Information Horizon, Portfolio Turnover, and Optimal Alpha Models

Optimizing multiple factors for the best IR.

Edward Qian, Eric H. Sorensen, and Ronald Hua

EDWARD QIAN
is the chief investment officer and head of research, Macro Strategies, at PanAgora Asset Management in Boston, MA.
eqian@panagora.com

ERIC H. SORENSEN
is president and chief executive officer of PanAgora Asset Management.
esorensen@panagora.com

RONALD HUA
is the chief investment officer and head of research, Equity Group, at PanAgora Asset Management.
rhua@panagora.com

The bottom-line value of an active investment process has two parts: the theoretical value of the alpha skill (the gross paper profit), minus the cost of implementation. The higher the former and the lower the latter, the happier the investor is. Clearly total assets under management (AUM) influence the latter. A strategy might be profitable with a low level of assets under management but unprofitable with more assets—as asset amounts grow, so do transaction costs.

Kahn and Shaffer [2005] note that one approach to remedy the size problem is to reduce portfolio turnover. While this is a sensible suggestion, Kahn and Shaffer's work is based on a hypothetical relation between turnover and expected alpha that might be too general to be applicable. In reality, any relation between turnover and expected alpha is not exogenous. It depends on alpha factors, their weights in an alpha model, and the rebalance horizon.

We propose an analytic framework for integrating alpha models with portfolio turnover. In practice, many alpha models are not constructed in such an integrated framework. Typically, managers first develop an alpha model (giving little consideration to turnover), and then throw the alpha model into an optimizer, setting turnover constraints to handle the transaction costs. There are two drawbacks to this two-step process: 1) It makes it hard to know the true effectiveness of the alpha model; and 2) it does not let managers adjust the alpha model along the way as AUM grow.

To integrate an alpha model with portfolio turnover, we extend work by Qian and Hua [2004] and Sorensen et al. [2004] that provides an optimal solution for

multifactor models with the objective of the highest information ratio. This work relies heavily on the time series correlation of information coefficients (IC) as well as contemporaneous correlations between factor signals.

We explore these critical correlation analyses to include serial correlations of factors and the concept of a factor *information horizon*. This allows us to evaluate implementation costs in finding the factor weights that optimize the information ratio (IR) with net returns.[1]

We first discuss the information horizon generally, and then derive an analytic formula for portfolio turnover conditioned on changes in forecasts. This solution allows us to estimate portfolio turnover for different quantitative alpha factors and related investment strategies. We find that portfolio turnover can be endogenous in a complete system, and that factor autocorrelation is the key exogenous ingredient.

Finally, we build a multifactor model by maximizing the IR under portfolio turnover constraints. A numerical example including a value and a momentum factor is used to illustrate the framework.

INFORMATION HORIZON

Information horizon is crucial to portfolio turnover. If a factor has a relatively short information horizon, it predicts security returns only for the very near term. The signal decays quickly. Momentum factors, especially the one-month reversal factor, behave this way. Such factors cause high turnover because their exposures must be constantly adjusted. If a factor has a relatively long information horizon, it predicts returns long after its information becomes available. The signal decays slowly. Valuation factors tend to behave this way. These factors will prompt low turnover—portfolios constructed on the basis of lagged factors can still generate excess returns.

In reality, no two factors (models) have the same profile in terms of information horizon. Typically, most factors lose all power by nine months; a few can keep going for as long as several years. Depending on the predictive power of different factors, it seems reasonable that turnover frequency will vary with the particular alpha model. The two should be in balance.[2]

Information Coefficient Terms

We study the general concept of information horizon through two specific expressions related to the information coefficient (IC): *lagged IC* and *horizon IC*. We denote IC as the cross-sectional correlation coefficient between the factor's value at the start of time t and the security returns over time period t, i.e., $IC_{t,t} = \text{corr}(\mathbf{F}_t, \mathbf{R}_t)$. Consider this the typical one-period IC measure, for a month or a quarter. An example is the first-quarter return IC. The factor values are observed December 31, and the return period is January through March.

The *lagged IC* is the correlation coefficient between time t factor values and *a later* period (by one, two, or more quarters) return vector, i.e., $IC_{t,t+l} = \text{corr}(\mathbf{F}_t, \mathbf{R}_{t+l})$, with lag l. The factor is lagged by l periods compared to the return. For example, we can correlate factor readings on December 31 with returns for later periods (second quarter [$l = 1$], third quarter [$l = 2$]), and so on. The IC will typically decay in power as the lag increases, but the rate of decline differs across different factors such as momentum and value. For simplicity, we assume the ICs are generated by stationary processes implying that $IC_{t,t+l} = IC_{t+s,t+l+s}$.

Exhibit 1 shows the average quarterly ICs and standard deviation of ICs of a specific price momentum (PM) factor based on nine-month returns. The average IC is high with no lag, but it declines steadily with greater lags. With a lag of three, the average IC is close to zero, indicating no predictability for the return.

Exhibit 2 shows the average quarterly ICs and standard deviation of ICs for a specific value factor based on trailing earnings yield (E2P). The average IC is lower compared to that of momentum, but shows less decay in power as the lag increases.

In both cases, the IC standard deviations are stable with respect to the change in lags. The standard deviation of IC is much higher for the momentum factor than the value factor. As a result, the IR of the value factor (annualized by multiplying by two) is higher than the IR

E X H I B I T   1
Momentum Factor Statistics for IC and Lagged ICs

Lag	0	1	2	3
Avg	0.073	0.047	0.026	0.003
Stdev	0.115	0.112	0.103	0.085
IR	1.269	0.847	0.505	0.078

IR is the annualized ratio of the average IC to the standard deviation of IC.

of the momentum factor at all lags. With a lag of three quarters, the IR of the momentum factor is nearly zero, while that of the value factor is still close to one.

Another important variant on the standard IC helps us understand the cumulative effect of multiple lagged ICs. We define *horizon IC* as an IC of a factor at a given time, t, for subsequent returns over multiperiod horizons. For example, if we have a factor available at December 31, we are interested in its correlations with *cumulative* returns of the next quarter, the next two quarters, or the next three quarters, and so on. We denote by $\mathbf{R}_{t,t+h}$ the risk-adjusted cumulative returns from period t through period $t + h$, and denote by $IC_t^h = \text{corr}(\mathbf{F}_t, \mathbf{R}_{t,t+h})$, $h = 0, 1, \ldots, H$, the horizon IC. For example, IC_t^o is the standard IC for the return in period t, and IC_t^1 is the correlation between the factor and the return vectors over the next six months, periods 1 and 2.

Relation Between Lagged IC and Horizon IC

Although the lagged IC typically decays with the lag, horizon IC often increases with the horizon, at least initially. By definition, the cumulative multiperiod return in the horizon IC is related to the single-period return by $\mathbf{R}_{t,t+h} = (1 + \mathbf{R}_t)(1 + \mathbf{R}_{t+1}) \cdots (1 + \mathbf{R}_{t+h}) - 1$. When the period's returns are low, we can approximate return by $\mathbf{R}_{t,t+h} \approx \mathbf{R}_t + \mathbf{R}_{t+1} + \cdots + \mathbf{R}_{t+h}$. Using this expression in the horizon IC yields:

$$IC_t^l \approx \frac{\text{cov}\left(\mathbf{F}_t, \mathbf{R}_t + \mathbf{R}_{t+1} + \cdots + \mathbf{R}_{t+l}\right)}{\text{dis}\left(\mathbf{F}_t\right)\text{dis}\left(\mathbf{R}_t + \mathbf{R}_{t+1} + \cdots + \mathbf{R}_{t+l}\right)} \quad (1)$$

If we further assume that the risk-adjusted returns in different periods are uncorrelated, the IC simplifies to:[3]

$$IC_t^l \approx \frac{IC_{t,t} + IC_{t,t+1} + \cdots + IC_{t,t+l}}{\sqrt{l+1}}$$
$$= \text{avg}(IC)\sqrt{l+1} \quad (2)$$

The horizon IC is an average of lagged ICs times the square root of the horizon length ($l + 1$). Note that the horizon IC covers returns of multiple periods, and the lagged ICs cover forecasts for single intervals for future periods. Suppose there is no information decay in

the lagged forecasts, i.e., $IC_{t,t} = IC_{t,t+1} = \cdots = IC_{t,t+l}$. Then from Equation (2) we have $IC_t^l = IC_{t,t}\sqrt{l+1}$. In this case, the horizon IC is the IC times the square root of the horizon length, and it therefore increases as the horizon lengthens.

Even when there is information decay, the horizon IC can still initially increase with the horizon length. It then declines as the horizon lengthens further or as the lagged ICs decline more rapidly.

Exhibit 3 plots one such case, in which the initial period IC is 0.10. The lagged IC is 0.08 with lag 1, 0.06 with lag 2, and so on. It reaches zero with lag 5, and turns negative thereafter. The horizon IC increases at first. For example, the IC is 0.128 for returns over the next two periods (1 on the horizontal axis), and 0.139 for returns over the next three periods (2 on the horizontal axis). Then, the horizon IC is eventually dragged down by the declining lagged ICs.

Exhibit 4 shows the horizon IC of the two factors in Exhibits 1 and 2. The horizon IC for the PM factor

E X H I B I T 2
Value Factor Statistics for IC and Lagged ICs

Lag	0	1	2	3
Avg	0.043	0.029	0.030	0.028
Stdev	0.066	0.061	0.060	0.057
IR	1.304	0.939	1.002	0.992

IR is the annualized ratio of the average IC to the standard deviation of IC.
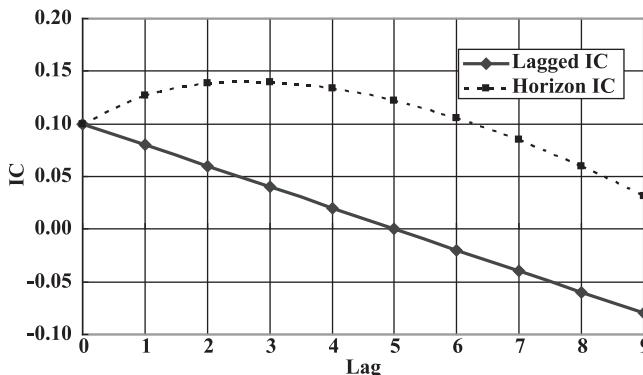
E X H I B I T 3
Lagged IC and Horizon IC of Signal

EXHIBIT 4
Horizon ICs for Two Factors

Horizon	0	1	2	3
PM	0.073	0.085	0.085	0.075
E2P	0.043	0.051	0.059	0.065

starts to decline at three quarters while that of E2P keeps increasing.

Horizon IC and Trading Horizon

The propensity for the horizon IC to increase initially with the horizon does not necessarily mean that we can increase the total IC with a longer trading horizon. Longer trading horizons allow fewer opportunities to rebalance, or fewer chances along the time dimension. Basically, the longer horizon reduces the *breadth* of the process, which adversely impacts the information ratio.

The reduced breadth leads to a higher variability of horizon IC. As a result, when we measure IR in terms of the ratio of average IC to the standard deviation of IC, the longer trading horizon may not confer any advantage. Suppose both forecasts and returns are of quarterly frequency. The quarterly IC has a mean of 0.1 and a standard deviation of 0.2. Then the quarterly IR is 0.5, and the annualized IR is $0.5\sqrt{4} = 1$.[4]

Let's assume that all subsequent one-period lagged ICs have the identical distribution, and are uncorrelated. Then, the horizon IC of one year, or four quarters, will have a mean of $4 \cdot 0.1 / \sqrt{4} = 0.2$, and a standard deviation of $\sqrt{4} \cdot 0.2 / \sqrt{4} = 0.2$. Hence, the annual IR is also 1—the same as the annualized IR of quarterly trading. There is no difference in terms of the performance, gross of any trading costs.

There could be differences in trading costs that arise from the implication of quarterly trading (rebalancing) rather than annual trading. In one case, we trade once a year. In the other case, we trade four times a year. The question is whether the total trading costs in the latter case exceeds the costs in the former case. The answer depends on the nature of the turnover induced by changes in alpha factors over each quarter versus the one-year holding period. Thus, the alpha model profile and trading costs become endogenous with respect to each other, and interact in determining the maximum deliverable net (after-cost) IR.

TURNOVER CAUSED BY CHANGE IN FORECASTS

Consider turnover over a single trading period, where the active weights change from w_i^t to w_i^{t+1}, for each security i. We define one-way turnover as one-half times the sum of the absolute value of weight changes:[5]

$$T = \frac{1}{2} \sum_{i=1}^{N} | w_i^{t+1} - w_i^t | \tag{3}$$

We assume the active weights for each security result from an unconstrained mean-variance optimization based on residual return and residual risk at times t and $t + 1$:

$$w_i^t = \frac{1}{\lambda_t} \frac{F_i^t}{\sigma_i}$$

$$w_i^{t+1} = \frac{1}{\lambda_{t+1}} \frac{F_i^{t+1}}{\sigma_i} \tag{4}$$

where F_i^t and F_i^{t+1} are risk-adjusted forecasts at t and $t + 1$, and λ_t and λ_{t+1} are risk aversion parameters. For simplicity, we assume that all stock-specific risks remain unchanged and that the number of stocks remains unchanged. If we hold the targeted tracking error σ_{model} for the portfolio constant, then the risk aversion parameter is given by:

$$\lambda_t = \frac{\sqrt{N-1}\text{dis}(\mathbf{F}^t)}{\sigma_{\text{model}}}$$

$$\lambda_{t+1} = \frac{\sqrt{N-1}\text{dis}(\mathbf{F}^{t+1})}{\sigma_{\text{model}}} \tag{5}$$

Substituting gives:

$$w_i^t = \frac{\sigma_{\text{model}}}{\sqrt{N-1}} \frac{\tilde{F}_i^t}{\sigma_i}$$

$$w_i^{t+1} = \frac{\sigma_{\text{model}}}{\sqrt{N-1}} \frac{\tilde{F}_i^{t+1}}{\sigma_i} \tag{6}$$

where $\tilde{F}_i^t$ and $\tilde{F}_i^{t+1}$ are now standardized with $\text{dis}(\tilde{\mathbf{F}}^t) = 1$, $\text{dis}(\tilde{\mathbf{F}}^{t+1}) = 1$. In other words, they reduce to simple *z*-scores. Equation (6) states that the active weight of a stock

is directly proportional to the portfolio target tracking error and its z-score, but is inversely proportional to the specific risk and the square root of the number of stocks. Using Equation (6) in (3) gives:

$$T = \frac{\sigma_{\text{model}}}{2\sqrt{N-1}} \sum_{i=1}^{N} \frac{\left| \tilde{F}_i^{t+1} - \tilde{F}_i^t \right|}{\sigma_i} \qquad (7)$$

The most difficult aspect of analyzing turnover is dealing with the absolute value function. Our solution to this problem is to approximate the turnover in Equation (3) as the expectation of the absolute difference of two *continuous* variables underlying two sets of forecasts. When we do this, we can rely on standard statistical theory to evaluate various expectations. In the appendix, we show that portfolio turnover is given by:

$$T = \sqrt{\frac{N}{\pi}} \sigma_{\text{model}} \sqrt{1 - \rho_f}\, E\left(\frac{1}{\sigma} \right) \qquad (8)$$

Equation (8) represents our solution to forecast-induced turnover for an unconstrained portfolio.[6] It depends on four elements. Turnover is higher:

- The higher the tracking error.
- The higher the number of stocks—proportional to the square root of N.
- The lower the forecast autocorrelation-cross-sectional correlation between consecutive forecasts, $\rho_f = \text{corr}(\tilde{F}^{t+1}, \tilde{F}^t)$.
- The lower the average stock-specific risk.

The forecast autocorrelation becomes the most relevant for our analysis of turnover. There is considerable intuition behind this. For example, consider two extremes. If the correlation between the consecutive forecasts is equal to one, then the weights are identical and there is no turnover. When the correlation is less than unity, it will be advantageous to incur turnover. At the other extreme, with a correlation of −1, turnover will be at the maximum. Indeed, all weights flip signs and the portfolio reverses itself.

Suppose stock-specific risks are the same for all stocks; in this case, the turnover is reduced to:

$$T = \sqrt{\frac{N}{\pi}} \frac{\sigma_{\text{model}}}{\sigma_0} \sqrt{1 - \rho_f} \qquad (9)$$

For an active portfolio (which could be either long-short, market-neutral, or active versus a benchmark) with $N = 500$, $\sigma_{\text{model}} = 5\%$, $\sigma_0 = 30\%$, and $\rho_f = 0.9$, the one-time turnover would be

$$T = \sqrt{\frac{500}{3.1415}} \frac{5\%}{30\%} \sqrt{1 - 0.9} = 66\%$$

Exhibit 5 plots the function $\sqrt{1 - \rho_f}$. This is the dependence of turnover on the forecast autocorrelation. Turnover is a declining function of forecast autocorrelation. The function behaves much like a linear function for most of the range, but it drops more precipitously when ρ_f is greater than 0.8.

Among the common quantitative factors practitioners use, value factors as a category generally have the highest forecast autocorrelation, and thus the lowest turnover. They also have the slowest information decay, characterized by high lagged ICs. Their forecast autocorrelation can be as high as 0.95. Among value factors, those based on cash flow have slightly lower forecast autocorrelation.

The momentum factors have the lowest forecast autocorrelation, and thus the highest turnover. Momentum factors' forecast autocorrelations generally lie between 0.6 and 0.7. For price momentum factors, the autocorrelation typically increases as the time window used for return calculation lengthens, up to 12 months. Therefore, one should use a longer time window to measure price momentum in order to reduce turnover.

E X H I B I T 5
Dependence Function of Turnover on Forecast Autocorrelation

THE JOURNAL OF PORTFOLIO MANAGEMENT **31**
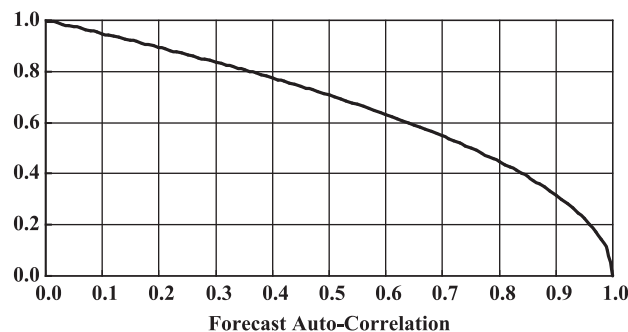
TURNOVER OF MULTIFACTOR MODELS

Multifactor models offer diversification among factors and can improve the information ratio of the overall model (see Sorensen et al. [2004]). Drawing on the results so far, we analyze the turnover of multifactor models by studying their forecast autocorrelation. Multifactor models can depend on both cross-sectional and time series averages of multiple factors. We can reduce portfolio turnover if the composite model has higher forecast autocorrelation than individual factors.

Cross-Sectional Average

In the cross-sectional dimension, the average consists of different factors whose performance is evaluated over the same time interval. To illustrate, consider a two-factor case, where the composite forecasts are linear combinations $\mathbf{F}_c = v_1 \mathbf{F}_1 + v_2 \mathbf{F}_2$; both $\mathbf{F}_1$ and $\mathbf{F}_2$ are standardized factors; and v_1, v_2 are their weights. The autocorrelation of the composite factor is:

$$\rho_{f_c} = \frac{\text{cov}\left(\mathbf{F}_c^t, \mathbf{F}_c^{t+1}\right)}{\text{var}\left(\mathbf{F}_c^t\right)}$$

$$= \frac{v_1^2 \rho_{11}^{0,1} + v_2^2 \rho_{22}^{0,1} + v_1 v_2 \left(\rho_{12}^{0,1} + \rho_{21}^{0,1}\right)}{v_1^2 + v_2^2 + 2 v_1 v_2 \rho_{12}^{0,0}} \quad (10)$$

where $\rho_{1,2}^{0,0}$ is the contemporaneous correlation between the two factors, and $\rho_{ij}^{0,1}$ is the correlation between $\mathbf{F}_i^t$ and $\mathbf{F}_j^{t+1}$. For $i = j$, $\rho_{ij}^{0,1}$, is the serial autocorrelation of a single factor, and for $i \neq j$, $\rho_{ij}^{0,1}$ is the serial cross-correlation between two different factors.

The autocorrelation of the composite factor depends on this correlation structure and on factor weights. All else equal, the autocorrelation of the composite factor will be high if the two factors have high serial auto- and cross-correlation, and the composite factor has low volatility. The low composite volatility depends, in turn, on a low contemporaneous factor correlation. Generally, low contemporaneous correlation also means better factor score diversification and therefore may imply a lower volatility for the IC.

For a numerical example, suppose the serial autocorrelations of two factors—PM and E2P—are $\rho_{11}^{0,1} = 0.68$ and $\rho_{22}^{0,1} = 0.94$; the serial cross-correlations are $\rho_{12}^{0,1} = -0.09$ and $\rho_{21}^{0,1} = 0.00$; and the contemporaneous correlation $\rho_{12}^{t,t} = -0.08$. Then:

$$\rho_{f_c} = \frac{0.68 v_1^2 + 0.94 v_2^2 - 0.09 v_1 v_2}{v_1^2 + v_2^2 - 0.16 v_1 v_2}$$

Exhibit 6 plots the autocorrelation as a function of v_1, and we have $v_2 = 1 - v_1$. When $v_1 = 0$, it is equal to the autocorrelation of factor 2, while when $v_1 = 1$, it is equal to the autocorrelation of factor 1. When v_1 is low, there is a slight increase in the composite autocorrelation.

Mathematically, all the correlation coefficients discussed above can be neatly placed into a single symmetric non-singular matrix—the correlation matrix for the stacked vector $(\mathbf{F}_1^{t+1}, \mathbf{F}_2^{t+1}, \mathbf{F}_1^t, \mathbf{F}_2^t)$:

$$\mathbf{C} = \begin{array}{c} \mathbf{F}_1^{t+1} \\ \mathbf{F}_2^{t+1} \\ \mathbf{F}_1^t \\ \mathbf{F}_2^t \end{array} \begin{pmatrix} 1 & \rho_{12}^{0,0} & \rho_{11}^{1,0} & \rho_{12}^{1,0} \\ \rho_{12}^{0,0} & 1 & \rho_{21}^{1,0} & \rho_{22}^{1,0} \\ \rho_{11}^{1,0} & \rho_{21}^{1,0} & 1 & \rho_{12}^{0,0} \\ \rho_{12}^{1,0} & \rho_{22}^{1,0} & \rho_{12}^{0,0} & 1 \end{pmatrix} \quad (11)$$

The autocorrelation of composite factors can be calculated using this correlation matrix.

Forecast Autocorrelation of Moving Averages

When signals are volatile, we can smooth them using moving averages. In a multifactor model, moving

---

E X H I B I T 6
Autocorrelation of Composite Forecast as Function of Weight in Factor 1
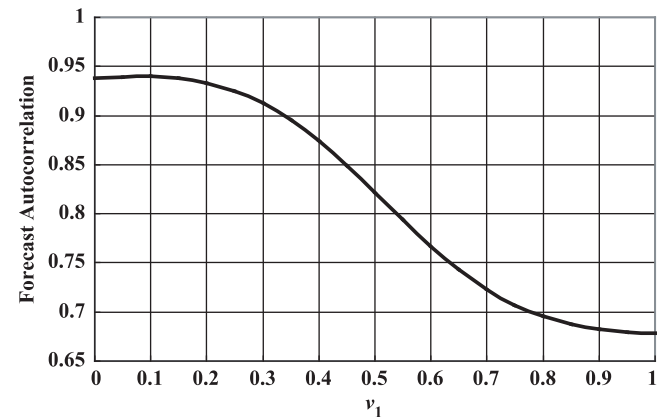
averages are also considered composite factors—a linear combination of new and past information. A natural question is why we would use outdated information in forecasts, as one tends to think that a forecast based on the most recent information is better than a lagged forecast, or has more predictive power for subsequent returns, i.e., a higher IR. This may be true, but one should verify this empirically.

The primary reason to use lagged forecasts for many alpha factors is that moving averages increase the autocorrelation, thus lowering turnover. Despite possible information decay in the lagged forecasts, a trade-off between potential profit reduction and improved transaction costs may favor the inclusion of lagged factors in a multifactor model. Our next step is to provide a mathematical framework for analyzing this important trade-off.

Autocorrelations of moving averages are calculated similarly to cross-sectional averages. Given forecast series $(\mathbf{F}^t, \mathbf{F}^{t-1}, \mathbf{F}^{t-2}, \ldots)$, we form a moving average of order L as $\mathbf{F}_{ma}^t = \Sigma_{l=0}^{L-1} v_l \mathbf{F}^{t-l}$. For $L = 2$ $\mathbf{F}_{ma}^t = v_0 \mathbf{F}^t + v_1 \mathbf{F}^{t-1}$. The serial autocorrelation is given by:

$$\rho_{f_{ma}} = \frac{\text{cov}\left(v_0 \mathbf{F}^t + v_1 \mathbf{F}^{t-1}, v_0 \mathbf{F}^{t+1} + v_1 \mathbf{F}^t\right)}{\text{var}\left(v_0 \mathbf{F}^t + v_1 \mathbf{F}^{t-1}\right)}$$

$$= \frac{v_0 v_1 + \left(v_0^2 + v_1^2\right)\rho_f(1) + v_0 v_1 \rho_f(2)}{v_0^2 + v_1^2 + 2 v_0 v_1 \rho_f(1)} \quad (12)$$

where $\rho_f(h)$ is the serial autocorrelation function of $\mathbf{F}^t$, with $\rho_f(0) = 1$.

For given serial autocorrelations $\rho_f(h)$, $h = 1, 2$, the correlation is a function of the weights, v_0 and v_1. Since the correlation is invariant to a scalar, we can require $v_0 + v_1 = 1$. Exhibit 7 plots Equation (12) as a function of v_1—the weight of the lagged forecast for the E2P factor, with $\rho_f(1) = 0.94$ and $\rho_f(2) = 0.84$. When $v_1 = 0$, the moving average is identical to the original factor, so the serial autocorrelation is 0.94. As v_1 increases, the lagged forecast is added to the moving average, and the serial autocorrelation of $\mathbf{F}_{ma}^t$ increases; it reaches a maximum close to 0.96 at $v_1 = 0.5$ when the two terms are equally weighted. As v_1 changes from 0.5 to 1, the autocorrelation declines from the maximum to 0.94.

Exhibit 8 graphs Equation (12) for the PM factor, with $\rho_f(1) = 0.68$ and $\rho_f(2) = 0.40$. As v_1 increases, the lagged forecast is added to the moving average, and the serial autocorrelation of $\mathbf{F}_{ma}^t$ increases; it reaches a maximum close
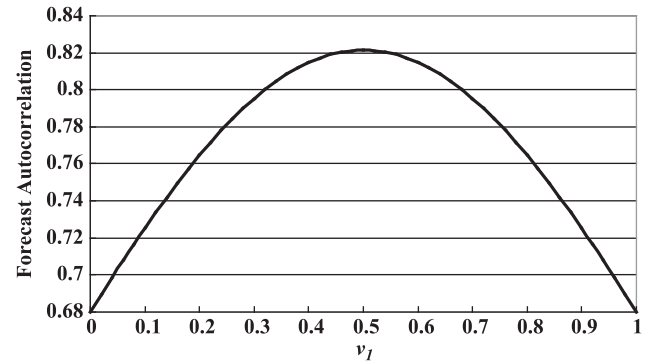
EXHIBIT 7
Serial Autocorrelation of Forecast Moving Average
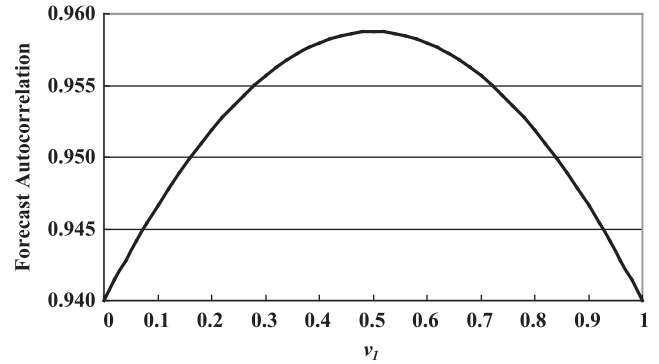
$L = 2$, $\rho_f(1) = 0.94$, $\rho_f(2) = 0.84$.

EXHIBIT 8
Serial Autocorrelation of Forecast Moving Average

$L = 2$, $\rho_f(1) = 0.68$, $\rho_f(2) = 0.40$.

to 0.82 at $v_1 = 0.5$ when the two terms are equally weighted. For both factors, the composite forecast is made more stable by including the lagged factors leading to reduced portfolio turnover. This is generally the case in practice.[7]

Cross-Sectional and Time Series Averages

A comprehensive multifactor model should have both cross-sectional and times series averages. We write such models as

$$\mathbf{F}_{c,ma}^t = \sum_{j=1}^{M} \sum_{l=0}^{L-1} v_{lj} \mathbf{F}_j^{t-l}$$

where there are M factors each with a moving average of order L. For clarity, we consider here the case of two factors and one lag, all standardized with standard deviation one:

$$\mathbf{F}_{c,ma}^{t} = v_{01}\mathbf{F}_1^t + v_{02}\mathbf{F}_2^t + v_{11}\mathbf{F}_1^{t-1} + v_{12}\mathbf{F}_2^{t-1} \tag{13}$$

Although it is still possible to calculate the serial autocorrelation of (13) algebraically as before, the expression becomes cumbersome and intractable with more factors and more lags. It is much more succinct to derive the autocorrelation by matrix notation instead. To this end, we denote the weights in (13) as a vector, $\mathbf{v} = (v_{01} \quad v_{02} \quad v_{11} \quad v_{12})'$. We consider the stacked vector $(\mathbf{F}_1^{t+1}, \mathbf{F}_2^{t+1}, \mathbf{F}_1^t, \mathbf{F}_2^t, \mathbf{F}_1^{t-1}, \mathbf{F}_2^{t-1})$ and denote its correlation matrix by:

$$\mathbf{C} = \begin{array}{c} \mathbf{F}_1^{t+1} \\ \mathbf{F}_2^{t+1} \\ \mathbf{F}_1^t \\ \mathbf{F}_2^t \\ \mathbf{F}_1^{t-1} \\ \mathbf{F}_2^{t-1} \end{array} \begin{pmatrix} 1 & \rho_{12}^{0,0} & \rho_{11}^{1,0} & \rho_{12}^{1,0} & \rho_{11}^{2,0} & \rho_{12}^{2,0} \\ \rho_{12}^{0,0} & 1 & \rho_{21}^{1,0} & \rho_{22}^{1,0} & \rho_{21}^{2,0} & \rho_{22}^{2,0} \\ \rho_{11}^{1,0} & \rho_{21}^{1,0} & 1 & \rho_{12}^{0,0} & \rho_{11}^{1,0} & \rho_{12}^{1,0} \\ \rho_{12}^{1,0} & \rho_{22}^{1,0} & \rho_{12}^{0,0} & 1 & \rho_{21}^{1,0} & \rho_{22}^{1,0} \\ \rho_{11}^{2,0} & \rho_{21}^{2,0} & \rho_{11}^{1,0} & \rho_{21}^{1,0} & 1 & \rho_{12}^{0,0} \\ \rho_{12}^{2,0} & \rho_{22}^{2,0} & \rho_{12}^{1,0} & \rho_{22}^{1,0} & \rho_{12}^{0,0} & 1 \end{pmatrix} \tag{14}$$

where $\rho_{ij}^{l,k} = \mathrm{corr}(\mathbf{F}_i^{t+l}, \mathbf{F}_j^{t+k})$. The autocorrelation of the composite (13) is the ratio of the covariance between the composite and its lagged value to its variance. To calculate the covariance and variance, we denote the 4×4 matrix in the upper left-hand corner of $\mathbf{C}$ as $\mathbf{C}_4$ and the 4×4 matrix in the upper right-hand corner of $\mathbf{C}$ as $\mathbf{D}_4$:

$$\mathbf{C}_4 = \begin{pmatrix} 1 & \rho_{12}^{0,0} & \rho_{11}^{1,0} & \rho_{12}^{1,0} \\ \rho_{12}^{0,0} & 1 & \rho_{21}^{1,0} & \rho_{22}^{1,0} \\ \rho_{11}^{1,0} & \rho_{21}^{1,0} & 1 & \rho_{12}^{0,0} \\ \rho_{12}^{1,0} & \rho_{22}^{1,0} & \rho_{12}^{0,0} & 1 \end{pmatrix}$$

$$\mathbf{D}_4 = \begin{pmatrix} \rho_{11}^{1,0} & \rho_{12}^{1,0} & \rho_{11}^{2,0} & \rho_{12}^{2,0} \\ \rho_{21}^{1,0} & \rho_{22}^{1,0} & \rho_{21}^{2,0} & \rho_{22}^{2,0} \\ 1 & \rho_{12}^{0,0} & \rho_{11}^{1,0} & \rho_{12}^{1,0} \\ \rho_{12}^{0,0} & 1 & \rho_{21}^{1,0} & \rho_{22}^{1,0} \end{pmatrix} \tag{15}$$

Then the variance of $\mathbf{F}_{c,ma}^t$ is

$$\mathrm{var}\left(\mathbf{F}_{c,ma}^t\right) = \mathbf{v}' \cdot \mathbf{C}_4 \cdot \mathbf{v} \tag{16}$$

And the covariance is

$$\mathrm{cov}\left(\mathbf{F}_{c,ma}^t, \mathbf{F}_{c,ma}^{t-1}\right) = \mathbf{v}' \cdot \mathbf{D}_4 \cdot \mathbf{v} \tag{17}$$

Combining Equations (16) and (17) yields the serial autocorrelation of $\mathbf{F}_{c,ma}^t$:

$$\rho_{f_{c,ma}} = \frac{\mathrm{cov}\left(\mathbf{F}_{c,ma}^t, \mathbf{F}_{c,ma}^{t-1}\right)}{\mathrm{var}\left(\mathbf{F}_{c,ma}^t\right)} = \frac{\mathbf{v}' \cdot \mathbf{D}_4 \cdot \mathbf{v}}{\mathbf{v}' \cdot \mathbf{C}_4 \cdot \mathbf{v}} \tag{18}$$

For a given factor correlation matrix $\mathbf{C}$, the autocorrelation is an analytic function of the weight vector $\mathbf{v}$.

OPTIMAL ALPHA MODEL WITH INCLUSIVE TURNOVER CONSTRAINTS

We have laid the groundwork for building optimal alpha models that explicitly consider turnover constraints, using multifactor linear models that include both current and lagged factors. The autocorrelation of the model sets the constraint on portfolio turnover while the model's IR is optimized according to the average ICs and covariances of ICs, based on the framework developed first in Qian and Hua [2004] and extended in Sorensen at al. [2004].

The key insight from this constrained optimization is the optimal use of lagged forecasts as part of the composite alpha model even if the lagged ICs are sometimes weaker than the contemporaneous ICs. Including lagged forecasts increases composite forecast autocorrelation, and thus reduces portfolio turnover, giving rise to savings in transaction costs. The equilibrium trade-off between the lagged ICs and the forecast autocorrelations determines the optimal model weights in the current as well as the lagged factors.

Constrained Optimization

Continuing with the case of two factors and one lag, Equation (13) describes an alpha model, and we are interested in the model weights $\mathbf{v} = (v_{01} \quad v_{02} \quad v_{11} \quad v_{12})'$ that maximize the IR while controlling the turnover. The autocorrelation of the composite is given by Equation (18) and the matrices $\mathbf{C}_4$ and $\mathbf{D}_4$ are submatrices of $\mathbf{C}$, given by Equation (15). As turnover is a function of $\rho_{f_{c,ma}}$, we thus constrain the autocorrelation to be at a targeted value.

The IR of the alpha model is approximately the ratio of average IC to the standard deviation of IC (Qian and Hua [2004]). Denote the average IC of $(\mathbf{F}_1^t, \mathbf{F}_2^t, \mathbf{F}_1^{t-1}, \mathbf{F}_2^{t-1})$ by $\overline{\mathbf{IC}}$ and the IC covariance matrix by Σ_{IC}. The average IC of the model is $\mathbf{v}' \cdot \overline{\mathbf{IC}}$, and the standard deviation of IC is $\sqrt{\mathbf{v}' \cdot \Sigma_{IC} \cdot \mathbf{v}}$. The constrained optimization problem is:

$$\text{Maximize:} \quad IR = \frac{\mathbf{v}' \cdot \overline{\mathbf{IC}}}{\sqrt{\mathbf{v}' \cdot \Sigma_{IC} \cdot \mathbf{v}}}$$

$$\text{subject to:} \quad \rho_{f_{c,ma}} = \frac{\mathbf{v}' \cdot \mathbf{D}_4 \cdot \mathbf{v}}{\mathbf{v}' \cdot \mathbf{C}_4 \cdot \mathbf{v}} = \rho_t \qquad (19)$$

The target autocorrelation is denoted by ρ_t. The autocorrelation constraint is quadratic in nature. This implies that (19) is a non-linear optimization with a quadratic constraint for which no analytic solution is readily available. It is easy to solve the problem numerically, as in the example below. We also note that the problem can be easily extended to include more factors and multiple lags.

Example—Inputs

We describe a numerical example using the two factors, PM and E2P. To make the example more realistic, we consider models with three lags.

First, we consider the IC inputs associated with the factors in Exhibits 1 and 2. These are quarterly data from the Russell 3000 universe for 1987-2004, and include the average ICs and standard deviations of IC for the two factors and their lagged factors. Thus, with two factors and three lags, we have eight different sources of alpha. Next, we compute the IR of the composite model using factor IC correlations (Exhibit 9). The subscripted numbers denote lags.

The most notable feature of Exhibit 9 is the negative IC correlation between the two factors.[8] For instance, the ICs of PM_0 and E2P_0 have a correlation of −0.42, indicating significant diversification benefit. That is, when the PM factor has a higher IC adding more alpha, the E2P factor tends to have a lower IC adding less alpha.

The diversification carries over to ICs of the lagged forecasts. For example, the ICs

of PM_1 and E2P_1 have a correlation of −0.45, and the ICs of PM_0 and E2P_1 have a correlation of −0.37.

We also should note that the IC correlations among the same factors but of different lags tend to be high, indicating less diversification of information, although the correlation drops with an increasing time period between forecasts. For instance, for the PM factor, the correlation is 0.86 between PM_0 and PM_1, 0.78 between PM_0 and PM_2, and 0.61 between PM_0 and PM_3. For the value factor, the correlations are even higher: 0.92 between E2P_0 and E2P_1, 0.84 between E2P_0 and E2P_2, and 0.78 between E2P_0 and E2P_3.

These IC correlations determine the composite alpha model's strategy risk. The factor correlations determine the alpha model autocorrelation.

Exhibit 10 presents the average factor correlations between factors of different lags that are needed to determine the forecast autocorrelation. Notice there are four lags in Exhibit 10, instead of the three lags in Exhibit 9. This is because we need to consider serial autocorrelation

E X H I B I T 9
IC Correlation Matrix of Current and Lagged Values for Momentum and Value Factors

	PM_0	E2P_0	PM_1	E2P_1	PM_2	E2P_2	PM_3	E2P_3
PM_0	1.00	-0.42	0.86	-0.37	0.78	-0.26	0.61	-0.19
E2P_0	-0.42	1.00	-0.44	0.92	-0.31	0.84	-0.29	0.78
PM_1	0.86	-0.44	1.00	-0.45	0.88	-0.36	0.71	-0.30
E2P_1	-0.37	0.92	-0.45	1.00	-0.33	0.94	-0.30	0.86
PM_2	0.78	-0.31	0.88	-0.33	1.00	-0.28	0.83	-0.22
E2P_2	-0.26	0.84	-0.36	0.94	-0.28	1.00	-0.28	0.94
PM_3	0.61	-0.29	0.71	-0.30	0.83	-0.28	1.00	-0.30
E2P_3	-0.19	0.78	-0.30	0.86	-0.22	0.94	-0.30	1.00

E X H I B I T 1 0
Factor Correlation Matrix of Current and Lagged Values for Momentum and Value Factors

	PM_0	E2P_0	PM_1	E2P_1	PM_2	E2P_2	PM_3	E2P_3	PM_4	E2P_4
PM_0	1.00	-0.08	0.68	0.00	0.40	0.05	0.09	0.08	0.07	0.09
E2P_0	-0.08	1.00	-0.09	0.94	-0.06	0.84	0.01	0.73	0.03	0.61
PM_1	0.68	-0.09	1.00	-0.08	0.68	0.00	0.40	0.05	0.09	0.08
E2P_1	0.00	0.94	-0.08	1.00	-0.09	0.94	-0.06	0.84	0.01	0.73
PM_2	0.40	-0.06	0.68	-0.09	1.00	-0.08	0.68	0.00	0.40	0.05
E2P_2	0.05	0.84	0.00	0.94	-0.08	1.00	-0.09	0.94	-0.06	0.84
PM_3	0.09	0.01	0.40	-0.06	0.68	-0.09	1.00	-0.08	0.68	0.00
E2P_3	0.08	0.73	0.05	0.84	0.00	0.94	-0.08	1.00	-0.09	0.94
PM_4	0.07	0.03	0.09	0.01	0.40	-0.06	0.68	-0.09	1.00	-0.08
E2P_4	0.09	0.61	0.08	0.73	0.05	0.84	0.00	0.94	-0.08	1.00

(with one lag) of forecasts that are made of factors of three lags.

Note that correlations among the same factors are high, for E2P in particular. This is not surprising, as earnings yields or PE multiples change slowly. Essentially, high serial autocorrelation of value factors is consistent with their minimal information decay.

The factor correlations are much lower for the PM factor; the lag 1 correlation is 0.68; the lag 2 correlation is 0.40; and the lag 3 and lag 4 correlations drop nearly to zero. This implies that winners and losers defined by price momentum change drastically over time—winners today have little resemblance to the winners nine months earlier. Thus, the construction process will incur more turnover in maintaining momentum exposure as the PM factor updates frequently.

Finally, note that the correlations between PM and E2P of different lags are low and not quite as negative as the IC correlations.

Optimal Alpha Models— Results and Insights

Given the inputs, we solve the optimization problem in Equation (19) for a series of forecast autocorrelation targets ρ_t, ranging from 0.85 to 0.97. The optimal model weights for each autocorrelation target, together with the corresponding IR, are presented in Exhibit 11.

First, notice that as ρ_t goes from 0.85 to 0.97, the optimal IR increases from 2.30 to 2.39 when ρ_t is 0.89 and then drops to 1.88 when ρ_t reaches 0.97. The highest IR occurs when the optimal weights are 36% for PM_0 and 64% for E2P_0, with no lagged factors. This represents the unconstrained optimal alpha model using the current factors only.

Second, we can see that for targeted autocorrelation above 0.89, the optimal model begins to add weight to lagged factors, which will slow turnover—but this is at the expense of the IR. PM_1 is the first to appear, and it is followed by E2P_1, E2P_2, and E2P_3. The other two lagged momentum factors, PM_2 and PM_3, never attain any significant weight in the model. The reason for their absence is both PM_2 and PM_3 have very low average ICs. All E2P factors, however, have consistent ICs and high autocorrelation. Their optimal weights near 10% each as the target autocorrelation gets very high.

EXHIBIT 11
Optimal Weights for Different Levels of Autocorrelation and Optimal IR

ρ_f	IR	PM_0	E2P_0	PM_1	E2P_1	PM_2	E2P_2	PM_3	E2P_3
0.85	2.30	45%	55%	0%	0%	0%	0%	0%	0%
0.86	2.33	43%	57%	0%	0%	0%	0%	0%	0%
0.87	2.36	41%	59%	0%	0%	0%	0%	0%	0%
0.88	2.38	39%	61%	0%	0%	0%	0%	0%	0%
0.89	**2.39**	36%	64%	0%	0%	0%	0%	0%	0%
0.90	2.38	34%	65%	2%	0%	0%	0%	0%	0%
0.91	2.37	31%	65%	4%	0%	0%	0%	0%	0%
0.92	2.36	28%	65%	7%	0%	0%	0%	0%	0%
0.93	2.33	24%	65%	10%	0%	0%	0%	0%	1%
0.94	2.28	21%	58%	12%	4%	0%	1%	0%	4%
0.95	2.21	18%	50%	12%	8%	0%	4%	0%	8%
0.96	2.09	15%	42%	11%	10%	2%	7%	2%	10%
0.97	1.88	11%	32%	8%	14%	5%	12%	5%	14%

EXHIBIT 12
Aggregated Optimal Weights with Autocorrelation Targets and Associated IRs

ρ_f	IR	PM	E2P	w_0	w_1	w_2	w_3
0.85	2.30	45%	55%	100%	0%	0%	0%
0.86	2.33	43%	57%	100%	0%	0%	0%
0.87	2.36	41%	59%	100%	0%	0%	0%
0.88	2.38	39%	61%	100%	0%	0%	0%
0.89	**2.39**	36%	64%	100%	0%	0%	0%
0.90	2.38	35%	65%	98%	2%	0%	0%
0.91	2.37	35%	65%	96%	4%	0%	0%
0.92	2.36	35%	65%	93%	7%	0%	0%
0.93	2.33	34%	66%	88%	10%	0%	1%
0.94	2.28	33%	67%	79%	15%	1%	4%
0.95	2.21	30%	70%	68%	20%	4%	8%
0.96	2.09	30%	70%	57%	21%	9%	13%
0.97	1.88	28%	72%	42%	23%	16%	19%

Finally, we see that the weight increase in the lagged factors comes first at the expense of the current momentum factor PM_0 and then the current value factor E2P_0.

To understand the quantitative trade-off between IR and turnover, consider a comparison. The maximum IR 2.39 occurs when the forecast autocorrelation ρ_t is at 0.89, and the model IR drops to 2.33 when ρ_f is at 0.93. Hence, the IR drops by roughly 2.5%. At the same time, the turnover changes according to the function $\sqrt{1-\rho_f}$, dropping from $\sqrt{1-0.89}$ to $\sqrt{1-0.93}$, or by as much as 20%.

In Exhibit 12 we aggregate optimal weights into weights for PM and E2P, and into weights for factors of lags 0, 1, 2, and 3. As ρ_t increases from 0.85 to 0.97, the PM weight drops from 45% to 28%, while the E2P

weight increases from 55% to 72%. At the same time, the weight with lagged factors increases from 0% to 58%, offset by reductions in the weights for zero-lagged factors.

Transaction Costs and Net Returns

One last piece to the puzzle is an assumption for trading costs. First, we measure the effect of autocorrelation on turnover by calculating turnover according to Equation (9), on an annual basis for a long-short portfolio. The inputs are N = 3,000, target risk σ_{model} = 4%, and stock-specific risk σ_0 = 30%. The turnover results and the IR are graphed in Exhibit 13.

First note the extremely high turnover when auto-correlation is low; it is nearly 550% when ρ_t is 0.89.[9] But the most important feature of the graph is that the rate of decline is markedly different for the IR and the turnover, as the autocorrelation ρ_t increases. While the turnover drops consistently, the IR changes rather slowly except when the autocorrelation reaches a very high level. Since the turnover drops more rapidly than the IR over a wide range, it is entirely feasible that net expected return—expected return less transaction costs—is higher for alpha models more highly autocorrelated with lagged factors.

We compute net expected return by imposing different levels of transaction costs that represent simply a linear proportion of the portfolio turnover. For example, at 50 basis points or 0.5%, a turnover of 100% would cost 0.5% and a turnover of 200% would cost 1%, and so on. With increasing assets under management, one can use a higher multiple of turnover as transaction costs.

Exhibit 14 shows the gross return given by IR times the target tracking error, turnover, and net returns with different transaction assumptions. We also include transaction costs of 1.0% and 1.5%.

As expected, the gross return is maximized at ρ_f = 0.89, where the IR reaches its maximum. But the net return attains its maximum at a higher factor model autocorrelation ρ_f, corresponding to a different alpha model with lagged factors. When the cost of 100% turnover is 0.5%, the maximum net return is at ρ_f = 0.93. In this instance, the IR is 2.33, but the turnover drops to 436% from 547%. When the transaction cost is higher at 1.0%, the maximum net return is at ρ_f = 0.95 where the paper IR is 2.21 and the expected net return is higher than the model with ρ_f = 0.89, by over 1.0%. At the even higher transaction cost of 1.5% for 100% turnover, the optimal

EXHIBIT 13
IR and Portfolio Turnover of Optimal Alpha Models with Given Forecast Autocorrelation
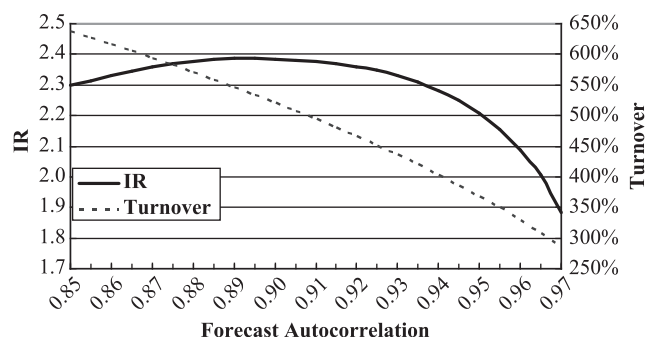
IR scale is on the left axis, and turnover scale is on the right axis.

EXHIBIT 14
Gross Excess Return and Net Excess Return under Different Transaction Cost Assumptions

| ρ_f	IR	Gross Return	Turn-over	Net Return (0.5%)	Net Return (1.0%)	Net Return (1.5%)
0.85	2.30	9.19%	638%	6.00%	2.81%	0.38%
0.86	2.33	9.32%	617%	6.24%	3.15%	0.07%
0.87	2.36	9.43%	594%	6.46%	3.49%	0.52%
0.88	2.38	9.51%	571%	6.66%	3.80%	0.95%
0.89	**2.39**	**9.55%**	547%	6.81%	4.08%	1.35%
0.90	2.38	9.53%	521%	6.93%	4.32%	1.71%
0.91	2.37	9.50%	494%	7.03%	4.56%	2.08%
0.92	2.36	9.44%	466%	7.11%	4.78%	2.45%
0.93	2.33	9.33%	436%	**7.15%**	4.97%	2.79%
0.94	2.28	9.13%	404%	7.11%	5.09%	3.07%
0.95	2.21	8.83%	369%	6.98%	**5.14%**	3.30%
0.96	2.09	8.35%	330%	6.70%	5.06%	**3.41%**
0.97	1.88	7.53%	285%	6.10%	4.68%	3.25%

Portfolios with N = 3,000, target risk σ_{model} = 4%, and stock-specific risk σ_0 = 30%.

model for net return would be at ρ_f = 0.96. Alpha models with these autocorrelation targets include significant weights of lagged factors (see Exhibits 11 and 12).

In Exhibit 15 we plot the data in Exhibit 14. The square on each curve denotes the model with the maximum net return. As the transaction costs increase, the net return gets lower and lower. This is especially true for the left sides of the return curves, due to higher turnover. The right side of the curves drops to a lesser extent because the turnover is lower (higher model autocorrelation). On each cost curve, the point of maximum net return shifts to the right as the autocorrelation increases. We also note

E X H I B I T   1 5
Gross Excess Return and Net Excess Return under Different Transaction Cost Assumptions
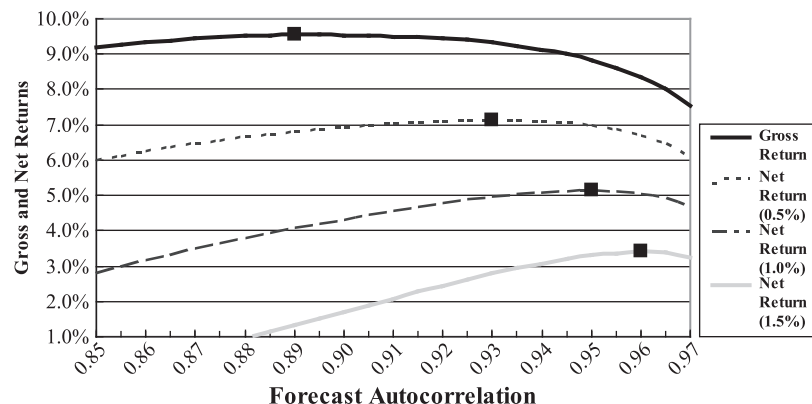
Portfolios with N = 3,000, target risk σ_{model} = 4%, and stock-specific risk σ_0 = 30%.

that when transaction costs are high, there is a more rapid increment in net return as autocorrelation increases—optimal models with high autocorrelation have more of a chance to yield positive net returns as trading costs rise.

These results have strong implications for the construction of optimal alpha models. First, using the model with maximum gross IR can be suboptimal in terms of net return when transaction costs are taken into account. For example, with 1% cost, the net return of that model is lower than that of the optimal model with ρ_f = 0.95, by more than 1 percentage point.

Second, the optimal model in terms of highest net return changes as transaction costs rise. This indicates the need to adjust the alpha model as assets under management grow. Our results reveal that one way to do this is to add lagged factors based on the risk/return/turnover trade-off.

Third, both the net return and the optimal model are sensitive to the IR assumption. If the IRs are lower than those in the example, then for a given level of transaction costs, the maximum net return would be achieved with even higher ρ_f models. In other words, when the information content of the factors is lower, we need to pay even more attention to constrain portfolio turnover to reduce transaction costs. This inevitably leads to more weight for the lagged factors, especially lagged value factors.[10]

SUMMARY

Realistic alpha models necessitate the direct endogenous analysis of implementation costs. This means that we should build optimal alpha models that integrate transaction costs in the modeling process. We provide an analytic framework to do so by maximizing the information ratio of alpha models under portfolio turnover constraints.

The solution we provide is quite general. We show that portfolio turnover is an algebraic function of forecast autocorrelation. Hence, factor autocorrelation is a key diagnostic in evaluating single-factor efficacy and in creating optimal composite models. For linear composite forecasts, the autocorrelation depends on auto- and cross-factor correlations. Optimal models by definition must depend on factor correlations for contemporaneous and lagged values as well as average ICs, standard deviation of ICs, and IC correlations.

These results should be useful to practitioners in several ways. First, autocorrelation is an important diagnostic for evaluating factors. Second, comprehensive alpha models necessitate the direct integration of transaction costs. Third, the analyses here lead directly to reasonable estimations of strategy capacity associated with increasing assets under management levels. They suggest adjusting weights in favor of lagged factors to reduce turnover by increasing the target autocorrelation, thus maximizing the net return.

APPENDIX

Portfolio Turnover

In this appendix we derive the theoretical solution of portfolio turnover. We first rewrite Equation (7) as an expectation:

$$T = \frac{\sigma_{model}\sqrt{N}}{2} \frac{1}{N} \sum_{i=1}^{N} \frac{\left| \tilde{F}_i^{t+1} - \tilde{F}_i^t \right|}{\sigma_i}$$
$$= \frac{\sigma_{model}\sqrt{N}}{2} E\left(\frac{\left| \tilde{F}^{t+1} - \tilde{F}^t \right|}{\sigma} \right) \tag{A-1}$$

To evaluate the expectation, we further assume that the change in the risk-adjusted forecast and the stock-specific risk are independent. Hence, Equation (A-1) can be written as:

$$T = \frac{\sigma_{model}\sqrt{N}}{2} E(| \tilde{F}^{t+1} - \tilde{F}^t |) E\left(\frac{1}{\sigma} \right) \tag{A-2}$$

Both sets of forecasts have a standard deviation of one. We further assume they form a bivariate normal distribution with mean zero, and the cross-sectional correlation between the two sets of consecutive forecasts is ρ_f. This is simply the lag 1 auto-correlation of the risk–adjusted forecasts. If the correlation is high, the change in forecasts is minimal, and the turnover is low. Conversely, if the correlation is low, the forecast change is significant, and turnover will be high.

Since both forecasts are normally distributed, the change $\tilde{F}^{t+1} - \tilde{F}^t$ is still a normal distribution with zero mean and standard deviation $d = \sqrt{2(1 - \rho_f)}$. For a random variable x with distribution $x \cdot N(0, d^2)$, we can show that:

$$E(|x|) = \sqrt{\frac{2}{\pi}} d$$

Therefore:

$$E(|\tilde{F}^{t+1} - \tilde{F}^t|) = \frac{2\sqrt{1 - \rho_f}}{\sqrt{\pi}} \qquad \text{(A-3)}$$

Substituting Equation (A-3) into (A-2) yields:

$$T = \sqrt{\frac{N}{\pi}} \sigma_{\text{model}} \sqrt{1 - \rho_f}\, E\left(\frac{1}{\sigma}\right) \qquad \text{(A-4)}$$

Since targeted tracking error is linked to the portfolio leverage, we can derive a relation between leverage and forecast-induced turnover, using expectations. We have:

$$L = \sum_{i=1}^{N} |w_i| = \frac{\sigma_{\text{model}}}{\sqrt{N-1}} \sum_{i=1}^{N} \left| \frac{\tilde{F}_i^t}{\sigma_i} \right|$$

$$= \sigma_{\text{model}} \sqrt{N}\, E\left(\left| \frac{\tilde{F}^t}{\sigma} \right| \right)$$

$$= \sigma_{\text{model}} \sqrt{N}\, E\left(|\tilde{F}^t| \right) E\left(\frac{1}{\sigma} \right) \qquad \text{(A-5)}$$

It should be apparent from Equation (A-5) that the leverage of the portfolio should remain the same if the targeted tracking error and the specific risks remain constant.

Since $\tilde{F}^t$ is a standard normal variable, we have $E(|\tilde{F}^t|) = \sqrt{2/\pi}$. Using this result in Equation (A-5) yields:

$$L = \sqrt{\frac{2N}{\pi}} \sigma_{\text{model}} E\left(\frac{1}{\sigma} \right) \qquad \text{(A-6)}$$

Combining (A-6) and (A-4) yields:

$$T = \frac{L\sqrt{1 - \rho_f}}{\sqrt{2}} \qquad \text{(A-7)}$$

Therefore, turnover is directly proportional to leverage times the square root of 1 minus forecast autocorrelation.

ENDNOTES

[1]We assume that on an aggregated portfolio level the implementation cost is proportional to the amount of trading, or portfolio turnover. The majority of implementation cost is related to trading. These costs could be commission, bid-ask spread, and market impact. The trading cost may vary from stock to stock. For an approximation, we simply assume that trading cost is a fixed multiple of portfolio turnover.

[2]Turnover can also be caused by flows in and out of portfolios. These forced turnovers are not due to portfolio rebalancing, and they are easy to analyze. We ignore them in our analysis.

[3]If there is short-term reversion between consecutive-period returns, the horizon IC will be higher.

[4]See Qian and Hua [2004] for a development of the IR expression: $\text{IR} \approx \text{Expected IC} / \sigma(\text{IC})$.

[5]Our definition of turnover measures the percentage change of the portfolio versus total portfolio capital, which is most relevant in terms of amount of trading and costs. There are other variations that use total portfolio leverage or notational exposures as denominators, which tend to reduce the turnover percentage.

[6]For constrained portfolios such as long-only portfolios, turnover can be substantially less, since constraints work to suppress changes in portfolio weights (see Qian, Hua, and Tilney [2004]). Turnover can also be reduced through other means such as cutting weights proportionally or ignoring small trades. For details, see Grinold and Stuckelman [1993] and Qian, Hua, and Tilney [2004].

[7]Our analysis shows that the inclusion of a lagged forecast would increase the serial autocorrelation as long as autocorrelation of lag 2 is above a certain threshold. The value of the threshold is given by two times the lag 1 correlation squared minus one. These values are easily exceeded for most factors encountered in practice.

[8] While factor correlation has some bearing on the corresponding IC correlation, the two are not the same. In building

an alpha model, it is the IC correlation not the factor correlation that plays a crucial role in determining the weights of factors.

[9]In practice, the actual turnover will be lower when the transactions are incorporated into the portfolio optimization.

[10]It is not hard to see this situation might apply to market segments that are relatively less inefficient, such as U.S. large-cap stocks.

REFERENCES

Grinold, Richard C., and M. Stuckelman. "The Value-Added/Turnover Frontier." *The Journal of Portfolio Management*, vol. 19, no. 4 (Summer 1993), pp. 8-17.

Kahn, R. N., and J. S. Shaffer. "The Surprisingly Small Impact of Asset Growth on Expected Alpha." *The Journal of Portfolio Management*, vol. 32, no. 1 (Fall 2005), pp. 49-60.

Qian, Edward, and Ronald Hua. "Active Risk and Information Ratio." *The Journal of Investment Management*, vol. 2, no. 3 (2004), pp. 20-34.

Qian, Edward, Ronald Hua, and John Tilney. "Turnover of Quantitatively Managed Portfolios." Working paper, PanAgora Asset Management, 2004.

Sorensen, Eric H., Ronald Hua, Edward Qian, and Robert Schoen. "Multiple Alpha Sources and Active Management." *The Journal of Portfolio Management*, vol. 31, no. 2 (Winter 2004), pp. 39-45.

To order reprints of this article, please contact Dewey Palmieri at dpalmieri@iijournals.com or 212-224-3675.