.33 |
| India | 2.62 | 2.94 | 0.20 | 0.17 | 0.17 | 0.17 | 0.19 | 0.26 |
| Indonesia | 4.58 | 4.36 | 0.20 | 0.22 | 0.28 | 0.29 | 0.33 | 0.33 |
| Ireland | 3.80 | 3.03 | 0.18 | 0.19 | 0.21 | 0.21 | 0.23 | 0.30 |
| Israel | 3.78 | 4.98 | 0.15 | 0.23 | 0.19 | 0.27 | 0.30 | 0.32 |
| Italy | 4.14 | 5.34 | 0.18 | 0.25 | 0.22 | 0.29 | 0.25 | 0.42 |
| Japan | 2.81 | 3.72 | 0.22 | 0.23 | 0.23 | 0.25 | 0.23 | 0.23 |
| Jordan | 5.99 | 4.53 | 0.26 | 0.24 | 0.25 | 0.24 | 0.29 | 0.22 |
| Korea | 6.62 | 10.28 | 0.21 | 0.27 | 0.31 | 0.39 | 0.41 | 0.49 |
| Kuwait | 3.34 | 3.46 | 0.13 | 0.23 | 0.15 | 0.23 | | |
| Malaysia | 8.15 | 6.98 | 0.21 | 0.18 | 0.29 | 0.30 | 0.47 | 0.46 |
| Mexico | 6.16 | 4.53 | 0.24 | 0.26 | 0.23 | 0.27 | 0.30 | 0.34 |
| Morocco | 4.23 | 4.62 | 0.27 | 0.27 | 0.26 | 0.26 | 0.27 | 0.26 |
| Netherland | 3.91 | 4.47 | 0.22 | 0.30 | 0.30 | 0.32 | 0.36 | 0.36 |
| New Zealand | 3.64 | 3.25 | 0.19 | 0.21 | 0.22 | 0.23 | 0.24 | 0.23 |
| Norway | 4.42 | 2.93 | 0.26 | 0.29 | 0.26 | 0.30 | 0.24 | 0.28 |
| Oman | 3.62 | 3.41 | 0.28 | 0.34 | 0.31 | 0.35 | 0.24 | 0.36 |
| Pakistan | 4.54 | 4.96 | 0.23 | 0.25 | 0.23 | 0.27 | 0.27 | 0.30 |
| Peru | 4.77 | 3.44 | 0.33 | 0.23 | 0.30 | 0.21 | 0.38 | 0.40 |
| Philippines | 7.25 | 6.11 | 0.19 | 0.17 | 0.26 | 0.20 | 0.32 | 0.26 |
| Poland | 4.36 | 3.47 | 0.20 | 0.21 | 0.14 | 0.15 | 0.16 | 0.21 |
| Portugal | 3.25 | 2.98 | 0.21 | 0.18 | 0.27 | 0.24 | 0.31 | 0.35 |
| Qatar | 7.09 | 6.22 | 0.32 | 0.34 | 0.34 | 0.40 | | |
| Russia | 7.95 | 7.08 | 0.22 | 0.21 | 0.22 | 0.20 | 0.42 | 0.45 |
| Saudi Arabia | 3.02 | 3.04 | 0.25 | 0.19 | 0.27 | 0.21 | 0.31 | 0.27 |
| Singapore | 5.31 | 6.07 | 0.26 | 0.21 | 0.32 | 0.31 | 0.36 | 0.41 |
| South Africa | 4.80 | 4.55 | 0.26 | 0.24 | 0.25 | 0.26 | 0.22 | 0.30 |
| Spain | 3.48 | 2.81 | 0.17 | 0.16 | 0.23 | 0.20 | 0.25 | 0.27 |
| Sweden | 3.61 | 4.21 | 0.20 | 0.21 | 0.24 | 0.21 | 0.23 | 0.26 |
| Switzerland | 4.28 | 3.41 | 0.23 | 0.21 | 0.27 | 0.26 | 0.29 | 0.34 |
| Taiwan | 3.80 | 4.59 | 0.20 | 0.22 | 0.26 | 0.26 | 0.24 | 0.29 |
| Thailand | 4.26 | 4.51 | 0.24 | 0.26 | 0.24 | 0.25 | 0.31 | 0.34 |
| Turkey | 5.09 | 5.29 | 0.28 | 0.23 | 0.25 | 0.22 | 0.24 | 0.27 |
| UAE | 5.88 | 6.50 | 0.38 | 0.31 | 0.39 | 0.37 | | |
| UK | 2.96 | 3.10 | 0.17 | 0.26 | 0.20 | 0.28 | 0.23 | 0.35 |
| US | 3.48 | 3.43 | 0.21 | 0.25 | 0.28 | 0.27 | 0.28 | 0.26 |
| Average (Equal Wt) | 4.48 | 4.45 | 0.23 | 0.23 | 0.25 | 0.26 | 0.29 | 0.33 |

### Table 7.2

Industry 12-month RAD and monthly kurtosis values for sample period January 1997 to June 2008. Portfolios are cap-weighted carve-outs of GEM2 estimation universe. Active portfolios are obtained by running portfolios against MSCI ACWI IMI as the benchmark.

| | Kurtosis | | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|---|---|
| Portfolio Name | Portfolio | Active | Portfolio | Active | Portfolio | Active | Portfolio | Active |
| Energy Equipment & Services | 4.10 | 4.55 | 0.25 | 0.24 | 0.27 | 0.26 | 0.35 | 0.38 |
| Oil, Gas & Consumable Fuels | 3.22 | 2.86 | 0.25 | 0.32 | 0.26 | 0.31 | 0.23 | 0.23 |
| Oil & Gas Exploration & Production | 3.23 | 2.90 | 0.23 | 0.23 | 0.27 | 0.26 | 0.38 | 0.48 |
| Chemicals | 3.55 | 3.85 | 0.19 | 0.26 | 0.23 | 0.25 | 0.24 | 0.30 |
| Construction, Containers, Paper | 4.01 | 5.68 | 0.26 | 0.23 | 0.29 | 0.26 | 0.30 | 0.34 |
| Aluminum, Diversified Metals | 3.12 | 3.56 | 0.26 | 0.28 | 0.24 | 0.29 | 0.21 | 0.26 |
| Gold, Precious Metals | 4.63 | 4.52 | 0.24 | 0.20 | 0.22 | 0.21 | 0.22 | 0.24 |
| Steel | 2.89 | 2.70 | 0.25 | 0.24 | 0.23 | 0.22 | 0.17 | 0.16 |
| Capital Goods | 3.70 | 2.90 | 0.22 | 0.19 | 0.28 | 0.19 | 0.28 | 0.34 |
| Commercial & Professional Services | 4.03 | 3.00 | 0.23 | 0.27 | 0.27 | 0.28 | 0.28 | 0.33 |
| Transportation Non-Airline | 3.61 | 4.00 | 0.15 | 0.25 | 0.20 | 0.28 | 0.21 | 0.24 |
| Airlines | 6.64 | 5.95 | 0.26 | 0.26 | 0.30 | 0.32 | 0.33 | 0.40 |
| Automobiles & Components | 4.05 | 4.19 | 0.24 | 0.16 | 0.28 | 0.19 | 0.24 | 0.30 |
| Consumer Durables & Apparel | 3.63 | 3.38 | 0.21 | 0.17 | 0.26 | 0.16 | 0.24 | 0.22 |
| Consumer Services | 5.08 | 4.68 | 0.28 | 0.25 | 0.32 | 0.23 | 0.31 | 0.31 |
| Media | 3.98 | 3.88 | 0.21 | 0.20 | 0.30 | 0.21 | 0.40 | 0.34 |
| Retailing | 3.03 | 3.40 | 0.17 | 0.26 | 0.22 | 0.23 | 0.28 | 0.23 |
| Food & Staples Retailing | 4.26 | 5.74 | 0.22 | 0.23 | 0.29 | 0.26 | 0.33 | 0.40 |
| Food, Beverage & Tobacco | 3.98 | 4.44 | 0.23 | 0.24 | 0.28 | 0.28 | 0.25 | 0.41 |
| Household & Personal Products | 6.63 | 3.97 | 0.25 | 0.28 | 0.32 | 0.35 | 0.29 | 0.49 |
| Health Care Equipment & Services | 5.79 | 3.43 | 0.26 | 0.21 | 0.31 | 0.24 | 0.23 | 0.44 |
| Biotechnology | 7.20 | 9.17 | 0.32 | 0.35 | 0.42 | 0.44 | 0.60 | 0.69 |
| Pharmaceuticals, Life Sciences | 2.85 | 4.56 | 0.20 | 0.40 | 0.25 | 0.42 | 0.20 | 0.44 |
| Banks | 5.16 | 5.00 | 0.21 | 0.26 | 0.31 | 0.36 | 0.31 | 0.40 |
| Diversified Financials | 4.52 | 4.87 | 0.23 | 0.30 | 0.32 | 0.35 | 0.32 | 0.39 |
| Insurance | 4.29 | 4.96 | 0.24 | 0.26 | 0.33 | 0.29 | 0.33 | 0.42 |
| Real Estate | 3.58 | 2.78 | 0.25 | 0.22 | 0.25 | 0.21 | 0.22 | 0.23 |
| Internet Software & Services | 3.57 | 4.27 | 0.23 | 0.23 | 0.30 | 0.27 | 0.60 | 0.69 |
| IT Services, Software | 3.74 | 4.12 | 0.24 | 0.23 | 0.33 | 0.31 | 0.46 | 0.55 |
| Communications Equipment | 3.85 | 5.07 | 0.20 | 0.19 | 0.27 | 0.23 | 0.41 | 0.46 |
| Computers, Electronics | 3.15 | 3.76 | 0.26 | 0.27 | 0.35 | 0.35 | 0.37 | 0.46 |
| Semiconductors | 3.17 | 3.19 | 0.19 | 0.14 | 0.26 | 0.21 | 0.41 | 0.46 |
| Telecommunication Services | 3.47 | 4.32 | 0.20 | 0.24 | 0.29 | 0.31 | 0.39 | 0.46 |
| Utilities | 3.66 | 3.38 | 0.20 | 0.20 | 0.25 | 0.22 | 0.26 | 0.30 |
| Average (Equal Weight) | 4.10 | 4.21 | 0.23 | 0.24 | 0.28 | 0.27 | 0.31 | 0.38 |

### Table 7.3

Style 12-month RAD and monthly kurtosis values for sample period January 1997 to June 2008. Portfolios are obtained by cap-weighting the 20 percent of stocks with maximum and minimum exposure to the factor. Active portfolios are obtained by running portfolios against MSCI ACWI IMI as the benchmark.

| | Kurtosis | | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|---|---|
| Portfolio Name | Portfolio | Active | Portfolio | Active | Portfolio | Active | Portfolio | Active |
| Volatility (top 20%) | 3.54 | 5.95 | 0.17 | 0.33 | 0.22 | 0.32 | 0.26 | 0.34 |
| Volatility (bottom 20%) | 3.75 | 5.57 | 0.26 | 0.29 | 0.33 | 0.37 | 0.42 | 0.64 |
| Momentum (top 20%) | 3.04 | 4.27 | 0.21 | 0.26 | 0.27 | 0.30 | 0.29 | 0.31 |
| Momentum (bottom 20%) | 4.48 | 4.59 | 0.22 | 0.25 | 0.30 | 0.25 | 0.32 | 0.29 |
| Size (top 20%) | 4.20 | 6.38 | 0.30 | 0.39 | 0.40 | 0.47 | 0.51 | 0.68 |
| Size (bottom 20%) | 3.60 | 5.41 | 0.30 | 0.43 | 0.32 | 0.43 | 0.31 | 0.48 |
| Value (top 20%) | 5.22 | 4.17 | 0.24 | 0.19 | 0.31 | 0.19 | 0.28 | 0.25 |
| Value (bottom 20%) | 3.25 | 5.32 | 0.26 | 0.30 | 0.34 | 0.35 | 0.39 | 0.59 |
| Growth (top 20%) | 3.79 | 4.66 | 0.23 | 0.17 | 0.30 | 0.22 | 0.35 | 0.35 |
| Growth (bottom 20%) | 3.54 | 3.05 | 0.21 | 0.24 | 0.29 | 0.20 | 0.29 | 0.31 |
| NL Size (top 20%) | 4.27 | 6.52 | 0.24 | 0.27 | 0.31 | 0.30 | 0.35 | 0.40 |
| NL Size (bottom 20%) | 3.88 | 3.43 | 0.23 | 0.25 | 0.31 | 0.20 | 0.31 | 0.29 |
| Liquidity (top 20%) | 3.80 | 7.82 | 0.25 | 0.32 | 0.33 | 0.42 | 0.40 | 0.67 |
| Liquidity (bottom 20%) | 4.52 | 5.45 | 0.24 | 0.29 | 0.32 | 0.35 | 0.33 | 0.39 |
| Leverage (top 20%) | 3.08 | 4.40 | 0.17 | 0.25 | 0.24 | 0.27 | 0.20 | 0.36 |
| Leverage (bottom 20%) | 3.25 | 3.76 | 0.25 | 0.29 | 0.34 | 0.36 | 0.45 | 0.65 |
| Average (Equal Weight) | 3.83 | 5.05 | 0.24 | 0.28 | 0.31 | 0.31 | 0.34 | 0.44 |

**Table 7.4**

Summary 12-month rolling bias statistics results for country, industry and style portfolios contained in Tables 7.1, 7.2, and 7.3. The full sample period, January 1997 to June 2008, is divided into two sub-periods, A and B.

A. January 1997 to June 2002 (66 months)

| | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|
| Item | Portfolio | Active | Portfolio | Active | Portfolio | Active |
| Mean Rolling Bias Statistic | 1.11 | 1.10 | 1.14 | 1.10 | 1.16 | 1.17 |
| Mean Rolling Absolute Deviation | 0.25 | 0.25 | 0.28 | 0.27 | 0.30 | 0.37 |
| Over-forecast | 1.7% | 2.1% | 2.0% | 3.5% | 3.1% | 6.4% |
| Within Confidence Interval | 81.4% | 82.9% | 76.0% | 78.8% | 73.8% | 70.0% |
| Under-forecast | 16.9% | 15.0% | 22.1% | 17.7% | 23.0% | 23.6% |

B. July 2002 to June 2008 (72 months)

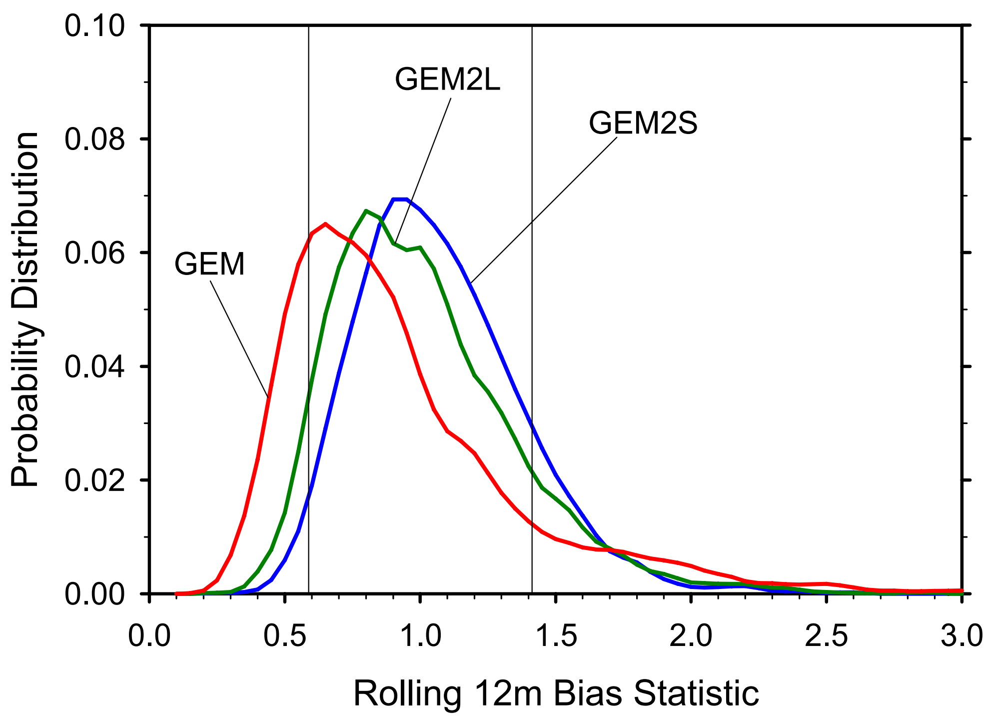
| | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|
| Item | Portfolio | Active | Portfolio | Active | Portfolio | Active |
| Mean Rolling Bias Statistic | 0.97 | 0.98 | 0.89 | 0.89 | 0.78 | 0.70 |
| Mean Rolling Absolute Deviation | 0.21 | 0.23 | 0.27 | 0.26 | 0.32 | 0.36 |
| Over-forecast | 4.5% | 4.4% | 11.5% | 11.5% | 27.2% | 37.0% |
| Within Confidence Interval | 91.7% | 89.2% | 85.3% | 84.3% | 70.9% | 61.7% |
| Under-forecast | 3.8% | 6.4% | 3.2% | 4.2% | 1.9% | 1.3% |

C. January 1997 to June 2008 (138 months)

| | GEM2S | | GEM2L | | GEM | |
|---|---|---|---|---|---|---|
| Item | Portfolio | Active | Portfolio | Active | Portfolio | Active |
| Mean Rolling Bias Statistic | 1.04 | 1.04 | 1.01 | 0.99 | 0.96 | 0.92 |
| Mean Rolling Absolute Deviation | 0.23 | 0.24 | 0.27 | 0.27 | 0.31 | 0.36 |
| Over-forecast | 3.2% | 3.3% | 7.0% | 7.7% | 15.7% | 22.4% |
| Within Confidence Interval | 86.9% | 86.2% | 80.9% | 81.7% | 72.3% | 65.7% |
| Under-forecast | 10.0% | 10.4% | 12.1% | 10.6% | 11.9% | 11.9% |

**Figure 7.1**

Frequency distribution of 12-month rolling bias statistics for country, industry and style active portfolios (long/short) contained in Tables 7.1, 7.2, and 7.3. Sample period is January 1997 to June 2008. Vertical lines indicate the confidence interval [0.59, 1.41].



## Appendix A: Descriptors by Style Factor

### Volatility (Global Relative)

*HBETA* — Historical beta ($\beta_n$)

Computed as the slope coefficient in a time-series regression of local excess stock returns $r_{nt}$ against the cap-weighted local excess return of the estimation universe $R_{nt}$,

$$r_{nt} = \alpha_n + \beta_n R_{nt} + e_{nt} . \tag{A1}$$

The regression coefficients are estimated on the trailing 104 weeks of returns.

*HSIGMA* — Historical sigma ($\sigma_n$)

Computed as the volatility of residual returns in Equation A1,

$$\sigma_n = \mathrm{std}(e_{nt}) . \tag{A2}$$

The volatility is estimated over the trailing 104 weeks of returns.

*DASTD* — Daily standard deviation

This descriptor differentiates stocks based on recent volatility, and is computed as the standard deviation of daily local returns over the past 65 trading days.

*CMRA* — Cumulative range

This descriptor differentiates stocks that have experienced wide swings over the last 12 months from those that have traded in a narrow range. Let $Z(T)$ be the cumulative local excess logarithmic return over the past $T$ months,

$$Z(T) = \sum_{t=1}^{T} \ln(1+r_{nt}) - \sum_{t=1}^{T} \ln(1+r_{ft}) . \tag{A3}$$

where $r_{nt}$ is the local return of stock $n$ for month $t$, and $r_{ft}$ is the risk-free return of the local currency. The cumulative range is given by

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

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

### Momentum (Global Relative)

*HALPHA* — Historical alpha ($\alpha_n$)

Given by the intercept term $\alpha_n$ in Equation A1. Regression coefficients are estimated on trailing 104 weeks of returns.

*RSTR6* <u>*Relative strength (6m)*</u>

Computed as the cumulative local excess logarithmic return over the previous six months,

$$RSTR6 = \sum_{t=1}^{6} \ln(1 + r_{nt}) - \sum_{t=1}^{6} \ln(1 + r_{ft}), \tag{A5}$$

with a one-month lag. Here, $r_{nt}$ is the local return of stock $n$ for month $t$, and $r_{ft}$ is the risk-free return of the local currency.

*RSTR12* <u>*Relative strength (12m)*</u>

Computed as the cumulative local excess logarithmic return over the previous 12 months,

$$RSTR12 = \sum_{t=1}^{12} \ln(1 + r_{nt}) - \sum_{t=1}^{12} \ln(1 + r_{ft}), \tag{A6}$$

with a one-month lag.

**Size (Country Relative)**

*LNCAP* <u>*Log of market cap*</u>

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

**Value (Country Relative)**

*EPFWD* <u>*Predicted earnings-to-price ratio*</u>

Given by the 12-month forward-looking earnings per share (*EPS12F*) divided by the current price. Forward-looking earnings per share are defined as a weighted average between the average analyst-predicted earnings per share for the current and next fiscal years. For details, refer to Section 1.3 of *MSCI Barra Fundamental Data Methodology*.

*ETOP* <u>*Trailing earnings-to-price ratio*</u>

Given by the trailing 12-month earnings per share (*EPS*) divided by the current price. For details on the calculation of the *EPS*, refer to Section 2.1 of *MSCI Barra Fundamental Data Methodology*.

*YIELD* <u>*Dividend-to-price ratio*</u>

Given by the annualized dividend-per-share (DPS) divided by the current price. For details, refer to Section 1.2.4 of *MSCI Barra Fundamental Data Methodology*.

*BTOP* Book-to-price ratio

Computed using the latest book value per share and the price. For details, refer to Section 1.2.5 of *MSCI Barra Fundamental Data Methodology*.

*CETOP* Cash earnings-to-price ratio

Computed using the trailing 12-month cash earnings per share (*CEPS*) divided by the current price. For details, refer to Section 1.2.5 of *MSCI Barra Fundamental Data Methodology*.

**Growth (Country Relative)**

*EGRLF* Long-term predicted earnings growth

Long-term (3-5 years) earnings growth rate forecasted by analysts. For details, refer to Section 2.2.5 of *MSCI Barra Fundamental Data Methodology*.

*SGRO* Sales growth (trailing five years)

Annual reported sales per share are regressed against time over the past five fiscal years. The slope coefficient is then divided by the average annual sales per share to obtain the sales growth. For details, refer to Section 2.2.1 of *MSCI Barra Fundamental Data Methodology*.

*EGRO* Earnings growth (trailing five years)

Annual reported earnings per share are regressed against time over the past five fiscal years. The slope coefficient is then divided by the average annual earnings per share to obtain the earnings growth. For details, refer to Section 2.2.1 of *MSCI Barra Fundamental Data Methodology*.

**Non-linear Size (Country Relative)**

*NLSIZE* Cube of log of market cap

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

## Liquidity (Country Relative)

*STOM* *Share turnover, one month*

Computed as the logarithm

$$STOM = \log\left(\frac{V_t}{S_t}\right), \tag{A7}$$

where $V_t$ is the trading volume of the asset for the month, and $S_t$ is the corresponding number of shares outstanding.

*STOQ* *Average share turnover, trailing 3 months*

Computed as the logarithm

$$STOQ = \log\left(\frac{1}{3}\sum_t \frac{V_t}{S_t}\right), \tag{A8}$$

where $V_t$ is the trading volume of the asset for month $t$, and $S_t$ is the number of shares outstanding. The sum runs over the past 3 months.

*STOA* *Average share turnover, trailing 12 months*

Computed as the logarithm

$$STOA = \log\left(\frac{1}{12}\sum_t \frac{V_t}{S_t}\right), \tag{A9}$$

where $V_t$ is the trading volume of the asset for month $t$, and $S_t$ is the number of shares outstanding. The sum runs over the past 12 months.

## Leverage (Country Relative)

*MLEV* *Market leverage*

Computed as

$$MLEV = \frac{MCAP + PREF + LD}{MCAP}, \tag{A10}$$

where $MCAP$ is the market value of common equity at previous month-end, $PREF$ is the most recent book value of preferred equity, and $LD$ is the most recent book value of long-term debt. For details, refer to Section 5 of *MSCI Barra Fundamental Data Methodology.*

*BLEV* *Book leverage*

Computed as

$$BLEV = \frac{BV + PREF + LD}{BV}, \tag{A11}$$

where $BV$ is the most recent book value of common equity, $PREF$ is the most recent book value of preferred equity, and $LD$ is the most recent book value of long-term debt. For details, refer to Section 5 of *MSCI Barra Fundamental Data Methodology.*

*DTOA* *Debt-to-assets*

Computed as

$$DTOA = \frac{TD}{TA}, \tag{A12}$$

where $TD$ is the book value of total debt (long-term debt, $LD$, and current liabilities, $CL$), and $TA$ is most recent book value of total assets. For details, refer to Section 5 of *MSCI Barra Fundamental Data Methodology.*

# Appendix B: Review of Bias Statistics

## B1. Single-Window Bias Statistics

To assess a model's predictive performance, it is not enough to consider its explanatory power; we must also test how well its risk forecasts do. In this section, we describe how to evaluate the accuracy of risk forecasts.

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

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

$$b_{nt} = \frac{r_{nt}}{\sigma_{nt}}, \tag{B1}$$

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

$$B_n = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(b_{nt} - \bar{b}_n\right)^2}, \tag{B2}$$

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

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

$$B_n \in \left[1-\sqrt{2/T},\, 1+\sqrt{2/T}\right]. \tag{B3}$$

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

If returns are not normally distributed, however, then fewer than 95 percent of the observations will fall within the confidence interval, *even for perfect risk forecasts*. In Figure B1, we show simulated results for the percentage of observations actually falling within this interval, plotted versus observation-window length $T$, for several values of kurtosis $k$.

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

### B2. Rolling-Window Bias Statistics

The purpose of bias-statistic testing is to assess the accuracy of risk forecasts, typically over a long sample period. Let $T$ be the length of the observation window, which corresponds to the number of months in the sample period. One possibility is to select the entire sample period as a single window, and to compute the bias statistic as in Equation B2. This would be a good approach if financial data were stationary, as sampling error is reduced by increasing the length of the window. In reality, however, financial data are not stationary. It is possible to significantly over-predict risk for some years, and under-predict it for others, while ending up with a bias statistic very close to 1.

A more relevant question is to study how accurate the risk forecasts were during any 12-month period. For this purpose, we define the rolling 12-month bias statistic for portfolio $n$,

$$B_n^{\tau} = \sqrt{\frac{1}{11}\sum_{t=\tau}^{\tau+11}(b_{nt} - \bar{b}_n)^2}\ , \tag{B4}$$

where $\tau$ denotes the first month of the 12-month window. The 12-month windows are rolled forward one month at a time until reaching the end of the observation window. If $T$ is the number of periods in the observation window, then each portfolio will have $T-11$ (overlapping) 12-month windows. It is often informative to consider the frequency distribution of rolling 12-month bias statistics. These histograms quickly indicate whether there were any 12-month periods for which portfolio risk was significantly either over-forecast or under-forecast. Figure 7.1 is an example of such a distribution. If there are $N$ portfolios, then the histogram will contain $N(T-11)$ data points, resulting in fairly smooth distributions.

It is also useful to consider the mean of the rolling 12-month bias statistics, given by

$$\tilde{B} = \frac{1}{N(T-11)}\sum_{n,\tau} B_n^{\tau}\ . \tag{B5}$$

This number is reported in Table 6.1 and Table 7.4, and indicates whether, on average, risk was over-forecast or under-forecast during the observation window.

Again, it is possible to over-forecast the risk of some portfolios within certain time periods, and to under-forecast the risk of other portfolios during different time periods, while producing a mean rolling bias statistic $\tilde{B}$ close to 1. What is needed is a way to penalize every deviation away from the ideal bias statistic of 1. The 12-month *rolling absolute deviation* (RAD) for portfolio $n$, defined as

$$\mathrm{RAD}_n = \frac{1}{T-11}\sum_{\tau=1}^{T-11}\left| B_n^{\tau} - 1 \right|, \tag{B6}$$

captures this effect. Smaller RAD numbers, of course, are preferable to larger ones. To assess whether an observed RAD is statistically significant or not, we must consider the properties of the RAD distribution.

In Figure B2, we plot the probability distributions for 12-month RAD over observation windows of 12, 36, and 150 months. Results were obtained via numerical simulation using normally distributed returns. The plot labeled by $T = 12$ corresponds to a single period. The most likely RAD occurs near zero in this case, and the mean of the distribution is approximately $\sqrt{1/12\pi} = 0.163$. A more precise computation, which accounts for small sample sizes, gives a mean of 0.17.

An observation window of $T = 36$ with rolling 12-month windows corresponds to roughly three independent (i.e., non-overlapping) observations. The peak of the probability distribution shifts to the right, and the probability of observing an RAD near zero practically vanishes. This makes intuitive sense, since an RAD of zero would require bias statistics of 1 during *every* 12-month sub-period.

Increasing the observation window to 150 months leads to more than 12 completely independent observations. In this case, the distribution becomes narrower, and the peak shifts further to the right. As $T$ increases further still, the central limit theorem assures that the distribution becomes more sharply peaked about the mean value 0.17, with an approximately normal distribution.

In Figure B2, we indicate by vertical lines the critical mean RAD values, denoted $\tilde{Q}_T^C$, for rolling 12-month windows. These represent the 95$^{\text{th}}$ percentile of the cumulative RAD distributions. In other words, assuming perfect risk forecasts and normally distributed returns, over a 36-month observation window we would expect to find an RAD above 0.30 only five percent of the time. If we observe an RAD above this value, we reject the null hypothesis that the risk forecasts were accurate.

In reality, of course, returns are not normally distributed. In Figure B3, we plot the critical RAD versus size of observation window $T$, for various levels of kurtosis. Not surprisingly, the normal distribution $(k = 3)$ has the smallest critical values. One useful case to consider is the $T \to \infty$ limit, which converges to a critical value of 0.17 for kurtosis $k = 3$. As kurtosis increases, however, the critical values also increase. For instance, using the same 36-month observation window as before but now assuming a kurtosis $k = 5$, RAD values in excess of 0.38 would occur about five percent of the time.

Often we are interested in the mean RAD of a collection of portfolios. This is defined as

$$\text{Mean RAD} = \sum_n (w_n)(\text{RAD}_n), \tag{B7}$$

where $w_n$ is the weight given to observation $n$. Usually we report equal-weighted mean RAD, as in Figure 5.1 or Table 7.4, although sometimes cap-weighted mean RAD are also reported (e.g., Figure 5.2 or Table 6.1).

**Figure B1**

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

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

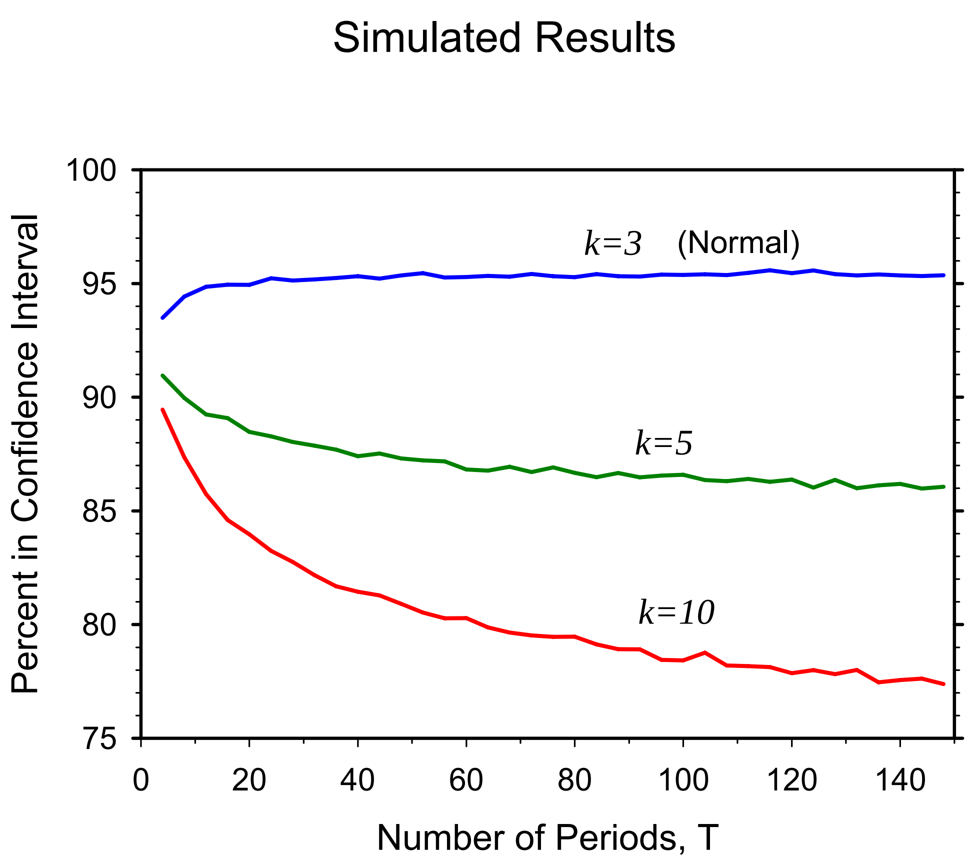


**Figure B2**

Probability distribution of 12-month rolling absolute deviation (RAD). Results were simulated using normally distributed returns, for *observation* windows of 12, 36, and 150 months.

Since a 12-month rolling window within a 12-month observation window corresponds to a single fixed window, the curve labeled $T = 12$ is peaked near zero. A 36-month observation window, however, corresponds to roughly three independent observations, and the peak of the distribution shifts to the right. As the number of periods $T$ becomes large, the central limit theorem assures that the distribution converges to a normal distribution centered at approximately 0.17.

Critical values $\tilde{Q}_T^C$ are indicated by the vertical lines. They indicate the 95th percentile of the cumulative distribution.

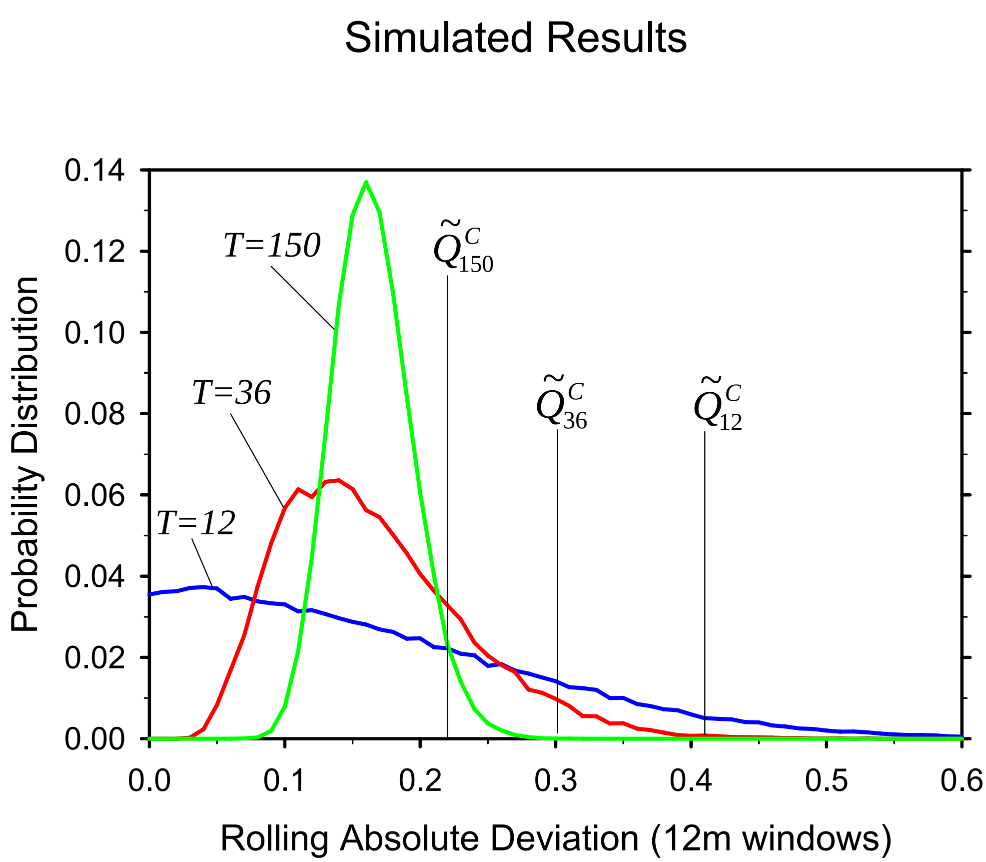
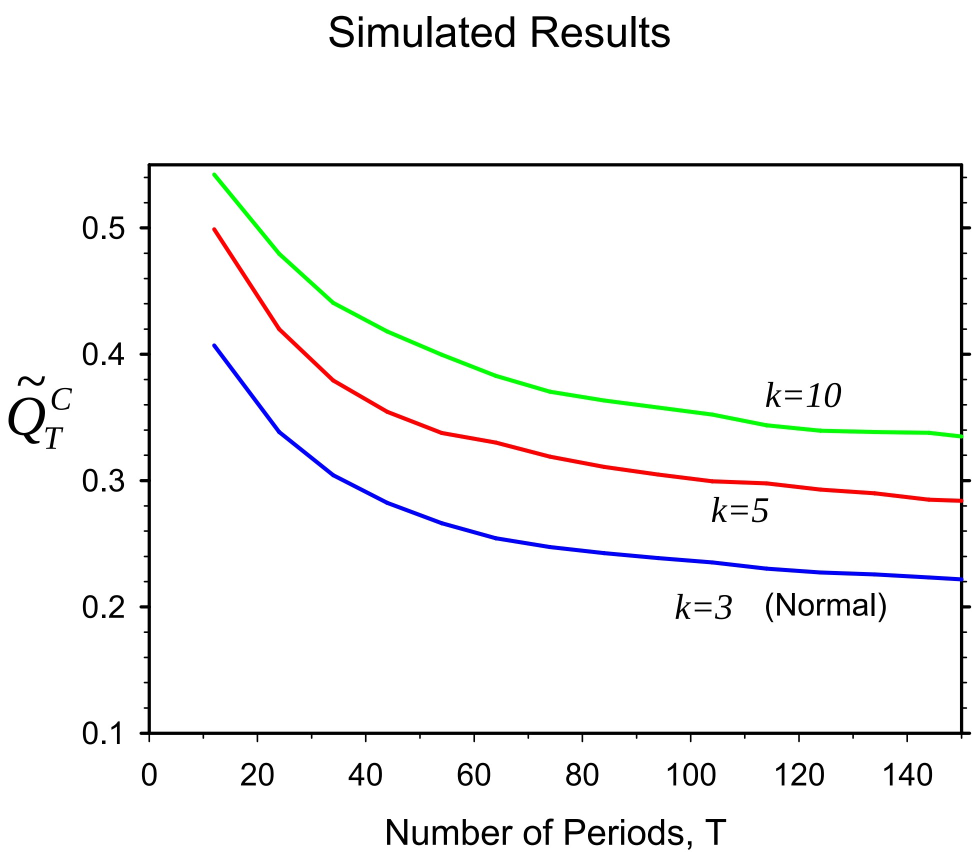


**Figure B3**

Critical values of rolling absolute deviation (RAD), for rolling 12-month windows, versus the number of periods in the observation window. For perfect risk forecasts and normally distributed returns, over a 60-period observation window we expect to observe an RAD above 0.26 only five percent of the time. For a kurtosis of 10, however, the critical value at 60 periods rises to about 0.40.



## Contact Information

clientservice@mscibarra.com

**Americas**

| | |
|---|---|
| Americas | 1.888.588.4567 (toll free) |
| Atlanta | + 1.404.949.4529 |
| Boston | + 1.617.856.8716 |
| Chicago | + 1.312.706.4999 |
| Montreal | + 1.514.847.7506 |
| New York | + 1.212.762.5790 |
| San Francisco | + 1.415.576.2323 |
| Sao Paulo | + 55.11.3706.1360 |
| Toronto | + 1.416.943.8390 |

**Europe, Middle East & Africa**

| | |
|---|---|
| Amsterdam | + 31.20.462.1382 |
| Cape Town | + 27.21.683.3245 |
| Frankfurt | + 49.69.2166.5325 |
| Geneva | + 41.22.817.9800 |
| London | + 44.20.7618.2222 |
| Madrid | + 34.91.700.7275 |
| Milan | + 39.027.633.5429 |
| Paris | 0800.91.59.17 (toll free) |
| Zurich | + 41.44.220.9300 |

**Asia Pacific**

| | |
|---|---|
| China Netcom | 10800.852.1032 (toll free) |
| China Telecom | 10800.152.1032 (toll free) |
| Hong Kong | + 852.2848.7333 |
| Singapore | + 65.6834.6777 |
| Sydney | + 61.2.9033.9333 |
| Tokyo | + 81.3.5226.8222 |

www.mscibarra.com

## Notice and Disclaimer

- This document and all of the information contained in it, including without limitation all text, data, graphs, charts (collectively, the "Information") is the property of MSCI Inc., Barra, Inc. ("Barra"), or their affiliates (including without limitation Financial Engineering Associates, Inc.) (alone or with one or more of them, "MSCI Barra"), or their direct or indirect suppliers or any third party involved in the making or compiling of the Information (collectively, the "MSCI Barra Parties"), as applicable, and is provided for informational purposes only. The Information may not be reproduced or redisseminated in whole or in part without prior written permission from MSCI or Barra, as applicable.
- The Information may not be used to verify or correct other data, to create indices, risk models or analytics, or in connection with issuing, offering, sponsoring, managing or marketing any securities, portfolios, financial products or other investment vehicles based on, linked to, tracking or otherwise derived from any MSCI or Barra product or data.
- **Historical data and analysis should not be taken as an indication or guarantee of any future performance, analysis, forecast or prediction.**
- **None of the Information constitutes an offer to sell (or a solicitation of an offer to buy), or a promotion or recommendation of, any security, financial product or other investment vehicle or any trading strategy, and none of the MSCI Barra Parties endorses, approves or otherwise expresses any opinion regarding any issuer, securities, financial products or instruments or trading strategies. None of the Information, MSCI Barra indices, models or other products or services is intended to constitute investment advice or a recommendation to make (or refrain from making) any kind of investment decision and may not be relied on as such.**
- The user of the Information assumes the entire risk of any use it may make or permit to be made of the Information.
- NONE OF THE MSCI BARRA PARTIES MAKES ANY EXPRESS OR IMPLIED WARRANTIES OR REPRESENTATIONS WITH RESPECT TO THE INFORMATION (OR THE RESULTS TO BE OBTAINED BY THE USE THEREOF), AND TO THE MAXIMUM EXTENT PERMITTED BY LAW, MSCI AND BARRA, EACH ON THEIR BEHALF AND ON THE BEHALF OF EACH MSCI BARRA PARTY, HEREBY EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES (INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTIES OF ORIGINALITY, ACCURACY, TIMELINESS, NON-INFRINGEMENT, COMPLETENESS, MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE) WITH RESPECT TO ANY OF THE INFORMATION.
- **Without limiting any of the foregoing and to the maximum extent permitted by law, in no event shall any of the MSCI Barra Parties have any liability regarding any of the Information for any direct, indirect, special, punitive, consequential (including lost profits) or any other damages even if notified of the possibility of such damages. The foregoing shall not exclude or limit any liability that may not by applicable law be excluded or limited, including without limitation (as applicable), any liability for death or personal injury to the extent that such injury results from the negligence or wilful default of itself, its servants, agents or sub-contractors.**
- Any use of or access to products, services or information of MSCI or Barra or their subsidiaries requires a license from MSCI or Barra, or their subsidiaries, as applicable. MSCI, Barra, MSCI Barra, EAFE, Aegis, Cosmos, BarraOne, and all other MSCI and Barra product names are the trademarks, registered trademarks, or service marks of MSCI, Barra or their affiliates, in the United States and other jurisdictions. The Global Industry Classification Standard (GICS) was developed by and is the exclusive property of MSCI and Standard & Poor's. "Global Industry Classification Standard (GICS)" is a service mark of MSCI and Standard & Poor's.



**About MSCI Barra**

MSCI Barra is a leading provider of investment decision support tools to investment institutions worldwide. MSCI Barra products include indices and portfolio risk and performance analytics for use in managing equity, fixed income and multi-asset class portfolios.

The company's flagship products are the MSCI International Equity Indices, which are estimated to have over USD 3 trillion benchmarked to them, and the Barra risk models and portfolio analytics, which cover 56 equity and 46 fixed income markets. MSCI Barra is headquartered in New York, with research and commercial offices around the world. Morgan Stanley, a global financial services firm, is the controlling shareholder of MSCI Barra.