INFORMATION COEFFICIENTS AND LINEAR FACTOR MODELS

Oliver Buckley
INVESCO

This paper focuses on techniques for evaluating the effectiveness of stock selection factors in a multiple factor framework and for combining factors together to form an overall "alpha" model. These techniques are evaluated using the assumption of a linear model – that is, expected returns are a linear function of each asset's exposure to each factor. To make comparisons across factors, the paper assumes that the exposures have a cross-sectional mean of zero and a cross-sectional standard deviation of one.

The paper starts with the simplest concept: one factor over a single holding period (typically one month). The three techniques used are fractile analysis, information coefficients, and regression coefficients. Fractile analysis, most commonly performed with deciles or quintiles, divides stocks into ranked groups based on exposures to the factor at the beginning of the holding period and calculates the return to each group over the holding period. The information coefficients (IC) is the correlation between exposures and holding period returns. The regression coefficient is the coefficient estimated by regressing holding period returns against exposures and an intercept. The IC is just the square root of the regression R-squared (or R).

It is easy to show that regression coefficient is simply the information coefficient scaled by the cross-sectional volatility of holding period returns divided by the cross-sectional standard deviation of exposures:

$$\text{Regression\_Coefficient} = \frac{\text{Return\_Volatility}}{\text{Factor\_Std.\_Deviation}} * IC$$

The regression coefficient can also be shown to be the return to a "factor" portfolio that is the minimum variance portfolio with zero weight and a factor exposure of 1. Since we have assumed unit standard deviation for our exposures, we can think of this factor portfolio return as arising from two sources: skill (the factor IC) and opportunity (the cross-sectional return volatility). A factor with a high IC may not have much value if applied in a market with low cross-sectional volatility, because stock returns are then all quite similar within the market.

To show the relationship between ICs and fractile returns, I performed a simulation in which I generated many samples from a bivariate normal distribution. One variable in the distribution represents returns, the other exposures, and the correlation between the two represents the IC. By varying the IC, the cross-sectional volatility of returns, and the number of assets in the market, it is possible to understand how these parameters change the return spread between the first and last quintile or decile of the factor distribution.

These simulations demonstrate that over a reasonable range, the quintile spread is linear in IC and linear in the cross-sectional return volatility. This is identical to the result shown for the regression parameter, which is the product of the IC and cross-sectional volatility.

Multiple factor linear models create a composite factor that is just the weighted sum of the simple factors. For a single holding period, simple algebraic manipulation shows that the IC of the composite factor is the weighted average of the simple factor ICs, scaled by a constant related to the weights and the cross-sectional correlation matrix of factor exposures. The correlation of the exposures measures how similar the factor attributes are at a point in time. Because we have assumed that all factors have an exposure standard deviation of 1, the constant is always less than one, so that the composite IC is always at least as great as the weighted average IC. Thus, there is a "diversification" benefit related to expected return by having multiple factors in the model.

As an example, suppose we have two uncorrelated factors, each with an IC of 0.03, and we create an equally weighted composite factor. In this case, the IC of the composite factor is 0.03 times the square root of two (0.042). If the factors have an exposure correlation of 0.9, the equally weighted composite has an IC of only 0.031, which reflects the fact that the two factors are capturing similar attributes, and so little additional value can be gained by combining them.

A natural idea is to find the set of weights that maximizes the IC. Unfortunately, there is no unique solution to the problem, because one can multiply one optimal sets of weights by any positive scalar and not change the IC. However, if the weights are constrained to sum to 1, the solution is unique, and is a scalar multiple of the weights from the corresponding multiple regression of returns against the factor exposures and an intercept. Thus, finding the maximum IC is equivalent mathematically to performing this regression, and the maximum IC is (as in the single factor case) the square root of the regression R-squared.

Once again, the regression coefficient can be thought of as the return to a "factor" portfolio. Here, the factor portfolio not only has zero weight and unit exposure to the factor, but has zero exposure to all the other factors included in the model. This last property means that the multiple regression coefficient can differ substantially from the simple regression coefficient, and that the optimal weights from maximizing the IC are not proportional to the individual ICs. The structure of the exposure correlation matrix can have a significant effect on the optimal weights.

The concept of maximizing the IC can be used to construct factor weights for calculating an overall alpha. To do this, we need a forecast of the IC for each factor (the example just uses the historical time series of ICs, with an exponential smoothing parameter of 0.97) and the current exposure correlation matrix, since it is only the current exposures that determine how the factors will interact next period. Since this is a straightforward optimization, it is easy to include constraints such as all non-negative weights or upper bounds on weights.

This approach compares favorably with using historical regression coefficients to determine forecast weights. First, it is easy to add or drop factors at any point in the process, because only the IC is computed for each factor without knowledge of the other factors. Second, the regression coefficients embed the historical correlation matrix, which may have changed dramatically through time. Third, the regression coefficients embed any historical relationships between the performance of the factor and market cross-sectional volatility. Finally, this approach makes it easy to add and to understand the impact of constraints.

Information Coefficients and Linear Factor Models

7th Annual Northfield Client Conference

October 25, 1996

Presented by

Oliver Buckley

INVESCO®

Overview

- Background
- Analysis Techniques
- Single Factors
- Multiple Factors
- Forecasting

Background

- Factor Exposures capture expected return response (ex-ante)
- Linear model implies expected return response is linear
- Equal weighting of stocks (corresponds to OLS regression)

Analysis Techniques

- Single Holding Period
- Fractile Analysis (Quintiles, Deciles)
  - Divide stocks into ranked groups based on exposures (equal number per group)
  - Measure holding period returns for each group
- Information Coefficients
  - Correlation between exposures and holding period returns (IC_k)
  - Regression R (square root of R-squared)
- Simple Regression
  - Coefficient in regression of holding period returns against exposures (β_k)

Single Factors - Information and Regression Coefficients

- Regression coefficient is scaled Information Coefficient

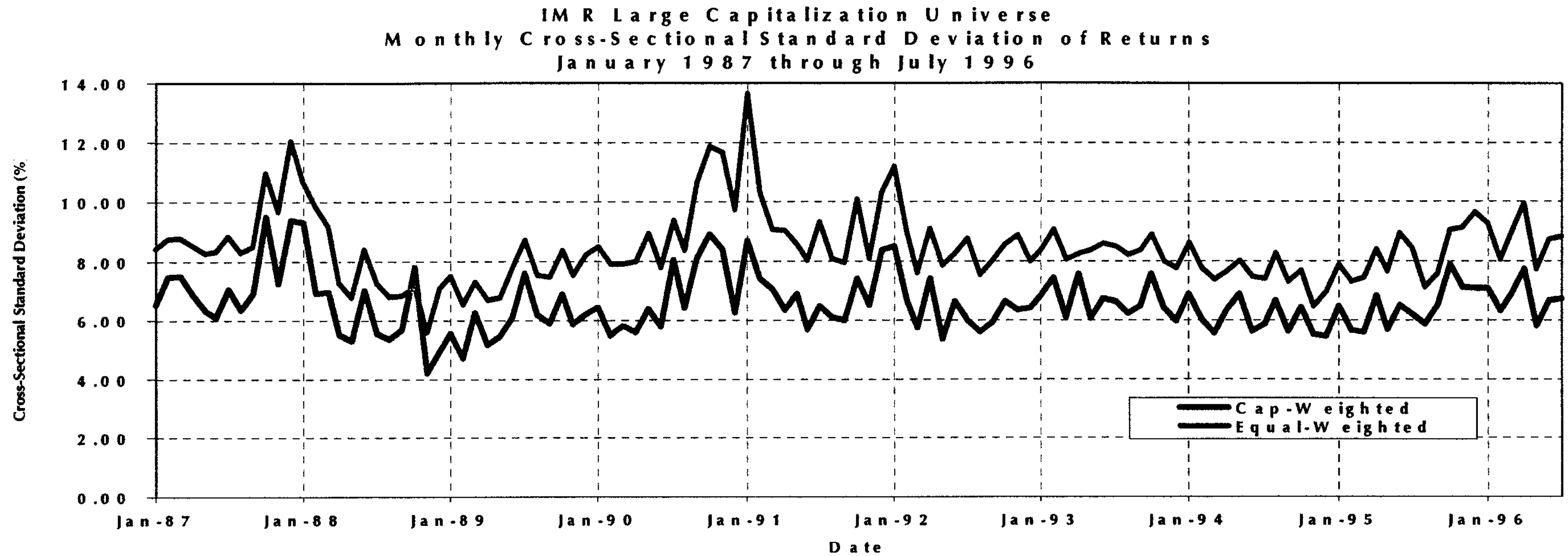
$$\beta_k = \frac{\sigma_r}{\sigma_k} \cdot IC_k$$

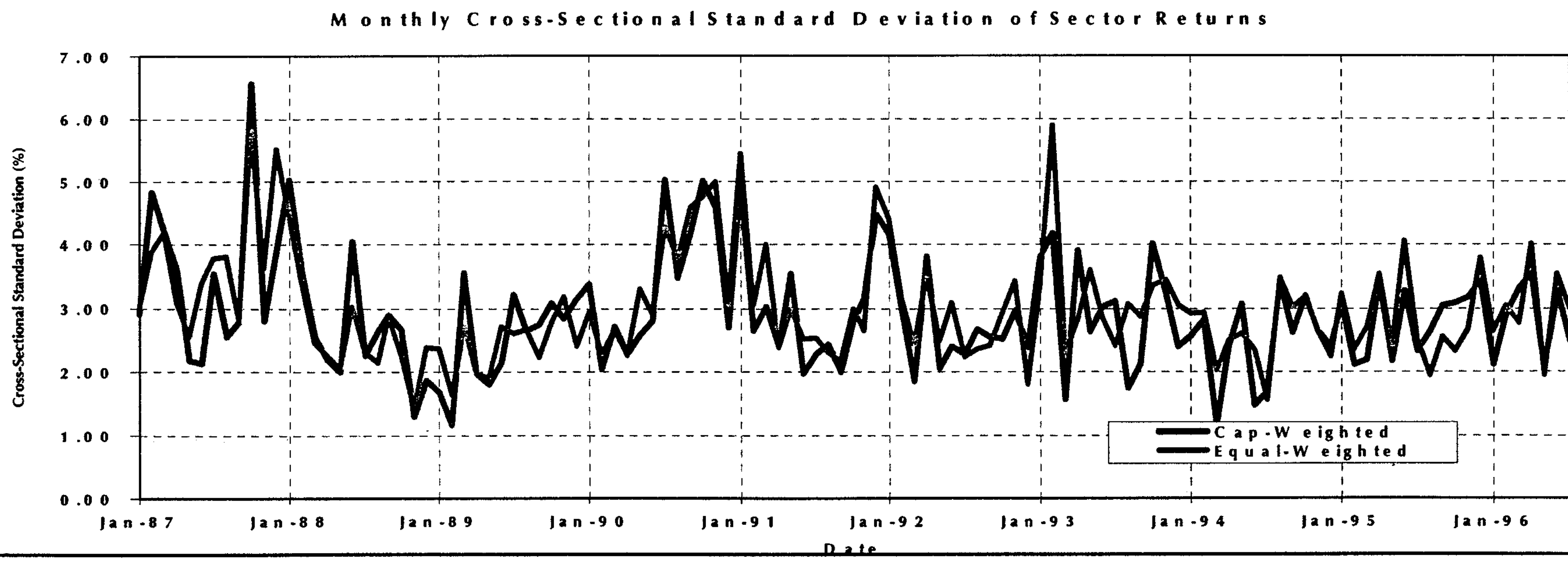
- Scale factor exposures to have unit standard deviation
- Regression coefficients can be interpreted as return to "factor portfolio" which has:
  - Zero weight
  - Factor exposure of 1
  - Minimum (residual) variance portfolio with those characteristics
- Factor Portfolio Return = Skill * Opportunity

Single Factors - Information Coefficients and Quintile Returns

- Assume asset returns are normally distributed
- Generate samples from bivariate normal distribution with given correlation
  - One variable represents returns (must supply volatility)
  - One variable represents exposures
  - Correlation between variables is IC
- Divide into quintiles based on exposures
- Calculate spread between first and last quintile
- Repeat often

Single Factors - U.S. Largecap Stock and Sector Volatilities





INVESCO®

Single Factors - Information Coefficient Simulation Results

Information Coefficient Simulations Equal-Weighted Quintile1/Quintile 5 Spreads						
	1000 Assets/8% Volatility 3000 Months		20 Assets/8%Volatility 21000 Months		20 Assets/3% Volatility 21000 Months	
Information Coefficient	Average Monthly Return (BP)	Percent Positive Months	Average Monthly Return (BP)	Percent Positive Months	Average Monthly Return (BP)	Percent Positive Months
0.00	0.6	50.9	-5.7	49.2	-2.1	49.2
0.01	22.9	61.3	17.8	51.1	6.7	51.1
0.02	45.2	70.6	45.5	53.5	17.1	53.5
0.03	63.8	78.2	64.6	54.3	24.2	54.3
0.05	112.6	92.3	109.0	57.5	40.9	57.5
0.07	157.2	97.7	146.0	60.4	54.7	60.4
0.10	224.5	99.6	214.4	64.8	80.4	64.8
0.20	443.9	100.0	430.8	78.0	161.5	78.0
Avg. Volatility	80.2		562		210.7	

- Quintile spread is linear in IC
- Quintile spread is linear in cross-sectional volatility
- Sample volatility shrinks with the square root of number of assets

INVESCO®

Single Factors - Time Series Comparison of IC and Regression

Sample "Value" Factor
Information Coefficient and Regression Statistics
January 1987 through July 1996

Economic Sector	Average Monthly Cross-Sectional Standard Deviation (%)	Average Monthly Information Coefficient	Average Value Added (BP)	Average Regression Coefficient (BP)	Larger Value	IC/Sector Volatility Correlation
Non-Energy Minerals	8.1	0.020	16.2	20.6	REG	0.096
Energy Minerals	7.1	0.056	39.7	41.4	REG	0.050
Producer Manufacturing	7.6	0.021	16.0	15.9	IC	-0.002
Process Industries	7.2	0.016	11.5	14.3	REG	0.140
Commercial Services	7.6	-0.001	-0.5	-3.1	IC	-0.056
Electronic Technology	10.9	0.004	4.8	2.3	IC	-0.061
Health Technology	9.2	0.010	9.4	12.2	REG	0.066
Technology Services	11.4	-0.013	-14.5	-2.2	REG	0.131
Health Services	9.7	-0.013	-12.3	-9.8	REG	0.033
Consumer Durables	8.5	0.014	11.7	21.1	REG	0.220
Consumer Non-Durables	7.0	-0.008	-5.8	-4.3	REG	0.067
Consumer Services	8.5	0.003	2.3	-1.0	IC	-0.097
Retail Trade	9.2	-0.003	-2.5	-3.6	IC	-0.037
Transportation	8.5	0.033	28.4	25.9	IC	-0.046
Non-Bank Finance	7.1	0.021	14.7	16.1	REG	0.055
Utilities	4.5	0.021	9.5	10.0	REG	0.025
Banks	6.2	0.040	24.6	20.8	IC	-0.085
Defense/Aerospace	7.3	0.070	51.3	51.3	IC	0.000
Telephones	6.8	0.003	1.8	2.1	REG	0.004

- Forecast volatility?

Multiple Factors - Combining ICs

- Create composite factor

$$F = \Sigma \, w_k \bullet x_k$$

- Algebraic manipulation

$$IC(F) = \frac{\Sigma \, w_k \bullet IC_k}{\sqrt{\Sigma\Sigma \, w_i w_k \rho_{ik}}} = \frac{w^T (IC)}{\sqrt{w^T C w}}$$

- Diversification benefit with multiple factors

- Model of expected return, not risk!

Multiple Factors - Combining ICs

- Equally weight two uncorrelated factors, each with IC of 0.03

$$IC = \frac{0.5 \bullet 0.03 + 0.5 \bullet 0.03}{\sqrt{0.25 \bullet 1.0 + 0.50 \bullet 0.0 + 0.25 \bullet 0.0}} = 0.03 \bullet \sqrt{2}$$

- Equally weight N uncorrelated factors, with average IC of $\overline{IC}$

$$IC = \sqrt{N} \bullet \overline{IC}$$

- Equally weight two factors with 0.9 correlation, each with IC of 0.03

$$IC = \frac{0.03}{\sqrt{0.25 + 0.25 + 0.5 \bullet 0.9}} = 0.031$$

- Equally weight two factors with -0.5 correlation, each with IC of 0.03

$$IC = \frac{0.03}{\sqrt{0.25 + 0.25 - 0.5 \bullet 0.5}} = \frac{0.03}{\sqrt{0.25}} = 0.06$$

Multiple Factors - Maximize IC

- No unique solution to maximum IC
- Constrain weights to sum to 1
- Optimal weights are scalar multiples of regression coefficients
- Composite IC is square root of regression R-squared

Simplified Five Factor Model Cross-Sectional Correlations for December 1995								
	Correlations							
	Mom 1	Value 1	Mom 2	Value 2	Other	IC	Weights	Reg Coeff
Mom 1	1.00	-0.23	0.43	-0.23	0.02	0.027	0.410	0.327
Value 1	-0.23	1.00	-0.19	0.46	0.02	0.053	0.761	0.606
Mom 2	0.43	-0.19	1.00	-0.33	-0.17	0.012	0.040	0.032
Value 2	-0.23	0.46	-0.33	1.00	-0.02	0.006	-0.181	-0.144
Other	0.02	0.02	-0.17	-0.02	1.00	-0.001	-0.030	-0.024
Total						0.069	1.000	0.798

Multiple Factors - Maximize IC

- No unique solution to maximum IC
- Constrain weights to sum to 1
- Optimal weights are scalar multiples of regression coefficients
- Composite IC is square root of regression R-squared

Simplified Five Factor Model
In-Sample Cross-Sectional Correlations for December 1995

	Correlations							
	Mom 1	Value 1	Mom 2	Value 2	Other	IC	Weights	Reg Coeff
Mom 1	1.00	-0.23	0.43	-0.23	0.02	0.027	0.410	0.327
Value 1	-0.23	1.00	-0.19	0.46	0.02	0.053	0.761	0.606
Mom 2	0.43	-0.19	1.00	-0.33	-0.17	0.012	0.040	0.032
Value 2	-0.23	0.46	-0.33	1.00	-0.02	0.006	-0.181	-0.144
Other	0.02	0.02	-0.17	-0.02	1.00	-0.001	-0.030	-0.024
Total						0.069	1.000	0.798

Multiple Factors - Single Period Analysis

- Regression coefficients can be interpreted as return to "factor portfolio" which has:
  - Zero weight
  - Factor exposure of 1
  - Zero exposure to other factors
  - Minimum (residual) variance portfolio with those characteristics

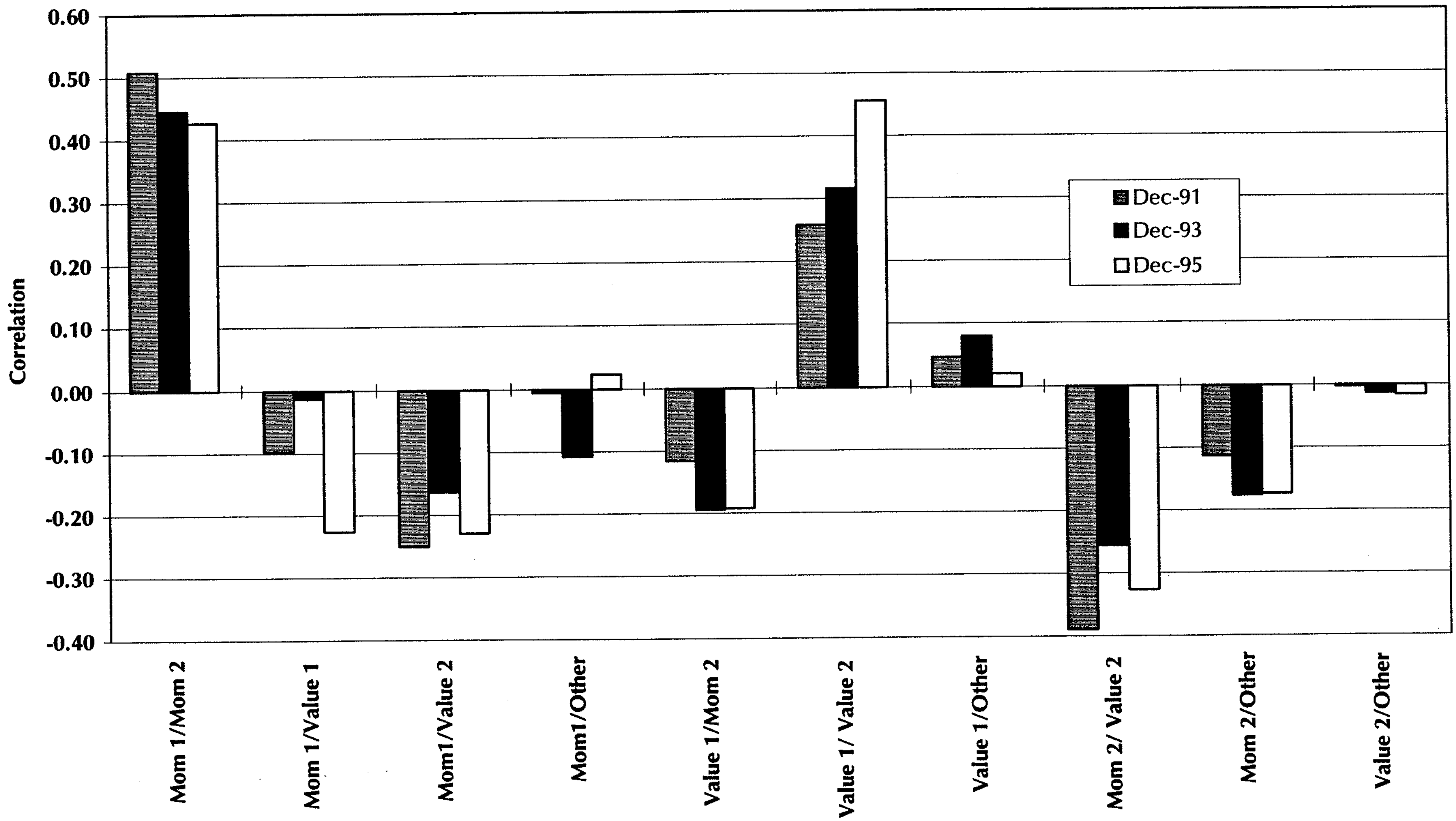
- Weighted Factor Portfolio Return = Skill * Opportunity

- Optimal weights are *not* proportional to IC

- Cross-sectional correlation matrix has potentially large impact on weights

Multiple Factors - Time Series Analysis

Factor Pair Cross-Sectional Correlations
December 1991 December 1993 December 1995
IMR Large Capitalization Universe



INVESCO®

Forecasting

December 1995 Factor Weights
Out-of-Sample Optimal Solution with Constraints

	Forecast		Unconstrained		Non-Negative Weight		40% Maximum	
Factor	IC	Regression Coefficient	IC Optimal Weight	Scaled Regression Coefficient	IC Optimal Weight	Scaled Regression Coefficient	IC Optimal Weight	Scaled Regression Coefficient
Mom 1	0.033	0.285	0.673	0.641	0.568	0.602	0.400	0.400
Value 1	0.015	0.090	0.300	0.203	0.265	0.191	0.263	0.287
Mom 2	0.006	-0.029	-0.090	-0.065	0.000	0.000	0.086	0.000
Value 2	0.011	0.097	0.164	0.219	0.166	0.205	0.199	0.309
Other	-0.001	0.001	-0.047	0.003	0.000	0.002	0.052	0.004
IC			0.0412	0.0408	0.0409	0.0406	0.0389	0.0370

- Forecast IC or regression coefficient with exponential smoothing parameter of 0.97
- Same covariance matrix as before
- "Nearly optimal" solutions

Forecasting - Advantages of IC Approach

- Treat factors independently until combined at last step
  - Add or drop factors without re-estimation
- Eliminate problem of changing correlation matrix through time
  - Only use current correlation matrix for forecasting
- Eliminate problem of correlations between factor performance and cross-sectional volatility
- More naturally add and understand impact of constraints

Forecasting - Other Issues

- Model of expected return, not risk!
- Time series properties of factors necessary to capture risk/return tradeoff
- Stability of weights through time
  - How fast does correlation matrix change?
  - How fast do IC forecasts change?
  - "Nearly optimal" solutions