MSCI

Model Insight

The Barra China International Equity Model

Empirical Notes

Jun Wang, Jay Yao, and Mehmet K. Bayraktar

jun.wang@msci.com

jay.yao@msci.com

mehmet.bayraktar@msci.com

December 2014

1. Introduction

This paper provides a methodology overview and the empirical results for the Barra China International Equity Model (CXE1)[1]. This model covers both Chinese assets for non-domestic investors, including China B Shares, H Shares, Red Chips, P Chips, and depository receipts of Chinese companies, and assets with Hong Kong as the country of exposure. This coverage universe is referred to as the China International Market in this paper.

This paper includes extensive information about the factor structure, commentary on the performance of select factors, an analysis of the explanatory power of the model, and an examination of the statistical significance of the factors. It also includes a side-by-side comparison of the forecasting accuracy and backtesting performance of the Barra China International Equity Model and the Barra Hong Kong Equity Model (HKE1) for Hong Kong-based firms.

The Barra China International Equity Model captures the dynamics of the China International Market through a comprehensive factor set, including the expanded Systematic Equity Strategy (SES) factors. The model leverages the same methodological advances used in the latest generation of Barra equity models, including the latest versions of the Global Equity Model, the Europe Equity Model, and the US and Japan Equity Models. Detailed descriptions of these methodological enhancements can be found in Menchero, Orr, and Wang (2011). In addition, the Barra China International Equity Model employs a Cross-Market Shrinkage methodology to shrink industry and style factor returns for the China International Market toward the factor returns estimated for a Total Market, which is comprised of 11 local markets in the Asia-Pacific region. As many Asian markets are becoming more integrated, this methodology provides the ability to use the same factor structure across multiple single-country models. The benefits are a richer factor structure with reduced estimation error and deep histories even for smaller markets. The shrinkage methodology was first employed by the Barra US Sector Model (USSM1) to determine the style factor returns.

The Barra China International Equity Model offers three versions: Long, Short, and Daily, each with full daily updates. This allows managers to construct and analyze portfolios across different investment horizons. These three versions have identical factor exposures and factor returns, but differ in their factor covariance matrices and specific risk forecasts. The Short horizon model is designed to be more responsive and provide more accurate monthly risk forecasts. The Long horizon model is designed for investors who have longer investment horizons of over six months and are willing to trade some degree of accuracy for greater stability in monthly risk forecasts. The Daily horizon model provides investors with a tactical risk forecast for a daily horizon.

Main advances of the Barra China International Equity Model include:

- Refined style and industry factor set adopted from the latest Barra Asia Pacific Equity Model (ASE2)
- New and enhanced style factors reflecting the latest research on Systematic Equity Strategies (SES), which are aligned with investment strategies frequently used by investment practitioners
- Cross-Market Shrinkage methodology to estimate the factor returns for the China International Market, in order to reduce the estimation errors in the factor returns
- Long, Short, and Daily-horizon versions to serve multiple investment horizons
- Daily update for all three versions of the model

[1] We use the CXE1 abbreviation to refer to the Barra China International Equity Model throughout this document. CXE1D, CXE1S, CXE1L refer to the daily-horizon, short-horizon, and long-horizon model versions, respectively.

- Volatility Regime Adjustment designed to calibrate factor volatilities and specific risk forecasts to current market levels
- Innovative Optimization Bias Adjustment designed to improve risk forecasts for optimized portfolios by reducing the effects of sampling error in the factor covariance matrix
- New Market factor to separate pure industry effect from the overall market, and provide timelier correlation forecasts
- New specific risk model based on daily asset-level specific returns with a Bayesian adjustment technique to reduce biases due to sampling error

2. Methodology Highlights

2.1. Systematic Equity Strategies as Risk Factors

The concept of Systematic Equity Strategies (SES) is discussed by Bayraktar, et al (2013c) and is implemented in the latest generation of Barra equity models. The Barra China International Equity Model also includes these strategies as risk factors. The HKE1 Model included four Systematic Equity Strategies: Momentum (called Success in HKE1), Book-to-Price, Earnings Yield (called Earnings-to-Price in HKE1), and Dividend Yield (called Yield in HKE1). The Barra China International Equity Model expands this list to Earnings Quality, Industry Momentum, News Sentiment, Seasonality, and Short-Term Reversal strategies, which are commonly employed by investment practitioners.

SES factors allow investors to measure their exposure to popular but potentially-crowded investment strategies. Furthermore, asset managers are able to attribute realized returns to these factors, and as a consequence obtain more meaningful insights into the risk and return drivers of their proprietary strategies.

2.2. Cross-Market Shrinkage Methodology for Factor Returns

When constructing a single country model, one often encounters the situation where the estimation universe does not have enough depth, at least for an extended period near the beginning of the model history. This results in large estimation errors in the factor returns estimated from the multi-factor regression. The effect is even more serious for thinly-populated industries, which are industries that are dominated by a small number of assets. To limit the impact of the estimation errors on the factor returns, and in turn on the factor covariance matrices, one often chooses to limit the number of factors contained in the factor structure by adopting a coarse industry structure, by reducing the number of style factors or by combining both methods.

The Barra China International Equity Model addresses this issue by using a novel methodology, namely the Cross-Market Shrinkage methodology, which is based on a methodology first introduced by the Barra US Sector Model (USSM1) to estimate the style factor returns. In this approach, we define a Total Market comprised of 11 countries in the Asia-Pacific region to create a set of factors returns that shrink industry and style factor returns for the China International Market towards the factor returns estimated for the Total Market. This approach allows us to adopt a rich factor structure even for smaller markets, while reducing the estimation errors in the factor returns.

The methodology works as follows:

1. First, we define a Total Market as a combination of 11 well-integrated, local markets in the Asia-Pacific region. These countries are China International/Hong Kong, Australia, Indonesia, India, South Korea, Malaysia, New Zealand, Philippines, Singapore, Thailand, and Taiwan. Countries like Japan and China are less integrated with the other markets, and therefore are not a part of the Total Market. These countries are also selected because they form the core of the estimation universe in the latest Barra Asia Pacific Equity Model (ASE2). In addition, we also retain the same estimation universe, factor structure and raw factor exposures as the Barra Asia Pacific Equity Model, but we re-standardize the style factor exposures within each local market to market capitalization-weighted mean of zero, and equal-weighted standard deviation of one.
2. Second, we estimate the factor returns for the Total Market. The obtained industry and style factor returns serve as the priors for the factor returns for the China International Market.

3. Third, we estimate the factor returns for the 11 local markets using their respective local estimation universes. For each industry or style factor, we thus have a cross-section of 11 estimates for the factor return.
4. Finally, we use the Bayesian Shrinkage methodology to determine the final factor return estimates for the China International Market. For an industry or style factor, the factor return estimate from the regression for the China International Market in the third step is our likelihood. Its associated estimation error is compared to the cross-sectional dispersion of the 11 estimates to determine the relative weights for the likelihood and the prior. At each time point, the final estimate for the factor return is the weighted average of the likelihood and the prior. The exception is the China International Market factor return. The Market factor return from the regression for the China International Market is directly used as the final Market factor return; the regressions for the other ten local markets do not provide an estimate for this factor return.

More details about the Cross-Market Shrinkage methodology can be found in Appendix A.

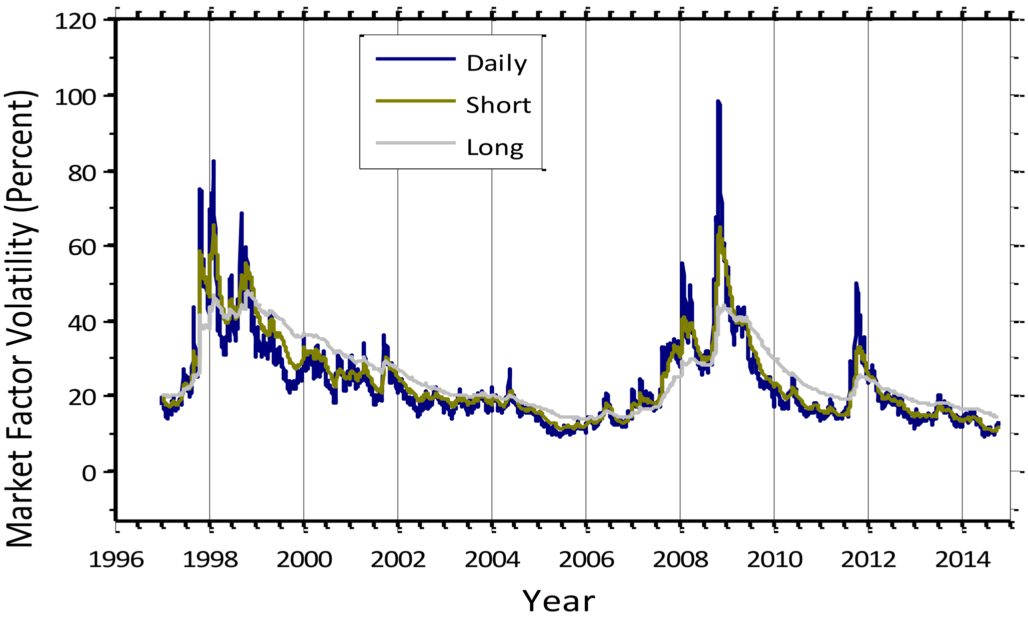
2.3. Multiple-Horizon Risk Forecasts

Risk levels in financial markets change continuously. How investors respond to this challenge depends partly on their investment horizon, i.e., how much time typically elapses between major revisions in portfolio positions. The risk that usually matters most to the investors is the average level of risk a portfolio will experience during the intended investment horizon. Typically, risk levels during a shorter horizon display more pronounced clustering and are more volatile than risk levels during a longer horizon. Risk forecasts aimed at investors with shorter horizons, therefore, should be more volatile than risk forecasts aimed at long-term investors. The short-horizon forecast must react more quickly to rapid, and typically short-lived changes in risk levels than its long-horizon counterpart. In view of this, the Barra China International Equity Model provides monthly risk forecasts for both short and long horizons.

The new daily-horizon version is designed to be more responsive and to provide accurate forecasts at a daily prediction horizon. This is achieved primarily by using shorter half-lives in the factor covariance matrix and specific risk estimators, which makes the model more responsive to volatility changes in recent history.

In **Figure 2.1,** we plot the time series of annualized volatility forecast for the Market factor by model horizon.

Figure 2.1: Market factor risk forecast: Daily, Short, and Long horizons



From the above figure, we see that, immediately after each significant market shock, the forecast for the Daily horizon model is significantly higher than the Long or Short horizon forecasts. As volatility subsides in the subsequent periods, the forecast for the Daily horizon version drops more quickly.

2.4. Volatility Regime Adjustment

A major source of risk model bias is the unstable nature of volatilities over time—a characteristic known as *non-stationarity*. Because risk models must look backward to make predictions about the future, they tend to underpredict risk in times of rising volatility and to overpredict risk in times of falling volatility.

An important innovation, first introduced in the Barra US Equity Model (USE4) and applied here in the Barra China International Equity Model, is the Volatility Regime Adjustment for estimating factor volatilities (see Appendix B). As described in Menchero, Wang, and Orr (2011), the Volatility Regime Adjustment relies on the concept of a cross-sectional bias statistic, which may be interpreted as an *instantaneous* measure of risk model bias for that particular period. By taking a weighted average of this measure over a suitable interval, the non-stationarity bias can be significantly reduced.

Just as factor volatilities are not stable across time, the same holds true for specific risk. We therefore also apply the Volatility Regime Adjustment to the Barra China International Equity Model specific risk model.

2.5. Optimization Bias Adjustment

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

Another important innovation, first introduced in the USE4 Model and applied here in the Barra China International Equity Model, is the identification of portfolios that capture these biases and the correction for them directly within the factor covariance matrix. As shown by Menchero, Wang, and Orr (2011), the *eigenfactors* of the sample covariance matrix are systematically biased. More specifically, the sample covariance matrix tends to underpredict the risk of low-volatility eigenfactors, while overpredicting the risk of high-volatility eigenfactors.

In Optimization Bias Adjustment, we estimate the biases of the eigenfactors via Monte Carlo simulation, and then adjust the predicted eigenvalues of the eigenfactors to correct for these biases (see Appendix C). This procedure helps to improve factor risk forecasts of optimized portfolios. Furthermore, it has the benefit of building the corrections directly into the factor covariance matrix, while fully preserving the meaning and intuition of the pure factors. Lee, et al (2011) demonstrates the effectiveness of this approach by backtesting active portfolios.

2.6. Market Factor

Traditionally, single country models like the previous Barra Hong Kong Equity Model (HKE1) have included industry and style factors, but no Market factor. An important improvement with the Barra China International Equity Model is to explicitly include a Market factor, which is analogous to the World factor introduced in the Barra Global Equity Model (starting with GEM2), as described by Menchero, Morozov, and Shepard (2008, 2010).

One significant benefit of the Market factor is the insight and intuition that it affords to active portfolio managers. For instance, Menchero, Orr, and Wang (2011) show that the Market factor portfolio can be cleanly interpreted as the cap-weighted market portfolio. Furthermore, the Market factor disentangles the pure industry effect from the overall market effect, thus providing a cleaner interpretation of the industry factors.

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

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

Another benefit of the Market factor pertains to improvements in risk forecasting. Intuitively and empirically, we know that industries tend to become more highly correlated in times of financial crisis. As shown in Menchero, Orr, and Wang (2011), the Market factor is able to capture these changes in industry correlation in a timelier fashion. The underlying mechanism for this effect is that net-long industry portfolios have common exposure to the Market factor, and when the volatility of the Market factor rises during times of market stress, it explains the increased correlations for the industries.

2.7. Specific Risk Model with Bayesian Shrinkage

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

A potential shortcoming of the pure time-series approach is that specific volatilities may not fully persist out-of-sample. In fact, as shown in Menchero, Orr, and Wang (2011), there is a tendency for time-series volatility forecasts to overpredict the specific risk of high-volatility stocks, and underpredict the risk of low-volatility stocks.

To reduce these biases, a Bayesian shrinkage technique (see Appendix D) was introduced in the USE4 Model. Stocks are segmented into deciles based on their market capitalization. Within each size bucket, the mean and standard deviation of the specific risk forecasts are computed. The volatility forecast is then pulled or 'shrunk' to the mean within the size decile, with the shrinkage intensity increasing with the number of standard deviations away from the mean.

3. Factor Structure Overview

3.1. Estimation Universe

The estimation universe is a subset of the coverage universe used to standardize the style factor exposures and estimate the factor returns. As discussed in Section 2.2, the estimation universe for the new Barra Asia Pacific Equity Model (ASE2) is used as the estimation universe for the Total Market. Specifically, the estimation universe is constructed country by country, according to a consistent set of rules:

- Illiquid assets and micro-caps (market capitalization < USD 10m) are excluded.
- Cross-listed issues and depositary receipts are excluded.
- Assets traded outside the Asia-Pacific region are excluded.
- One issuer cannot have more than one issue included.
- All members of the MSCI All Country (AC) Asia Pacific Investable Market Index (IMI) are included after June 2002.
- Exceptions are made for important local index constituents to be included. For example, HSBC is an important constituent in the Hang Seng Index, thus is included in the estimation universe for the China International Market.

The estimation universe is based primarily on the MSCI All Country (AC) Asia Pacific Investable Market Index (IMI). Until June 2002, the back-calculated history of the index included a broad group of small-cap assets. From July 2002 onwards, more restrictive liquidity filtering rules are applied to the index, and this results in a sudden drop of about 1600 constituents. To ensure continuity, the estimation universe adds small-cap stocks after June 2002, and conversely excludes some small-cap index stocks before that date.

3.2. Industry Factors

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

As discussed in Section 2.2, the Barra China International Equity Model retains the industry structure from the new Barra Asia Pacific Equity Model (ASE2). There are 27 industries in the industry structure, as presented in **Table 3.1**. Industries that qualify as factors tend to exhibit high volatility and have significant weight. Also reported in Table 3.1 are the average weights from December 31, 1996 to September 30, 2014, and the end-of-period weights.

Note that weights are determined within the estimation universe using market capitalization. Averages are computed over the sample period from December 31, 1996 to September 30, 2014.

Table 3.1: Industry factors for the Barra China International Equity Model

Sector	Industry Factor Name	Average Weight (%)	Sep 30, 2014 Weight (%)
Energy	Energy	4.73	5.76
Materials	Chemicals	0.41	0.42
	Construction Materials and Packaging	1.15	0.92
	Metals and Mining ex Gold and Steel	0.51	0.62
	Gold	0.10	0.11
	Steel	0.28	0.49
Industrials	Building Products and Construction	0.43	0.69
	Capital Goods ex Building and Machinery	5.79	7.41
	Machinery	0.48	0.70
	Commercial Services and Transportation	2.60	2.10
	Airlines and Marine	1.30	0.70
Consumer Discretionary	Automobiles and Components	0.64	1.36
	Consumer Durables and Apparel	1.73	2.26
	Consumer Services	2.10	5.19
	Media and Retailing	3.00	2.35
Consumer Staples	Consumer Staples	2.22	3.90
Health Care	Health Care	0.42	1.70
Financials	Banks	28.58	20.25
	Diversified Financials	5.45	2.12
	Insurance	1.98	4.56
	Real Estate	16.23	13.35
Information Technology	Software and Services	1.43	5.12
	Electronic Equipment and Components	1.04	1.18
	Computers and Peripherals	0.48	0.63
	Semiconductors and Semiconductor Equipment	0.34	0.91
Telecommunication Services	Telecommunication Services	11.28	10.20
Utilities	Utilities	5.31	5.01

In **Table 3.2**, we report the underlying GICS codes that map to each of the industry factors for the Barra China International Equity Model. This table illustrates the customization that takes place within each sector.

Table 3.2: Mapping of industry factors to GICS codes

Industry Factor Name	GICS Code
Energy	10
Chemicals	151010
Construction Materials and Packaging	151020, 151030, 151050
Metals and Mining ex Gold and Steel	15104010, 15104020, 15104040, 15104045
Gold	15104030
Steel	15104050
Building Products and Construction	201020, 201030
Capital Goods ex Building and Machinery	201010, 201040, 201050, 201070
Machinery	201060
Commercial Services and Transportation	2020, 203010, 203040, 203050
Airlines and Marine	203020, 203030
Automobiles and Components	2510
Consumer Durables and Apparel	2520
Consumer Services	2530
Media and Retailing	2540, 2550
Consumer Staples	30
Health Care	35
Banks	4010
Diversified Financials	4020
Insurance	4030
Real Estate	4040
Software and Services	4510
Electronic Equipment and Components	452010, 452030, 452040
Computers and Peripherals	452020
Semiconductors and Semiconductor Equipment	452050, 4530
Telecommunication Services	50
Utilities	55

3.3. Style Factors

Investment style represents another major source of systematic risk for equity portfolios. Style factors are constructed from financially intuitive stock attributes called *descriptors*, which serve as effective predictors of equity return covariance. Appendix E summarizes the descriptor definitions for each style factor. The descriptor weights are retained as proprietary.

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

A summary of all style factors is as follows:

- The *Beta* factor captures market risk that cannot be explained by the Market factor. We compute Beta by a time-series regression of stock excess returns against the cap-weighted return of the estimation universe. To better understand how Beta relates to the Market factor, consider a fully invested long-only portfolio that is tilted toward high-beta stocks. Intuitively, this portfolio has greater market risk than a portfolio with a beta of one. This additional market risk is captured through positive exposure to the Beta factor. Since the time-series correlation between the Market factor and the Beta factor is typically very high, these two sources of risk are additive in this example. If, by contrast, the portfolio is invested primarily in low-beta stocks, then the risk from the Beta and the Market factors would be partially offset, as expected.
- The *Book-to-Price Ratio* factor captures the portion of asset return that is derived from the book value relative to the market capitalization of the firm.
- The *Developed Markets Sensitivity* factor measures how sensitive the assets are to the performance of the developed markets.
- The *Dividend Yield* factor associates the asset return to the dividend payout by the firm.
- The *Downside Beta* factor captures sensitivity of asset returns to negative market performance. This factor is orthogonalized to Beta factor to reduce collinearity.
- The *Earnings Quality* factor explains return differences between high quality and low quality stocks. Companies with a large accrual component to their earnings are more likely to disappoint investors in the future than are companies with strong cash earnings. The descriptors in this factor are accruals and cash-earnings-to-earnings ratio.
- The *Earnings Yield* factor describes return differences based on a company's earnings relative to its price. Earnings Yield is considered by many investors to be a strong value signal. The descriptors in this factor are analyst-predicted 12-month forward earnings-to-price ratio, cash-earnings-to-price ratio, and earnings-to-price ratio.
- The *Growth* factor differentiates stocks based on their prospects for sales or earnings growth. This factor contains forward-looking descriptors in the form of long/short-term analyst predicted earnings growth as well as historical descriptors for sales and earnings growth over the trailing five years.
- The *Industry Momentum* factor differentiates stocks based on both their performance over the trailing six months and the industry performance over the same time. This factor is orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Downside Beta and Mid Capitalization.
- The *Leverage* factor captures return differences between high-leverage and low-leverage stocks. The descriptors within this style factor include market leverage, book leverage, and debt-to-assets ratio.

- The *Liquidity* factor describes return differences due to relative trading activity. The descriptors are based on the fraction of total shares outstanding that trade over a recent window.
- The *Momentum* factor differentiates stocks based on their performance over the trailing 12 months. It consists of two descriptors, relative strength and historical alpha.
- The *Mid-Capitalization* factor captures the premium for the mid-cap companies. This factor is based on a single descriptor—the cube of the Size exposure. However, because this descriptor is highly collinear with the Size factor, it is orthogonalized with respect to Size. This procedure does not affect the fit of the model, but does mitigate the confounding effects of collinearity, while preserving an intuitive meaning for the Size factor. As described by Menchero (2010), this factor roughly captures the risk of a “barbell portfolio” that is long mid-cap stocks and short small-cap and large-cap stocks.
- The *News Sentiment* factor captures sensitivity of asset returns to company-related news announcements. This factor is orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Downside Beta, and Mid Capitalization.
- The *Oil Sensitivity* factor captures the sensitivity of the assets to changes in oil price.
- The *Residual Volatility* factor captures the volatility of the residual return in the times series regression of the *Beta* factor. It is composed of three descriptors: (a) the volatility of daily excess returns, (b) the volatility of daily residual returns, and (c) the cumulative range of the stock over the last 12 months. Since these descriptors tend to be highly collinear with the Beta, Size and Liquidity factors, the Residual Volatility factor is orthogonalized with respect to those factors, as described by Menchero (2010).
- The *Seasonality* factor captures historical seasonal variation and time-of-the-year effects in stock performance. This factor is orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Downside Beta, and Mid-Capitalization.
- The *Short-Term Reversal* factor captures how stocks under/over-perform the market over the recent past as they are expected to over/under-perform in the near future. This factor is orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Downside Beta, and Mid Capitalization.
- The *Size* factor represents the strongest source of equity return covariance, and captures return differences between large-cap stocks and small-cap stocks. We measure Size by the logarithm of the issuer’s market capitalization.

4. Model Characteristics and Properties

4.1. Shrinkage Intensity

As discussed in Section 2.2, in the Cross-Market Shrinkage methodology for factor returns, we shrink the industry and style factor returns from the regression for the China International Market toward those from the Total Market regression. The shrinkage intensity measures the weights on the Total Market factor returns. In general, the shrinkage intensity is higher for the factors with higher estimation errors from the regression for the China International Market.

Figure 4.1: Shrinkage intensities for the Energy (ENERGY), Chemicals (CHEMICAL), and Construction Materials and Packaging (CNSTRP) factors

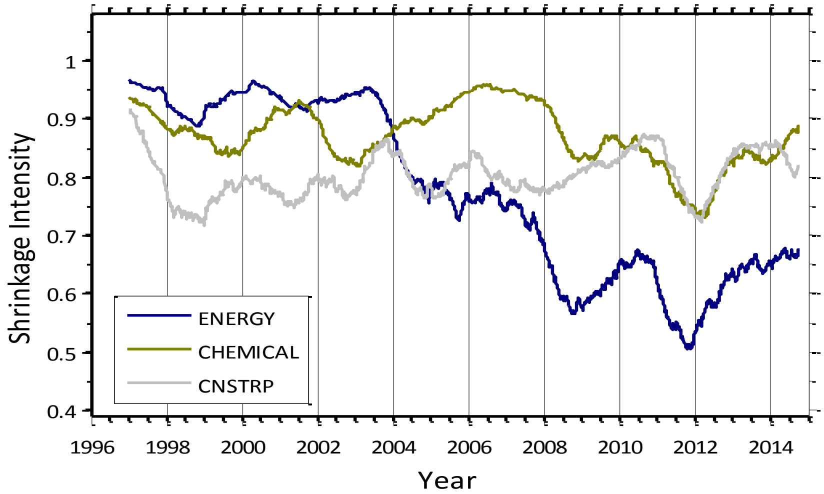


Figure 4.2: Shrinkage intensities for the Metals and Mining ex Gold and Steel (METMIN), Gold (GOLD), and Steel (STEEL) factors

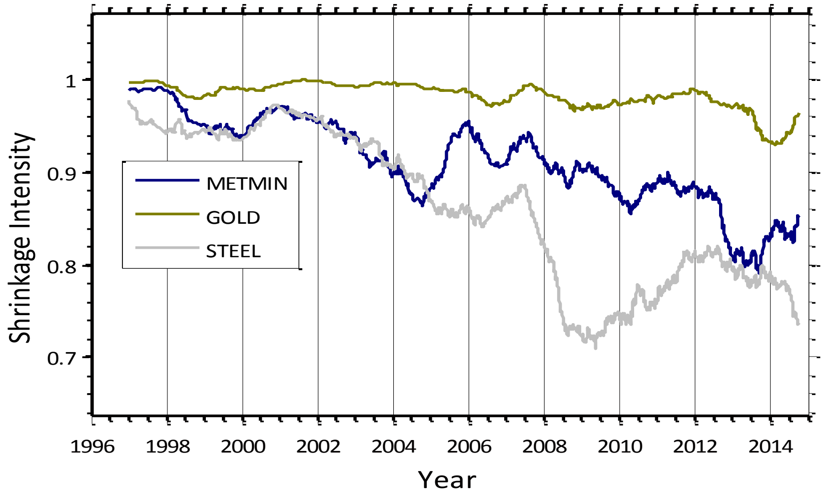


Figure 4.3: Shrinkage intensities for the Building Products and Construction (BUILD), Capital Goods ex Building and Machinery (CAPGOODS), and Machinery (MACHINE) factors

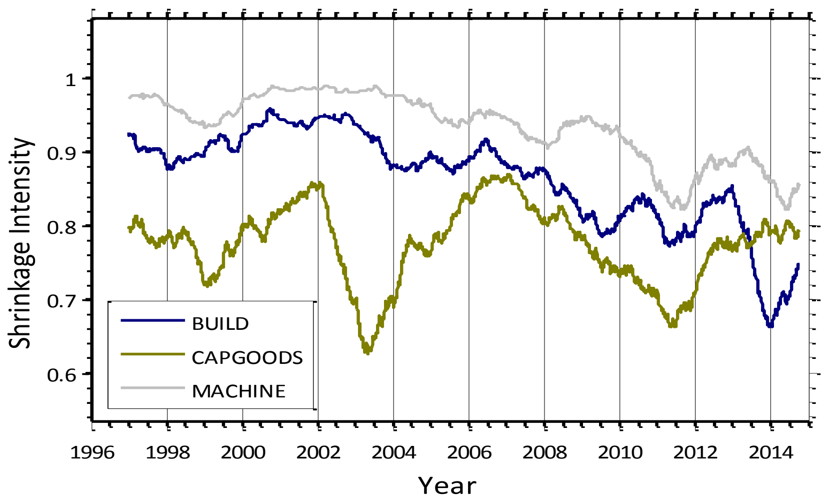


Figure 4.4: Shrinkage intensities for the Commercial Services and Transportation (COMTRA), Airlines and Marine (AIRMARNE), and Automobiles and Components (AUTOCOMP) factors

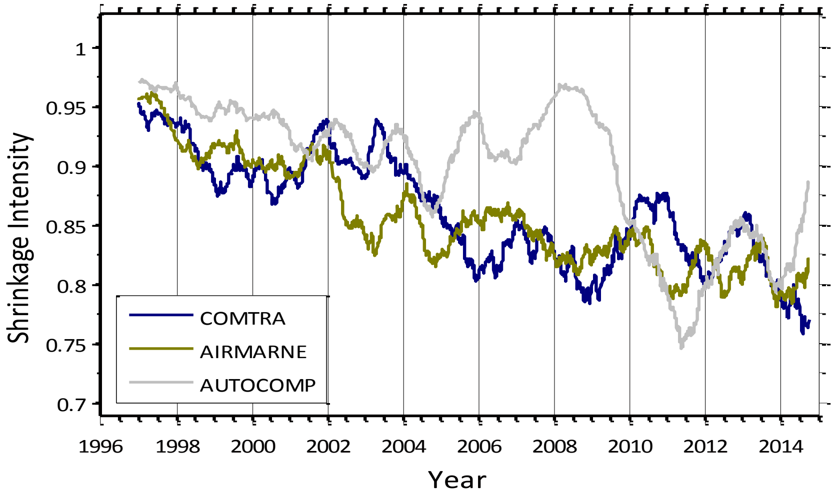


Figure 4.5: Shrinkage intensities for the Consumer Durables and Apparel (CONDUR), Consumer Services (CONSRV), and Media and Retailing (MEDIARET) factors

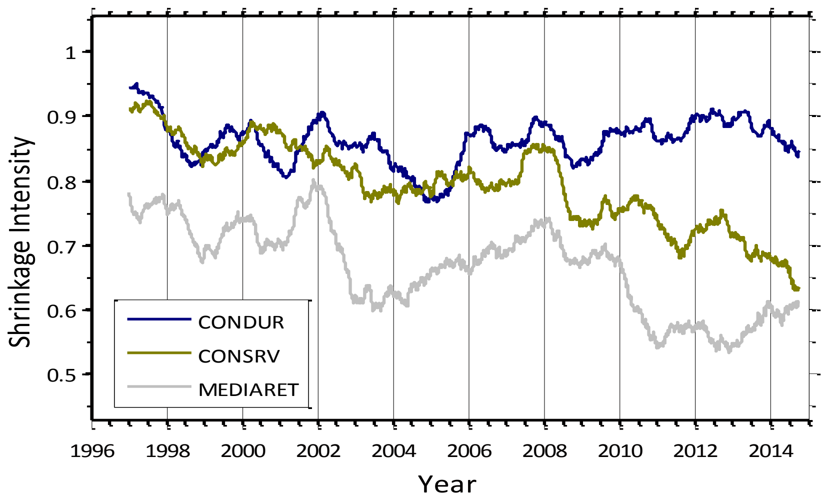


Figure 4.6: Shrinkage intensities for the Consumer Staples (CONSTAP), Health Care (HEALTH), and Banks (BANKS) factors

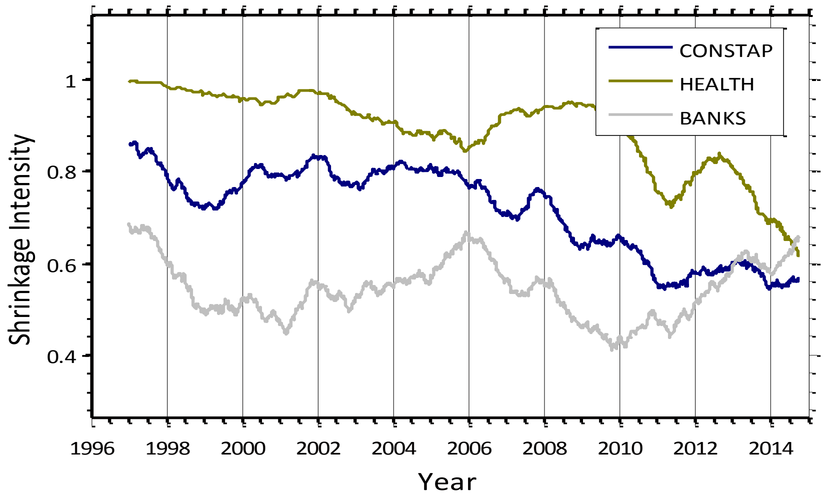


Figure 4.7: Shrinkage intensities for the Diversified Financials (DIVFINAN), Insurance (INSURAN), and Real Estate (REALEST) factors

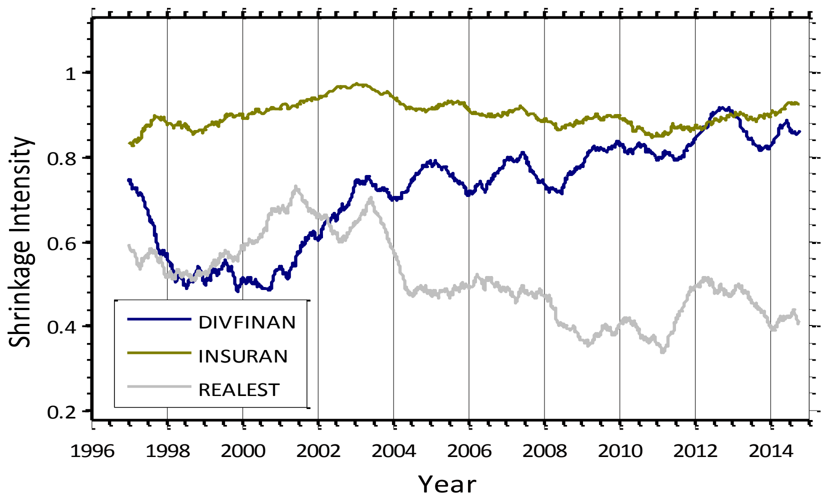


Figure 4.8: Shrinkage intensities for the Software and Services (SOFTWARE), Electronics Equipment and Components (ELECTRON), and Computers and Peripherals (COMPUTER) factors

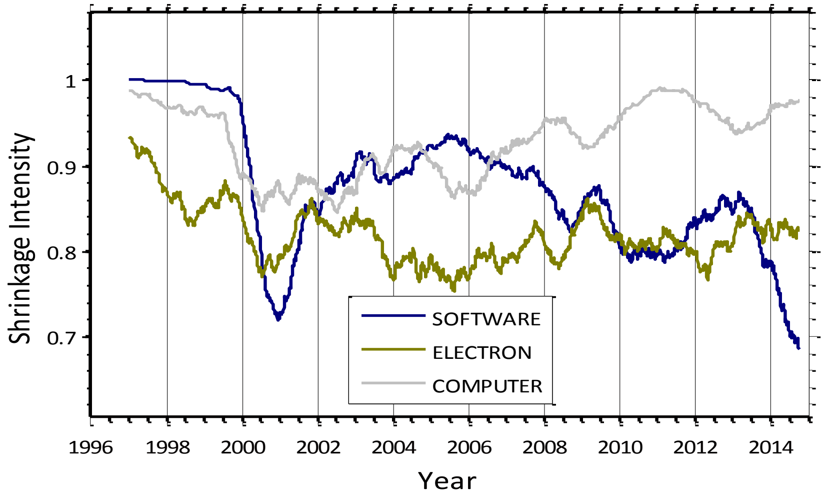


Figure 4.9: Shrinkage intensities for the Semiconductors and Semiconductor Equipment (SIMICOND), Telecommunication Services (TELECOMM), and Utilities (UTILITY) factors

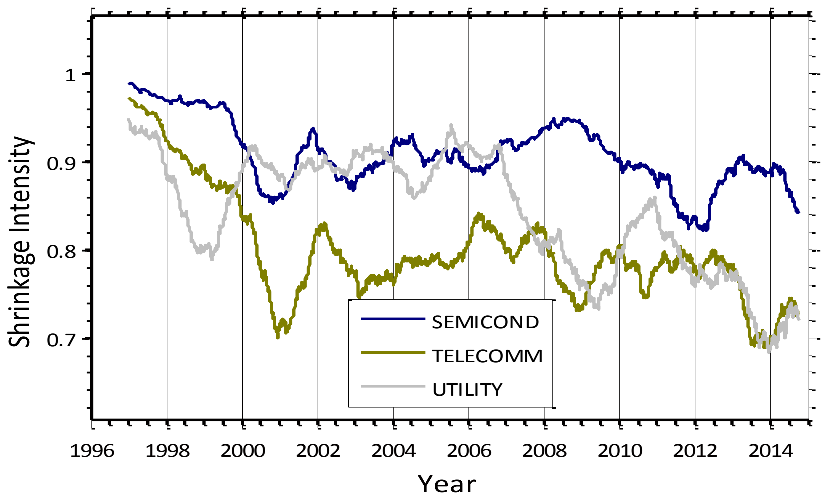


Figure 4.10: Shrinkage intensities for the Beta, Size, and Residual Volatility factors

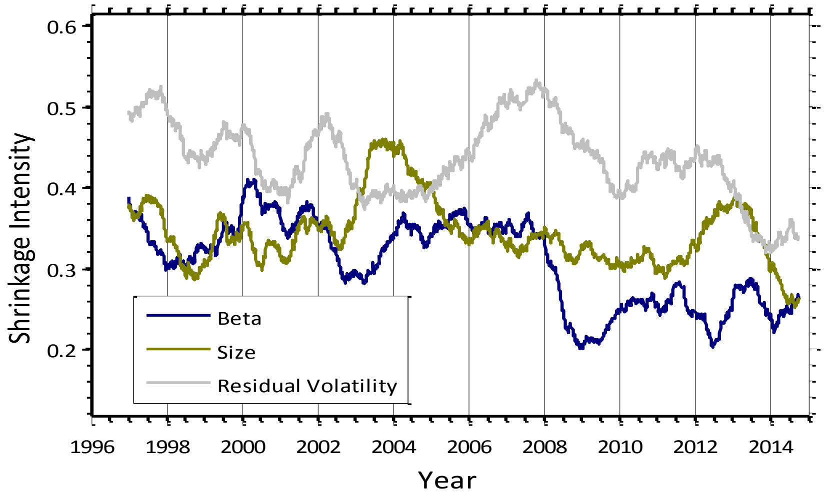


Figure 4.11: Shrinkage intensities for the Leverage, Liquidity, and Mid Capitalization factors

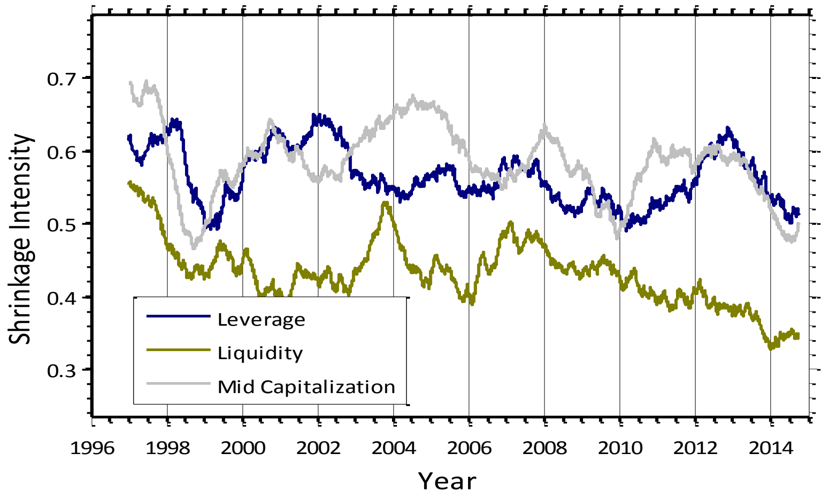


Figure 4.12: Shrinkage intensities for the Downside Beta, Oil Sensitivity, and Developed Markets Sensitivity factors

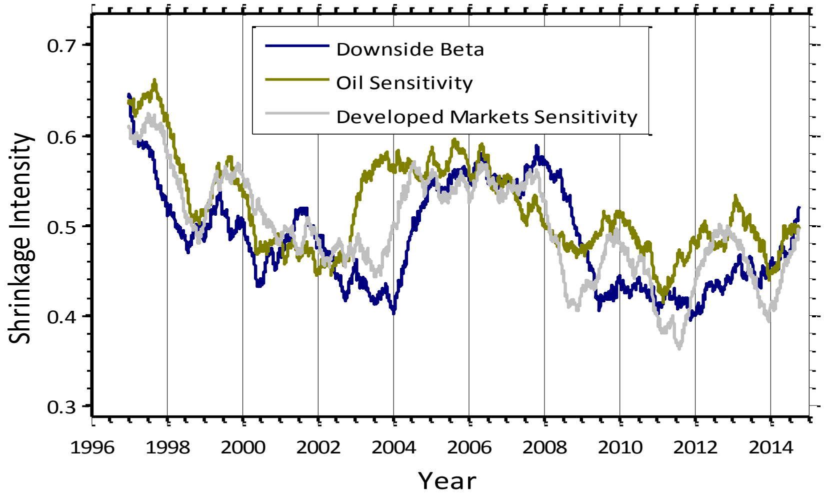


Figure 4.13: Shrinkage intensities for the Momentum, Industry Momentum, and Short-Term Reversal factors

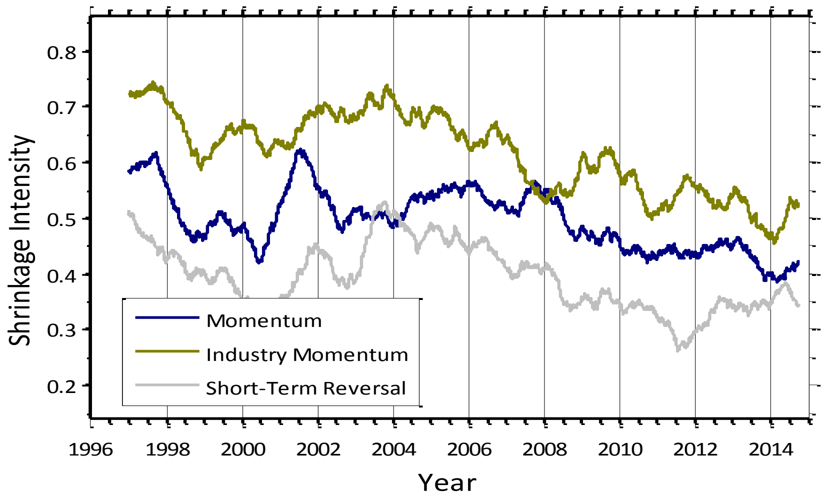


Figure 4.14: Shrinkage intensities for the Earnings Yield, Growth, and Dividend Yield factors

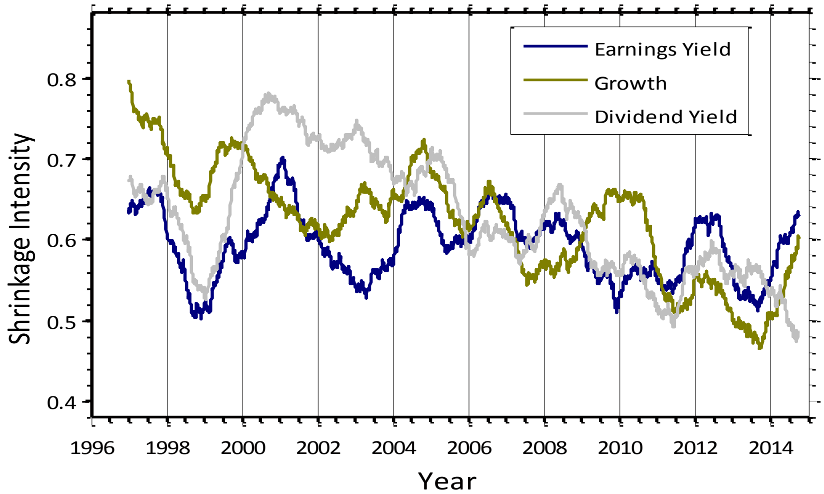


Figure 4.15: Shrinkage intensities for the Book-to-Price Ratio, Earnings Quality, and Seasonality factors

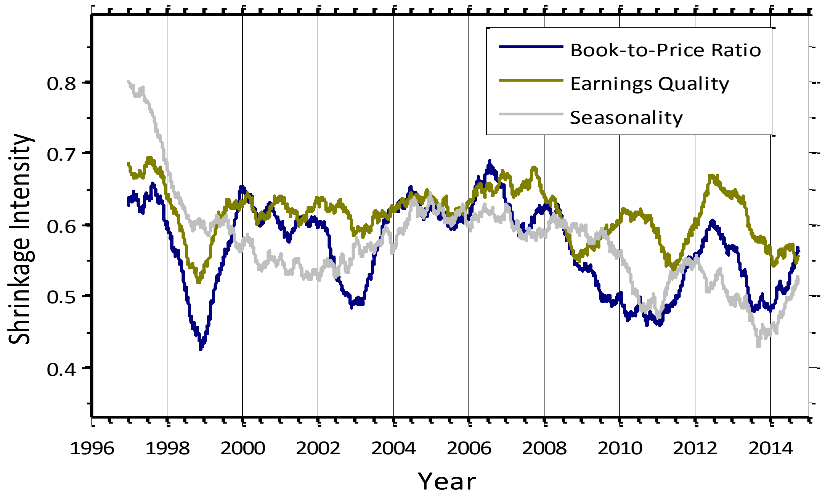
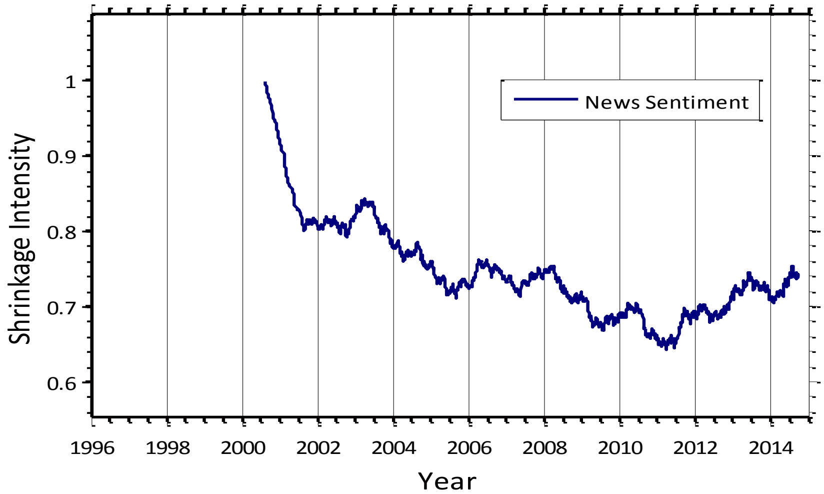


Figure 4.16: Shrinkage intensity for the News Sentiment factor



4.2. Factor Performance

In the following figures, we report cumulative returns to individual factors in the Barra China International Equity Model. The cumulative returns in the charts are calculated by summing rather than compounding daily returns. One immediate consequence of this is that cumulative returns may drop below -100% in some cases.

Figure 4.17: Cumulative return of the Market factor

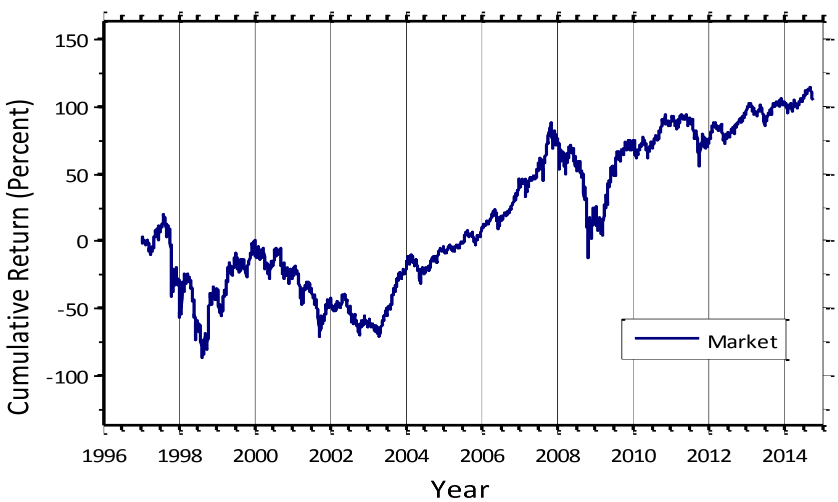


Figure 4.18: Cumulative returns of the Energy (ENERGY), Chemicals (CHEMICAL), and Construction Materials and Packaging (CNSTRP) factors

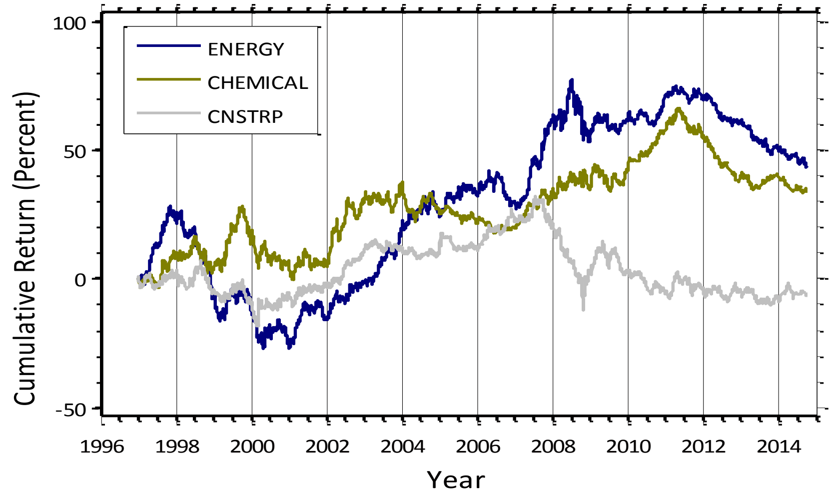


Figure 4.19: Cumulative returns of the Metals and Mining ex Gold and Steel (METMIN), Gold (GOLD), and Steel (STEEL) factors

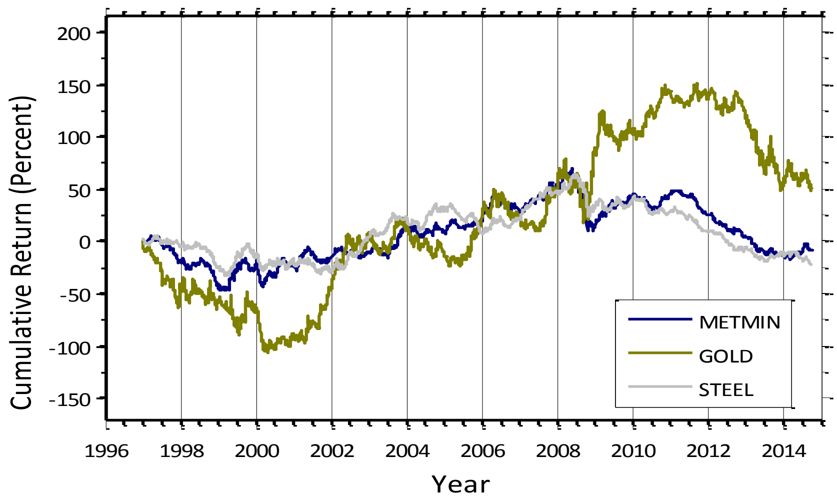


Figure 4.20: Cumulative returns of the Building Products and Construction (BUILD), Capital Goods ex Building and Machinery (CAPGOODS), and Machinery (MACHINE) factors

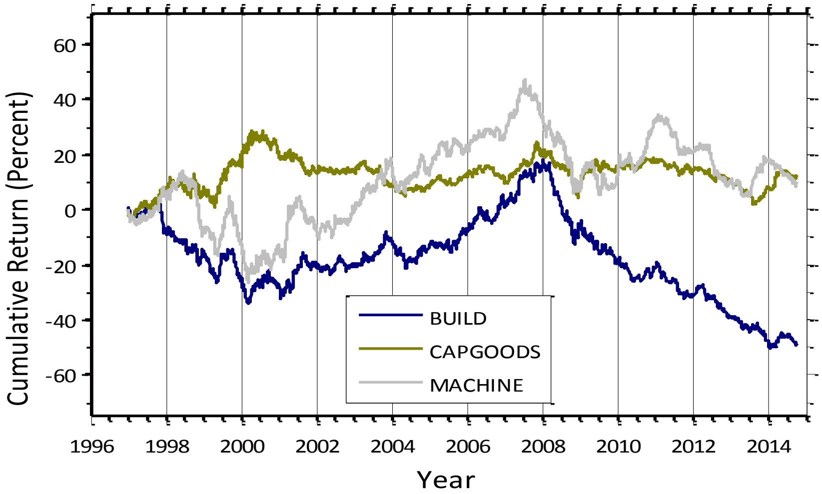


Figure 4.21: Cumulative returns of the Commercial Services and Transportation (COMTRA), Airlines and Marine (AIRMARNE), and Automobiles and Components (AUTOCOMP) factors

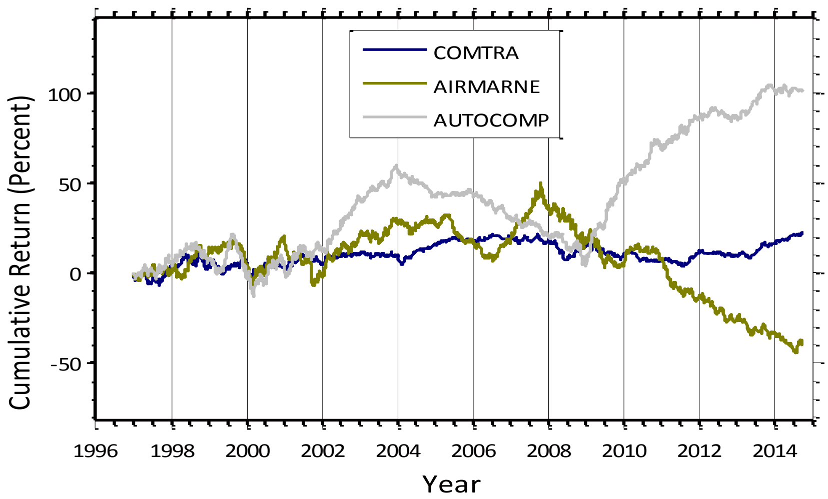


Figure 4.22: Cumulative returns of the Consumer Durables and Apparel (CONDUR), Consumer Services (CONSRV), and Media and Retailing (MEDIARET) factors

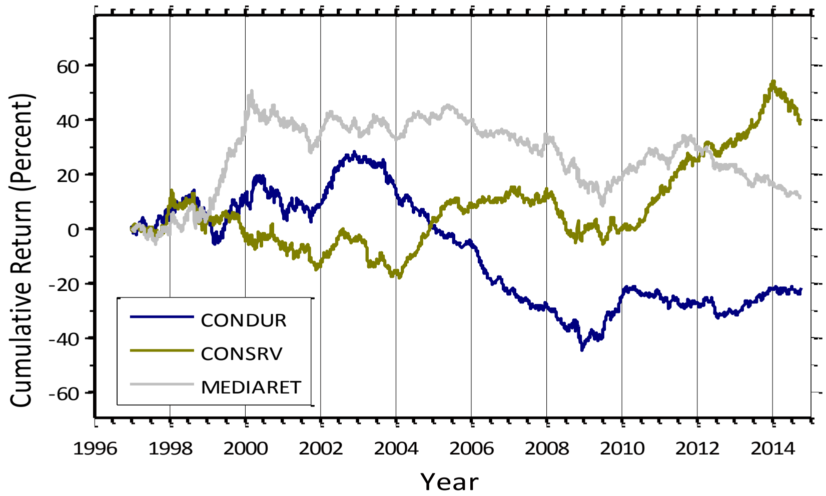


Figure 4.23: Cumulative returns of the Consumer Staples (CONSTAP), Health Care (HEALTH), and Banks (BANKS) factors

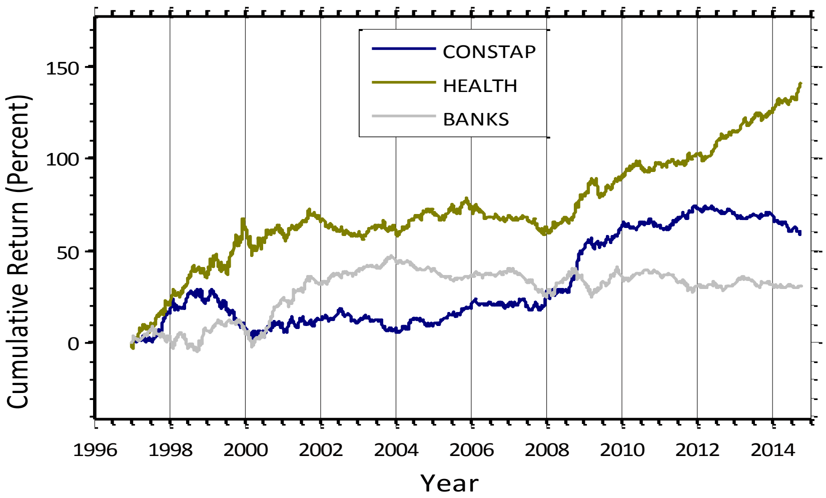


Figure 4.24: Cumulative returns of the Diversified Financials (DIVFINAN), Insurance (INSURAN), and Real Estate (REALEST) factors

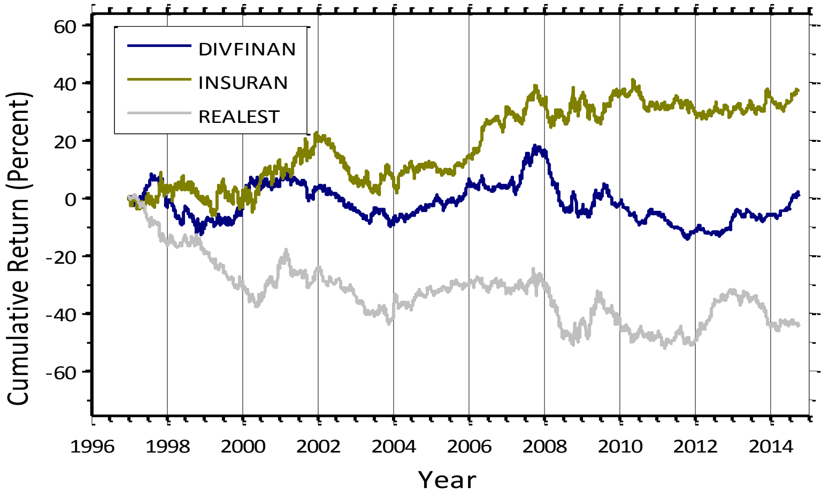


Figure 4.25: Cumulative returns of the Software and Services (SOFTWARE), Electronics Equipment and Components (ELECTRON), and Computers and Peripherals (COMPUTER) factors

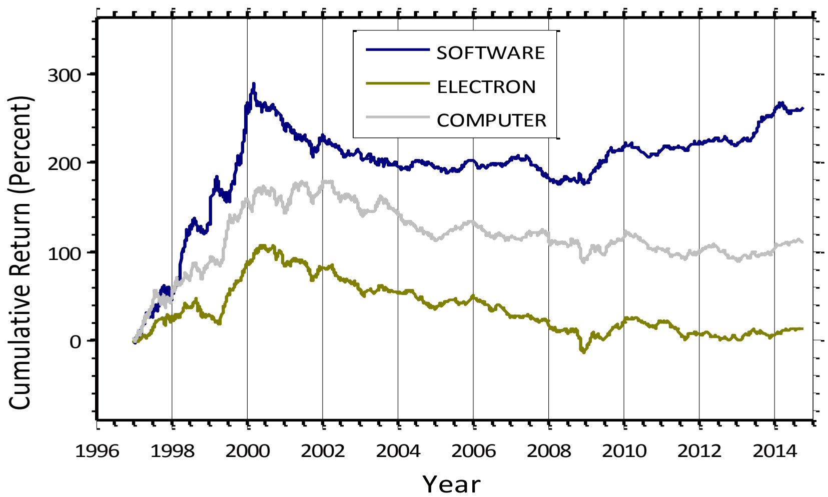


Figure 4.26: Cumulative returns of the Semiconductors and Semiconductor Equipment (SEMICOND), Telecommunication Services (TELECOMM), and Utilities (UTILITY) factors

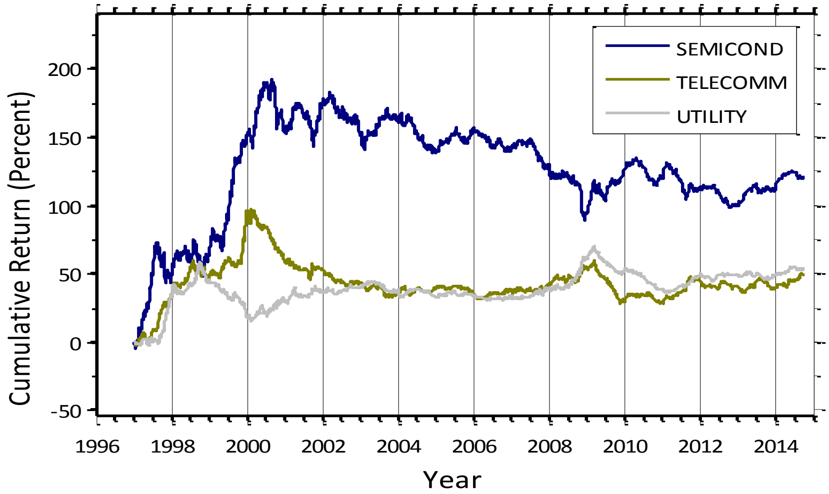


Figure 4.27: Cumulative returns of the Beta, Size, and Residual Volatility factors

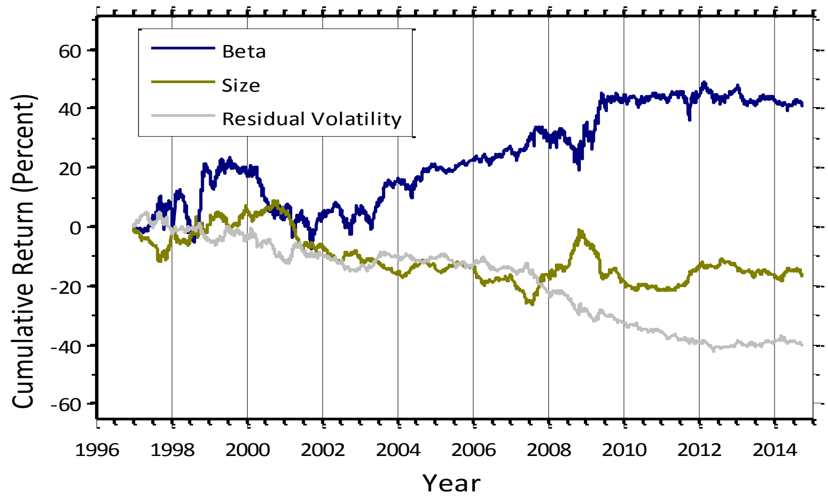


Figure 4.28: Cumulative returns of the Leverage, Liquidity, and Mid Capitalization factors

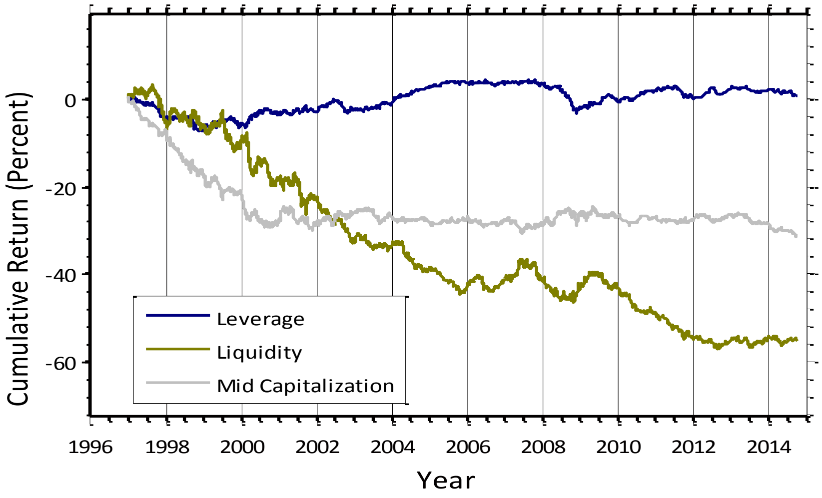


Figure 4.29: Cumulative returns of the Downside Beta, Oil Sensitivity, and Developed Markets Sensitivity factors

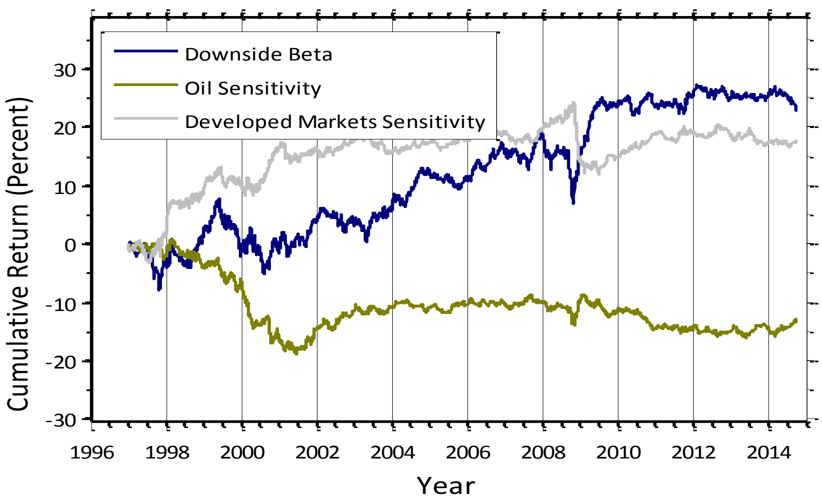


Figure 4.30: Cumulative returns of the Momentum, Industry Momentum, and Short-Term Reversal factors

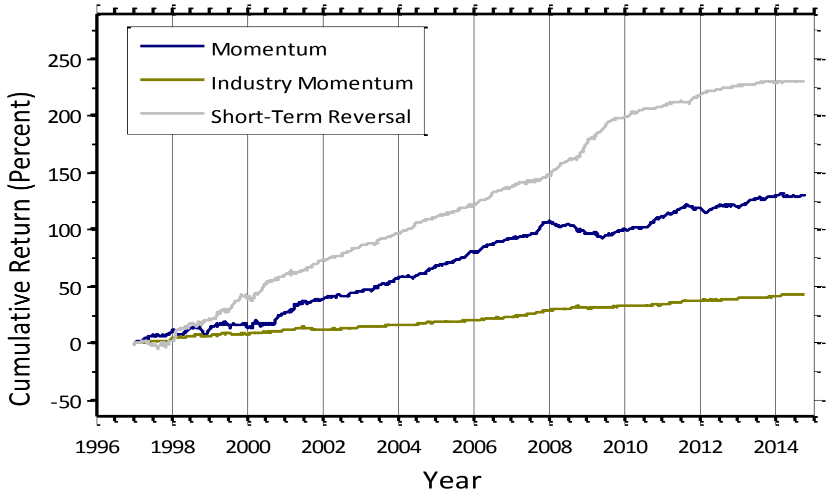


Figure 4.31: Cumulative returns of the Earnings Yield, Growth, and Dividend Yield factors

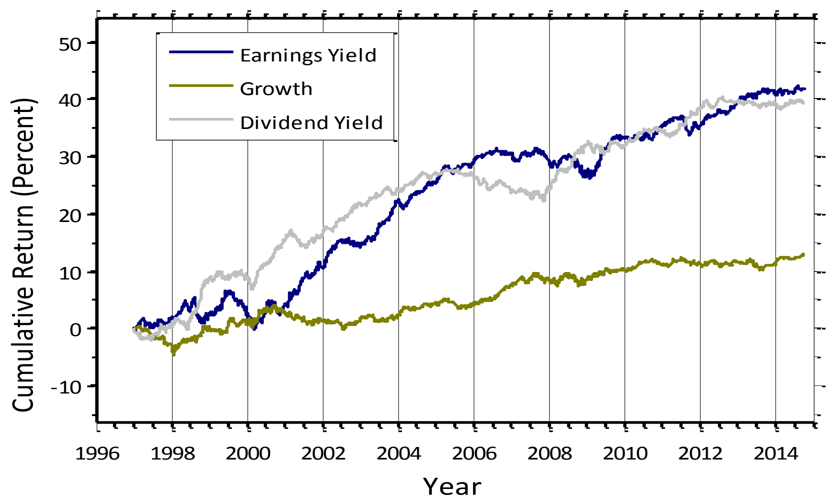


Figure 4.32: Cumulative returns of the Book-to-Price Ratio, Earnings Quality, and Seasonality factors

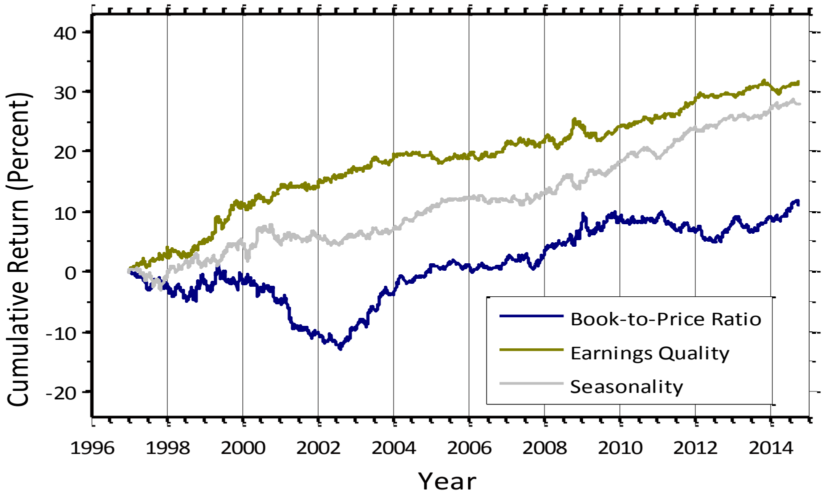
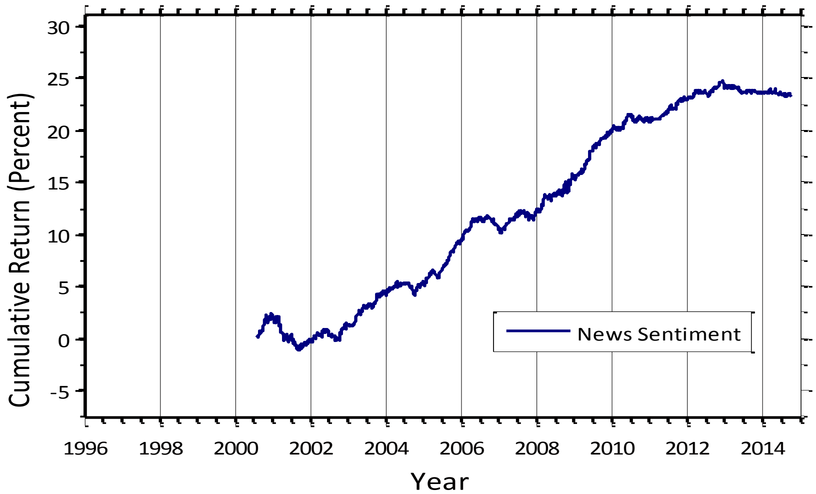


Figure 4.33: Cumulative return of the News Sentiment factor



4.3. Country and Industry Factors

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

In **Table 4.1**, we report mean absolute *t*-statistics for the CXE1 Market factor and industry factors, as well as the percentage of observations with $|t|>2$. We also report the annualized returns and volatilities, and the Information Ratios (IR) for the factors, and the correlations of the daily factor returns with the return of the estimation universe (ESTU).

Table 4.1: Market and industry factor summary statistics

Factor Name	Average Absolute t-stat	Percent Observ. $\|t\| > 2$	Annual Return	Annual Volatility	IR	Correl. With ESTU
Market	5.91	73	6.14	25.54	0.24	0.99
Energy	1.85	36	2.48	8.89	0.28	0.02
Chemicals	1.52	27	1.95	7.70	0.25	-0.02
Construction Materials and Packaging	1.45	26	-0.39	7.59	-0.05	-0.01
Metals and Mining ex Gold and Steel	2.07	41	-0.54	11.93	-0.05	0.08
Gold	2.41	44	2.60	27.70	0.09	-0.01
Steel	1.63	31	-1.32	10.36	-0.13	0.06
Building Products and Construction	1.32	22	-2.83	6.77	-0.42	-0.04
Capital Goods ex Building and Machinery	1.50	28	0.69	5.99	0.12	0.00
Machinery	1.52	28	0.54	8.58	0.06	-0.05
Commercial Services and Transportation	1.20	18	1.33	5.57	0.24	-0.04
Airlines and Marine	1.67	32	-2.20	9.59	-0.23	0.05
Automobiles and Components	1.72	34	5.81	8.64	0.67	-0.05
Consumer Durables and Apparel	1.48	28	-1.27	6.68	-0.19	-0.16
Consumer Services	1.60	29	2.26	7.58	0.30	-0.09
Media and Retailing	1.54	29	0.61	7.27	0.08	-0.07
Consumer Staples	1.37	24	3.43	6.16	0.56	-0.11
Health Care	1.43	26	8.05	8.05	1.00	-0.03
Banks	1.44	25	1.77	5.32	0.33	0.02
Diversified Financials	1.44	26	0.08	6.17	0.01	0.03
Insurance	1.55	29	2.17	8.47	0.26	0.05
Real Estate	1.74	34	-2.54	6.54	-0.39	0.11
Software and Services	2.04	41	15.11	16.12	0.94	0.03
Electronic Equipment and Components	2.05	42	0.71	10.32	0.07	-0.08
Computers and Peripherals	2.19	43	6.32	13.76	0.46	-0.10
Semiconductors and Semiconductor Equipment	2.74	54	6.98	15.53	0.45	-0.11
Telecommunication Services	1.75	35	2.77	8.50	0.33	-0.07
Utilities	1.52	29	3.11	6.82	0.46	-0.21

4.4. Style Factors

In **Table 4.2**, we report summary statistics for the style factors for the Barra China International Equity Model, along with the factor stability coefficient. The factor stability coefficient is computed as the cross-sectional correlation of factor exposures from one month-end to the next, using a square root of market capitalization weighting.

Table 4.2: Style factor summary statistics

Factor Name	Average Absolute t-stat	Percent Observ |t| > 2	Annual Return	Annual Volatility	IR	Correl. With ESTU	Stability Coeff
Beta	2.75	53	2.34	8.20	0.29	0.78	0.95
Book-to-Price Ratio	1.22	19	0.65	2.58	0.25	0.09	0.96
Developed Markets Sensitivity	1.44	26	1.01	2.61	0.38	-0.03	0.80
Dividend Yield	1.21	19	2.28	1.73	1.32	0.06	0.95
Downside Beta	1.45	28	1.31	3.47	0.38	0.15	0.83
Earnings Quality	1.19	18	1.81	1.84	0.98	-0.01	0.93
Earnings Yield	1.26	21	2.41	2.12	1.14	0.05	0.93
Growth	1.24	20	0.74	1.61	0.46	0.08	0.94
Industry Momentum	1.22	19	2.48	1.51	1.65	-0.01	0.53
Leverage	1.17	18	0.03	1.75	0.02	0.01	0.98
Liquidity	1.64	32	-3.17	3.85	-0.82	0.42	0.98
Mid Capitalization	1.29	22	-1.81	2.68	-0.68	-0.04	0.99
Momentum	1.59	31	7.46	4.18	1.78	-0.01	0.90
News Sentiment	1.18	17	1.67	1.24	1.34	0.02	0.32
Oil Sensitivity	1.34	23	-0.75	2.23	-0.34	0.03	0.80
Residual Volatility	1.56	30	-2.31	3.51	-0.66	0.28	0.94
Seasonality	1.25	20	1.60	1.94	0.83	0.01	-0.03
Short-Term Reversal	1.82	37	13.23	3.90	3.39	0.19	0.22
Size	1.87	37	-0.93	4.59	-0.20	0.28	0.99

4.5. Explanatory Power

As discussed in Section 2.2, and also in more details in Appendix A, we use the Cross-Market Shrinkage methodology to determine the final estimates for the factor returns for the China International Market. Two other choices for the factor return estimates were also considered:

- The factor returns given by the Total Market regression—these are the priors in the Cross-Market Shrinkage
- The factor returns given by the regression for the China International Market—these are the likelihood in the Cross-Market shrinkage

To compare the performance of the three choices for the factor return estimates in capturing the underlying drivers for the cross-sectional variance of excess returns, we use the leave-one-out cross-validation residual R-squared measure. During model selection, this is often used to measure the over-fitting or under-fitting risk of a model. One generally selects the model that minimizes this measure.

Specifically, the comparison is conducted as follows:

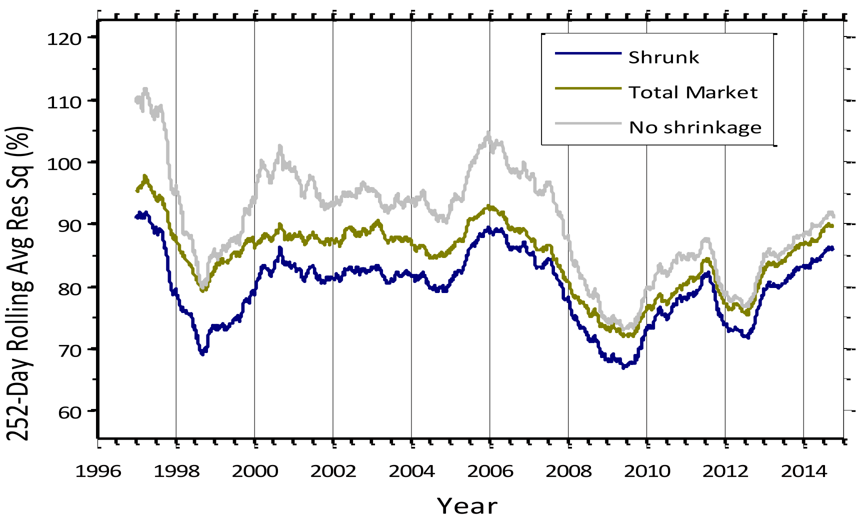
1. First, for each period, we randomly select 20 percent of the assets in the estimation universe for the China International Market.

2. Second, we drop each of the randomly selected assets from the estimation universe iteratively, and for each dropped asset, we re-generate the three sets of factor return estimates using the rest of the estimation universe, and then calculate the residual returns for the dropped asset using the three sets of factor return estimates. After iterating through the assets, we calculate the residual R-squared for the period:
$$R^2 = \frac{\sum_i w_i u_i^2}{\sum_i w_i r_i^2}$$
where r_i is the excess return, u_i is the residual return, and w_i is the regression weight of the *i*-th dropped asset. The square-root of market cap is used as the regression weight. We thus obtain three residual R^2 for the period, one for each of the three choices of the factor return estimates.

3. Third, we repeat the above two steps through the entire model history. We thus obtain three time series of residual R^2.

In **Figure 4.34**, we plot the 252-day rolling average of daily residual R^2 from December 31, 1996 to September 30, 2014. The factor return estimates given by the Cross-Market Shrinkage consistently results in a lower residual R^2 than the other two choices for the factor return estimates. Therefore, we decide to adopt this set of factor return estimates for the China International Equity Model.

Figure 4.34: 252-day rolling average of the leave-one-out cross-validation residual R-squared



5. Risk Forecasting Accuracy

In this section, we compare the risk forecasting accuracy of the Barra China International Equity Model and its predecessor, the Barra Hong Kong Equity Model (HKE1). Our methodology for evaluating and comparing the accuracy of risk model forecasts is based on the Q-statistic and bias statistic (see Appendix F). We use the Q-statistic to quantify the differences between models, and use the bias statistic to build intuition about the periods when a model under-forecasts or over-forecasts risk.

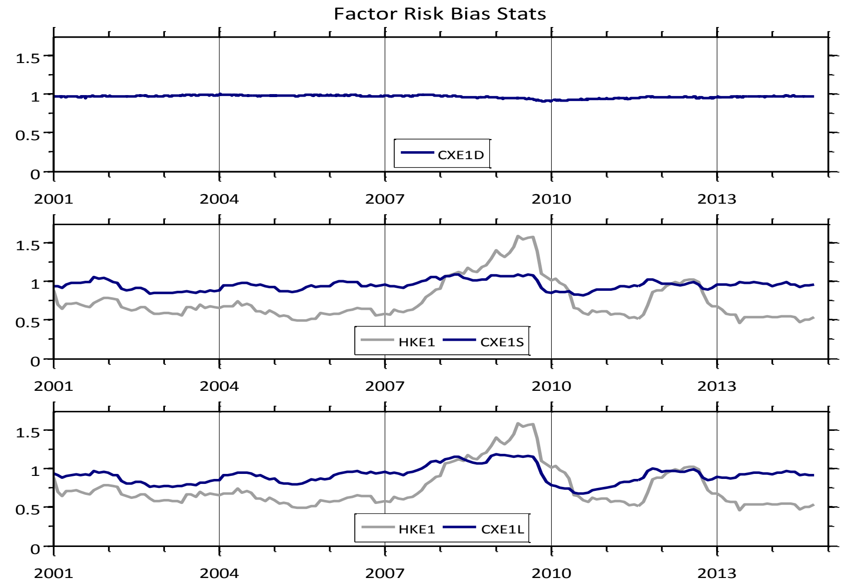
Conceptually, the bias statistic is an out-of-sample measure that represents the ratio of realized risk to predicted risk. The ideal bias statistic for perfect risk forecasts should be close to 1. By plotting the mean rolling-window bias statistic across time for a collection of portfolios**,** we can quickly visualize the magnitude of the average biases and judge if they are persistent or regime-dependent.

One potential shortcoming of the bias statistic is that over a long term, we may have sub-periods of over-forecasting and under-forecasting, yet obtain a bias statistic close to 1 over the entire window. In other words, forecasting errors may cancel out over the long term, even though the risk forecasts may be poor over sub-periods. For a portfolio manager who may be devastated by a single year of poor performance, it is small consolation knowing that a risk forecast is good *on average*. For this reason, we focus on the mean Q-statistic. The Q-statistic provides a measure of the forecast error and grows with the error size. The mean Q-statistic is not prone to “error cancellation” and is minimized by having the right forecast for every portfolio for every time period. This gives us a tool to measure the improvements between models on the same set of portfolios. The better model will have a lower mean Q-statistic.

5.1. Factors

In **Figure 5.1,** we plot the rolling-window bias statistics for the forecast factor volatilities by the predecessor model Barra Hong Kong Equity Model (HKE1) and the new Barra China International Equity Model (CXE1). We use rolling windows of 63 days, 12 months, and 12 months for the Daily (CXE1D), Short (CXE1S) and Long (CXE1L) horizon models, respectively. We do so because users of a daily model often examine the bias statistics for a shorter period, while users of a longer-horizon model tend to study the bias statistics for a much longer window.

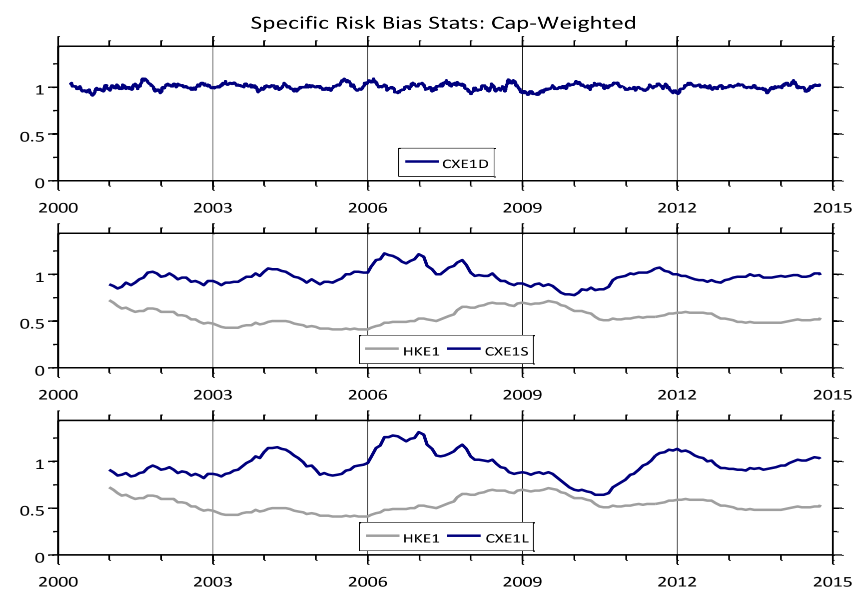
Figure 5.1: Rolling bias statistics for HKE1 and CXE1 pure factors



5.2. Specific Risk

In **Figure 5.2,** we plot the rolling-average cross-sectional bias statistics for the specific risk forecasts. The cross-sectional bias statistic is cap-weighted for all the assets that are in both the HKE1 and CXE1 estimation universes.

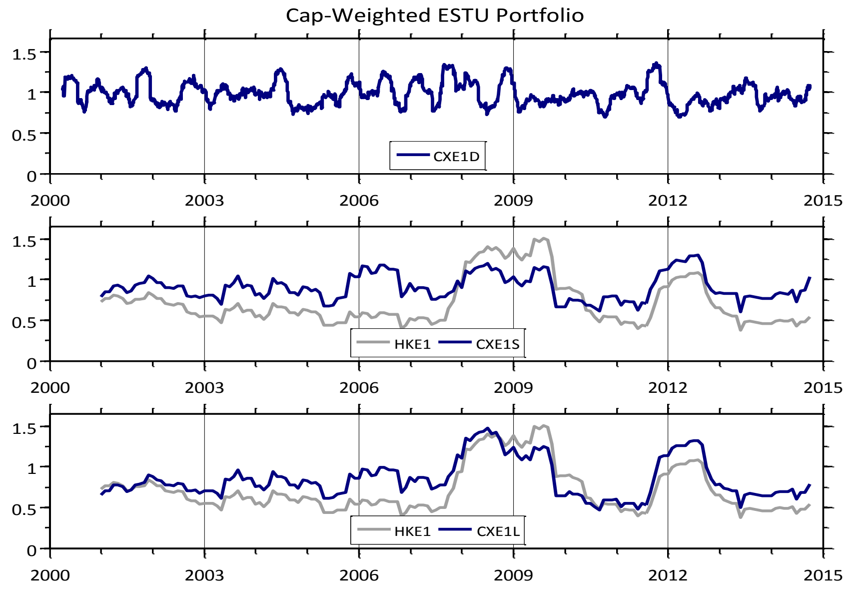
Figure 5.2: Rolling bias statistics for HKE1 and CXE1 specific risk



5.3. Estimation Universe Portfolio

In Figure 5.3, we plot the rolling-window bias statistics for the cap-weighted estimation universe portfolio. This portfolio contains the assets that are in both the HKE1 and CXE1 estimation universes.

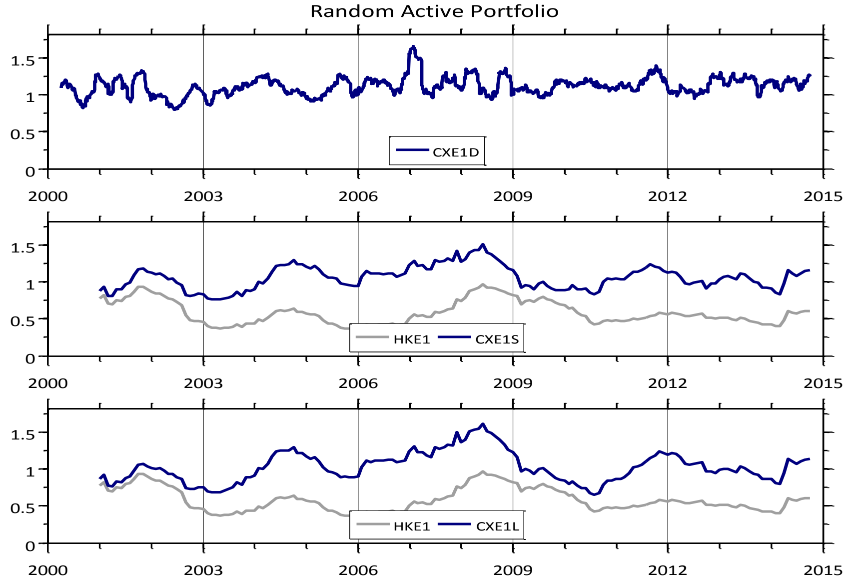
Figure 5.3: Rolling bias statistics for the cap-weighted estimation universe portfolio



5.4. Random Active Portfolios

In **Figure 5.4,** we plot the rolling-window bias statistics for the active risk for 100 random portfolios. The portfolios are constructed by going long 100 randomly selected stocks. The stocks are weighted by their market capitalization. The cap-weighted estimation universe portfolio in **Figure 5.3** is used as the benchmark. To reduce turnover, the list of stocks used to construct the random portfolios is fixed unless a stock drops out of the estimation universe, in which case it is replaced by another randomly selected estimation universe stock.

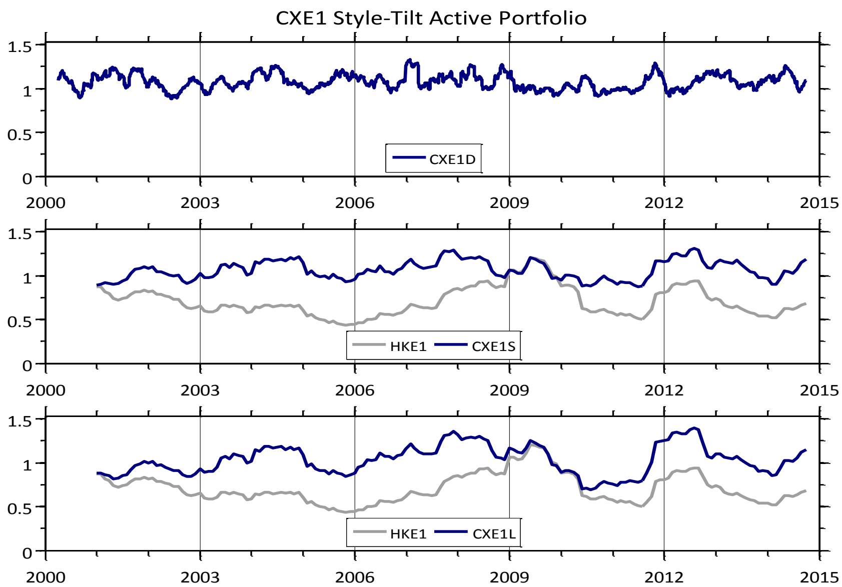
Figure 5.4: Rolling bias statistics for random active portfolios



5.5. Long-Only Style-Tilt Active Portfolios

In **Figure 5.5,** we plot rolling bias statistics for long-only style-tilt active portfolios. The portfolios are constructed by cap-weighting the assets in the top quintile of each of the style factors for the Barra China International Equity Model. The cap-weighted estimation universe is used as the benchmark. The rolling bias statistics are evaluated using the HKE1 and CXE1 models, respectively.

Figure 5.5: Rolling bias statistics for long-only CXE1 style-tilt active portfolios



5.6. Optimized Style-Tilt Portfolios

In **Figure 5.6,** we plot the rolling bias statistics for optimized style-tilt portfolios for the Barra China International Equity Model. These portfolios are constructed using the style factors as “alpha signals” and forming the unit-alpha, minimum risk portfolios. No other constraints are imposed in the portfolio construction. In **Figure 5.7,** we plot the rolling bias statistics for the optimized HKE1 style-tilt portfolios. Twenty random portfolios from Section 5.4 are used as the investment universes for constructing these optimized portfolios.

Figure 5.6: Rolling bias statistics for optimized CXE1 style-tilt portfolios

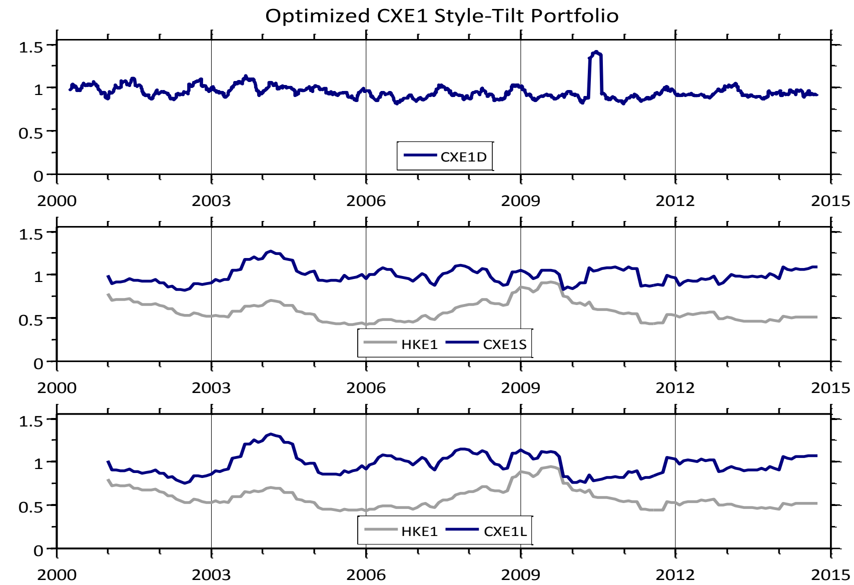
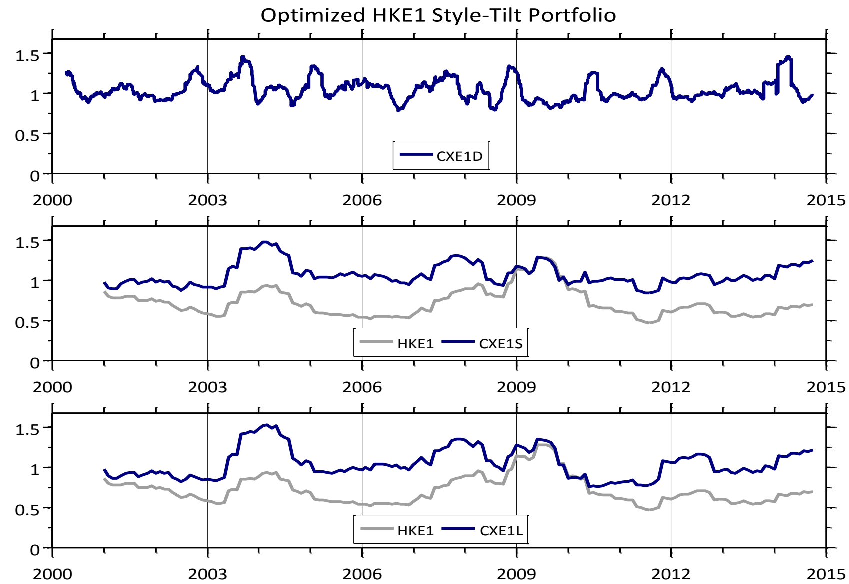


Figure 5.7: Rolling bias statistics for optimized HKE1 style-tilt portfolios



5.7. Analytical Minimum Risk Portfolios

In **Figure 5.8**, we plot the rolling bias statistics for the analytical minimum risk portfolios constructed using the CXE1S, CXE1L, and HKE1 models respectively. The investment universe is formed by the assets that are in both the HKE1 and CXE1 estimation universes. In **Figure 5.9**, we plot the rolling 12-month volatilities of these portfolios.

Figure 5.8: Rolling bias statistics for analytical minimum risk portfolios

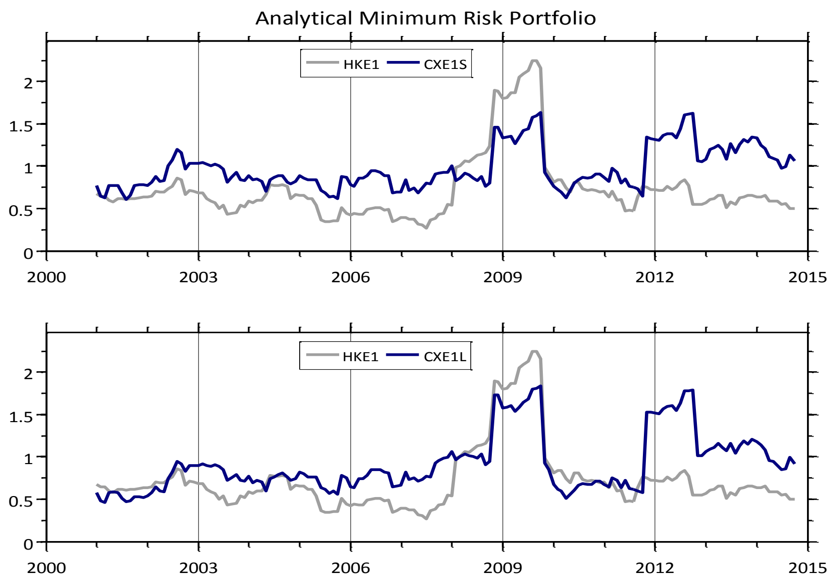
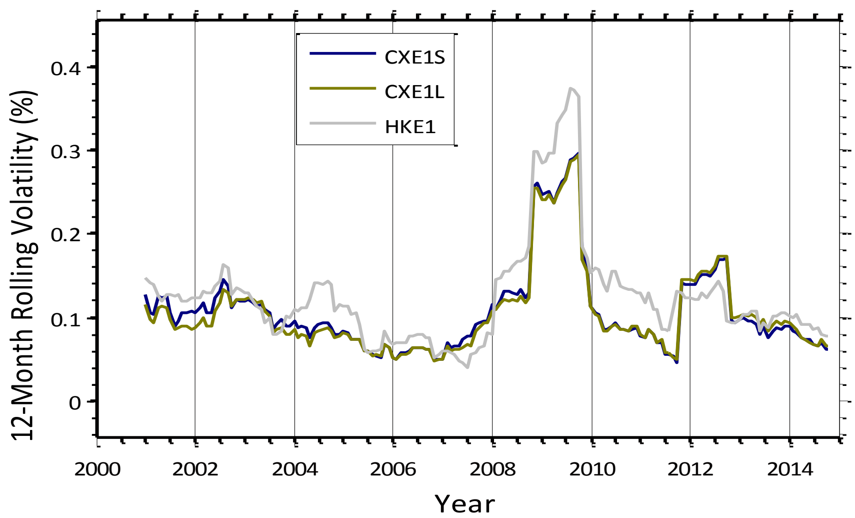


Figure 5.9: Rolling 12-month volatilities of analytical minimum risk portfolios



5.8. Summary of Risk Forecasting Bias Testing Results

The following tables summarize the bias and Q statistics for the test cases presented from Figures 5.1 to 5.8. As the tables illustrate, it is clear that the CXE1 Model provides more timely and accurate risk forecasts than the HKE1 Model does for the majority of the portfolios tested.

Table 5.1: Mean bias and Q-statistics for CXE1D

Portfolio Type	Figure	Bias Stat	Q
Pure Factor	5.1	0.96	2.3809
Specific Risk	5.2	1.00	2.5290
Estimation Universe	5.3	0.99	2.5058
Random Active	5.4	1.13	2.4574
Long-Only CXE1 Style-Tilt Active	5.5	1.08	2.4212
Optimized CXE1 Style-Tilt	5.6	0.92	2.4177
Optimized HKE1 Style-Tilt	5.7	1.01	2.4549

Table 5.2: Mean bias and Q-statistics for CXE1S and HKE1

Portfolio Type	Figure	CXE1S		HKE1		Q Diff (bps)
		Bias Stat	Q	Bias Stat	Q	
Pure Factor	5.1	1.00	2.3824	0.78	2.6798	2974
Specific Risk	5.2	0.97	2.4721	0.55	3.0005	5284
Estimation Universe	5.3	0.95	2.3527	0.75	2.5206	1679
Random Active	5.4	1.11	2.2857	0.61	2.6072	3215
Long-Only CXE1 Style-Tilt Active	5.5	1.11	2.3659	0.74	2.5299	1640
Optimized CXE1 Style-Tilt	5.6	1.04	2.4705	0.60	2.8237	3532
Optimized HKE1 Style-Tilt	5.7	1.11	2.5309	0.75	2.6770	1461
Analytical Minimum Risk	5.8	1.01	2.4326	0.76	2.6912	2586

Table 5.3: Mean bias and Q-statistics for CXE1L and HKE1

Portfolio Type	Figure	CXE1L		HKE1		Q Diff (bps)
		Bias Stat	Q	Bias Stat	Q	
Pure Factor	5.1	0.96	2.4163	0.78	2.6798	2635
Specific Risk	5.2	0.96	2.5043	0.55	3.0005	4962
Estimation Universe	5.3	0.88	2.3731	0.75	2.5206	1475
Random Active	5.4	1.08	2.3050	0.61	2.6072	3022
Long-Only CXE1 Style-Tilt Active	5.5	1.08	2.3928	0.74	2.5299	1371
Optimized CXE1 Style-Tilt	5.6	1.02	2.4831	0.61	2.8107	3276
Optimized HKE1 Style-Tilt	5.7	1.09	2.5566	0.75	2.6770	1204
Analytical Minimum Risk	5.8	0.96	2.6067	0.76	2.6912	845

6. Backtesting Results

In this section, we compare the Barra Hong Kong Equity Model (HKE1) and the Barra China International Equity Model (CXE1) on realistic portfolio construction cases. We focus on three types of long-only portfolios: minimum risk, benchmark tracking, and active strategy portfolios.

6.1. Minimum Risk Portfolio

In this set of tests, we construct long-only minimum risk portfolios from the investment universe of assets that are in both the HKE1 and CXE1 estimation universes. The portfolios are rebalanced monthly. The test period is January 2000 through September 2014. An upper bound of eight percent is set for monthly turnover.

In **Figure 6.1**, we plot the rolling 12-month volatilities of the minimum risk portfolios constructed using HKE1 and CXE1L, respectively. The upper subplot is for the portfolios with no constraint on forecast beta, whereas in the lower subplot we set a minimum forecast beta of 0.5. In **Figure 6.2**, we compare the rolling 12-month volatilities of the portfolios constructed using HKE1 and CXE1L, respectively.

Figure 6.1: Rolling 12-month volatilities of minimum risk portfolios constructed using HKE1 and CXE1S

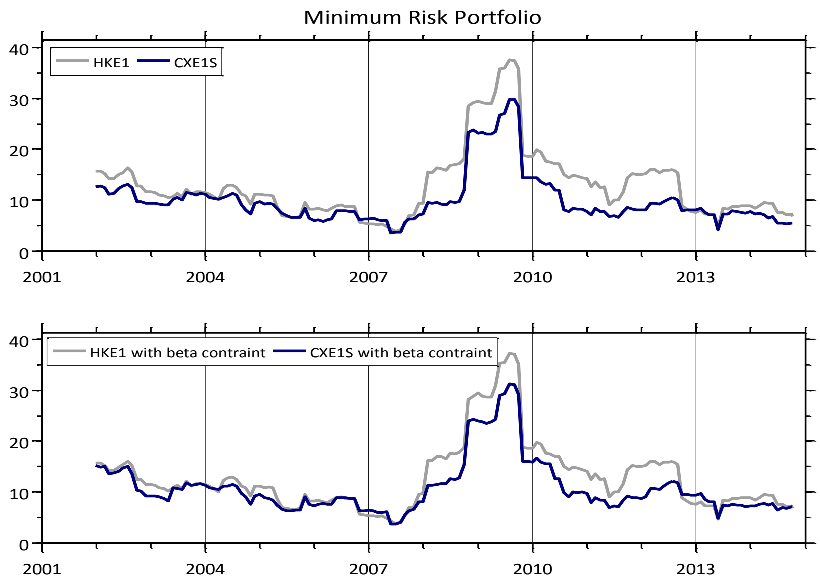
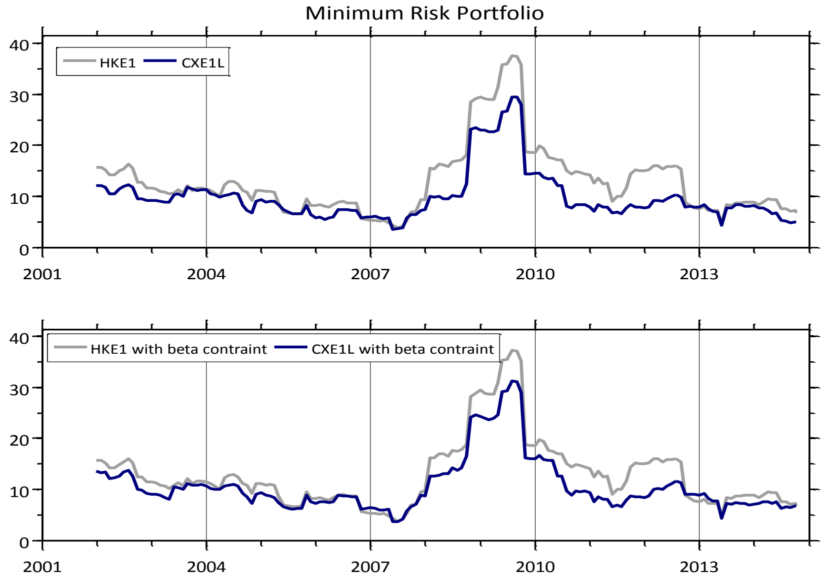


Figure 6.2: Rolling 12-month volatilities of minimum risk portfolios constructed using HKE1 and CXE1L



In **Table 6.1**, we summarize the realized volatilities, bias statistics, Q statistics, and annual turnovers for the minimum risk portfolios.

Table 6.1: Summary of minimum risk portfolios

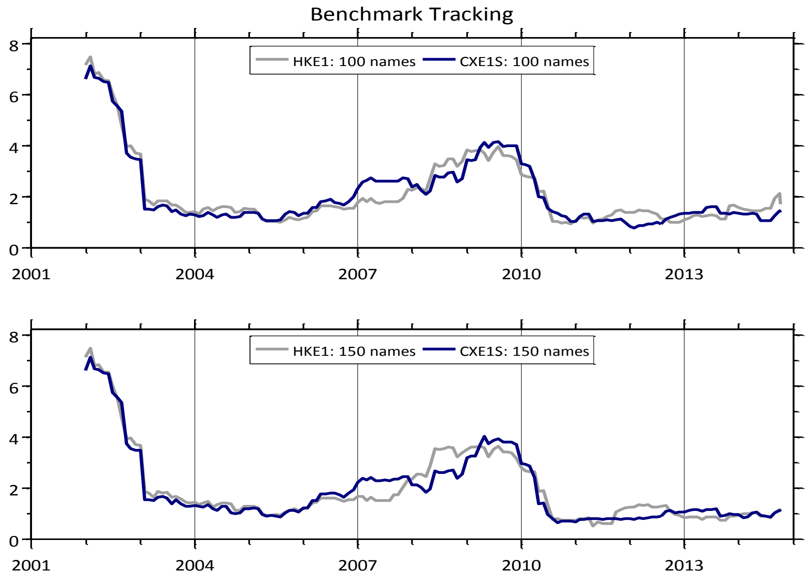
Beta Constraint?	Model	Realized Annual Volatility (%)	Bias Stat	Q	Annual Turnover (%)
No	HKE1	15.66	0.86	2.8598	96
	CXE1S	12.15	0.97	2.4694	96
	CXE1L	12.00	0.90	2.7256	96
Yes	HKE1	15.63	0.85	2.8359	96
	CXE1S	13.05	0.98	2.5287	96
	CXE1L	12.90	0.90	2.6765	96

6.2. Benchmark Tracking

In this set of tests, the market cap-weighted portfolio of assets that are in both the HKE1 and CXE1 estimation universes serves as both the benchmark portfolio and the investment universe. Long-only portfolios of 100 and 150 names are constructed to track the benchmark portfolio. The portfolios are rebalanced monthly. A monthly turnover upper bound of eight percent is used. The test period is from January 2000 to September 2014.

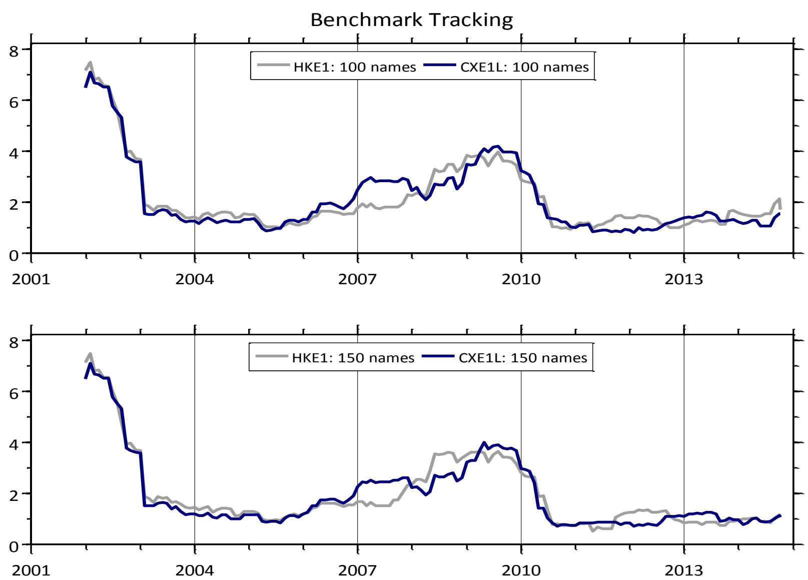
In **Figure 6.3**, we plot the rolling 12-month volatilities of the benchmark tracking portfolios constructed using HKE1 and CXE1S, respectively. The upper subplot is for the portfolios that have 100 names, and the lower subplot is for the portfolios that have 150 names.

Figure 6.3: Rolling 12-month volatilities of benchmark tracking portfolios constructed using HKE1 and CXE1S



In **Figure 6.4**, we compare the rolling 12-month volatilities of the portfolios constructed using HKE1 and CXE1L, respectively.

Figure 6.4: Rolling 12-month volatilities of benchmark tracking portfolios constructed using HKE1 and CXE1L



In **Table 6.2**, we summarize the realized tracking errors, bias statistics, Q statistics, and annual turnovers for the benchmark tracking portfolios.

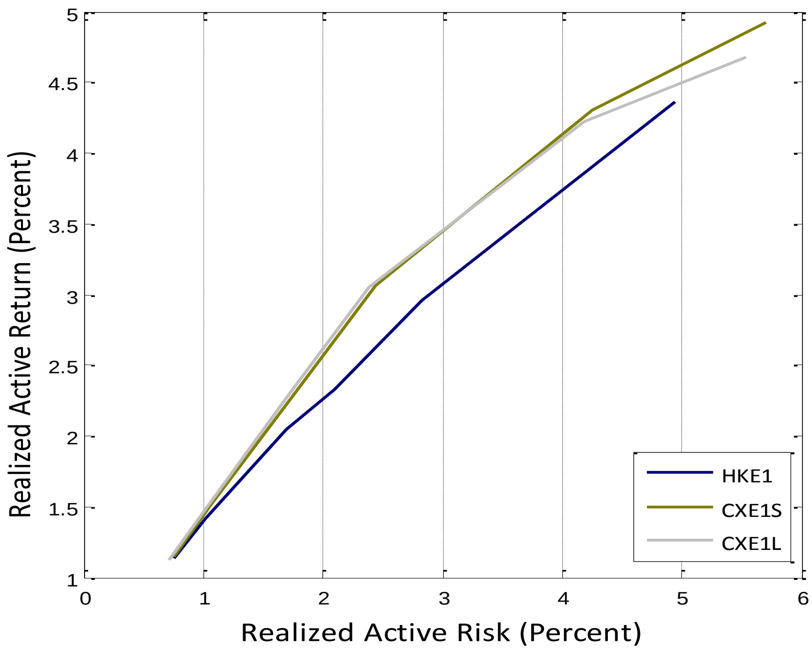
Table 6.2: Summary of benchmark tracking portfolios

Model	# of Names	Tracking Error (%)	Bias Stat	Q	Annual Turnover (%)
HKE1	100	2.92	0.71	2.4029	55
CXE1S	100	2.77	1.18	2.1301	67
CXE1L	100	2.76	1.14	2.1477	68
HKE1	150	2.82	0.66	2.6676	48
CXE1S	150	2.64	1.11	2.1441	62
CXE1L	150	2.64	1.11	2.3284	61

6.3. Active Strategy

In this set of tests, the market cap-weighted portfolio of assets that are in both the HKE1 and CXE1 estimation universes serves as both the benchmark portfolio and the investment universe. The alpha signal is constructed by equally combining the Book-to-Price, Earnings Quality, Earnings Yield, Growth, and Momentum factor exposures. Long-only active-strategy portfolios are constructed based on this alpha signal. The portfolios are rebalanced monthly, and a monthly turnover upper bound of eight percent is used. The risk aversion parameter is varied to construct the efficient frontier. The test period is from January 2000 to September 2014.In **Figure 6.5,** we plot the realized performance, before transaction costs, of the portfolios constructed using HKE1, CXE1S, and CXE1L, respectively. The horizontal axis is the realized annual active risk and the vertical axis the realized annual active return.

Figure 6.5: Realized performance, before transaction costs, of active portfolios



7. Conclusion

The Barra China International Equity Model incorporates innovations designed to address long-standing problems in risk modeling and portfolio construction. For instance, the Volatility Regime Adjustment calibrates volatilities to current market levels and represents a key determinant of risk forecasts, especially during times of market turmoil. The Optimization Bias Adjustment addresses the issue of underestimation of risk for optimized portfolios, and leads to a better conditioned covariance matrix. The introduction of the Market factor leads to more intuitive attribution of portfolio risk and return, while also providing more timely forecasts of industry correlations. Another enhancement is the use of a Bayesian adjustment technique in the specific risk model, which is designed to reduce biases in specific risk forecasts.

The Barra China International Equity Model incorporates the new Systematic Equity Strategy risk factors, which refer to the systematic (i.e., rules-based or computer-based) implementation of fundamental or technical investment anomalies or strategies.

The Barra China International Equity Model employs a Cross-Market Shrinkage methodology to estimate the factor returns for the China International Market. By leveraging additional information from ten other local markets in the Asia-Pacific region, this approach allows us to adopt a rich factor structure, while keeping the estimation errors in the factor returns under control.

In this Model Insight, we provided a thorough empirical analysis of the Barra China International Equity Model. We presented the factor structure in transparent detail, for both industries and styles. We reported key metrics at the individual factor level, including statistical significance, performance, volatility, and correlation with the estimation universe. In addition, we decomposed cross-sectional dispersion of asset returns into contributions from factors and stock-specific components, and further decomposed the factor contribution into Market, industry, and style components.

Over a test period from December 31, 1996 to September 30, 2014, we systematically compared the forecasting accuracy of the Barra China International Equity Model with its predecessor, the Barra Hong Kong Equity Model (HKE1). We considered several types of portfolios, including pure factors, random active portfolios, factor-tilt portfolios (both long-only and dollar-neutral), and factor-tilt optimized portfolios. We demonstrated the accuracy of specific risk forecasts of the Barra China International Equity Model compared to its predecessor.

Lastly, we showed the effectiveness of this model in portfolio construction using minimum risk and benchmark tracking portfolios, and active strategy portfolios composed of factors based on Systematic Equity Strategies.

Appendix A: Cross-Market Shrinkage Methodology for Factor Returns

The cross-market Bayesian Shrinkage of factor returns works as follows:

1. First, we define a Total Market as a combination of multiple local markets in the Asia Pacific region. The countries in the Total Market form the core of the estimation universe in the Barra Asia Pacific Equity Model (ASE2). We retain the same estimation universe, factor structure and raw factor exposures as the Barra Asia Pacific Equity Model but we re-standardize the style factor exposures within each local market to market capitalization-weighted mean of zero, and equal-weighted standard deviation of one.

2. Second, we estimate the factor returns for the Total Market from regression, using square root of USD market capitalization as the regression weight. Two constraints are posted to remove the exact two-fold multi-collinearity between the country and the Market factor exposures, and between the industry and the Market factor exposures. The first constraint requires the local market weighted sum of country factor returns to be zero. The second constraint is more subtle—it requires the weighted sum of industry factor returns to be zero, where the market capitalization of the industries within the China International Market, instead of within the entire Total Market, are used as the weights. This is important because we want to look at the industries through the same lens as for the China International Market. The obtained industry factor returns, together with the style factor returns, can serve as the Bayesian priors for the factor returns in the China International Market. Equation (A1) illustrates the regression of the Total Market where f_{TM} is the resulting vector of factor return.

$$\begin{aligned}&\min_{f_{TM}}\|r_{TM}-X_{TM}f_{TM}\|_2^2\\&subject\ to:\\&China\ International\ cap\ weighted\ industry\ return=0\\&Local\ market\ cap\ weighted\ country\ return=0\end{aligned} \tag{A1}$$

3. Third, we estimate the factor returns for the 11 individual local markets using their respective local estimation universes. For each regression, only one constraint is needed to remove the exact multi-collinearity between the industry and the Market factor exposures. Similar to the Total Market regression, we constrain the weighted sum of the industry factor returns to be zero, with the weights being the market capitalization of the industries within the China International Market. As a result, the industry factor returns obtained from these regressions can be compared on the same footing. For each industry or style factor, we thus have a cross market of 11 estimates for the factor return. Equation (A2) illustrates the regression of the each local market m, and f_m is the resulting vector of factor return. The $i-th$ factor return is thus denoted as $f_{m,i}$, and its estimation error as $\sigma_{m,i}^2$.Note m can be any one of the 11 local markets such as China International/Hong Kong, Taiwan, or India.

$$\begin{aligned}&\min_{f_m}\|r_m-X_mf_m\|_2^2\\&subject\ to:\\&China\ International\ cap\ weighted\ industry\ return=0\end{aligned} \tag{A2}$$

4. At the end, we use the Bayesian Shrinkage methodology to determine the final factor return estimates for the China International Market. For each industry or style factor i, the factor return estimate $f_{CXE1,i}$ from the regression for the China International Market in the third step

above is our likelihood. Its associated estimation error $\sigma_{i,CXE1}^2$ is compared to the cross-market dispersion $d_i$of the 11 estimates to determine the relative weights w_i for the likelihood and $1 - w_i$ for the prior. A low estimation error for the likelihood suggests a high confidence level, thus a higher weight for the likelihood. A low cross-sectional dispersion suggests that the factor return behaves consistently across different markets, thus a higher weight should be given to the prior. Conversely, a high estimation error for the likelihood or a high cross-sectional dispersion suggests the opposite. The final estimate for the factor return is the weighted mean of the likelihood and the prior.

$$d_i = max\left(0, var\left(f_{CXE1,i}, f_{TWE2,i}, f_{INE2,i}, \ldots\right) - mean\left(\sigma_{CXE1,i}^2, \sigma_{TWE2,i}^2, \sigma_{INE2,i}^2, \ldots\right)\right)$$

$$w_i = \frac{\sigma_{CXE1,i}^2}{\sigma_{CXE1,i}^2 + d_i}$$

The final estimate for the Market factor return for the China International Market is not shrunk, because the concept of cross-sectional dispersion does not apply for this factor return --- the regressions for the other ten local markets do not provide an estimate for it.

Appendix B: Volatility Regime Adjustment

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

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

$$B_t^F = \sqrt{\frac{1}{K}\sum_k \left(\frac{f_{kt}}{\sigma_{kt}}\right)^2} \tag{B1}$$

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

$$\lambda_F = \sqrt{\sum_t (B_t^F)^2 \, w_t} \tag{B2}$$

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

$$\tilde{\sigma}_k = \lambda_F \sigma_k \tag{B3}$$

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

Appendix C: Optimization Bias Adjustment

Let F_0 denote the KxK sample factor correlation matrix (FCM),

$$F_0 = cor(f, f) \quad \text{(C1)}$$

where f is the KxT matrix of realized factor returns, K is the number of factors and T is the number of periods. More detail on how to estimate F_0 is provided in Appendix H.

The sample FCM can be expressed in diagonal form as:

$$D_0 = U_0' F_0 U_0 \quad \text{(C2)}$$

where U_0 is the KxK rotation matrix whose columns are given by the eigenvectors ofF_0. The jth element of the kth column of U_0 gives the weight of pure factor j in eigenfactor k . The predicted eigenvalues of the eigenfactors are given by the diagonal elements of D_0. The fact that D_0 is diagonal indicates that the eigenfactors are mutually uncorrelated.

Although the true FCM is unobservable, we suppose for simulation purposes that the sample FCM F_0 governs the "true" return-generating process. We generate a set of factor returns for simulation m as:

$$f_m = U_0 b_m \quad \text{(C3)}$$

where b_m is a KxT matrix of simulated eigenfactor returns. The elements of row k of b_m are drawn from a random normal distribution with mean zero and eigenvalues given by the diagonal element $D_0(k)$ of matrix D_0. It can be easily verified that the simulated returns in Equation (C3) have a true FCM given by F_0. Due to sampling error, however, the *estimated* FCM:

$$F_m = cor(f_m, f_m) \quad \text{(C4)}$$

will differ from the true FCM F_0 . Nevertheless, F_m is unbiased in the sense that $\mathrm{E}[F_m] = F_0$. We diagonalize the simulated FCM:

$$D_m = U_m' F_m U_m \quad \text{(C5)}$$

where U_m denotes the simulated eigenfactors with estimated eignevalues given by the diagonal elements of D_m, i.e. $D_m(k)$.

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

$$\widetilde{D}_m = U_m' F_m U_m \quad \text{(C6)}$$

Note that since U_m is not composed of the "true" eigenfactors, the matrix $\widetilde{D}_m$ is not diagonal. Nevertheless, our current focus is on the diagonal elements of the matrix. We compute the *simulated* eigenvalue biases according to

$$v^2(k) = \frac{1}{M}\sum_m \frac{\widetilde{D}_m(k)}{D_m(k)} \quad \text{(C7)}$$

where M is the total number of simulations. The simulated eigenvalue bias is computed daily, and the average over the entire sample period.

We now assume that the sample FCM F_0, which uses the same correlation estimator as the simulated FCMF_m, also suffers from the same biases. Let $\widetilde{D}_0$ denote the diagonal FCM whose eigenvalues have been adjusted:

$$\widetilde{D}_0 = v^2 D_0 \quad \text{(C8)}$$

where v^2 is a diagonal matrix whose elements are given by $v^2(k)$. The FCM in Equation (C8) is now rotated from the diagonal basis to the pure factor basis using the sample eigenfactors. That is,

$$\tilde{F}_0 = U_0 \widetilde{D}_0 U_0' \quad \text{(C9)}$$

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

For further details, please refer to Menchero, Wang, and Orr (2011).

Appendix D: Specific Risk Bayesian Shrinkage

One potential problem with using a pure time-series approach is that specific volatilities may not fully persist out-of-sample. In particular, stocks with either extremely low or extremely high specific volatility forecasts tend to revert to the mean.

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

$$\sigma_n^{SH} = v_n \bar{\sigma}(s_n) + (1 - v_n)\,\hat{\sigma}_n \tag{D1}$$

where $\hat{\sigma}_n$ is the original forecast and v_n is the shrinkage intensity that determines the weight given to the Bayesian prior, also known as the shrinkage target,

$$\bar{\sigma}(s_n) = \sum_{n \in S_n} w_n \hat{\sigma}_n \tag{D2}$$

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

$$v_n = \frac{q|\hat{\sigma}_n - \bar{\sigma}(s_n)|}{\Delta_\sigma(s_n) + q|\hat{\sigma}_n - \bar{\sigma}(s_n)|} \tag{D3}$$

where q is an empirically determined shrinkage parameter and,

$$\Delta_\sigma(s_n) = \sqrt{\frac{1}{N(s_n)} \sum_{n \in S_n} (\hat{\sigma}_n - \bar{\sigma}(s_n))^2} \tag{D4}$$

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

Appendix E: Descriptors by Style Factor

This appendix defines the descriptors in the style factors. The descriptors are listed under the style factors to which they belong.

Style: **Beta**

Components: HBETA

Historical Beta

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

$$xr_t - r_{ft} = \alpha + \beta R_t + e_t \quad \text{(E1)}$$

The returns are aggregated over two-day windows to reduce the effects of non-synchronicity. The regression coefficients are estimated over the trailing 252 trading days of returns with a half-life of 63 trading days. The descriptor is then lagged by one trading day. This approach is robust for liquid stocks with sufficiently long history but tends to underestimate the Beta of illiquid assets that trade rarely and is not practical to estimate the Beta of recent IPOs. To circumvent those difficulties, Beta priors are introduced to blend the raw Beta estimates of recent IPOs and illiquid assets.

Style: **Book-to-Price**

Components: BTOP

Book-to-price

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

Style: **Developed Markets Sensitivity**

Components: DM_SENS

Developed markets sensitivity

This descriptor is designed to capture sensitivity of Asian market securities to changes in developed market performance. Computed as the sensitivity of asset residual return (in historical beta regression) to MSCI WORLD index return. The estimation is over the trailing window of 252 trading days with a half-life of 63 days and the lag of 1 day. Descriptor is orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Mid Capitalization, and Downside Beta.

Style: **Dividend Yield**

Components: YIELD

Dividend yield

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

Style: **Downside Beta**

Components: DSBETA

Downside historical beta

This descriptor is designed to capture sensitivity to negative market performance. Computed as the sensitivity of stock returns to the market return, similar to the Beta descriptor, but considers only those days when the two-day cumulative market performance is negative. The descriptor is then lagged by one trading day and orthogonalized to Beta to reduce collinearity.

Style: **Earnings Quality**

Components: CETOE

Cash earnings-to-earnings

This descriptor captures persistence of cash component of earnings of Asian market stocks. Computed as the difference between the cash-earnings-to-price and earnings-to-price ratios.

ABS

Accruals calculated using balance sheet data

This descriptor represents the measure of accounting accruals and computed as follows:

$$\left(\left(\Delta CA-\Delta Cash\right)-\left(\Delta CL-\Delta STD\right)+Dep\right)/TA \quad \text{(E2)}$$

where,

ΔCA – annual change in current assets

$\Delta Cash$ – annual change in cash and cash equivalents

ΔCL – change in current liabilities

ΔSTD – change in short-term debt

Dep – annual depreciation and amortization expense

TA – average total assets

Style: **Earnings Yield**

Components: EPIBS

Analyst-predicted earnings-to-price

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

CETOP

Cash earnings-to-price

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

ETOP

Earnings-to-price

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

Style: **Growth**

Components:

EGIBS — Analyst-predicted earnings growth
Long-term (3-5 years) earnings growth forecasted by analysts.

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

SGRO — Sales growth
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.

Style: **Industry Momentum**

Components:

INDMOM — Industry momentum
This descriptor measures industry relative strength. It is computed as the weighted average 126 days cumulative log return of all peer stocks in the same GICS® sub-industry with a half-life of 20 days. The descriptor is then computed as the equal-weighted average of the non-lagged values over the last 3-5 trading days and orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Mid Capitalization, and Downside Beta.

Style: **Leverage**

Components:

MLEV — Market leverage
Computed as:

$$MLEV = \frac{ME + PE + LD}{ME} \qquad \text{(E3)}$$

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

DTOA — Debt-to-assets
Computed as:

$$DTOA = \frac{TD}{TA} \qquad \text{(E4)}$$

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

BLEV

Book leverage

Computed as:

$$BLEV = \frac{BE + PE + LD}{BE} \quad \text{(E5)}$$

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

Style: **Liquidity**

Components: STOA

Annual share turnover

Let $STOM_\tau$ be the share turnover for month τ. The annual share turnover is defined by:

$$STOA = \ln\left[\frac{1}{T}\sum_{\tau=1}^{T}\exp\left(STOM_\tau\right)\right] \quad \text{(E6)}$$

where, $T = 12$ months.

STOM

Monthly share turnover

The monthly share turnover is computed as the log of the share turnover over the previous month:

$$STOM = \ln\left(\frac{V}{S}\right) \quad \text{(E7)}$$

here, V is the trading volume for the month, and S is the number of shares outstanding.

STOQ

Quarterly share turnover

Let $STOM_\tau$ be the share turnover for month τ. The quarterly share turnover is defined by:

$$STOQ = \ln\left[\frac{1}{T}\sum_{\tau=1}^{T}\exp\left(STOM_\tau\right)\right] \quad \text{(E8)}$$

where, $T = 3$ months.

Style: **Momentum**

Components: RSTR

Relative strength

First, non-lagged relative strength for day τ is computed as the sum of excess to the market log returns over the trailing $T = 252$ trading days:

$$RS(\tau) = \sum_{t=\tau-T+1}^{\tau} w_t[\ln(1 + r_t) - \ln(1 + r_{mt})] \quad \text{(E9)}$$

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

Relative strength is computed as the equal-weighted average of non-lagged relative strength over the previous 11 to 21 trading days:

$$RSTR = \frac{1}{11}\sum_{\tau=11}^{21} RS(\tau) \tag{E10}$$

This treatment introduces a lag while avoiding undue jumps in relative strength exposures when large returns enter the estimation window.

HALPHA

Historical alpha

First, non-lagged values of historical alpha are computed by the time-series regression of Equation E1. Historical alpha is then computed as the equal-weighted average of non-lagged values over the previous 11 to 21 trading days. This treatment introduces a lag while avoiding undue jumps in historical alpha exposures when large returns enter the estimation window.

Style: **Mid Capitalization**

Components: MIDCAP[2]

Cube of the logarithm of market capitalization

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

Style: **News Sentiment**

Components: CSS and ESS

Composite sentiment score and Event sentiment score

This descriptor captures recent uptick of positive news with respect to the trend, computed as equally weighted sum of 2 descriptors. Both descriptors are computed as follows:

$$\frac{S_T - L_T}{\sigma_L} \tag{E11}$$

where, S_T is computed as exponentially weighted average of 42 trading days of positive news counts with 126 days half-life. L_T represents news trend and computed similarly as exponentially weighted average with a longer window equal to 126 trading days of positive news counts with 126 days half-life. σ_L is standard deviation computed over a period of 126 trading days with 126 days half-life.

[2] Also referred to as SIZEN

The descriptor is then lagged by one trading day and orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Mid Capitalization and Downside Beta.

Style: **Oil Sensitivity**

Components: OILSEN

Oil sensitivity

This descriptor captures sensitivity of Asian market securities to changes in oil prices. Computed as the sensitivity of asset residual return (in historical beta regression) to changes in crude oil price. The estimation is over the trailing window of 252 trading days with a half-life of 63 days and a lag of 1 day. Descriptor is othogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Mid Capitalization and Downside Beta.

Style: **Residual Volatility**

Components: DASTD

Daily standard deviation

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

CMRA

Cumulative range

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

$$Z(T)=\sum_{\tau=1}^{T}\left[\ln(1+r_\tau)-\ln(1+r_{f\tau})\right] \qquad \text{(E12)}$$

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

$$CMRA = Z_{max} - Z_{\min} \qquad \text{(E13)}$$

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

HSIGMA

Historical sigma

Computed as the volatility of residual returns in Equation E1:

$$\sigma=\mathrm{std}\left(e_t\right) \qquad \text{(E14)}$$

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

Note: All the descriptors of Residual Volatility are lagged by one trading day. The Residual Volatility factor is orthogonalized to Beta, Size and Liquidity to reduce collinearity.

Style: **Seasonality**
Components: SEASON

Seasonality
This descriptor captures historical seasonal variation and time-of-the-year effects in stock performance. Computed as the sum of the same-month returns over the past five years. The descriptor is then lagged by one trading day and orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Mid Capitalization, and Downside Beta.

Style: **Short-Term Reversal**
Components: STREV

Short term reversal
This descriptor is computed in the same way as relative strength descriptor (RSTR), according to (E9), with the window of $T = 63$ trading days and half-life of 10 days, multiplied by -1. Equally-weighted average of non-lagged $RS(\tau)$ values is computed over the last 1-3 trading days and orthogonalized to the fundamental style factors and Momentum, Residual Volatility, Beta, Size, Mid Capitalization and Downside Beta.

Style: **Size**
Components: LNCAP

Logarithm of market capitalization
Given by the logarithm of the total market capitalization of the firm with a lag of one trading day.

Appendix F: Review of Bias Statistics

F.1. Single-Window Bias Statistics

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

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

$$b_{nt} = \frac{R_{nt}}{\sigma_{nt}}, \qquad \text{(F1)}$$

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

$$B_n = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(b_{nt} - \bar{b}_n\right)^2}, \qquad \text{(F2)}$$

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,

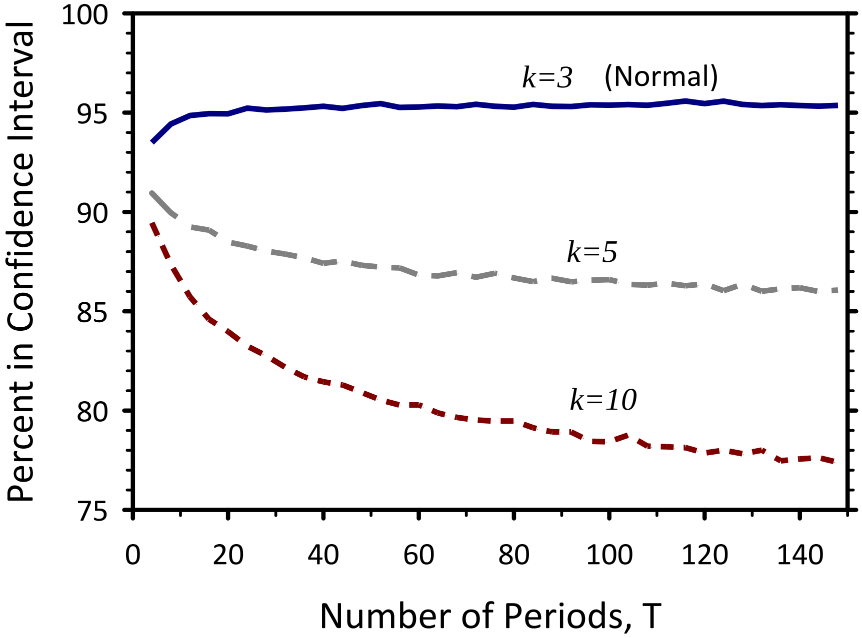
$$B_n \in \left[1-\sqrt{2/T}, 1+\sqrt{2/T}\right]. \qquad \text{(F3)}$$

If B_n falls outside this interval, we reject the null hypothesis that the risk forecast is 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 F.1**, we show simulated results for the percentage of observations actually falling within this interval plotted versus observation window length T, for several values of kurtosis k.

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

Figure F.1: Percent of observations falling within the confidence interval $1 \pm \sqrt{2/T}$, where T is the number of periods in the observation window. Results were simulated using a normal distribution $k = 3$, and using a *t*-distribution with kurtosis values $k = 5$ and $k = 10$. The standard deviations were equal to 1 in all cases. For the normal distribution, the percentage of observations inside the confidence interval quickly approaches 95 percent. As kurtosis is increased, however, the proportion within the confidence interval declines considerably.



F.2. Rolling-Window Bias Statistics

The purpose of bias-statistic testing is to assess the accuracy of risk forecasts, typically over a long sample period. One possibility is to select the entire sample period as a single window and to compute the bias statistic as in Equation F2. This would be a good approach if financial data were stationary, as sampling error is reduced by increasing the length of the window. In reality, however, financial data are not stationary. It is possible to significantly overpredict risk for some years, and underpredict it for others, while ending up with a bias statistic close to 1.

Often, a more relevant question is to study the accuracy of risk forecasts over a window of k observations. For this purpose, we define the rolling window bias statistic for portfolio n,

$$B_n^\tau = \sqrt{\frac{1}{k}\sum_{t=\tau-k+1}^{\tau}\left(b_{nt} - \bar{b}_n\right)^2}, \tag{F4}$$

where τ denotes the last observation of the window. The windows are rolled forward one observation at a time until reaching the end of the sample period. If T is the number of observations in the sample period, then each portfolio will have $T - k + 1$ (overlapping) k-observation windows.

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

$$\bar{B}^\tau = \frac{1}{N}\sum_n B_n^\tau \tag{F5}$$

We also define $B^{\tau}(5\%)$ and $B^{\tau}(95\%)$ to be the 5-percentile and 95-percentile values for the rolling window bias statistics at a given point in time.

F.3. Q-Statistic

The Q-Statistic is defined as $Q_{nt} = b_{nt}^2 - \ln b_{nt}^2$, where b_{nt} is a standardized return introduced in (F1). The Q-statistic penalizes both under and over forecast and is not prone to "error cancellation" when averaged across time and/or test portfolios. For averaging, we define the mean of Q-statistic as follows:

$$\overline{Q} = \sum_{n=1}^{N} \sum_{t=1}^{T} Q_{nt} \tag{F6}$$

where N is a number of portfolios and T is a sample size. Further information on Q-Statistic can be found in Patton (2011).

Appendix G. Covariance Matrix Estimation

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

For the Long and Short versions of the model, the prediction horizon is one month. The factor correlation matrix, however, is estimated from daily factor returns. We must therefore account for the possibility of serial correlation in factor returns, as these may affect risk forecasts over a longer horizon. In the case of the Barra Asia Pacific Equity model, serial correlation corrections are also applied to the Daily horizon over a two-day period to account for the asynchronous nature of currency and equity factor returns in the model.

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

With the correlation matrix computed, the next step is to calculate the factor volatilities. We use exponentially weighted averages, with half-life parameter τ_{σ}^{F}. In estimating monthly factor volatilities, we also scale daily volatility by a ratio of volatility estimates calculated using overlapping monthly returns and daily returns.

Next, we construct the initial covariance matrix by combining the factor volatilities and correlations. That is, the covariance between factors i and j is given by

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

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

References

Bayraktar, Mehmet K., Igor Mashtaler, Nicolas Meng, Jun Wang, Jay Yao. 2013a. "The Barra Japan Equity Model (JPE4): Empirical Notes" *MSCI Model Insight.*

Bayraktar, Mehmet K., Igor Mashtaler, Nicolas Meng, Jay Yao, Jun Wang. 2013b. "Systematic Equity Strategies: A Testcase Using Empirical Results from the Japan Equity Market." *MSCI Research Insight.*

Bayraktar, Mehmet K., Stan Radchenko, Kurt Winkelmann, Peter J. Zangari. 2013c. "Employing Systematic Equity Strategies: Distinguishing Important Sources of Risk from Common Sources of Return." *MSCI Research Insight.*

Bender, Jennifer, Jyh-Huei Lee, Dan Stefek and Jay Yao. 2009. "Forecast Risk Bias in Optimized Portfolios." *MSCI Research Insight.*

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

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

Lee, Jyh-Huei, Oleg Ruban, Dan Stefek and Jay Yao. 2013. "Alpha-Risk Factor Misalignment: Does it Pose a Problem?" *MSCI Research Insight.*

Lee, Jyh-Huei, Dan Stefek, Rong Xu and Jay Yao. 2011. "Mitigating Risk Forecast Biases of Optimized Portfolios." *MSCI Research Insight.*

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

Menchero, Jose. 2012. "Predicting Risk at Short Horizons." *MSCI Research Insight.*

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

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

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

Menchero, Jose, D.J. Orr, and Jun Wang. 2011. "The Barra US Equity Model (USE4): Methodology Notes." *MSCI Model Insight.*

Menchero, Jose, Jun Wang, and D.J. Orr. 2011. "Eigen-Adjusted Covariance Matrices." *MSCI Research Insight.*

Morozov, Andrei, Imre Balint, Simon Minovitsky, Mehmet Bayraktar. 2014a. "The Barra Asia Pacific Equity Model (ASE2): Empirical Notes" *MSCI Model Insight.*

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

Patton, Andrew. 2011. "Volatility Forecast Comparison using Imperfect Volatility Proxies", *Journal of Econometrics 160*, pp. 246-256.

Client Service Information is Available 24 Hours a Day

clientservice@msci.com

Americas

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

Europe, Middle East, Africa & India

Cape Town	+ 27.21.673.0100
Frankfurt	+ 49.69.133.859.00
Geneva	+ 41.22.817.9777
London	+ 44.20.7618.2222
Milan	+ 39.02.5849.0415
Mumbai	+91.22.6784.9160
Paris	0800.91.59.17 (toll free)

Asia Pacific

China North	10800.852.1032 (toll free)
China South	10800.152.1032 (toll free)
Hong Kong	+ 852.2844.9333
Seoul	00798.8521.3392 (toll free)
Singapore	800.852.3749 (toll free)
Sydney	+ 61.2.9033.9333
Taipei	008.0112.7513 (toll free)
Tokyo	+ 81.3.5290.1555

About MSCI

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

The company’s flagship product offerings are: the MSCI indexes with over USD 9.5 trillion estimated to be benchmarked to them on a worldwide basis1; Barra multi-asset class factor models, portfolio risk and performance analytics; RiskMetrics multi-asset class market and credit risk analytics; IPD real estate information, indexes and analytics; MSCI ESG (environmental, social and governance) Research screening, analysis and ratings; and FEA valuation models and risk management software for the energy and commodities markets. MSCI is headquartered in New York, with research and commercial offices around the world.

[1] *As of June 30, 2014, as reported on September 30 2014 by eVestment, Morningstar and Bloomberg* *Nov 2014*