MSCI

Quantitative Insight

What Makes a Good Statistical Model?

Case Studies in Passive Management

Rachael Smith and Imre Balint

rachael.smith@msci.com

imre.balint@msci.com

December 2012

ABSTRACT:

In this paper we investigate whether the new Barra Europe Stochastic Factor Model (EURS1) performs differently in portfolio construction when compared to other statistical models typically used by investors. In order to explore this, we built a typical principal component analysis (PCA) model and used both the PCA and EURS1 models to track two popular benchmarks using the Barra Open Optimizer. In all cases, we found EURS1 to have the lowest tracking error, and the EURS1 slow model to have the lowest turnover.

What Makes a Good Statistical Model?

Statistical factor models do not depend on any prior assumptions about the market structure, but require only a panel of asset returns data for the chosen market to 'learn' the factor exposures and returns. Statistical factors cannot generally be identified via fundamental investment attributes, such as an industry or country, and so they are labeled numerically.

The most commonly used approach to statistical factor modeling is ordinary principal component analysis (PCA). This technique detects the set of uncorrelated factors that most efficiently capture the variance of a given panel of data. We refer to Joliffe[1] for further information about PCA.

The Barra Europe Stochastic Factor Model (EURS1) is the first in a family of statistical factor models produced by MSCI designed to provide an alternative view on market risk and return that complements the insight provided by the Barra fundamental factor models. EURS1 avoids some of the well-known potential shortcomings of the PCA approach; in particular, the stability of factor exposures and the inability to capture fast changing volatility regimes. The typically unstable nature of factor exposures produced by PCA makes return and risk attribution difficult (for example, Factor 3 today could bear no resemblance to Factor 3 tomorrow). In addition, the inability to capture fast changing volatility regimes may mean that some short-lived correlations relevant for risk forecasting could be missed. The new model uses an innovative modeling approach that incorporates techniques taken from machine learning and signal processing to combat some of these issues. This is discussed in greater length in the EURS1 Research Notes.[2]

Given the investment and innovative modeling approach used to produce EURS1, we wanted to investigate how this new model performs compared to a conventional PCA model in the context of passive benchmark tracking.

Brief Summary of Statistical Models

Traditional Model

In order to make a comparison of EURS1 to a conventional PCA statistical model, we needed to build such a model. For the purposes of this investigation, we decided to use a weekly horizon model since this is the type of model we see in the marketplace. We estimate an ordinary PCA model using 180 weeks of equal-weighted weekly returns of these assets (where all assets have 180 valid weekly returns). We keep 20 principal components to be consistent with the EURS1 model.

EURS1

One of the highlights of the EURS1 model is the separation of the modeling of volatility and correlations. This follows from the intuitive standpoint that market correlations tend to evolve gradually, whereas

[1] Joliffe, I. T., 2002, Principal Component Analysis (2nd ed.), New York: Spring-Verlag.

[2] Ward, P., 2012, The Barra Europe Stochastic Factor model (EURS1) Research Notes.

volatilities can change more rapidly. This split in the modeling of volatilities and correlations means that we effectively have two models in one (meaning that more layers of insight can be found). In EURS1, the volatilities are modeled using wavelet decomposition. This generates smooth and more accurate volatility forecasts and overcomes the problem of over-forecasting following volatility shocks. In order to estimate the correlations, we use a factor learning algorithm designed to generate a stable factor structure and not over-fit noise. The model produces daily, weekly and monthly risk horizons, each with a fast and slow model version. The number of factors used to describe correlations is 20. In this paper, we focused on the weekly horizon model since this has the same forecast horizon as the PCA model we created. We will use both the fast and slow model versions in the case studies.

Portfolio Optimization

In this section, we describe the different case studies undertaken to compare the two models and their performance in benchmark tracking. We chose to focus on two benchmarks, the MSCI Europe Index and the more volatile MSCI Europe Value Index. MSCI Europe is a popular index amongst our passive management clients and has approximately 500 stocks as of March 2012. MSCI Europe Value is a subset of MSCI Europe and had 200 stocks as of March 2012. We chose this index since value investing is a popular strategy.

The case studies run were as follows:

- Minimize tracking error relative to the benchmark using a maximum number of assets = 50, 100, 150, and 200.
- All cases were run over the period June 2002 – March 2012.
- Universe chosen was the benchmark.
- Long only constraint.
- Turnover constraint equivalent to average maximum monthly turnover of 5 percent and 8 percent.
- Rebalance weekly.

We ran optimizations using Barra Open Optimizer using both the PCA and EURS1 models. All constraints and requirements (as set above) were kept the same, with only the model choice changing so that a fair comparison between models could be obtained.

In order to compare the results of the optimizations we considered the following measures:

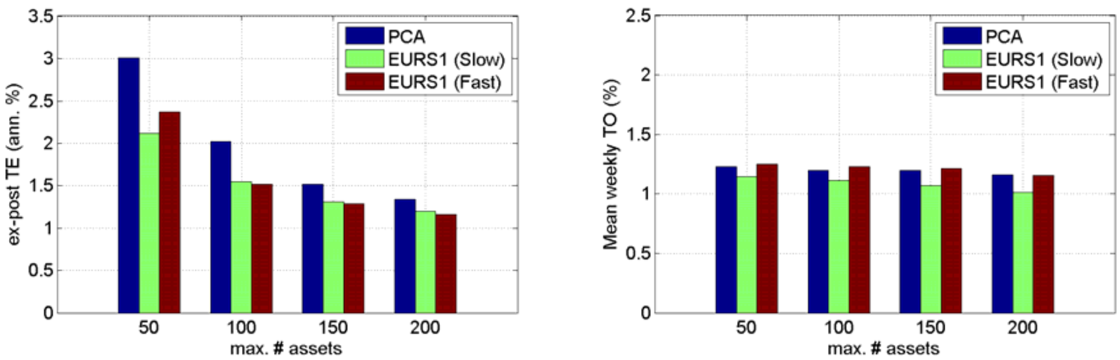
1. Optimized portfolio realized ex-post tracking error
2. Portfolio turnover
3. Comparison of realized to forecast risk
4. Stability of the portfolio exposure to the statistical factors

We use a simple measure of ex-post active risk: the equal-weighted standard deviation of the realized active returns over the period.

Ex-Post Tracking Error and Portfolio Turnover

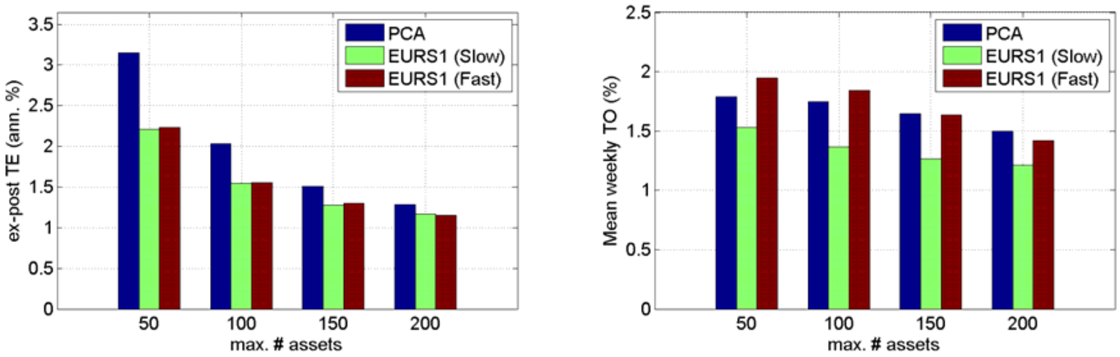
The first comparison we make between the two models is of the ex-post tracking error between the optimized portfolios and the benchmark; this measures how closely a portfolio has tracked its benchmark (not the accuracy of the forecasts) and should be as small as possible. In Figure 1, we show the results for the optimized portfolio when the benchmark is set to the MSCI Europe Index and the turnover constraint is set to the equivalent of a 5 percent average monthly turnover. On the left hand side, we show the realized tracking error results of the optimized portfolios generated using all three models (PCA, EURS1 slow, and EURS1 fast). We see that as the number of assets used to track the benchmark decreases, the EURS1 model has an increasing advantage over the PCA model. In Figure 1 we also show the portfolio turnover. We do not mind increased turnover if there is some return or tracking error improvement, but its desirable to have turnover as low as possible, since increased turnover results in higher transaction costs. Here we see that the slow model has a slight advantage over the other two models by having the lowest turnover.

Figure 1: Optimized portfolio tracking error and turnover, March 2002 – March 2012. Benchmark is MSCI Europe. Turnover constraint equivalent to 5 percent average monthly turnover.



If we increase the allowable turnover, we see the difference in results widen. In Figure 2, we show the equivalent results to Figure 1 but this time using turnover constraint equivalent to 8 percent monthly. Here, when we look at the ex-post tracking error, we see that the gap between the PCA model and the EURS1 model is more substantial with nearly a 1 percent difference in tracking error when using 50 stocks to track the benchmark. The difference in tracking error between the slow and fast EURS1 models is small. In addition, the turnover results show a marked difference, with the EURS1 slow model having the lowest turnover.

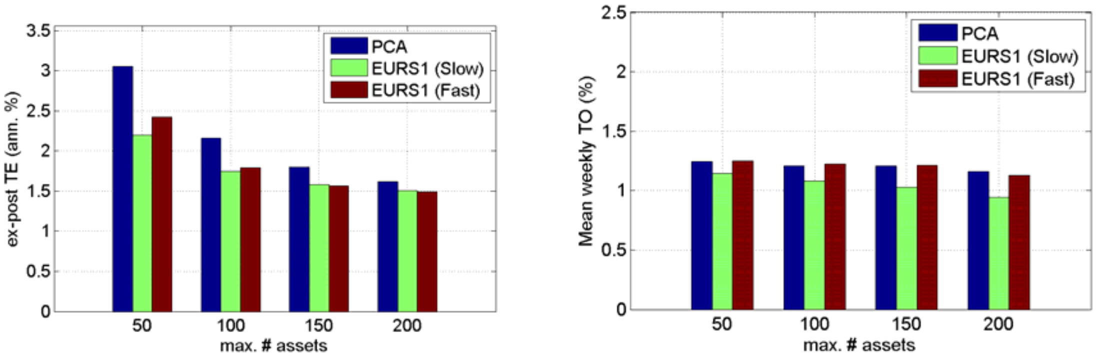
Figure 2: Optimized portfolio tracking error and turnover, March 2002 – March 2012. Benchmark is MSCI Europe. Turnover constraint equivalent to 8 percent average monthly turnover.



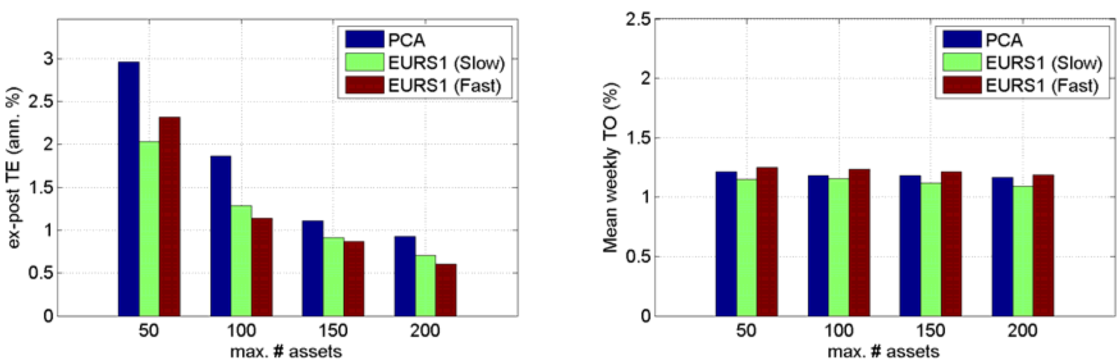
The data so far is an average over quite different volatility regimes. In Figure 3 we split up the date range over which we calculate the ex-post tracking error for the 5 percent turnover constraint to see the variation across time. There are a few interesting observations to be made.

Figure 3: Optimized portfolio tracking error and turnover. Benchmark is MSCI Europe. Turnover constraint equivalent to 5 percent average monthly turnover.

(a) 2002 – 2007



(b) 2007 – 2012



First of all, we see that EURS1 model performs better than the PCA model in all cases. In addition, the difference in tracking error produced by using the EURS1 slow or fast model during the period 2002-2007 is very small.

During the period 2007 – 2012, however, we see that the fast model version has an advantage over the slow model. When we consider the turnover produced by using the different models in 2002 – 2007, we see that the EURS1 slow model produces the lowest turnover of all the models.

The EURS1 fast model and the PCA model produce similar levels of turnover. During the period 2007-2012, we find the difference in turnover generated by the models is very small. Given the results presented so far, we would suggest that the EURS1 slow model performs best, considering both tracking error and portfolio turnover.

In the appendix, we show the corresponding results for the MSCI Europe Value Index. If we look at the overall picture from 2007 – 2012 in A1 and A2, we first observe that the EURS1 model performs better than the PCA model for all cases. When the turnover constraint of 5 percent monthly is applied, the EURS1 fast has a slight advantage over the EURS1 slow model for the 50 and 100 stocks cases, though with broader portfolios, the slow model has an advantage over the fast model. Similar to the results for MSCI Europe, comparing the turnover produced, we see that the EURS1 slow model produces the lowest turnover.

When the turnover constraint is increased to 8 percent, the difference between the EURS1 slow and fast models is reduced (in line with what we've seen for the MSCI Europe cases), and the performance gap between the EURS1 model and PCA model is increased. Again, we see that the slow EURS1 model produces the lowest turnover.

In A3-A7, we show these results when we split the period in two: June 2002-July 2007, and August 2007 – March 2012. The results from March 2002 – July 2007, in tables A3 and A5, show that the EURS1 fast model produces the lowest tracking error of all the models. This situation changes in the more volatile period of August 2007 – March 2012, when we see that if the number of assets allowed to track the benchmark is 50 or 100, the fast model has the advantage, while the slow model performs better in the 150 and 200 stock case. Throughout all these cases, the EURS1 slow model has the lowest turnover.

Forecast to Realized Risk

Up until now we have focused our attention on the realized tracking error. Another important consideration is the accuracy of the forecasts. In Table 1, we show a select summary of the results for tracking both the MSCI Europe and MSCI Europe Value indices.

We present the ex-post annualized active risk and compare this to the average ex-ante annualized active risk as measured by the model that we use for optimizing. It is important to note that this is not a like-for-like comparison between the models because the portfolios will be different (in particular, we see from the portfolio bias tests in the EURS1 Research Notes that the EURS1 fast model has more accurate forecasts than the slow model in all cases tested). The results in Table 1 show that, in all cases, EURS1 has more accurate forecasts than the PCA model. We also note that all models under-forecast the ex-post risk. This is in line with the results by Muller (1993)[3] that demonstrated empirically how risk models tend to underestimate risk of optimized portfolios.

[3] Muller, P. 1993, "Financial Optimization", Cambridge University Press (Zenious), 80-98.

Table 1: Summary comparing forecast and realized risk forecasts, March 2002 – March 2012. Turnover constraint equivalent to 5 percent average monthly turnover.

Benchmark	Num. Assets	Model	Ex-Post Ann. Active Risk (%)	Average Ex-Ante Ann. Active Risk (%)	Ex-Ante/Ex-Post Active Risk
MSCI Europe	50	EURS1 slow	2.12	1.70	0.80
		EURS1 fast	2.38	1.68	0.71
		PCA	3.15	1.41	0.45
	100	EURS1 slow	1.55	0.90	0.58
		EURS1 fast	1.52	0.86	0.56
		PCA	2.02	0.73	0.36
MSCI Europe Value	50	EURS1 slow	1.81	1.41	0.78
		EURS1 fast	1.73	1.34	0.77
		PCA	2.36	1.21	0.51
	100	EURS1 slow	1.00	0.68	0.68
		EURS1 fast	0.94	0.64	0.68
		PCA	1.35	0.73	0.54

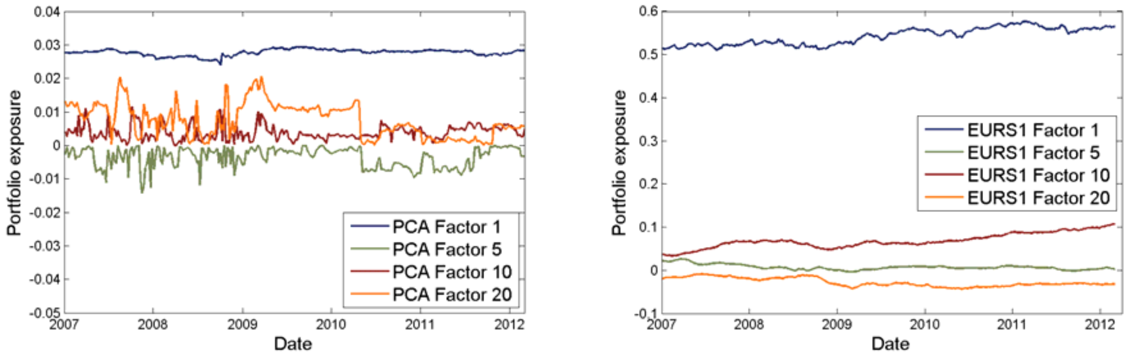
The full results are shown in the appendix. In all the cases shown, the EURS1 model performs better than the PCA model, and the portfolio generated using the EURS1 slow model have more accurate forecasts.

Exposure Stability

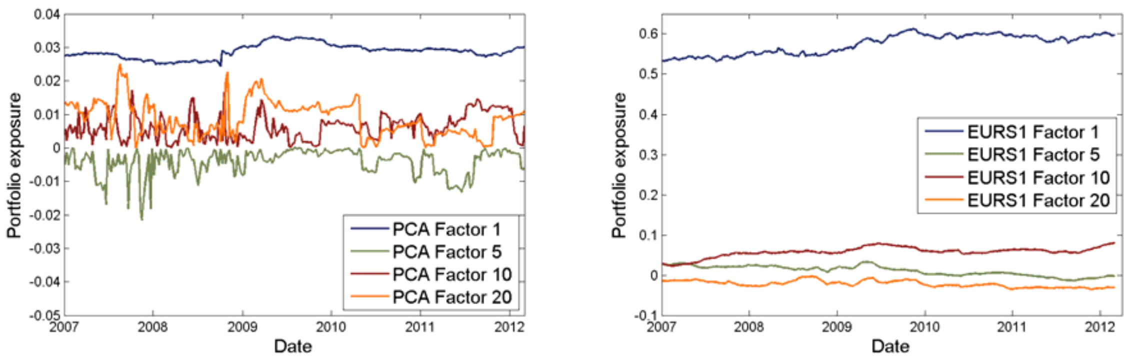
As we mentioned in the introduction, the EURS1 model has been designed to avoid some of the potential shortcomings of PCA; in particular, factor exposure stability. With a PCA model, the factors can move around so that, for example, factor 12 from one month bears no resemblance to that of the next month. The EURS1 model was designed to have stable, gradually changing exposures and factors across time. To illustrate this, we show in Figure 4 the optimized portfolio exposures to the PCA factors and the EURS1 factors between 2007 and 2012 of both the portfolios tracking the MSCI Europe and MSCI Europe Value indices. Here we can see that the PCA exposures are extremely unstable. This instability means that any portfolio risk or returns attribution in terms of the PCA factors will be virtually meaningless. In addition, we note that in order to make these figures more easily readable, we had to force the PCA portfolios not to flip sign. The sign does not matter in terms of risk attribution, but it does affect the returns attribution.

Figure 4: Exposure stability of portfolios tracking MSCI Europe and MSCI Europe Value generated by the PCA model (on left) and EURS1 portfolio (on right), August 2007 – March 2012.

(a) MSCI Europe tracking Portfolio, 50 stock portfolio



(b) MSCI Europe Value tracking portfolio, 50 stock portfolio



Conclusion

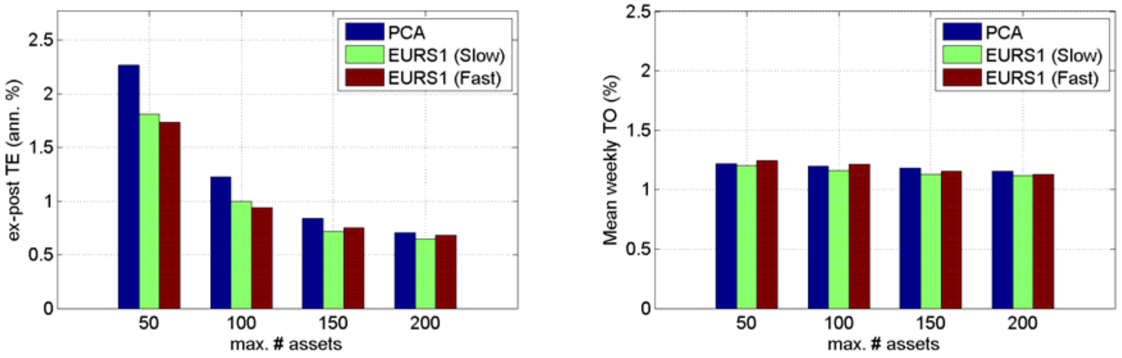
This study was conducted to investigate whether the new Barra Europe Stochastic Factor Model, EURS1, is an improvement on other statistical models that are typically used. Do the innovative methodologies translate into improved performance for the purpose of benchmark tracking?

In order to investigate this, we built a typical PCA model and used both the PCA and EURS1 models to track two popular benchmarks, the MSCI Europe and MSCI Europe Value indices, then used the Barra Open Optimizer.

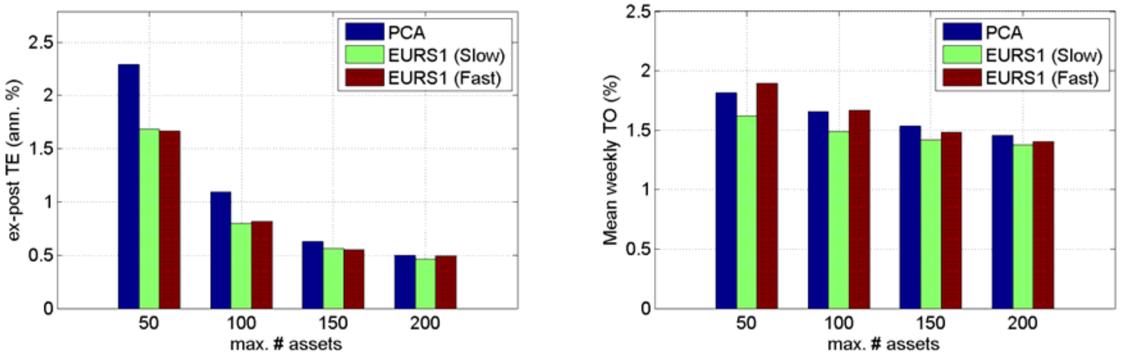
In all cases, we found the PCA model to have the worst realized tracking error; we also found the PCA model to have either similar or higher turnover than the EURS1 model. In addition, the portfolio turnover was least when using the EURS1 slow model. Both the fast and slow EURS1 model versions had very similar realized tracking error. When comparing the realized-to-forecast risk, we again found that the EURS1 model performed best. Overall, we would recommend the use of the EURS1 slow model for the purposes of portfolio optimization and the EURS1 model for purposes of risk and return attribution, due to the stability of the factor structure.

Appendix

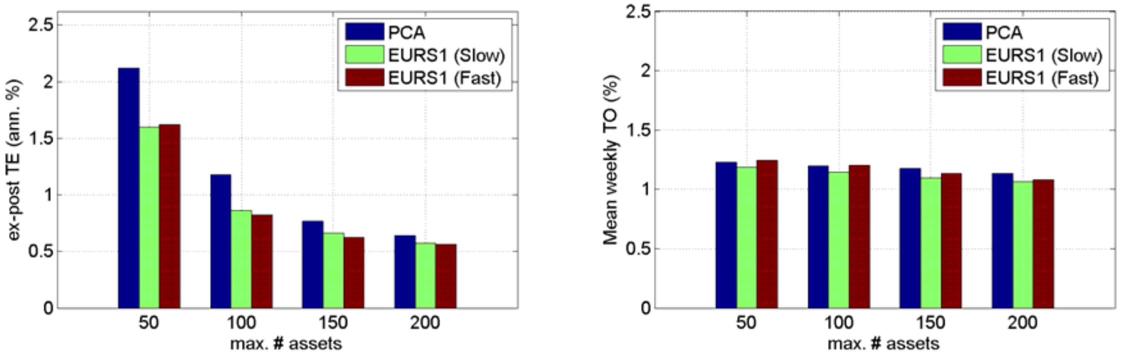
A1: Optimized portfolio tracking error and turnover, March 2002 – March 2012. Benchmark is MSCI Europe Value. Turnover constraint equivalent to 5 percent average monthly turnover.



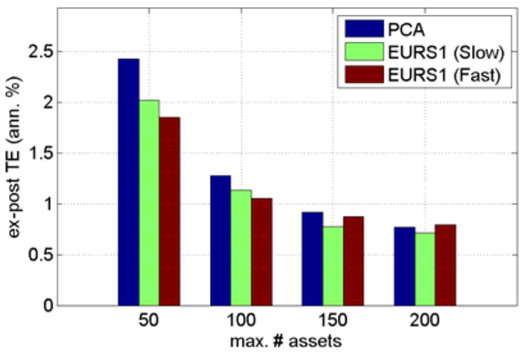
A2: Optimized portfolio tracking error and turnover, March 2002 – March 2012. Benchmark is MSCI Europe Value. Turnover constraint equivalent to 8 percent average monthly turnover.

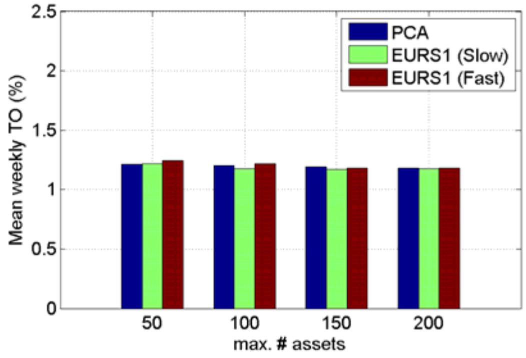


A3: Optimized portfolio tracking error and turnover March 2002 – July 2007. Benchmark is MSCI Europe Value. Turnover constraint equivalent to 5 percent average monthly turnover.

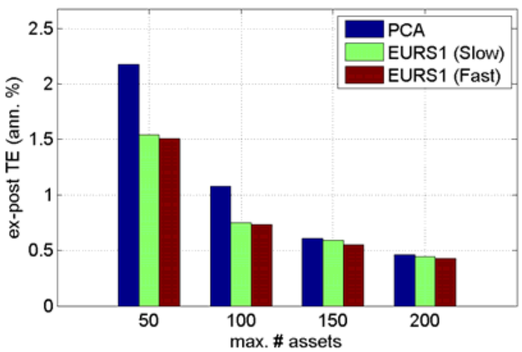


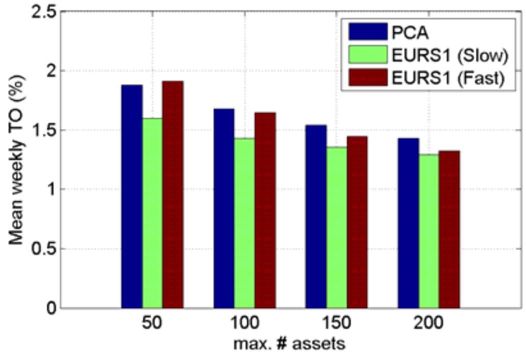
A4: Optimized portfolio tracking error and turnover August 2007 – March 2012. Benchmark is MSCI Europe Value. Turnover constraint equivalent to 5 percent average monthly turnover.



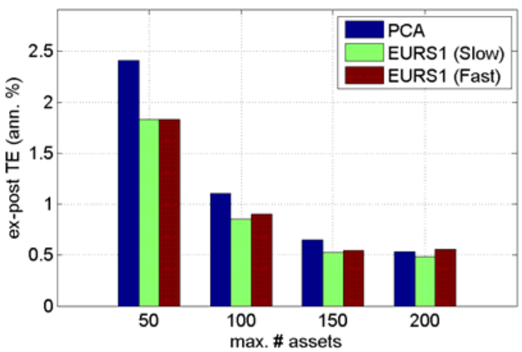


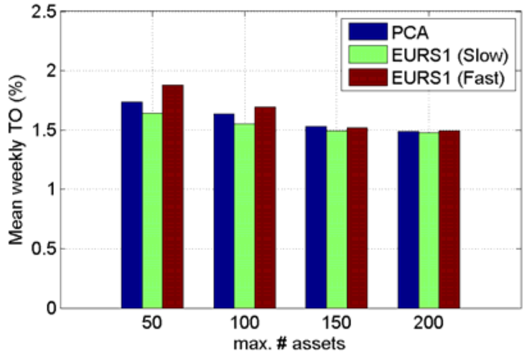
A5: Optimized portfolio tracking error and turnover March 2002 – July 2007. Benchmark is MSCI Europe Value. Turnover constraint equivalent to 8 percent average monthly turnover.





A6: Optimized portfolio tracking error and turnover August 2007 – March 2012. Benchmark is MSCI Europe Value. Turnover constraint equivalent to 8 percent average monthly turnover.





A7: Forecast and realized risk forecasts, March 2002 – March 2012. Benchmark is MSCI Europe. Turnover constraint equivalent to 5 percent average monthly turnover.

Num. Assets	Model	Ex-Post Ann. Active Risk (%)	Average Ex-Ante Ann. Active Risk (%)	Ex-Ante/Ex-Post Active Risk
50	EURS1 slow	2.12	1.70	0.80
	EURS1 fast	2.38	1.68	0.71
	PCA	3.15	1.41	0.45
100	EURS1 slow	1.55	0.90	0.58
	EURS1 fast	1.52	0.86	0.56
	PCA	2.02	0.73	0.36
150	EURS1 slow	1.31	0.56	0.43
	EURS1 fast	1.29	0.53	0.41
	PCA	1.44	0.44	0.31
200	EURS1 slow	1.20	0.38	0.32
	EURS1 fast	1.16	0.36	0.31
	PCA	1.25	0.30	0.24

A8: Forecast and realized risk forecasts, March 2002 – March 2012. Benchmark is MSCI Europe. Turnover constraint equivalent to 8 percent average monthly turnover.

Num. Assets	Model	Ex-Post Ann. Active Risk (%)	Average Ex-Ante Ann. Active Risk (%)	Ex-Ante/Ex-Post Active Risk
50	EURS1 slow	2.21	1.67	0.76
	EURS1 fast	2.23	1.58	0.71
	PCA	3.25	1.36	0.42
100	EURS1 slow	1.54	0.89	0.58
	EURS1 fast	1.56	0.82	0.53
	PCA	2.02	0.69	0.34
150	EURS1 slow	1.28	0.56	0.43
	EURS1 fast	1.30	0.52	0.40
	PCA	1.41	0.43	0.30
200	EURS1 slow	1.16	0.38	0.32
	EURS1 fast	1.15	0.35	0.30
	PCA	1.18	0.29	0.24

A9: Forecast and realized risk forecasts, March 2002 – March 2012. Benchmark is MSCI Europe Value. Turnover constraint equivalent to 5 percent average monthly turnover.

Num. Assets	Model	Ex-Post Ann. Active Risk (%)	Average Ex-Ante Ann. Active Risk (%)	Ex-Ante/Ex-Post Active Risk
50	EURS1 slow	1.81	1.41	0.78
	EURS1 fast	1.73	1.34	0.77
	PCA	2.36	1.21	0.51
100	EURS1 slow	1.00	0.68	0.68
	EURS1 fast	0.94	0.64	0.68
	PCA	1.35	0.64	0.47
150	EURS1 slow	0.72	0.45	0.63
	EURS1 fast	0.75	0.43	0.57
	PCA	0.91	0.44	0.48
200	EURS1 slow	0.65	0.36	0.55
	EURS1 fast	0.68	0.34	0.51
	PCA	0.79	0.37	0.46

A10: Forecast and realized risk forecasts, March 2002 – March 2012. Benchmark is MSCI Europe Value. Turnover constraint equivalent to 8 percent average monthly turnover.

Num. Assets	Model	Ex-Post Ann. Active Risk (%)	Average Ex-Ante Ann. Active Risk (%)	Ex-Ante/Ex-Post Active Risk
50	EURS1 slow	1.68	1.32	0.78
	EURS1 fast	1.67	1.24	0.74
	PCA	2.42	1.09	0.45
100	EURS1 slow	0.80	0.60	0.75
	EURS1 fast	0.82	0.55	0.68
	PCA	1.14	0.53	0.46
150	EURS1 slow	0.56	0.35	0.63
	EURS1 fast	0.55	0.33	0.60
	PCA	0.70	0.33	0.47
200	EURS1 slow	0.46	0.24	0.52
	EURS1 fast	0.49	0.23	0.47
	PCA	0.55	0.24	0.44

Client Service Information is Available 24 Hours a Day

clientservice@msci.com

Americas

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

Europe, Middle East & Africa

Amsterdam	+ 31.20.462.1382
Cape Town	+ 27.21.673.0100
Frankfurt	+ 49.69.133.859.00
Geneva	+ 41.22.817.9777
London	+ 44.20.7618.2222
Madrid	+ 34.91.700.7275
Milan	+ 39.02.5849.0415
Paris	0800.91.59.17 (toll free)
Zurich	+ 41.44.220.9300

Asia Pacific

China North	10800.852.1032 (toll free)
China South	10800.152.1032 (toll free)
Hong Kong	+ 852.2844.9333
Seoul	+827.0768.88984
Singapore	800.852.3749 (toll free)
Sydney	+ 61.2.9033.9333
Tokyo	+ 81.3.5226.8222

Notice and Disclaimer

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

About MSCI

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

The company's flagship product offerings are: the MSCI indices with approximately USD 7 trillion estimated to be benchmarked to them on a worldwide basis[1]; Barra multi-asset class factor models, portfolio risk and performance analytics; RiskMetrics multi-asset class market and credit risk analytics; MSCI ESG (environmental, social and governance) Research screening, analysis and ratings; ISS governance research and outsourced proxy voting and reporting services; FEA valuation models and risk management software for the energy and commodities markets; and CFRA forensic accounting risk research, legal/regulatory risk assessment, and due-diligence. MSCI is headquartered in New York, with research and commercial offices around the world.

[1]*As of June 30, 2011, based on eVestment, Lipper and Bloomberg data.*