

Facts about Factors

July 2015

Paula Cocoma
Megan Czasonis
Mark Kritzman
David Turkington

Contents

- Why use factors?
- Diversification and predictability
- Estimation error
- Classifications and data
- Results: asset classes versus fundamental factors and principal components
- Results: industries versus attributes and principal components
- Summary and conclusion

Why use factors?

- Factors may be less correlated than assets and therefore provide better diversification.
- Investors may be more skilled at relating current information to future factor behavior than to future asset behavior.
- Factors may have less estimation error than assets.
- Factors may be more effective than assets at reducing noise.

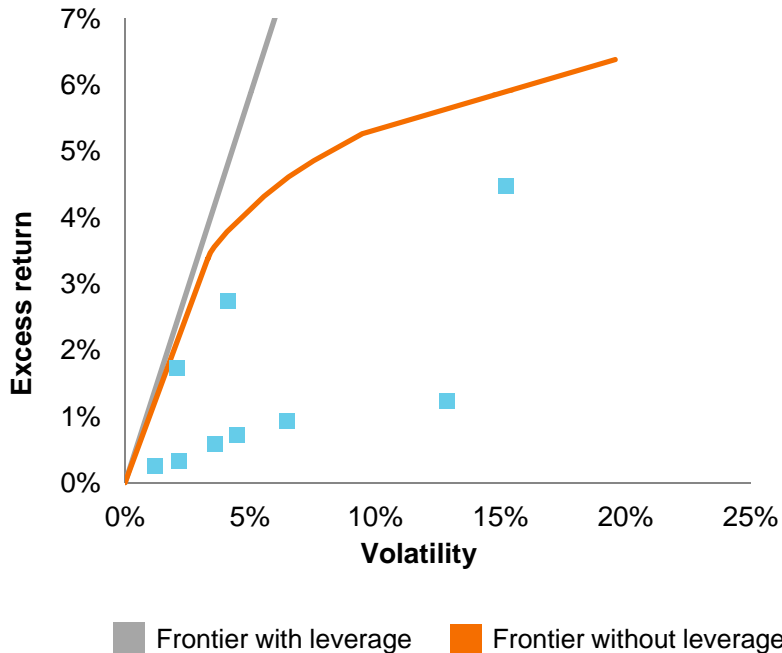
Diversification and Predictability



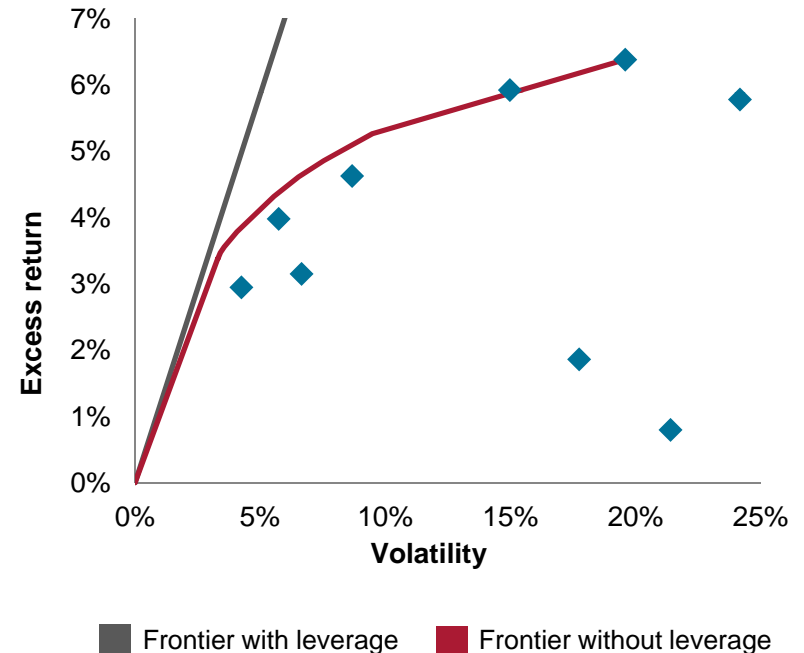
Diversification – a specious argument

Assets define the opportunity set. It is impossible to create a more efficient in-sample portfolio with the same constraints and of the same periodicity by regrouping assets into factors.

Principal components (all factors uncorrelated)



Asset classes (correlations range from -0.16 to 0.82)



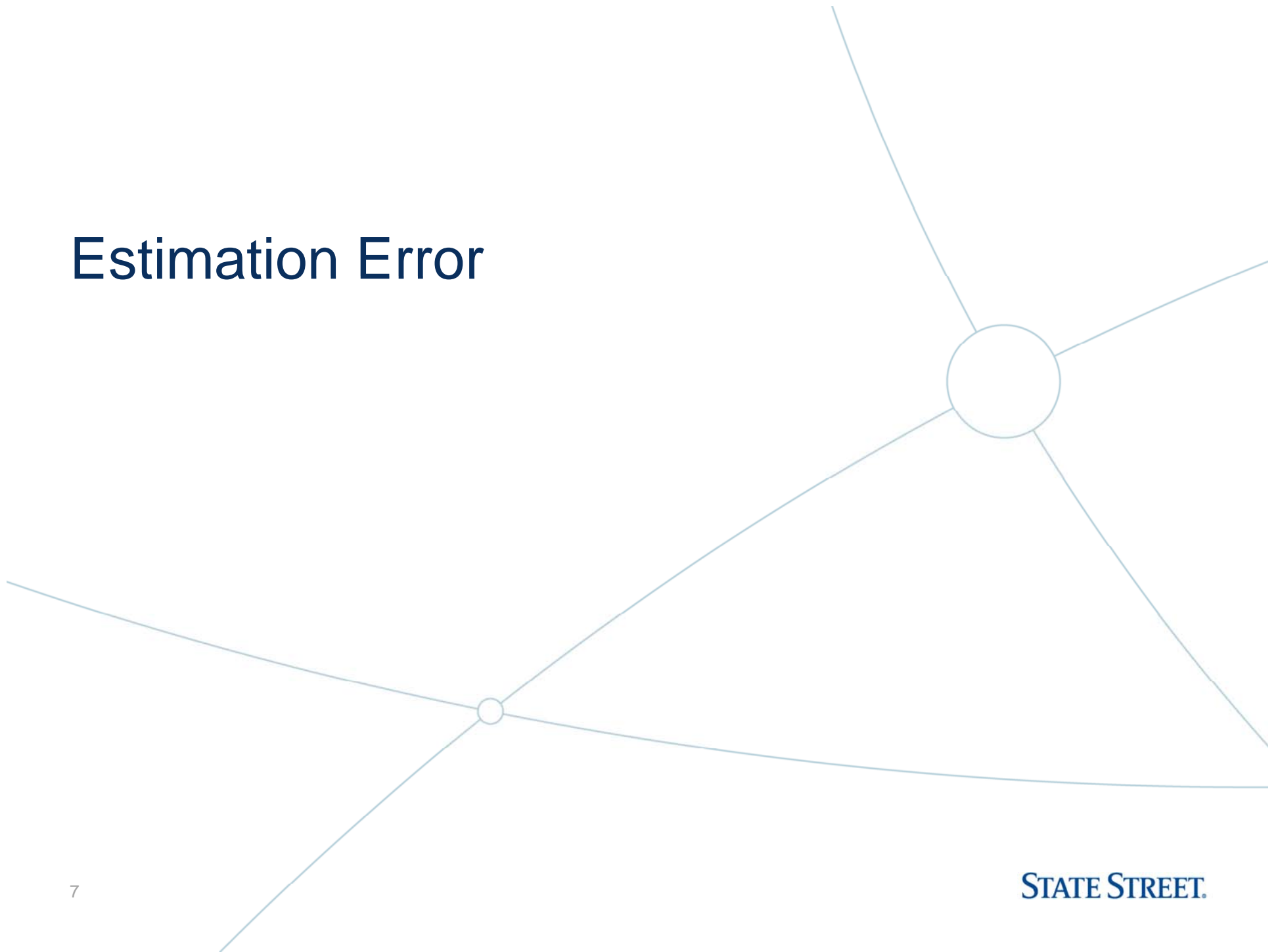
Source: State Street Global Exchange.

Notes: This analysis incorporates the following asset classes: U.S. large cap, U.S. small cap, EAFE equities, emerging equities, global sovereigns, U.S. government bonds, U.S. corporate bonds, commodities, and hedge funds. Based on monthly returns over the period Jan 1990 through Dec 2013. Excess returns represent the return over the risk-free rate. All data obtained from DataStream.

Superior predictability – an untestable hypothesis

- Some investors may be more skilled at relating past and current information to future factor behavior than to future asset behavior.
- However, some investors may be more skilled at relating past and current information to future asset behavior than to future factor behavior.
- It is impossible to test these hypotheses generally; they are investor specific.

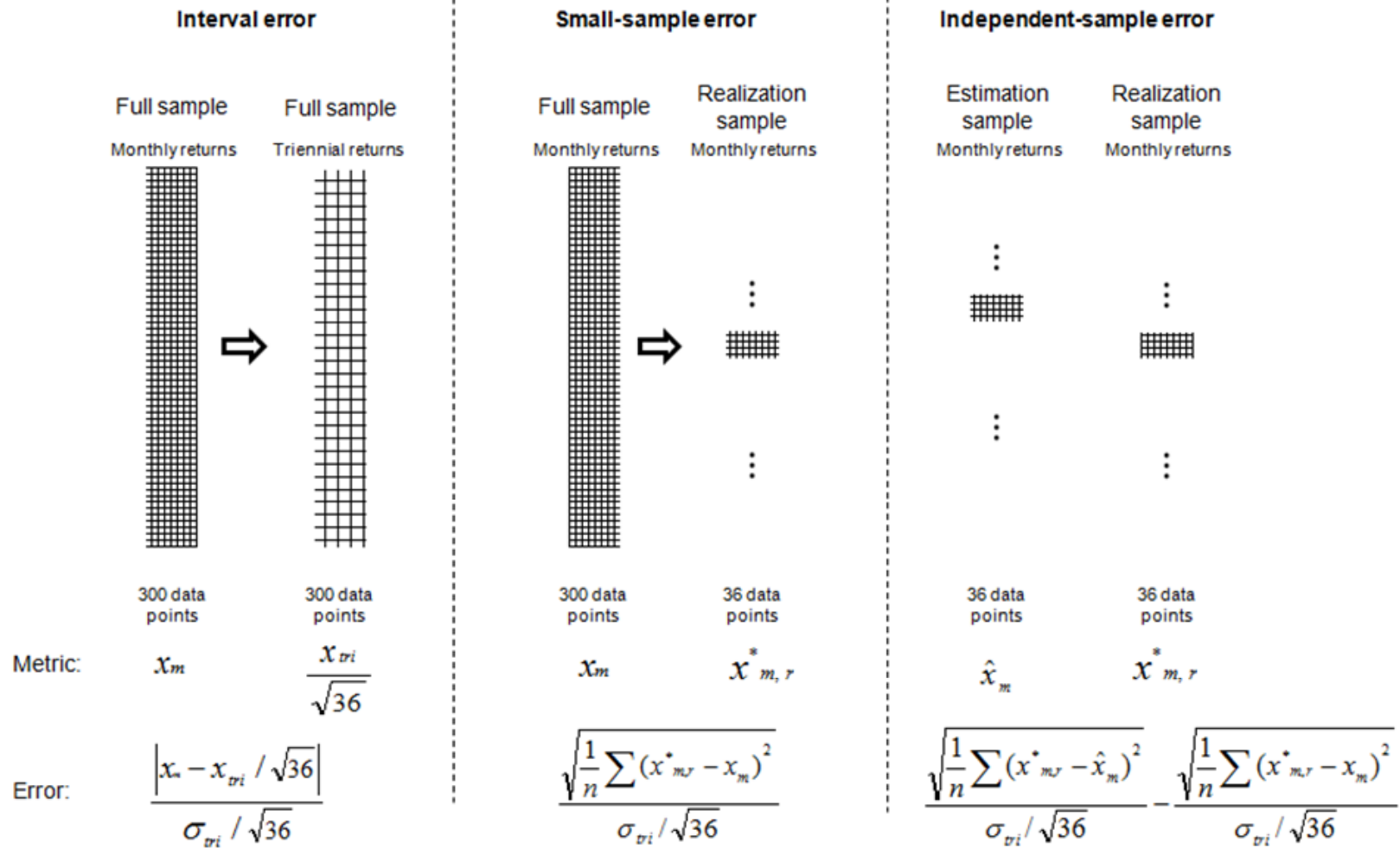
Estimation Error



Estimation error

- Interval error
- Small-sample error
- Independent-sample error

Estimation error



Source: State Street Global Exchange

1. In the case of measuring out-of-sample stationarity, we scale by triennial volatility for volatility and returns (not for correlation).

Interval error

Standard deviations and correlations estimated from monthly or higher-frequency returns often differ from those estimated from lower-frequency returns.

We call this interval error.

For more information, please refer to the following articles: “**The Divergence of High- and Low-Frequency Estimation: Implications for Performance Measurement**” by W. Kinlaw, M. Kritzman, and D. Turkington. Forthcoming in the Journal of Portfolio Management, Spring 2015, and “**The Divergence of High- and Low-Frequency Estimation: Causes and Consequences**” by W. Kinlaw, M. Kritzman, and D. Turkington. Journal of Portfolio Management, Special 40th Anniversary Issue, 2014.

The divergence of high- and low-frequency estimation

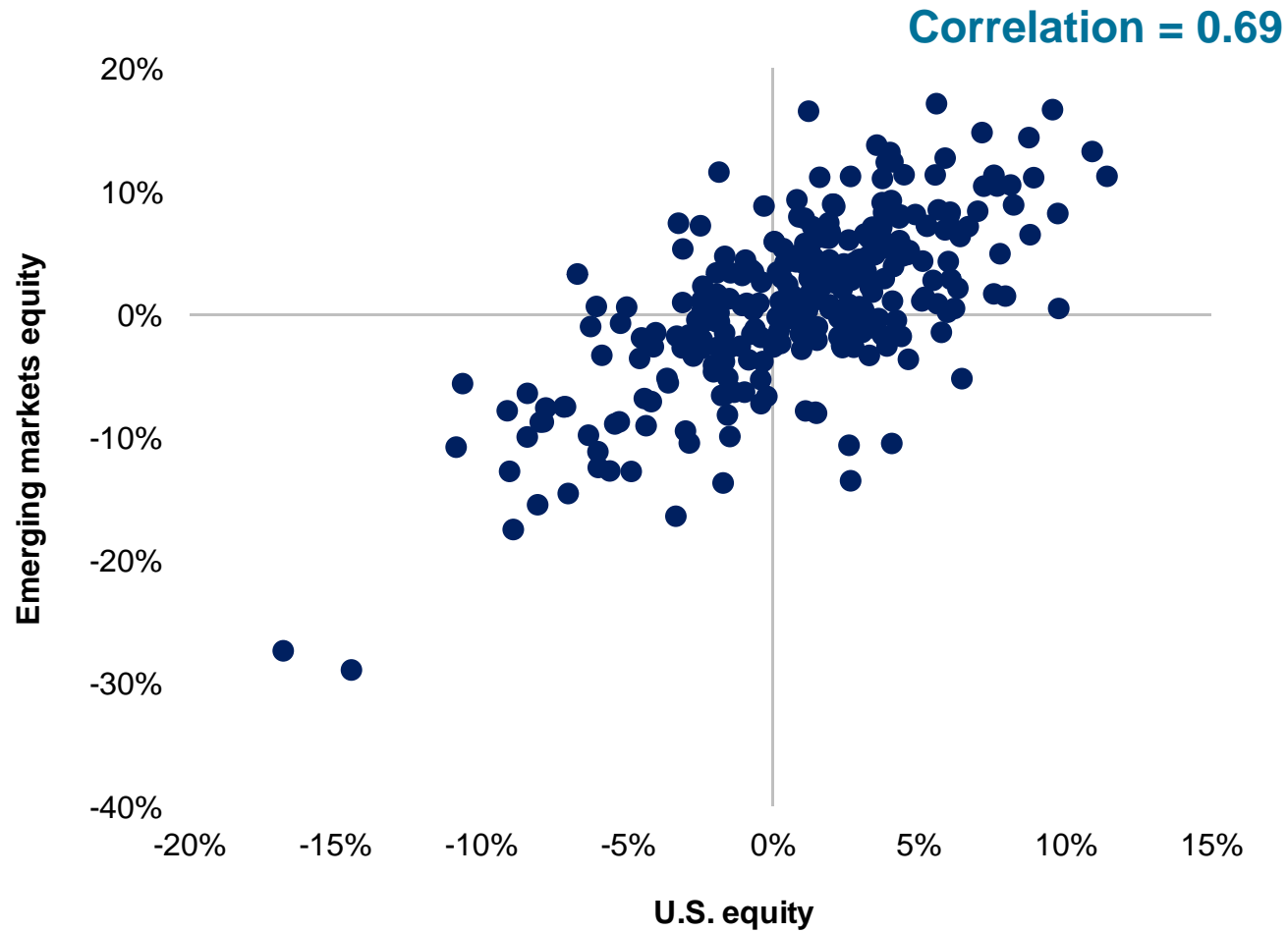
January 1990 to December 2013	
Emerging Markets Stock Return	9.30%
U.S. Stock Return	9.50%
Correlation of Monthly Returns	69%

January 2005 to December 2007	
Emerging Markets - US Stocks	121%

January 2011 to December 2013	
Emerging Markets - US Stocks	-62%

Interval error

U.S. and emerging markets stocks: monthly returns

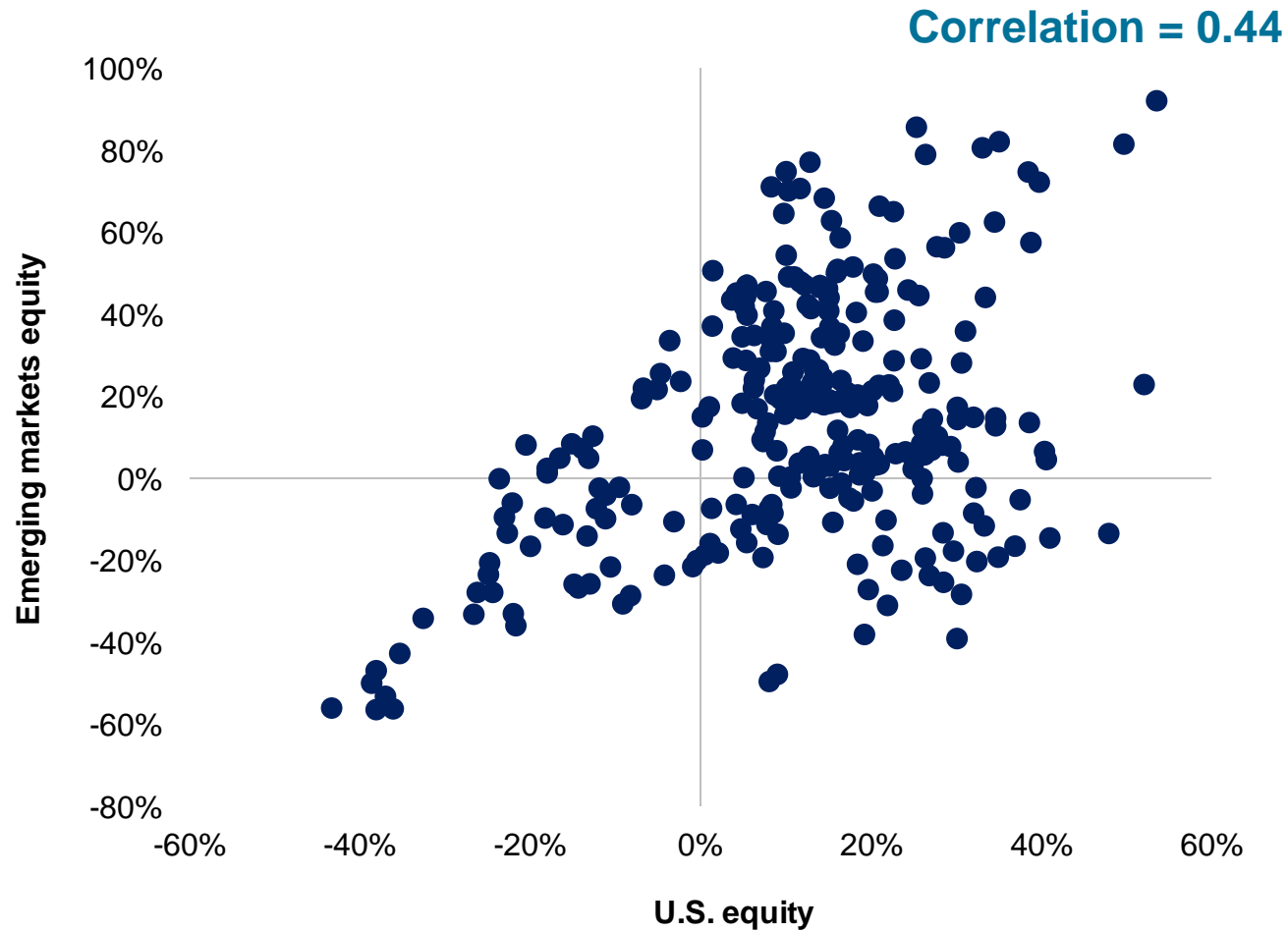


Source: State Street Global Exchange

Notes: Chart covers period Jan 1990 through Dec 2013. Data obtained from DataStream.

Interval error

U.S. and emerging markets stocks: annual returns

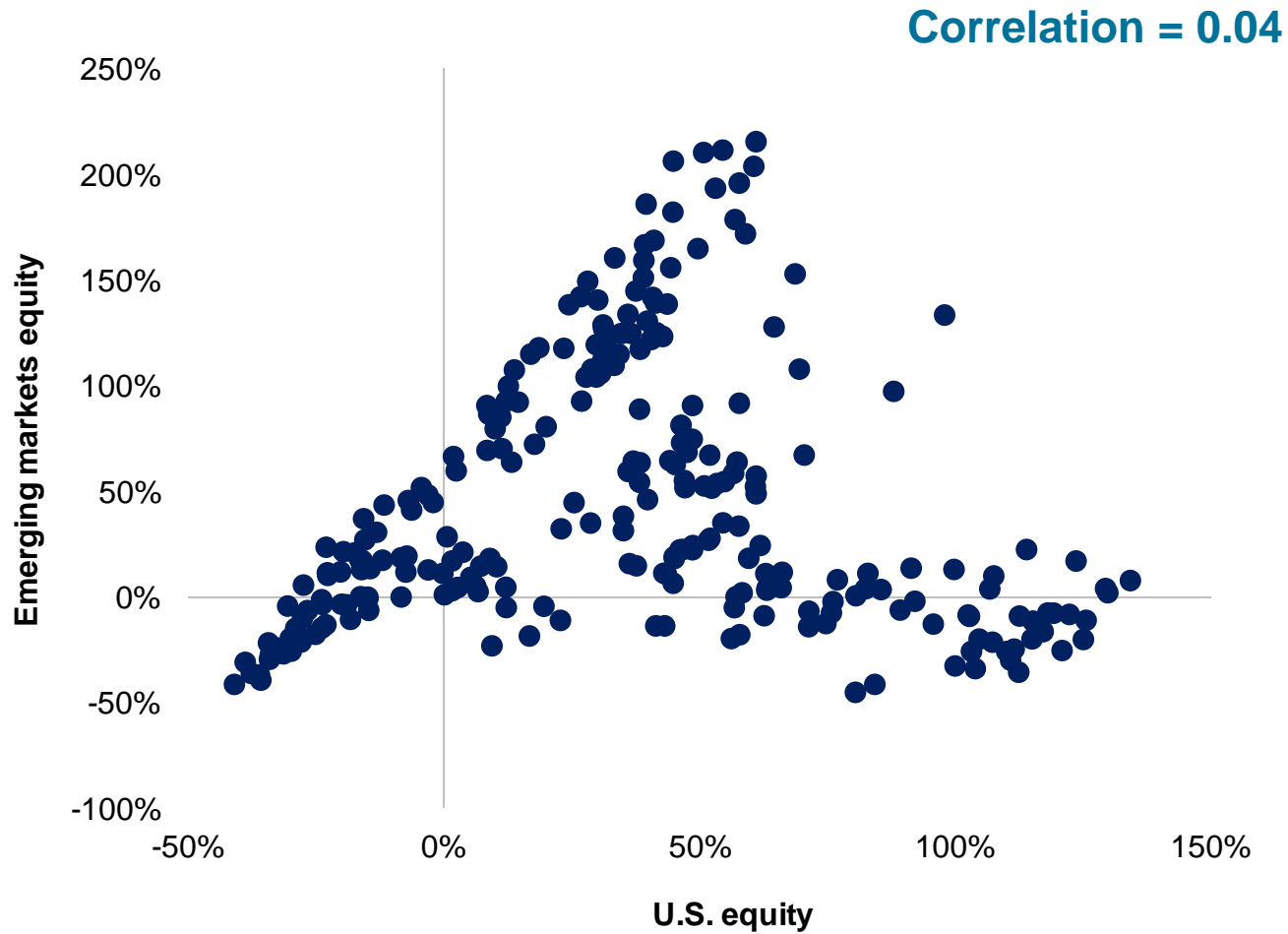


Source: State Street Global Exchange

Notes: Chart covers period Jan 1990 through Dec 2013. Data obtained from DataStream.

Interval error

U.S. and emerging markets stocks: triennial returns



Source: State Street Global Exchange

Notes: Chart covers period Jan 1990 through Dec 2013. Data obtained from DataStream.

The relation of high- and low-frequency standard deviation

The standard deviation of the cumulative continuous returns of x over q periods is given by:

$$\sigma(x_t + \dots + x_{t+q-1}) = \sigma_x \sqrt{q + 2 \sum_{k=1}^{q-1} (q-k) \rho_{x_t, x_{t+k}}}$$

The relation of high- and low-frequency standard deviation

The standard deviation of the cumulative continuous returns of x over q periods is given by:

$$\sigma(x_t + \dots + x_{t+q-1}) = \sigma_x \sqrt{q + 2 \sum_{k=1}^{q-1} (q-k) \rho_{x_t, x_{t+k}}}$$



This term reflects annualization in the absence of lagged effects

The relation of high- and low-frequency standard deviation

The standard deviation of the cumulative continuous returns of x over q periods is given by:

$$\sigma(x_t + \dots + x_{t+q-1}) = \sigma_x \sqrt{q + 2 \sum_{k=1}^{q-1} (q-k) \rho_{x_t, x_{t+k}}}$$



This term captures the impact of auto-correlation


The relation of high- and low-frequency correlation

The correlation between the cumulative returns of x and the cumulative returns of y over q periods is given by:

$$\rho(x_t + \dots + x_{t+q-1}, y_t + \dots + y_{t+q-1}) = \frac{q\rho_{x_t, y_t} + \sum_{k=1}^{q-1} (q-k)(\rho_{x_{t+k}, y_t} + \rho_{x_t, y_{t+k}})}{\sqrt{q + 2\sum_{k=1}^{q-1} (q-k)\rho_{x_t, x_{t+k}}} \sqrt{q + 2\sum_{k=1}^{q-1} (q-k)\rho_{y_t, y_{t+k}}}}$$

The relation of high- and low-frequency correlation

The correlation between the cumulative returns of x and the cumulative returns of y over q periods is given by:

$$\rho(x_t + \dots + x_{t+q-1}, y_t + \dots + y_{t+q-1}) = \frac{q\rho_{x_t, y_t} + \sum_{k=1}^{q-1} (q-k)(\rho_{x_{t+k}, y_t} + \rho_{x_t, y_{t+k}})}{\sqrt{q + 2\sum_{k=1}^{q-1} (q-k)\rho_{x_t, x_{t+k}}} \sqrt{q + 2\sum_{k=1}^{q-1} (q-k)\rho_{y_t, y_{t+k}}}}$$


This term captures the lagged cross-correlation between x and y

The relation of high- and low-frequency correlation

The correlation between the cumulative returns of x and the cumulative returns of y over q periods is given by:

$$\rho(x_t + \dots + x_{t+q-1}, y_t + \dots + y_{t+q-1}) = \frac{q\rho_{x_t, y_t} + \sum_{k=1}^{q-1} (q-k)(\rho_{x_{t+k}, y_t} + \rho_{x_t, y_{t+k}})}{\sqrt{q + 2\sum_{k=1}^{q-1} (q-k)\rho_{x_t, x_{t+k}}} \sqrt{q + 2\sum_{k=1}^{q-1} (q-k)\rho_{y_t, y_{t+k}}}}$$



This term captures the auto-correlation of x

The relation of high- and low-frequency correlation

The correlation between the cumulative returns of x and the cumulative returns of y over q periods is given by:

$$\rho(x_t + \dots + x_{t+q-1}, y_t + \dots + y_{t+q-1}) = \frac{q\rho_{x_t, y_t} + \sum_{k=1}^{q-1} (q-k)(\rho_{x_{t+k}, y_t} + \rho_{x_t, y_{t+k}})}{\sqrt{q + 2\sum_{k=1}^{q-1} (q-k)\rho_{x_t, x_{t+k}}} \sqrt{q + 2\sum_{k=1}^{q-1} (q-k)\rho_{y_t, y_{t+k}}}}$$



This term captures the auto-correlation of y

Measuring interval error

For a particular standard deviation or correlation:

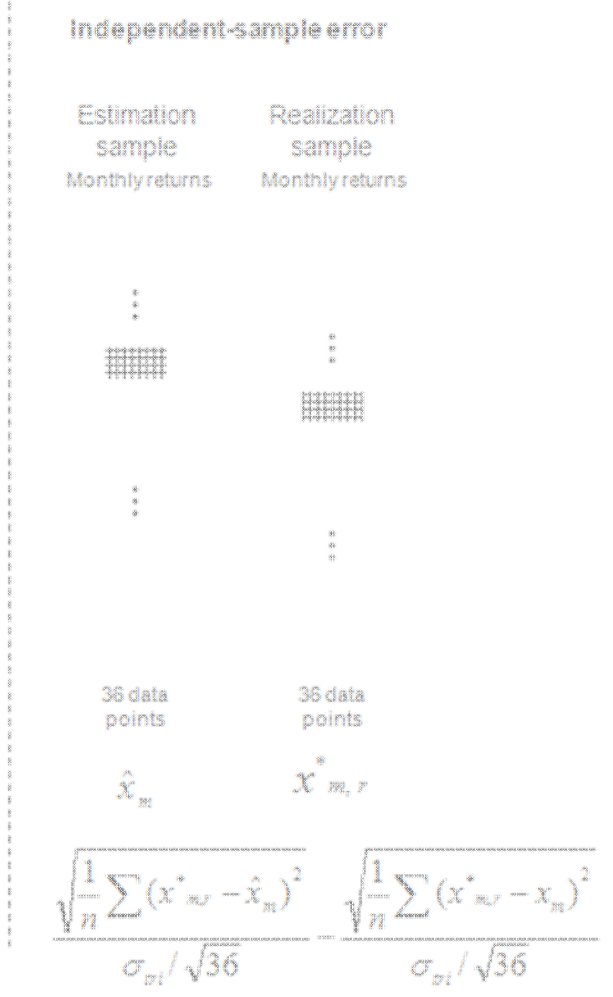
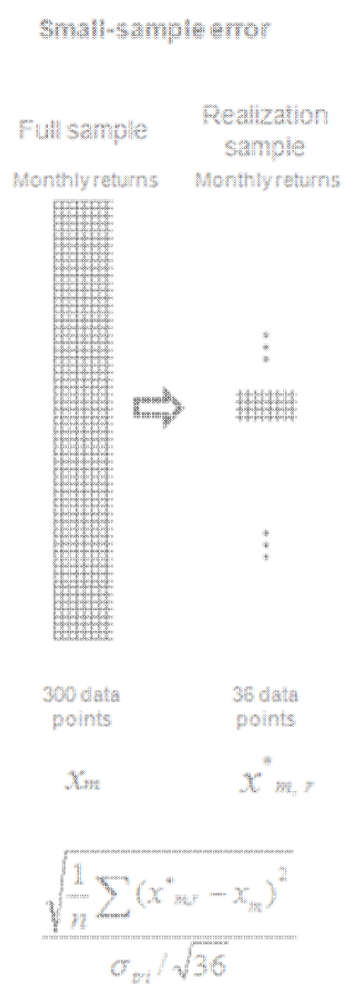
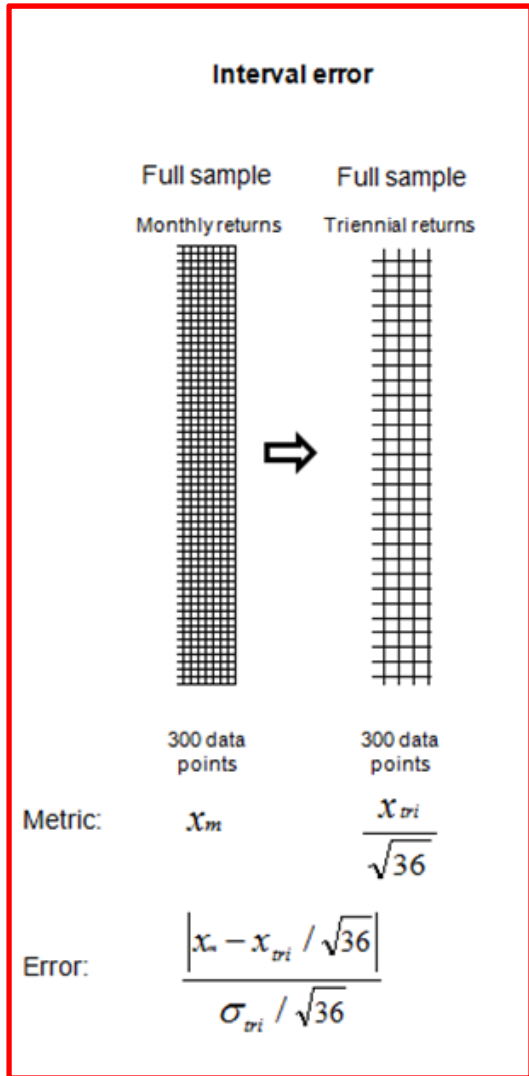
1. We estimate the parameter using the full sample of one-month returns: x_m
2. We estimate the parameter using three-year rolling returns¹: $x_{tri} / \sqrt{36}$
3. We measure interval error as the absolute difference between the one-month and three-year estimates (normalized by full sample triennial volatility)²:

$$\frac{\left| x_m - x_{tri} / \sqrt{36} \right|}{\sigma_{tri} / \sqrt{36}}$$

1. We divide by square root of 36 for volatility only (not for correlation).

2. We scale by triennial volatility for volatility only (not for correlation).

Measuring interval error



Measuring interval error

Example: Calculating interval error

Interval error of standard deviation	
Full sample, monthly standard deviation	4.00%
Full sample, triennial standard deviation $/\sqrt{36}$	5.00%
Absolute difference	$ 4.00\% - 5.00\% = 1.00\%$
Absolute difference, normalized	$1.00\% / 5.00\% = 20.0\%$

Interval error introduces a standardized error of 20% relative to the true parameter.

Small-sample error

The realization of parameters from a small sample will likely differ from the parameter values of a large sample from which it is selected.

We call this small-sample error.

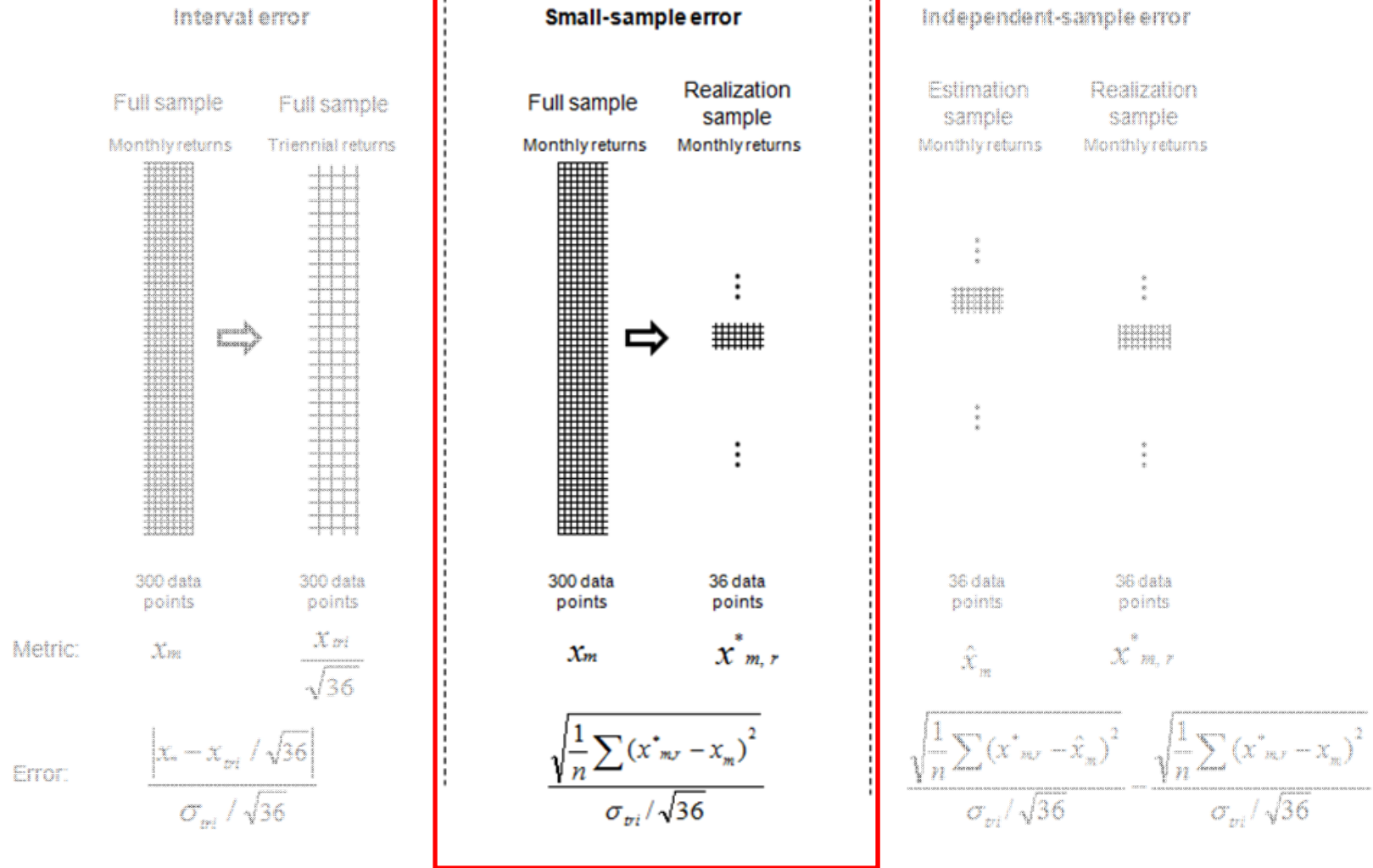
Measuring small-sample error

For a particular return, standard deviation, or correlation:

1. We estimate the parameter based on our full sample of data: x_m
2. We re-estimate the parameter from all realization sub-samples of 36 months: $x_{m,r}^*$
3. We then calculate the root of the average squared difference between the parameter estimated from the full sample and the parameter estimated from the realization samples (normalized by full sample triennial volatility)¹:

$$\frac{\sqrt{\frac{1}{n} \sum (x_{m,r}^* - x_m)^2}}{\sigma_{tri} / \sqrt{36}}$$

Measuring small-sample error



Source: State Street Global Exchange

Measuring small-sample error

Example: Calculating small-sample error

Small-sample error of standard deviation	
Full sample, monthly standard deviation	4.00%
Realization from sample A	6.00%
Error in sample A	$6.00\% - 4.00\% = 2.00\%$
Realization from sample B	3.50%
Error in sample B	$3.50\% - 4.00\% = -0.50\%$
Root of mean squared (RMS) error	$\sqrt{(2.00\% ^2 + -0.50\% ^2) / 2} = 1.45\%$
Full sample, triennial volatility / $\sqrt{36}$	5.00%
RMS error, normalized	$1.45\% / 5.00\% = 29.1\%$

Small-sample realizations introduce a standardized error of 29% relative to the true parameter.

Independent-sample error

The realization of parameters from a particular sample will likely differ from the parameter values of an independent, contiguous sample of the same size.

We call this independent-sample error.

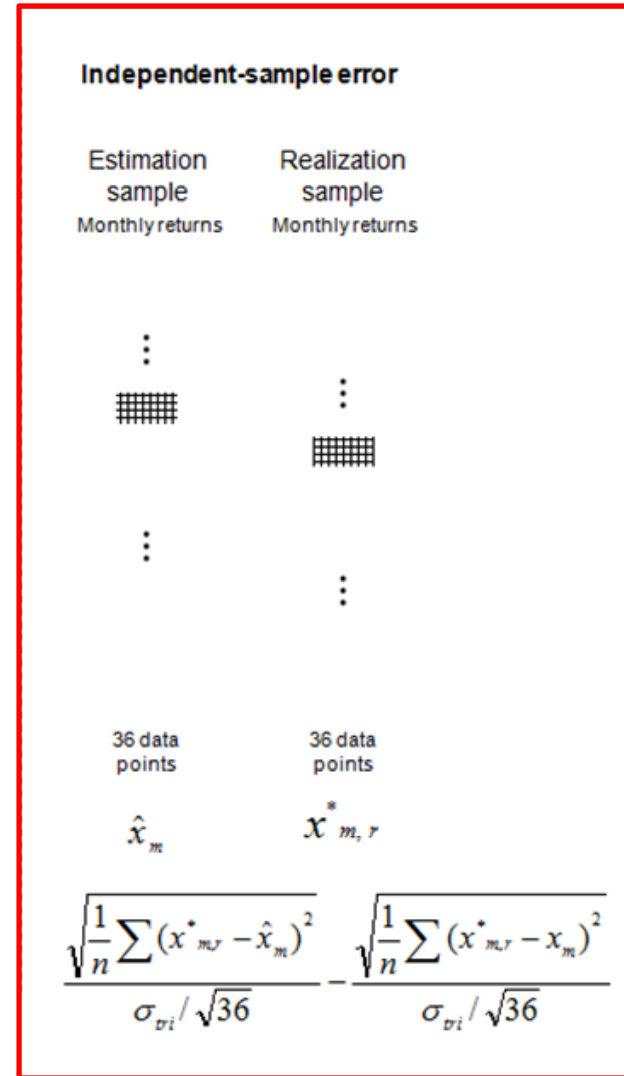
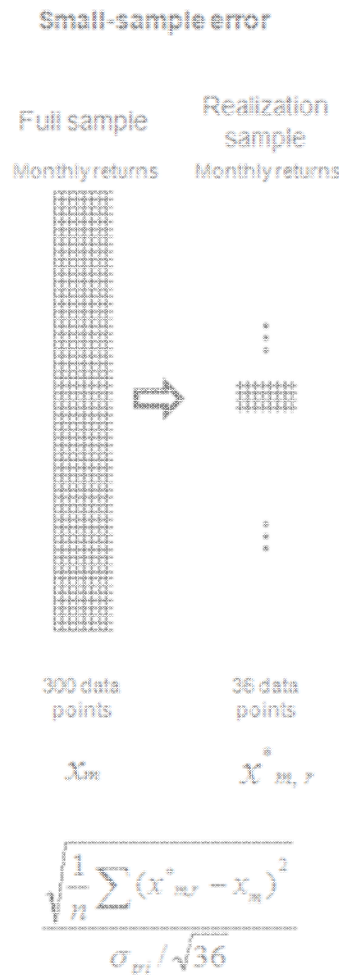
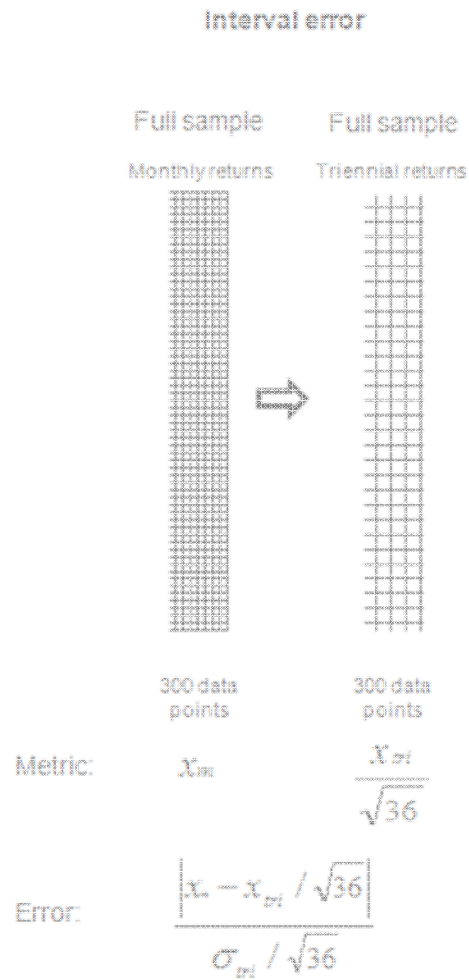
Measuring independent-sample error

For a particular return, standard deviation, or correlation:

1. We estimate the parameter from all estimation samples of 36 months: \hat{x}_m
2. We re-estimate the parameter from all independent, contiguous realization samples of 36 months: $x_{m,r}^*$
3. We then calculate the root of the average squared difference between the parameter estimated from the estimation samples and from the realization samples (normalized by full sample triennial volatility) and we subtract from it small-sample error¹:

$$\frac{\sqrt{\frac{1}{n} \sum (x_{m,r}^* - \hat{x}_m)^2}}{\sigma_{tri} / \sqrt{36}} - \frac{\sqrt{\frac{1}{n} \sum (x_{m,r}^* - x_m)^2}}{\sigma_{tri} / \sqrt{36}}$$

Measuring independent-sample error



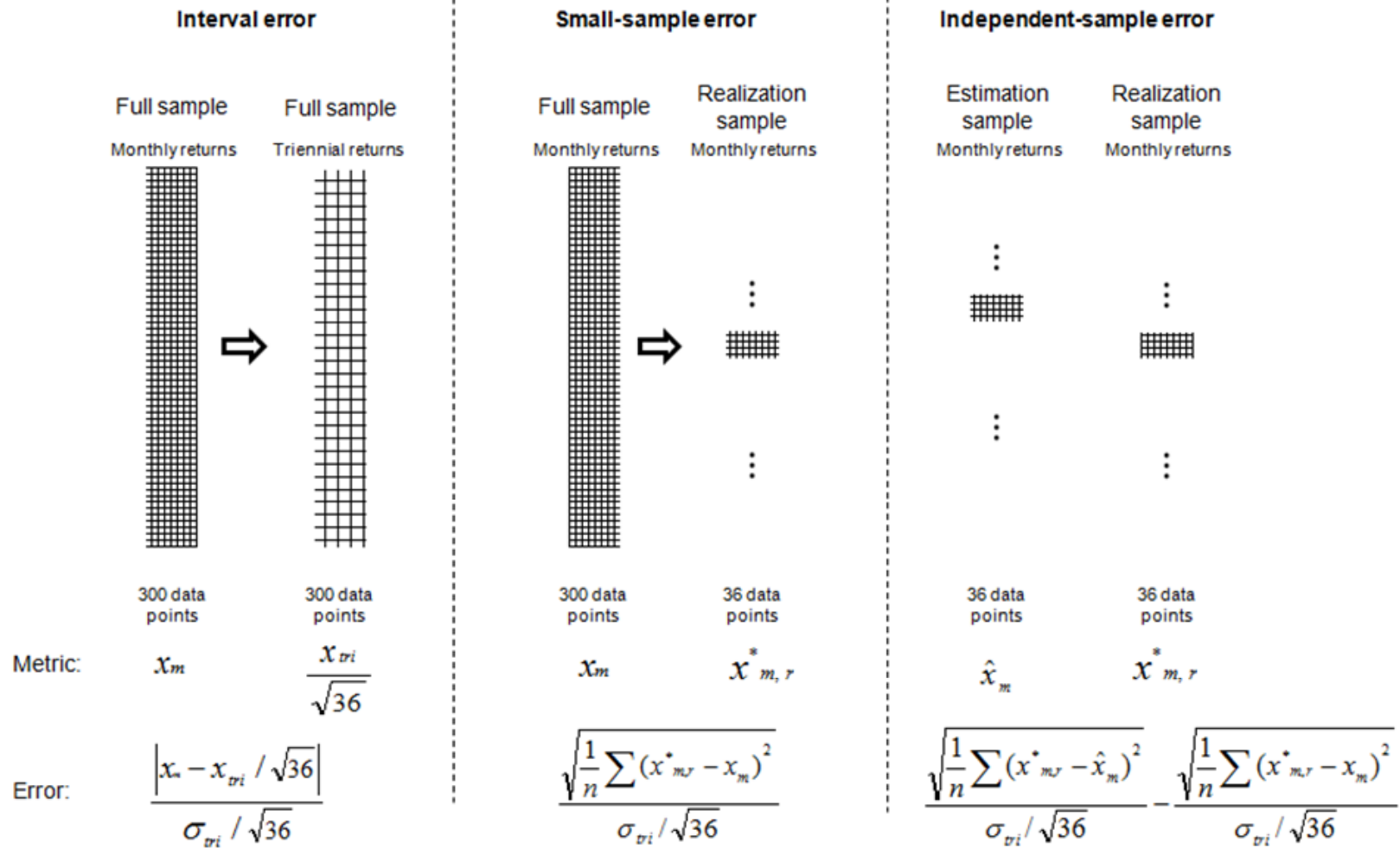
Measuring independent-sample error

Example: Calculating independent-sample error

Independent-sample error of standard deviation	
Estimation from sample A	4.00%
Realization from sample B	6.00%
Error in sample B	$6.00\% - 4.00\% = 2.00\%$
Estimation from sample C	5.50%
Realization from sample D	4.50%
Error in sample D	$4.50\% - 5.50\% = -1.00\%$
Root of mean squared (RMS) error	$\sqrt{(2.00\% ^2 + -1.00\% ^2) / 2} = 1.58\%$
Full sample, triennial standard deviation / $\sqrt{36}$	5.00%
RMS error, normalized	$1.58\% / 5.00\% = 31.6\% - 29.1\% = 2.46\%$

Independent-sample estimates introduce a standardized error of 2.5% relative to the true parameter.

Measuring estimation error

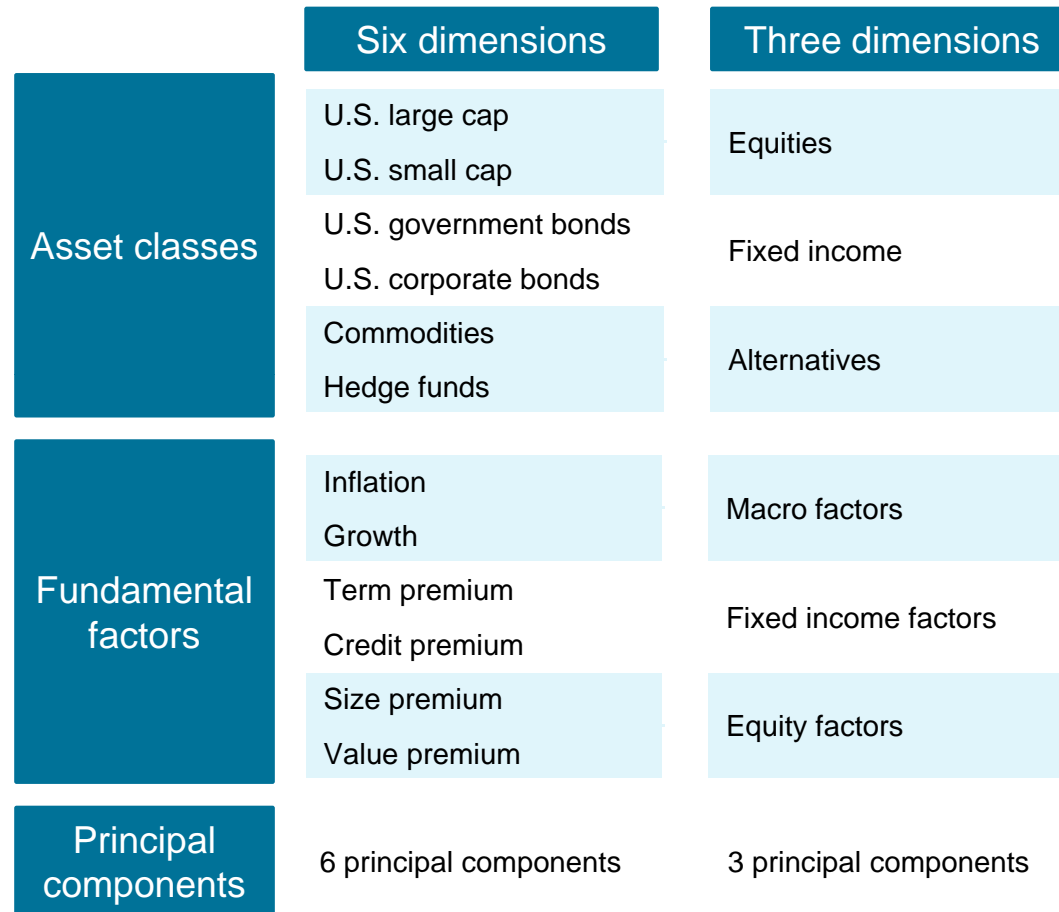


Source: State Street Global Exchange

Classifications and data



Asset classes versus fundamental factors and principal components



Source: State Street Global Exchange

Notes: To stratify using fundamental factors, we use regressions to form six portfolios, each of which is designed to mimic one of six fundamental factors. To stratify using statistical factors, we form six portfolios using principal components analysis (each portfolio corresponds to an eigenvector). To reduce the dimensionality of asset classes we group the six assets into three portfolios using asset class categories. To reduce the dimensionality of fundamental factors, we regress the six assets on three fundamental factors to form three portfolios. To reduce the dimensionality of statistical factors, we reduce the universe to three portfolios representing the top three eigenvectors.

Asset classes versus fundamental factors and principal components

Name	Source	Transformation	Period
Asset classes			
Equities			
U.S. Large Cap	S&P 500 Composite	Log return	Jan 1990 - July 2014
U.S. Small Cap	Russell 2000	Log return	Jan 1990 - July 2014
Fixed Income			
U.S. Government Bonds	Barclays U.S. Aggregate Government	Log return	Jan 1990 - July 2014
U.S. Corporate Bonds	Barclays U.S. Aggregate Corporate	Log return	Jan 1990 - July 2014
Alternatives			
Commodities	S&P GSCI Commodity	Log return	Jan 1990 - July 2014
Hedge Funds	HFRI Fund of Funds Composite	Log return	Jan 1990 - July 2014
Fundamental factors			
Macro			
Inflation	Simple return of U.S. SA CPI	Difference	Jan 1990 - July 2014
Growth	One year ahead U.S. GDP growth forecast	Difference	Jan 1990 - July 2014
Fixed Income			
Term premium	10-Year Minus 2-Year Treasury	Difference	Jan 1990 - July 2014
Credit premium	Baa Corporate Yield to 10-Year Treasury	Difference	Jan 1990 - July 2014
Equity			
Small minus Big	Fama-French SMB factor	Log return	Jan 1990 - July 2014
High minus Low	Fama-French HML factor	Log return	Jan 1990 - July 2014

Industries versus attributes and principal components

We construct a variety of industry and attribute groupings based on the current 400 stocks in the MSCI U.S. Index.

	49 dimensions	24 dimensions	10 dimensions
Industries	49 portfolios formed on GICS level III classifications.	24 portfolios formed on GICS level II classifications.	10 portfolios formed on GICS level I classifications.
Size	49 portfolios formed on market caps.	24 portfolios formed on market caps.	10 portfolios formed on market caps.
Value	49 portfolios formed on book-to-market ratios.	24 portfolios formed on book-to-market ratios.	10 portfolios formed on book-to-market ratios.
Momentum	49 portfolios formed on 1-year moving average returns.	24 portfolios formed on 1-year moving average returns.	10 portfolios formed on 1-year moving average returns.
Principal components	49 principal components	24 principal components	10 principal components

Source: State Street Global Exchange

Notes: We start with a universe of 400 stocks (based on constituents in the MSCI U.S. Index as of Jan 2015). We use average market cap, book-to-market ratios, and 1-year returns to form size, value, and momentum portfolios, respectively. Data covers the period Jan 1989 through Jan 2015.

Results

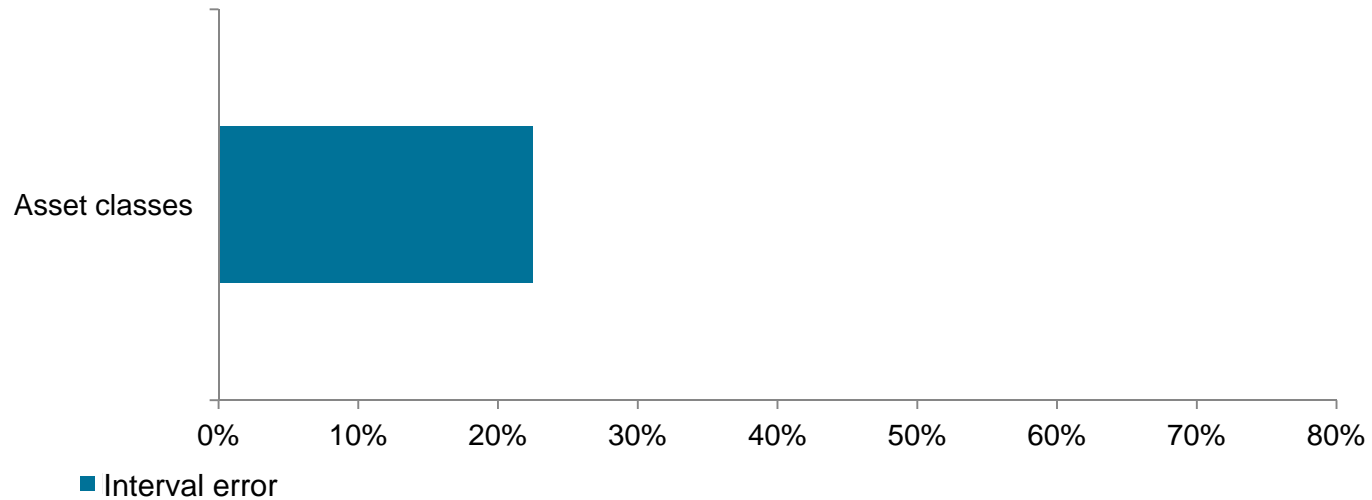
Asset classes versus fundamental factors and principal components



Interpreting the results:

Stationarity of standard deviation

Error due to intervals.



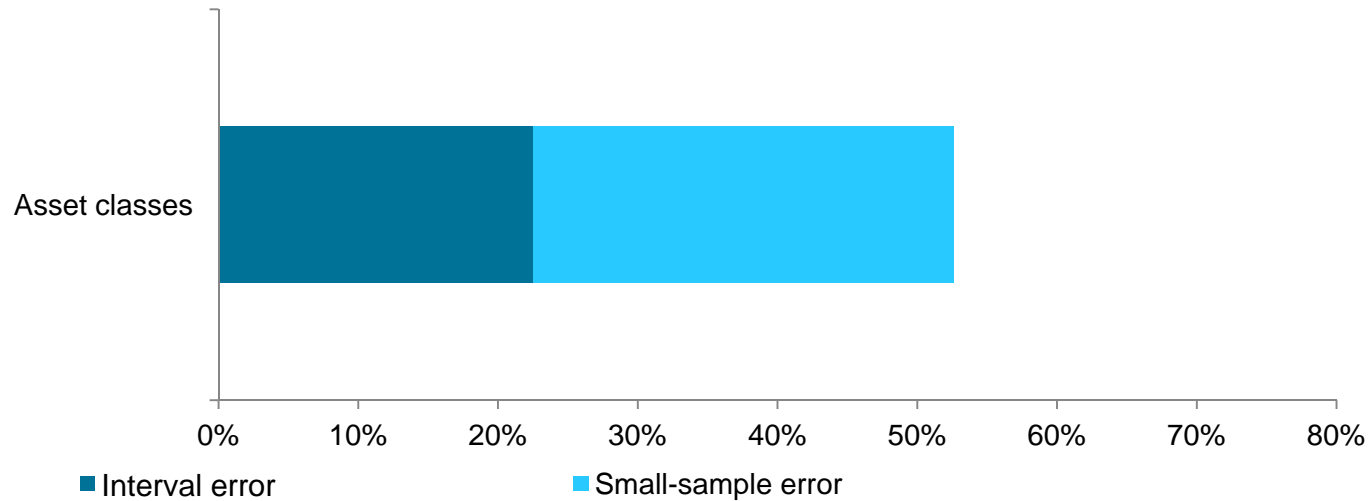
Source: State Street Global Exchange

Notes: Chart shows average result across six asset classes. Please refer to slides 30 and 31 for additional information on data.

Interpreting the results:

Stationarity of standard deviation

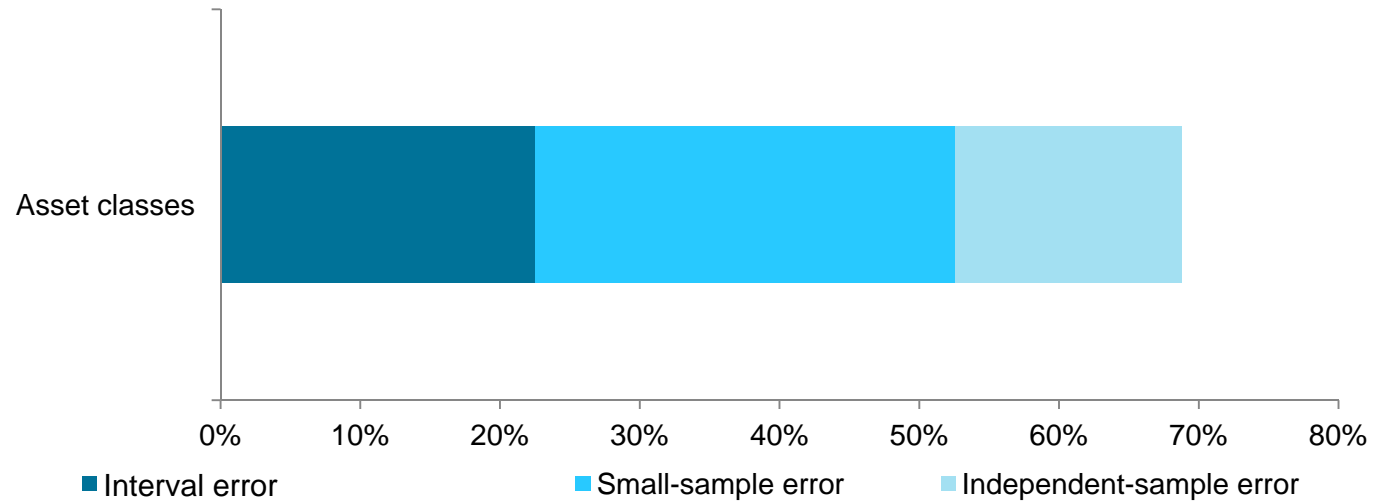
Error due to intervals and small-sample.



Interpreting the results:

Stationarity of standard deviation

Error due to intervals, small-sample, and independent-sample.

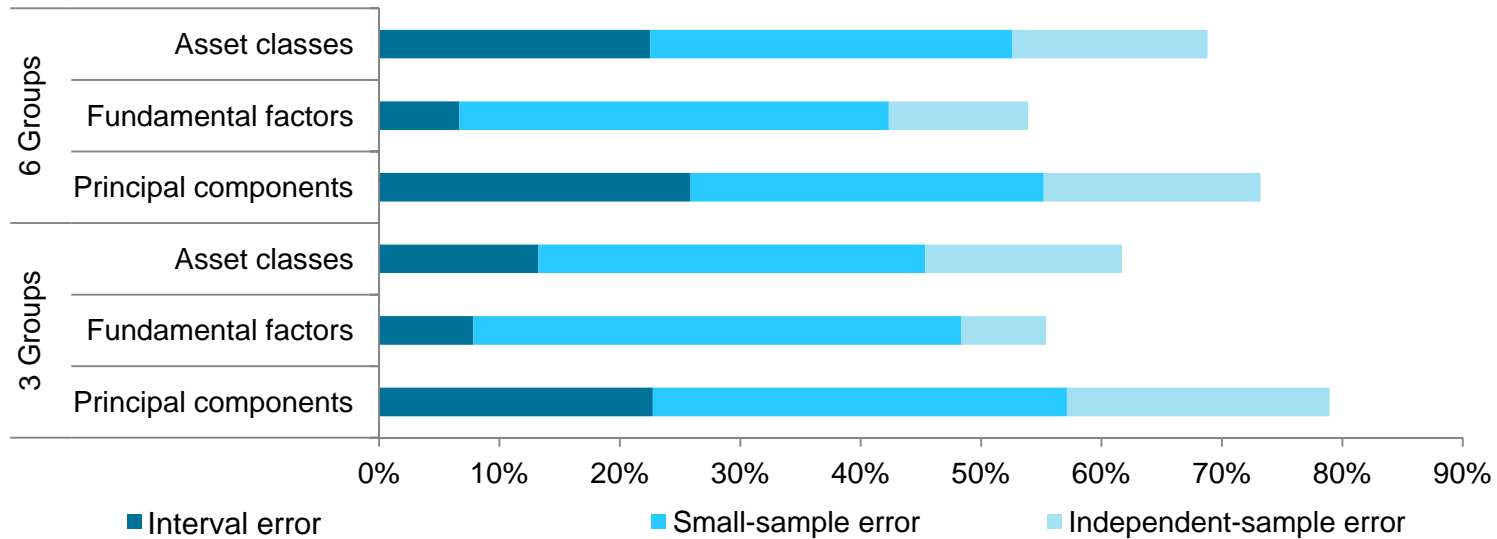


Source: State Street Global Exchange

Notes: Chart shows average result across six asset classes. Please refer to slides 30 and 31 for additional information on data.

Results: Stationarity of standard deviation

Asset classes versus fundamental factors and principal components

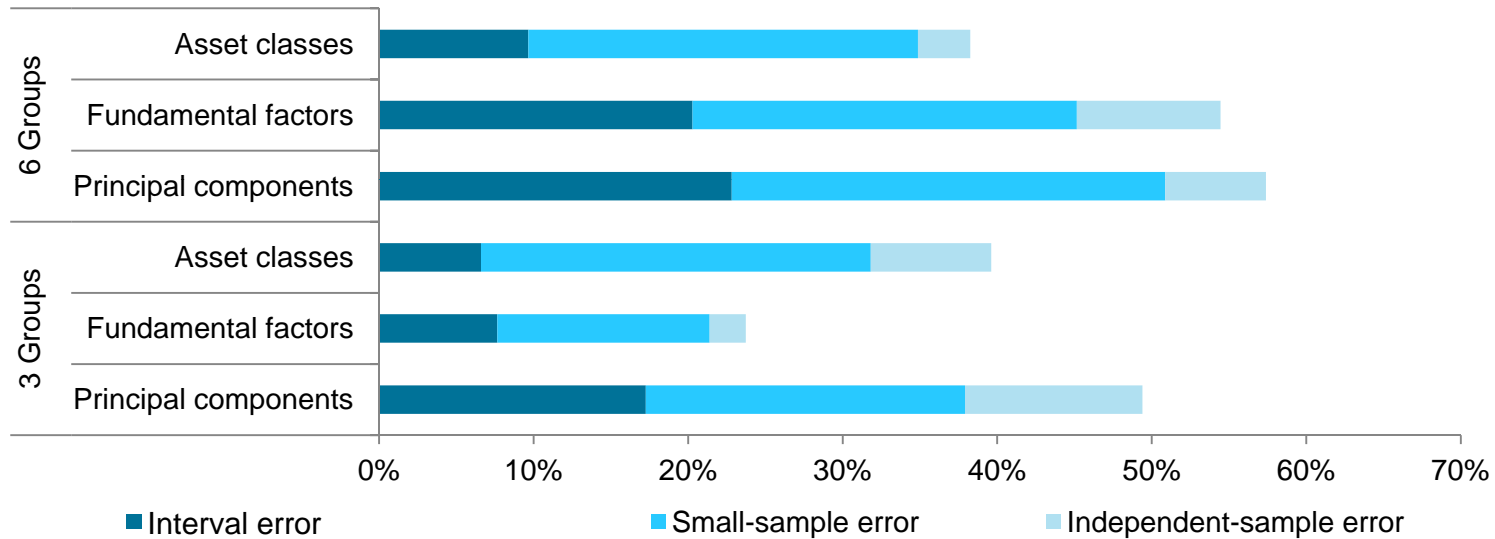


Source: State Street Global Exchange

Notes: Chart shows average result within each stratification (asset classes, fundamental factors, principal components). Please refer to slides 30 and 31 for additional information on data.

Results: Stationarity of correlation

Asset classes versus fundamental factors and principal components

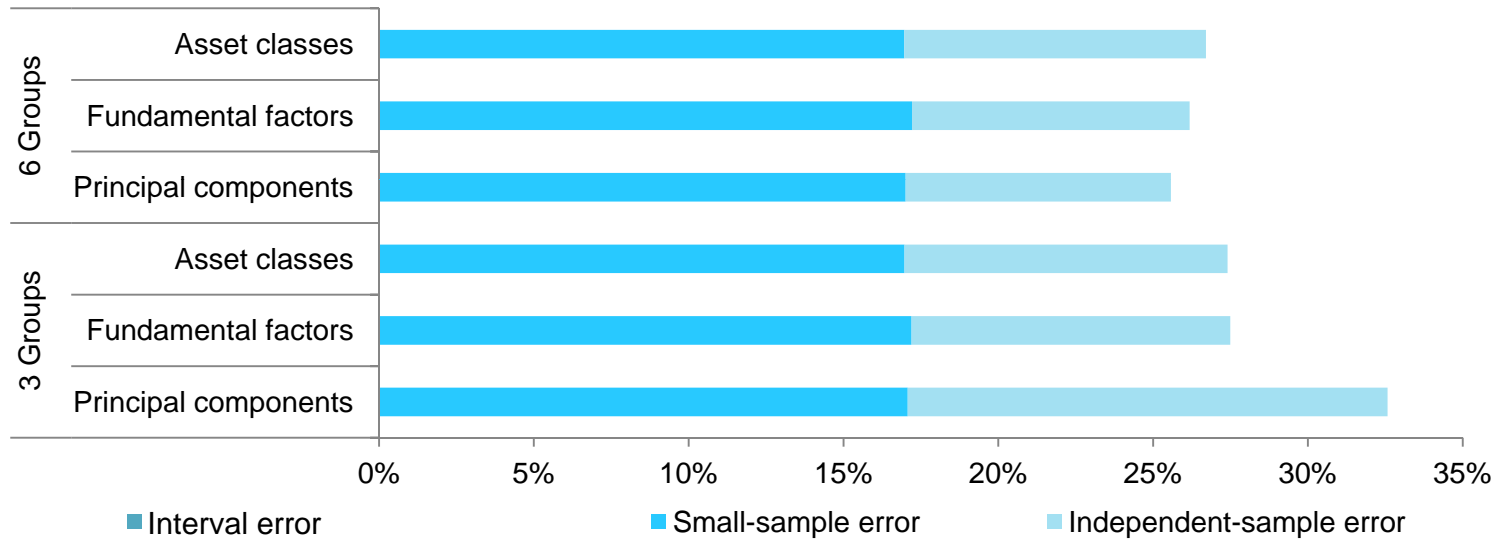


Source: State Street Global Exchange

Notes: Chart shows average result within each stratification (asset classes, fundamental factors, principal components). Please refer to slides 30 and 31 for additional information on data.

Results: Stationarity of returns

Asset classes versus fundamental factors and principal components



Source: State Street Global Exchange

Notes: Chart shows average result within each stratification (asset classes, fundamental factors, principal components). Please refer to slides 30 and 31 for additional information on data.

Results: Summary

Asset classes versus fundamental factors and principal components

Standard deviation

	6 dimensions	3 dimensions
Asset classes	69%	62%
Fundamental factors	54%	63%
Principal components	73%	79%

Correlation

	6 dimensions	3 dimensions
Asset classes	38%	40%
Fundamental factors	54%	24%
Principal components	57%	49%

Returns

	6 dimensions	3 dimensions
Asset classes	27%	27%
Fundamental factors	26%	29%
Principal components	26%	33%

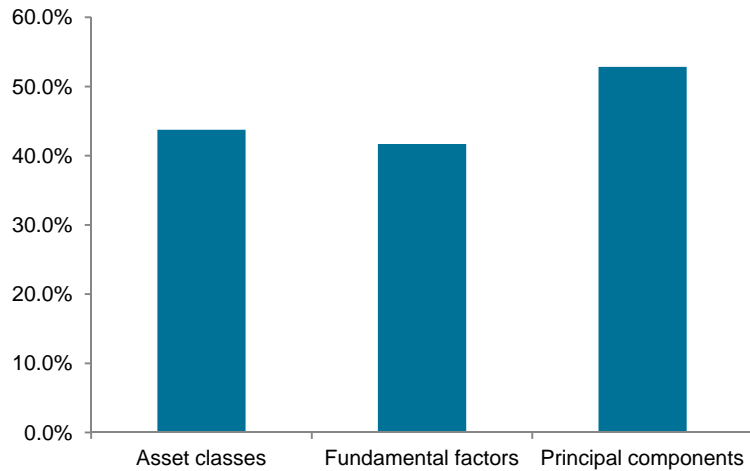
Source: State Street Global Exchange

Notes: Chart shows average total error within each stratification (asset classes, fundamental factors, principal components). Please refer to slides 30 and 31 for additional information on data.

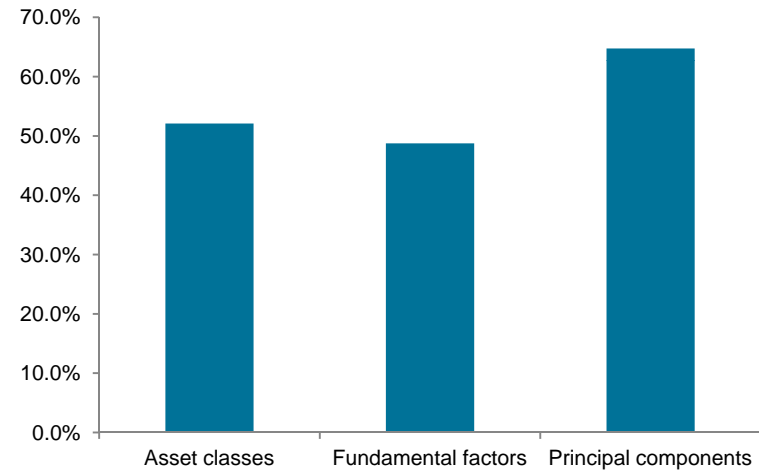
Results: Summary

Asset classes versus fundamental factors and principal components

Standard Deviation, Correlation, and Return



Standard Deviation and Correlation

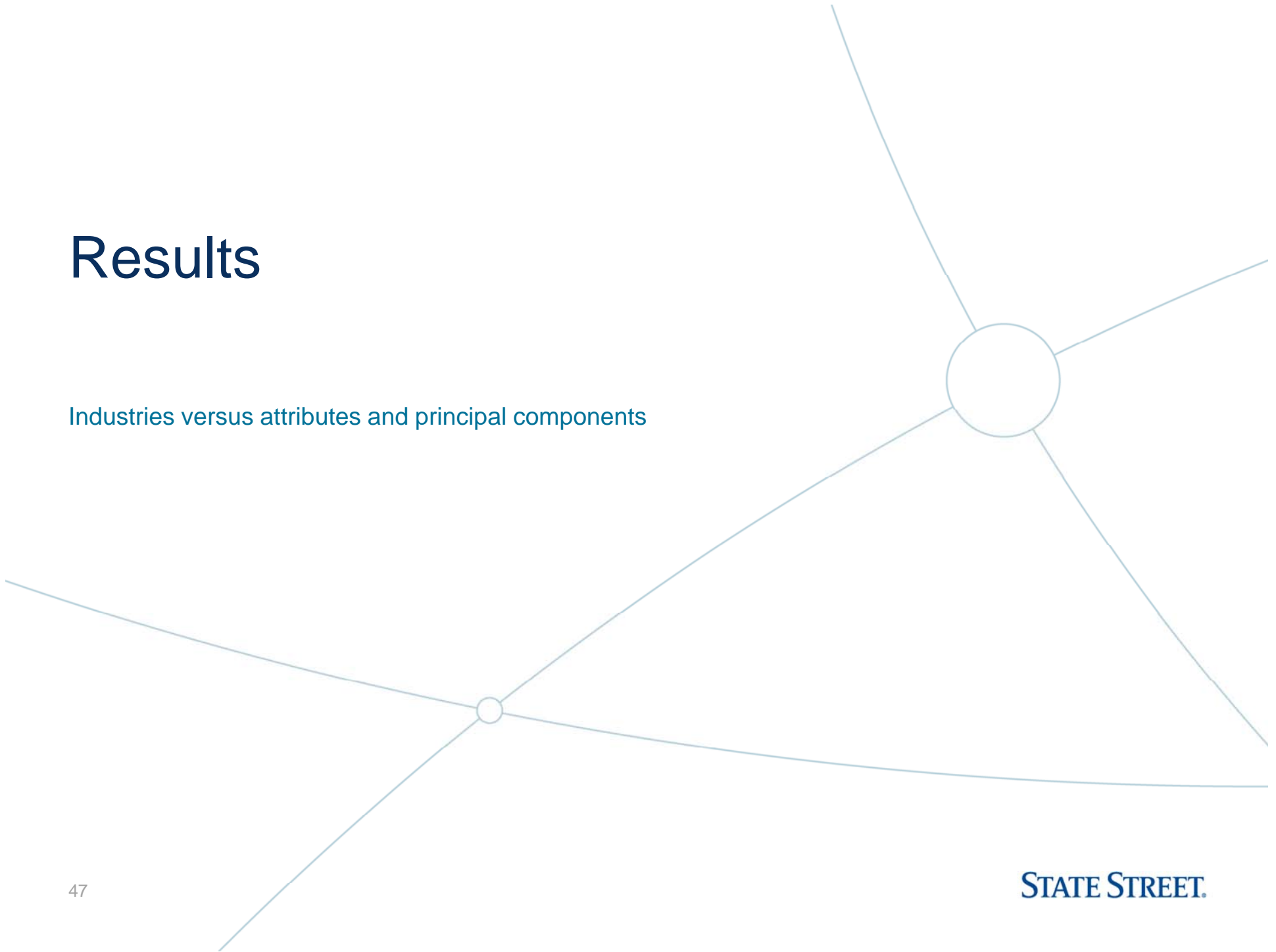


Source: State Street Global Exchange

Notes: Chart shows average total error within each stratification (asset classes, fundamental factors, principal components). Please refer to slides 30 and 31 for additional information on data.

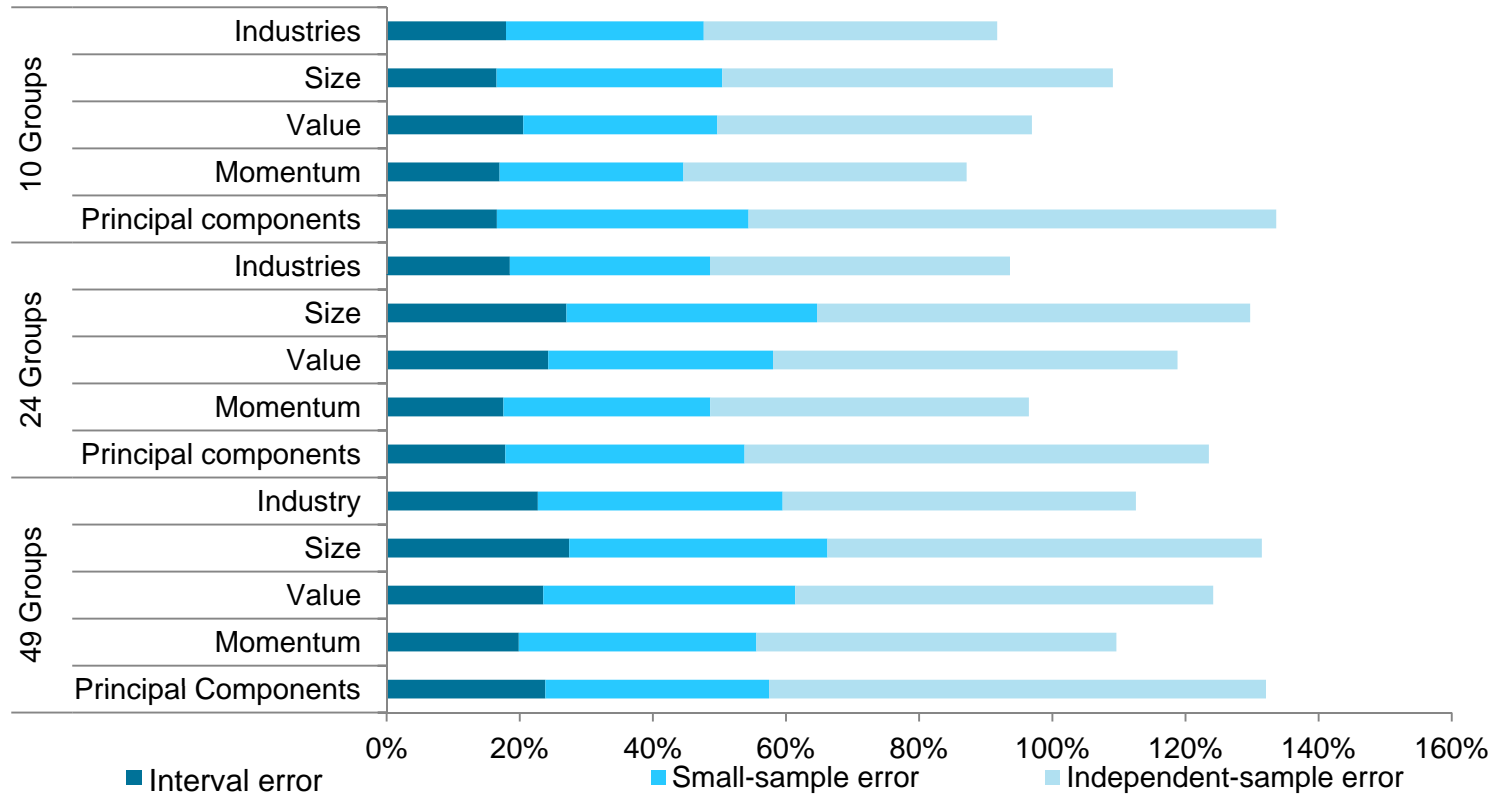
Results

Industries versus attributes and principal components



Results: Stationarity of standard deviation

Industries versus attributes and principal components

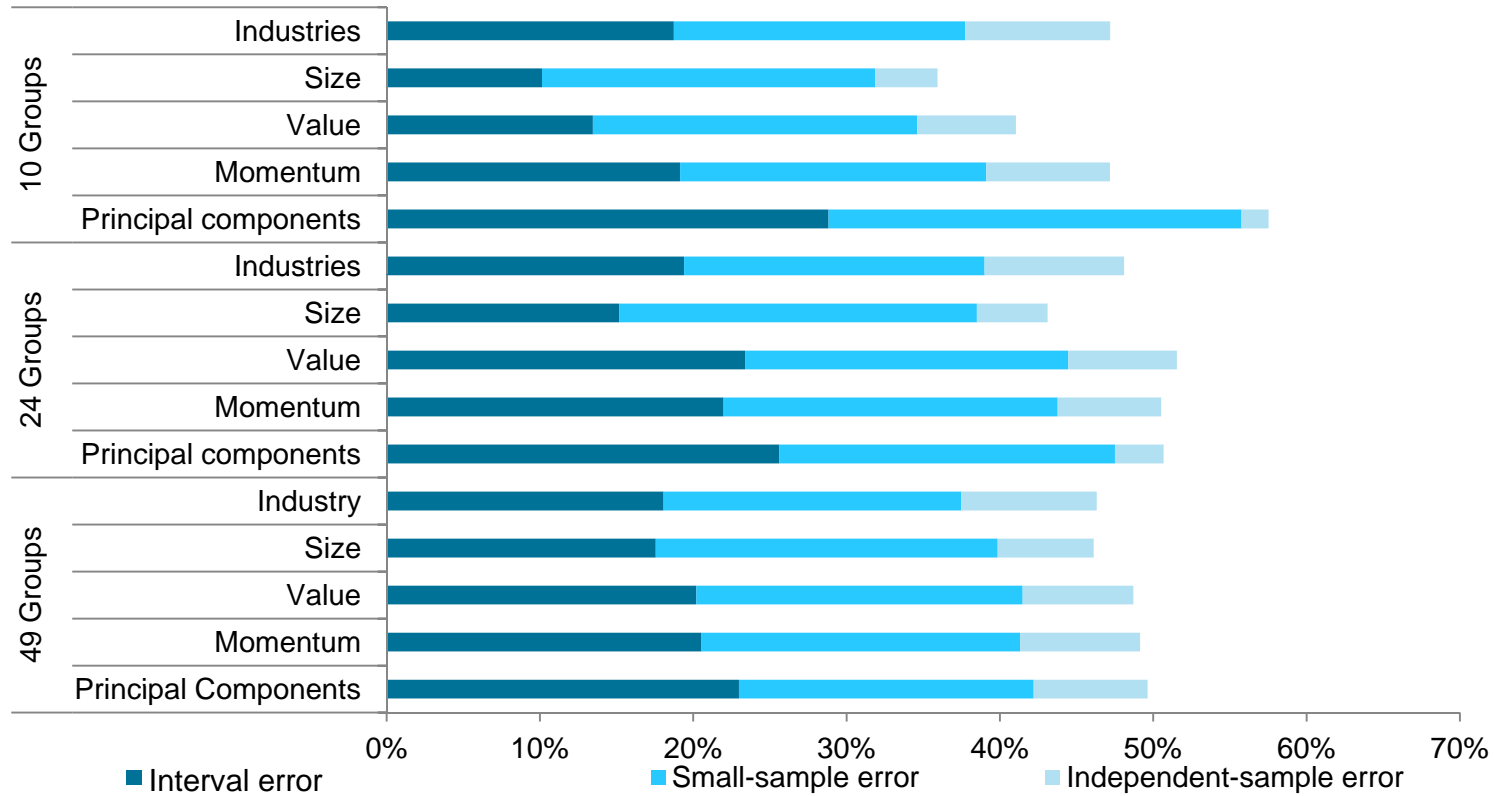


Source: State Street Global Exchange

Notes: Chart shows average result within each stratification (industries, size, value, momentum, principal components). Please refer to slide 32 for additional information on data.

Results: Stationarity of correlation

Industries versus attributes and principal components

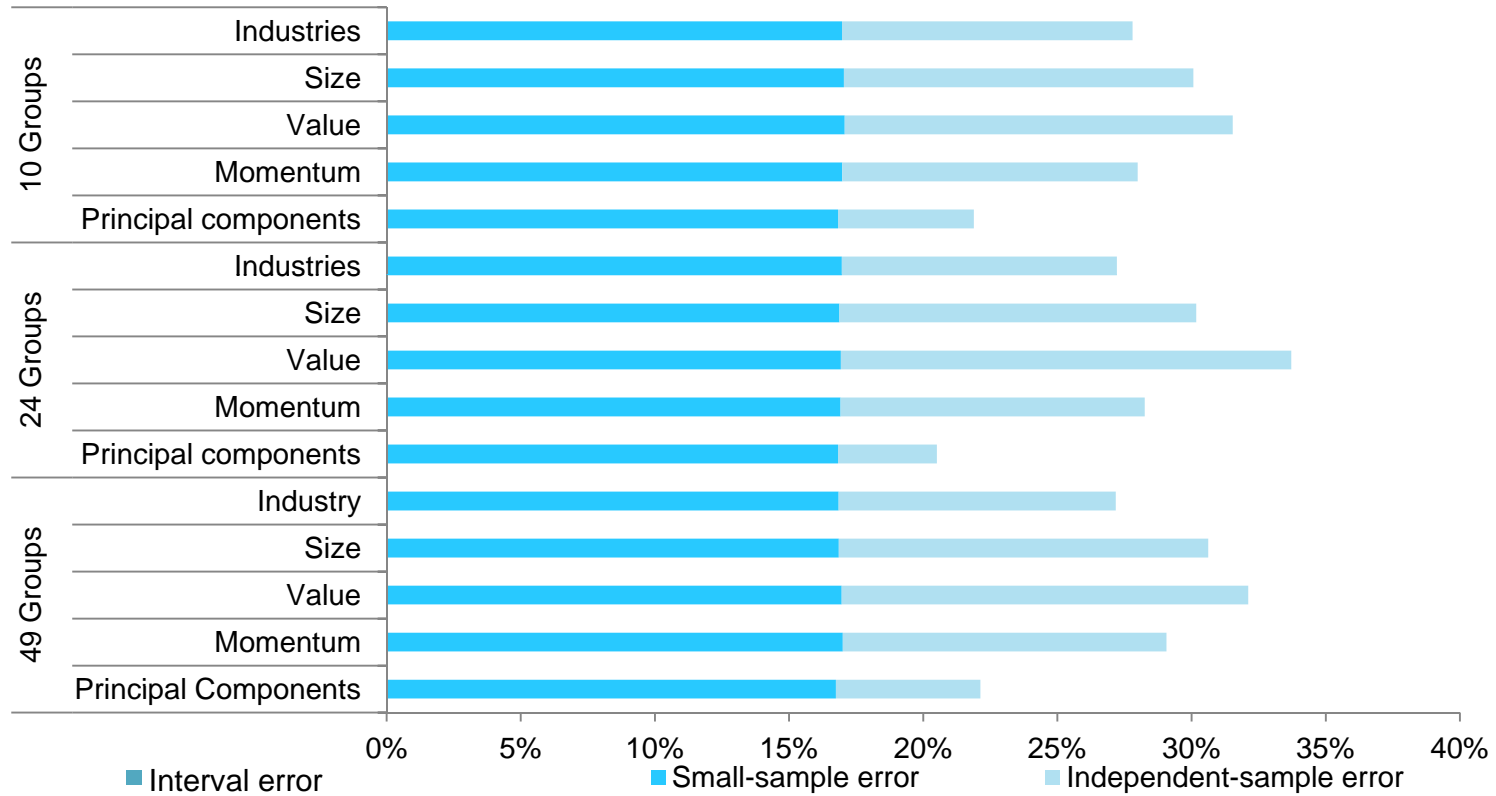


Source: State Street Global Exchange

Notes: Chart shows average result within each stratification (industries, size, value, momentum, principal components). Please refer to slide 32 for additional information on data.

Results: Stationarity of returns

Industries versus attributes and principal components



Source: State Street Global Exchange

Notes: Chart shows average result within each stratification (industries, size, value, momentum, principal components). Please refer to slide 32 for additional information on data.

Results: Summary

Industries versus attributes and principal components

Standard deviation (400 stocks: 78%)

	49 dimensions	24 dimensions	10 dimensions
Industries	78%	69%	65%
Size	102%	102%	105%
Value	82%	78%	60%
Momentum	69%	64%	60%
Principal components	98%	88%	96%

Correlation (400 stocks: 48%)

	49 dimensions	24 dimensions	10 dimensions
Industries	46%	48%	47%
Size	46%	43%	36%
Value	49%	52%	41%
Momentum	49%	51%	47%
Principal components	50%	51%	58%

Returns (400 stocks: 27%)

	49 dimensions	24 dimensions	10 dimensions
Industries	27%	27%	28%
Size	31%	30%	30%
Value	32%	34%	32%
Momentum	29%	28%	28%
Principal components	22%	21%	22%

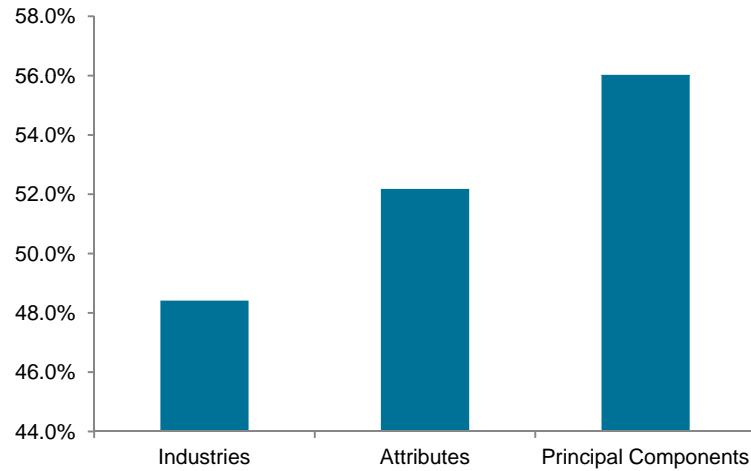
Source: State Street Global Exchange

Notes: Chart shows average total error within each stratification (industries, size, value, momentum, principal components). Please refer to slide 32 for additional information on data.

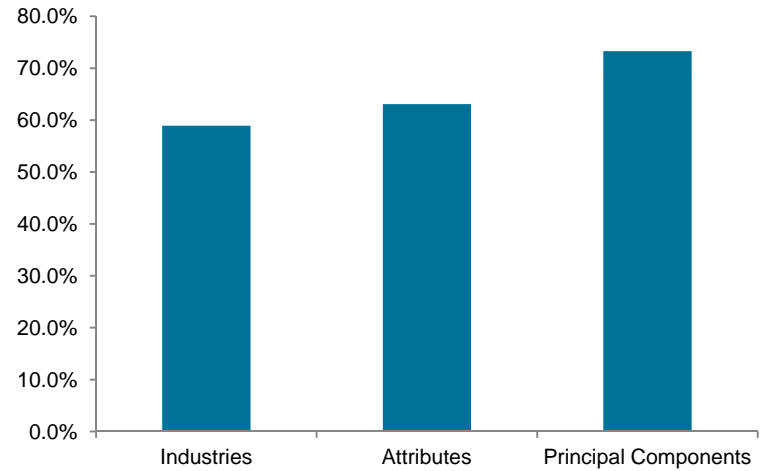
Results: Summary

Industries versus attributes and principal components

Standard Deviation, Correlation, and Return



Standard Deviation and Correlation



Source: State Street Global Exchange

Notes: Chart shows average total error within each stratification (industries, size, value, momentum, principal components). Please refer to slides 32 for additional information on data.

Noise reduction

- A priori, a more granular set of assets or factors will produce better results in-sample than a less granular set to the extent the additional assets or factors are not purely redundant.
- When we move out-of-sample, however, the more granular information may degrade more severely than the composite information because it is less stationary.
- Should we take a more granular approach to portfolio construction in order to capture additional information, noisy though it may be, or should we approach portfolio construction in a more consolidated way, thereby sacrificing information in favor of noise reduction?
- Does consolidating a larger set of assets into a smaller set of factors reduce noise more effectively than consolidating a larger set of assets into a smaller set of assets?

Noise reduction for industries and attributes

Standard Deviation, Correlation, and Return

	49 groups	24 groups	10 groups
Industries	4%	20%	27%
Attributes	12%	22%	30%

Standard Deviation and Correlation

	49 groups	24 groups	10 groups
Industries	3%	11%	19%
Attributes	8%	10%	22%

Source: State Street Global Exchange

Notes: Tables show the average percentage reduction in non-standardized errors. In this analysis, we do not divide the errors by standard deviation because doing so would obscure the level of noise, as the average standard deviation of individual securities is greater than the average standard deviation of industries and attributes. Please refer to slides 35, 36 and 37 for information on data.

Summary

- It has become fashionable to use factors instead of assets as the building blocks for forming portfolios.
- Some have argued that factors offer greater potential for diversification, but this argument is specious.
- Others argue that it is easier to relate past and current information to future factor behavior than to asset behavior, but this argument is generally untestable.
- Nonetheless, it may be the case that factor parameters are more stationary than asset parameters.
- We find no evidence that factors produce more stable results taking into account interval error, small-sample error, and independent-sample error.
- Finally, we find no compelling evidence that factors reduce security-level noise more effectively than assets.

Conclusion

Why use factors?

Conclusion

Why use factors?

- In our view, the case is yet to be made that investors should use factors rather than assets as building blocks for forming portfolios.
- However, investors may be able to gather useful insights about the performance of their portfolios by attributing performance to factor exposures in addition to asset exposures.
- And even if investors are no more skilled at forecasting factor behavior than asset behavior, understanding a portfolio's factor exposures may help investors to hedge exposures to particular factors more effectively.

Disclaimers and Important Risk Information

State Street Global ExchangeSM is a trademark of State Street Corporation (incorporated in Massachusetts) and is registered or has registrations pending in multiple jurisdictions.

This information is for general, marketing and/or informational purposes only and it does not constitute investment research or investment, legal, or tax advice, and it is not an offer or solicitation to buy or sell any product, service, or securities or any financial instrument, and it does not constitute any binding contractual arrangement or commitment of any kind. State Street Corporation and its subsidiaries and affiliates (“State Street”) provide products and services to professional and institutional clients, and this is not directed at retail clients. Any opinions expressed in this document are subject to change without notice. This information has been prepared and obtained from sources believed to be reliable at the time of publication, however it is provided “as-is” and State Street makes no guarantee, representation, or warranty of any kind as to its accuracy, suitability, timeliness, merchantability, fitness for a particular purpose, non-infringement of third-party rights, etc. This information is not intended to be relied upon by any person or entity. State Street disclaims all liability, whether arising in contract, tort or otherwise, for any losses, liabilities, damages, expenses or costs arising, either direct or consequential, from or in connection with the use of this document and/or the information herein. No permission is granted to reprint, sell, copy, distribute, or modify any material herein, in any form or by any means without the prior written consent of State Street.

This document contains certain statements that may be deemed forward-looking statements. Please note that any such statements are not guarantees of any future performance, and actual results or developments may differ materially from those projected. Investing involves risk, including the risk of loss of principal. Diversification does not ensure a profit or guarantee against loss. This document does not take into account any investor’s particular investment objectives, strategies, or tax and legal status, relevant regulations, nor does it purport to be comprehensive or intended to replace the exercise of an investor’s own careful independent review regarding any corresponding investment decision or related analysis. Any forecasted information in the document is not a reliable indicator for future performance.

Copyright © 2015 State Street Corporation, All rights reserved.