Analysis of Cross Sectional Equity Models

Edward Qian & Ronald Hua
Putnam Investments
One Post Office Square
Boston, MA 02109

February, 2002
1. Introduction

Generally, two classes of quantitative models are used in active portfolio management: time series and cross-sectional models. The time series models have proven to be very useful in the areas of tactical asset allocation and active currency management. This is understandable for at least two reasons. One, there are strong econometric basis for forecasting asset and currency return. Two, there exists lengthy collection of historical data in those areas – a necessary ingredient for building time series models. However, both of these conditions are often not met when it comes to quantitative stock selection. Stock selection is almost all about stock specific returns, which lies outside of the realm of econometric models. For many companies, there just is not enough time series data available. Moreover, the sheer number of stocks in a typical portfolio should prevent any serious attempt to build time series models on individual stocks. The DCF based models, on the other hand, require earning projection and discount rate assumption, which are mostly based on fundamental analysis rather than quantitative modeling. For the purpose of this paper, we don’t consider them as quantitative models.

As a result, quantitative equity analysts and portfolio managers increasingly rely upon cross-sectional models in the stock selection process. The intent of cross-sectional models is not to accurately forecast the future total return of individual stocks, but to predict the relative performance of all stocks. One advantage of these models is flexibility. Once constructed, we can apply them to almost all companies regardless of their ages, since the model inputs come from mostly current values of factors. Nevertheless, the construction of the models still requires an adequate history of cross-sectional data.

Despite its popularity, there seems to be a lack of better understanding concerning construction and performance of cross-sectional equity models. On a very basic level, it is intuitive to everyone that ability to rank future stock returns is of crucial importance in a cross-section perspective. A quantitative proxy for this ability is cross-sectional information coefficient (IC), defined as correlation coefficient between forecasts and actual returns of all stocks for a given time period. When the cross-sectional IC is positive, model ranking tends to be more in agreement with actual ranking of stock returns. Thus, we are able to generate positive alpha or excess return, by over-weighting
stocks with higher forecasts and at the same time under-weighting stocks with lower forecasts. While the cross-sectional IC is a primary factor in determining the sign of alpha, however, it is not the only factor that affects alpha. There are at least two additional factors. One is cross-sectional dispersion of the forecasts and the other is cross-sectional dispersion of the actual returns. Theoretically, the dispersion of the forecasts can be used to adjust the sizes of active positions, directly influencing magnitude of alpha. The dispersion of the actual returns is even more important because it ultimately determines the cross-sectional opportunity that actually exists.

Furthermore, conventional definition of cross-sectional IC does not fully reflect expected risk-adjusted alpha of active portfolios. In practice, one often conveniently computes the IC as the correlation coefficient between the forecasts and the actual total returns. This simple procedure ignores the fact that one often must “modify” the forecasts to accommodate various portfolio constraints. When this occurs, the IC so defined loses its connection to active alpha. Therefore, it is imperative that we seek to modify accordingly the definition of IC.

This paper provides an analytical framework to investigate these issues. We first prove that in general alpha is a product of IC and dispersions of forecast and actual return. We comment that the conventional definition of IC is appropriate only if the dollar neutrality is the sole constraint and if risks are the same for all stocks. We then derive a direct and rigorous relationship between the active alpha and cross-sectional IC for active portfolios with both dollar neutral and beta neutral constraints. When the beta neutral constraint is added, we redefine IC to preserve its relevance to and relationship with alpha. We next consider the long-term performance of a cross sectional model in terms of its information ratio. We demonstrate that IR depends not only on the strength of IC, but also on statistical relationships among IC and the dispersions of forecast and actual return over time. Finally, using a valuation factor based on free cash flow yield, we present an example to illustrate the practical implication of our results.

2. A framework of cross-sectional models

A cross-sectional equity model typically employs several factors, blending the effect of valuation, momentum, earning quality, etc. Here, we focus on a single factor model for the sake of simplicity. The blending of multiple factors entails further analysis,
which is out of the scope of the present paper. We also note that the analysis presented here is not limited to equity cross-sectional models, it applies to cross-sectional models in general.

2.1 Cross-sectional IC and active alpha

For any given time period $t$, the cross-sectional correlation of a factor with the actual returns is

$$
\rho_t = \text{corr}(f^t, r^t),
$$

where $f^t$ is the vector of factor values available at the beginning of the time period for all stocks and $r^t$ is the corresponding vector of subsequent actual returns for the time period.

As stated previously, if the IC is positive, then over-weighting stocks with high $f$'s and simultaneously under-weighting stocks with lower $f$'s should earn us positive alpha\(^1\). One simple way to come up with the active weight for stock $i$ could be

$$
w_{i,t} = \lambda^{-1} (f_{i,t} - \bar{f}_t),
$$

where $f_{i,t}$ is the factor value for the stock, $\bar{f}_t$ is the cross-sectional average of the factors for all stocks, and $\lambda^{-1}$ is the inverse of a risk-version parameter. In other words, the weights are proportional to the differences between the factors and the average. The weights are dollar neutral since they sum to zero. The products between the active weights and the actual returns produce the alpha for the time period

$$
\alpha_t = \sum_{i=1}^{N_t} w_{i,t} \cdot r_{i,t} = \lambda^{-1} \sum_{i=1}^{N_t} (f_{i,t} - \bar{f}_t) \cdot r_{i,t} = \lambda^{-1} \sum_{i=1}^{N_t} (f_{i,t} - \bar{f}_t) \cdot (r_{i,t} - \bar{r}_t).
$$

In equation (3), $r_{i,t}$ is the actual return of the $i$-th stock, and $N_t$ is the number of stocks included for the period. Since the weights are dollar neutral, we can use the relative returns versus the average $\bar{r}_t$ in the alpha calculation.

It is easy to recognize that equation (3) gives rise to the covariance between the factor and the actual return. We rewrite it using the relationship between the covariance and correlation and standard deviations,

\(^1\) Without loss of generality, we assume average IC is always positive.
Equation (4) proves that in this case, alpha is proportional to the cross-sectional IC, the cross-sectional standard deviations of the factors and the actual returns. Later on, we shall consider average and standard deviation with respect to time of the terms in equation (4). To avoid confusion, we shall refer cross-sectional standard deviation as dispersion and reserve the use of standard deviation for time series purpose only.

Several remarks can be made about equation (4). First, all terms are always positive except the IC. Therefore, the sign of IC determines profit or loss. Second, the magnitude of the alpha hinges upon the magnitude of IC as well as those of the two dispersions. But the two dispersions play different roles. The dispersion of the factors can be used to influence the size of the active weights. However, we could also choose to give up that control by standardizing it for all time so that the dispersion of the active weights is constant over time. The dispersion of the actual returns, on the other hand, is beyond anyone’s control. It will ultimately determine the size of the alpha.

2.2 Portfolio construction and cross-sectional IC

In practice, one rarely constructs a quantitative active portfolio so naïvely. The processes most likely involve some sort of mean-variance optimization. At least two inputs are needed for the MV optimization. First is predicted alpha of each stock. From now on, we assume \( f \)'s are the final predicted alphas. In the real world, the journey from factors to alphas is often a long one\(^2\). The covariance matrix is the second input. Here, the standard choice seems to be a full covariance matrix based on some risk models, such as those developed by BARRA. Additional inputs are needed when there are constraints on the active portfolio.

Given the complexity of the MV optimization with constraints, one simply has to ask under what circumstance, our active positions given by equation (2) are MV optimal. The answer is they are MV optimal when we use a diagonal covariance matrix that has

\[
\alpha_t = \lambda^{-1} (N_t - 1) \text{cov}(f_t, P_t) = \lambda^{-1} (N_t - 1) \text{corr}(f_t, P_t) \text{std}(f_t) \text{std}(P_t)
\]

\(^2\) The factors are often refined by procedures such as z-scoring and winsorizing to eliminate the outlier. Sometimes, an extra step is taken to convert the factor z-scores to forecasted alphas by multiplying a pre-selected a uniform IC for all stocks and the volatilities of stocks. While this last step is mathematically justified for time series models, it does not have a solid justification for cross-sectional models. As for the adjustment by IC and z-scoring, they are essentially equivalent to changing the risk-aversion parameter. Therefore, they have no effect on the efficient frontier.
equal diagonal elements and when the only constraint is the dollar neutral constraint. The special covariance matrix implies that all stocks have the same risk and their returns are independent of one another. When this is not true, as often in practice, one must question the validity of equation (4) and even the relevance of the cross-sectional IC itself.

Fortunately, it is possible to extend both the concept of cross-sectional IC and equation (4) for some optimal portfolios. We provide one such case here. In addition to the dollar neutral constraint, we add market or beta neutral constraint and construct portfolios using MV optimization. While the additional complexity does not include all the practical considerations encountered in an equity portfolio, it does make our portfolio more realistic. This is especially true for equity market-neutral hedge funds since these funds face few constraints beside the two considered here.

For portfolios with these assumptions, we can prove that the relationship (4) is preserved, provided that we adjust the factors and the actual returns to reflect both constraints and stock specific risks. The result in appendix shows

\[
\alpha_t = \lambda^{-1}(N_t - 1) \text{cov}(\tilde{F}_t, \tilde{R}_t) = \lambda^{-1}(N_t - 1) \text{corr}(\tilde{F}_t, \tilde{R}_t) \text{std}(\tilde{F}_t) \text{std}(\tilde{R}_t);
\]

where \( \tilde{F}_t \) and \( \tilde{R}_t \) are the vectors of refined factors and refined actual returns. Specifically,

\[
F_{i,t} = \frac{f_{i,t} - l_1 - l_2 \beta_{i,t}}{\sigma_{i,t}} = \frac{\tilde{f}_{i,t}}{\sigma_{i,t}}, \quad R_{i,t} = \frac{r_{i,t} - k_1 - r_{M,t} \beta_{i,t}}{\sigma_{i,t}} = \frac{\tilde{r}_{i,t}}{\sigma_{i,t}};
\]

where \( \beta_{i,t} \) is beta of the \( i \)-th stock, \( \sigma_{i,t} \) is its specific risk and \( r_{M,t} \) is the broad market return. The parameters \( l_1 \) and \( l_2 \) are required to make the portfolio both dollar and beta neutral, and the parameter \( k_1 \) ensures that the refined actual returns \( R \) is of cross sectional mean zero. The financial interpretation of equation (6) is straightforward: \( F \) and \( R \) are both dollar- and beta- adjusted factors and returns, standardized by the specific risk. We shall refer \( \tilde{f} \) and \( \tilde{r} \) as the adjusted factor and adjusted return, but \( F \), and \( R \) as the refined factor and refined return. Equation (5) asserts that active alpha is determined by the correlation between the refined factors and returns, and their respective dispersions.

The change from the raw factor and return to the refined ones is crucial for active portfolio with beta neutral constraint. The following table provides a simple 3-stock example to illustrate the difference. The first few columns are the factors, betas, specific risks, and the actual returns of the stocks. The next two columns contain the dollar and
beta adjusted factors and returns. The last two columns are the refined factors and returns. The cross-sectional IC between $f$ and $r$ is minus 0.5, because, by the raw factor alone, we would over-weight stock 3, under-weight stock 1, and be neutral on stock 2. This produces a negative alpha since stock 3 returns –10% and stock 1 returns 10%. On the other hand, the correlation between $F$ and $R$ turns out to be plus 0.76. This is because the beta neutral constraint leads to a significant overweight in stock 2 and only a slight overweight in stock 3, as reflected by values of $F$. Stock 2 has the best return and alpha from these new positions is positive. Thus, it would be wrong in this case to use the conventional IC to gauge the active alpha, and it highlights the importance of using the appropriate cross-sectional IC for different portfolio construct process.

Table 1. An example of three stocks

<table>
<thead>
<tr>
<th>Stock</th>
<th>$f$</th>
<th>$\beta$</th>
<th>$\sigma$</th>
<th>$R$</th>
<th>$\tilde{f}$</th>
<th>$\tilde{r}$</th>
<th>$F$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1.25</td>
<td>50%</td>
<td>10%</td>
<td>-0.7</td>
<td>-10%</td>
<td>-143%</td>
<td>-20%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1.50</td>
<td>50%</td>
<td>30%</td>
<td>0.6</td>
<td>20%</td>
<td>114%</td>
<td>40%</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.25</td>
<td>50%</td>
<td>-10%</td>
<td>0.1</td>
<td>-10%</td>
<td>29%</td>
<td>-20%</td>
</tr>
</tbody>
</table>

We have chosen the same specific risk for all three stocks to isolate the effect of beta neutral constraint. The effect of differing specific risk on active positions can also be important. But in general, unlike the beta neutral constraint, they only change the sizes of active positions, not their directions.

2.3 The long-term performance of cross-sectional models

Equation (4) and (5) show the relationship between the alpha and the IC and dispersions of factors and actual returns for a single time period. They also enable us to hypothesize that, the performance of the model over time shall depend on interactions among all three terms. For instance, the best case occurs when there are positive correlations among the three terms. Intuitively, a positive correlation between the IC and the dispersion of the actual returns means that when the IC is high and positive, the dispersion of the actual return is likely to be also high, leading to high positive alpha. Conversely, when the IC is low and negative, the dispersion of the actual return is likely to be small, thus limiting the loss. On the other hand, a negative correlation would have
the opposite effect. When the IC is high and positive, the dispersion often turns out to be low, capping the size of the positive alpha; when the IC is low and negative, the dispersion is likely to be large, magnifying the loss. The same argument applies to the correlation between the IC and the dispersion of the factors. On the other hand, when all three terms behave independently, the argument above no long apply, and the performance will become more dependent of the cross-sectional IC alone.

Information ratio of a cross-sectional model over time is the ratio of average alpha and the standard deviation of alpha over time,

\[
IR = \frac{\text{avg}_t(\alpha_t)}{\text{std}_t(\alpha_t)} = \frac{\text{avg}_t\{\text{corr}(F_t, R_t)\}\text{std}(F_t)\text{std}(R_t)}{\text{std}_t\{\text{corr}(F_t, R_t)\}\text{std}(F_t)\text{std}(R_t)}.
\]

The subscript \(t\) denotes expectation and standard deviation with respect to time. The approximation in equation (7) arises when the number of stocks stays approximately constant over time. If we standardize the dispersion of the factors\(^3\), then the information ratio is simply

\[
IR = \frac{\text{avg}_t\{\text{corr}(F_t, R_t)\}\text{std}(R_t)}{\text{std}_t\{\text{corr}(F_t, R_t)\}\text{std}(R_t)}.
\]

In the case where we have equation (4), then the original factors and actual returns should be used in equation (7) and (8).

Equation (7) and (8) might appear complicated at first glance. It helps readers to note that the correlation and the dispersions inside the bracket are cross-sectional and the average and standard deviation outside the brackets are time-series measures. Further simplification is possible for equation (7) and (8) under certain conditions. These equations essentially measure the information ratio of the products of the IC and the dispersions. It is often the case in practice that variability of IC relative to its mean is much higher than variability of the dispersions relative to their respective means. Then we can effectively treat the dispersions as constants, and as a result,

\[
IR = \frac{\text{avg}_t\{\text{corr}(F_t, R_t)\}}{\text{std}_t\{\text{corr}(F_t, R_t)\}}.
\]

\(^3\) Or equivalently, we adjust the risk aversion parameter, so it is always proportional to the forecast dispersion.
Equation (9) relates the performance of a cross-sectional model to the “information ratio” of the cross-sectional IC.

3. An example

This section provides a concrete example for the analysis in the previous sections. We study a one-factor cross-sectional model and its performance. The focus will on the statistical properties of the cross-sectional IC, the dispersion of factors and actual returns over time, as well as their relationships over time. Corresponding to equation (4) and (5), we shall build two versions of the model, one with the original factor for dollar neutral portfolios, the other with refined factor for dollar as well as beta neutral portfolios. We will show that the behaviors of the two versions are quite different.

3.1 The data

We use the Russell 3000 universe of stocks as the test case. The factor is of valuation type, based on company’s free cash flow. The beta and specific risk are based on the BARRA US equity risk model. The data are monthly and span from January 1987 to June 2001, and the models, portfolios and alphas are also updated monthly.

The actual number of stocks in the models is smaller than 3000 and it fluctuates from month to month. This has less to do with corporate actions, but more to do with the fact that not all data are available for all the stocks. Each month, we only include stocks for which all the data are available. For example, for June 2001, there were only 2489 stocks with all data available. To ensure factor accuracy and to prevent undue influence from outliers, we also trim the number of stocks by 5% on each side based on the sorted factor values. As a result, 2241 stocks remain for June 2001.

The valuation factor is based on the ratio of free cash flow to a company’s enterprise value, which can be thought of as a type of yield to shareholders and creditors. We first study the cross-sectional performance of this factor based on its raw form. The active alpha is then given by (4). Even though the information ratio is not affected by the choice of risk-aversion parameter $\lambda$ in our analysis. We shall select a value for it so to obtain the reasonable level of annual tracking error. The level is set at 5%.

3.2 Cross-sectional model with raw factor
Figure 1(a) shows the alpha stream with 5% ex-post tracking error over the entire period. For the period prior to 1996, the tracking error is remarkably low, probably well below 5%. From 1996 to the mid-1999, the alpha experiences a small increase in volatility. But from mid-1999 on, the volatility shoots up sharply, and alpha swing from month to month in excess of 5%. We note that the 12-month average is consistently positive until late 1999.

**Figure 1. Time series data for model with raw factors.**

Figure 1(b), 1(c), and 1(d) show the three components of the alpha. In Figure 1(b), the cross-sectional IC exhibits the same volatility pattern as in the alpha. And the 12-month average of IC also turns negative at the end of 1999. In Figure 1(c) and 1(d), the dispersions of the factors and actual returns are far less volatile compared to the IC. Interestingly, there is a declining trend of the dispersion of the factors from 1988 to 1999. One interpretation of this trend could be that the perceived cross-sectional opportunity declined because of investment arbitrage. Another possible interpretation is that this is actually due to economic competition as more companies enter business with high free cash flow yields and exit business with lower yield. Since we have used the dispersion of the factors in the alpha calculation, the sizes of the positions would be getting smaller.
over time. Note the trend began to reverse in mid-2000, indicating greater dispersion among companies in terms of the free cash flow yield.

On the other hand, the dispersion of the actual returns exhibits no such decline. The dispersion of the actual returns fluctuates between 10 to 20 per cent between 1987 to 1998. It escalated to a very high level since mid-1998, indicating the widening gap between winners and losers. In hindsight, one could claim a connection between the wide margin and the equity market bubble. The behavior of the IC, dispersions of the factors and actual returns during the period from mid-1999 to 2001 all contribute to the high volatility of alpha in that period.

Next, we analyze relationship among the three terms over the period. Table 2 summarizes the averages and standard deviations of the three terms, their ratios, and the correlation matrix. First, it is worth noting that the ratio for the IC is 0.49 while it is much higher for the two dispersions. This indicates that the dispersions are much more stable than the IC. As for the correlations, there exists a significant negative correlation between the IC and the dispersion of the actual returns. This will negatively impact the active alpha since the two terms are out of the sync with each other. As a result, the information ratio on a monthly basis is only 0.30, much smaller than the “information ratio” of the IC alone.

We also notice that there is little correlation between the dispersion of the factors and the other two terms. This implies that the perceived cross sectional opportunity so measured has no bearing on the actual opportunity in terms of the dispersion of the actual returns. Based on this observation and the fact that both dispersions are very stable over time, we conclude that standardizing the dispersion of the factors should not change investment performance. Indeed, the monthly IR remains at 0.30 if we do that.

Table 2. Statistical properties of model with raw factors.

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>std(f)</th>
<th>std(r)</th>
<th></th>
<th>IC</th>
<th>std(f)</th>
<th>std(r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg_t</td>
<td>0.05</td>
<td>7.0%</td>
<td>13.3%</td>
<td>IC</td>
<td>1.00</td>
<td>-0.06</td>
<td>-0.23</td>
</tr>
<tr>
<td>std_t</td>
<td>0.09</td>
<td>1.7%</td>
<td>3.8%</td>
<td>std(f)</td>
<td>-0.06</td>
<td>1.00</td>
<td>0.06</td>
</tr>
<tr>
<td>IR</td>
<td>0.49</td>
<td>4.12</td>
<td>3.54</td>
<td>std(r)</td>
<td>-0.23</td>
<td>0.06</td>
<td>1.00</td>
</tr>
</tbody>
</table>
3.3 Cross-sectional model with refined factors

We next investigate the cross-sectional model based on the same factor, but with a beta neutral constraint. The alpha is then given by equation (5), with refined $F$ and $R$ given by equation (6). Figure 2 presents the time series charts of alpha, IC, dispersions of the refined factors and the refined actual returns. Compared to Figure 1, the series in Figure 2 are more stationary. In particular, the active alpha [Figure (2a)] shows almost constant volatility throughout the entire period. The IC [Figure (2b)] does appear more volatile in the second half of the period, but the increase pales in comparison to Figure 1(b). Similar to Figure 1(c), the dispersion of the refined factor again declines from late 80’s to 2000 and it bottoms out in early 2000. Finally, the dispersion of the refined actual returns [Figure 2(d)] is rather stable, absent of the sharp increase in Figure 1(d) from late 1998 to 2001. Notice the refined factors and the refined actual returns are no longer in percentage, since both are ratios of adjusted “yield” or return over specific risk. Figure 2(d) seems to indicate that BARRA risk model has performed adequately in estimating the betas and the specific risks.

**Figure 2. Time series data for model with adjusted factors.**

Figure 3 provides cross-sectional average and dispersion of specific risks of the stocks. They were both rather stable until 1998. Both started to climb in 1998 and peaked
in January 2000 and have since declined. The pattern suggests that the declining of \( \text{std}(F) \) depicted in Figure 2(c) is more likely due to declining dispersion of free cash flow “yield” among stocks, not to the changes in the specific risks. It is also interesting to notice the high level of average and dispersion of stock specific risks in the late 90’s bull market.

**Figure 3. Time series of average and dispersion of specific risks.**

Table 3 summarizes the statistical properties of the IC and the two dispersions. Similar to Table 2, the ratio of average over standard deviation is much lower for the IC than for the two dispersions. The ratio for the dispersion of the refined returns is extremely high, almost rendering it a constant. Again, there exists a negative correlation between the IC and the dispersion of the refined returns. One different feature of this model is that the correlation between two dispersions is significantly positive. This is a desirable feature since the dispersion in the refined forecasts is indeed a predictor of the margin of opportunity in the refined returns. Aided by this positive correlation, the information ratio of the active alpha is 0.68, slightly higher than the “information ratio” of the IC alone.

**Table 3. Statistical properties of model with adjusted factor**

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>std(F)</th>
<th>std(R)</th>
<th></th>
<th>IC</th>
<th>std(F)</th>
<th>std(R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg_t</td>
<td>0.03</td>
<td>0.7%</td>
<td>1.1%</td>
<td>IC_t</td>
<td>1.00</td>
<td>-0.04</td>
<td>-0.16</td>
</tr>
<tr>
<td>std_t</td>
<td>0.06</td>
<td>0.2%</td>
<td>0.1%</td>
<td>std(F)</td>
<td>-0.04</td>
<td>1.00</td>
<td>0.24</td>
</tr>
<tr>
<td>IR</td>
<td>0.61</td>
<td>3.24</td>
<td>8.35</td>
<td>std(R)</td>
<td>-0.16</td>
<td>0.24</td>
<td>1.00</td>
</tr>
</tbody>
</table>
What if we standardize the dispersion of the refined forecasts over time? In that case, we lose the positive impact from the positive correlation between $\text{std}(F)$ and $\text{std}(R)$ and we suffer from the negative correlation between IC and $\text{std}(R)$. The information ratio decreases to 0.58. Hence, in this version of cross-sectional model, we benefit from using $\text{std}(F)$ to determine the amount of risk we should take each month.

4. Summary

This paper presents an analytic framework for quantitative cross-sectional equity models and addresses several theoretical and practical issues. First, we point out that besides cross-sectional IC, the cross-sectional dispersions of forecasts and actual returns also affect alpha of each time period. Second, we emphasize that when the active portfolio is constructed with risk-adjusted returns and additional constraints, which is typical of MV optimization, we must modify the definition of IC to accommodate the portfolio construction procedures. We give one such definition for portfolios with dollar neutral and beta neutral constraints.

Our analysis also shows that the long-term performance of cross-sectional models depends not only on the strength of IC, but also on pair-wise correlations between IC and two dispersions over time. Different relationship among the three terms leads to different outcomes in terms of model’s information ratio. This is reflected in our example of one factor model with free cash flow yield.

In the model with the raw factor, negative correlations exist between the IC and the two dispersions and as a result, the model IR is less than the IR of the cross-sectional IC alone. In the model with the refined factor, which is used for portfolios with beta neutral constraint, the two negative correlations are more than offset by the positive correlation between the two dispersions. As a result, the model IR is higher than IR of the IC. The comparison of the two models also strongly indicates that for free cash flow yield, the refined factor is superior to the raw factor.

Further work can be pursued in several areas. First, we should expand the analysis to other factors used in cross-sectional models and to other market segments in terms of size, style, and sector. For each combination, we can study the relationship between IC and dispersions. Second, an analytical framework of multi-factor cross-sectional models is desired. One focus of such analysis would be on how to combine several factors.
together in light of the new insights regarding a single-factor model. Finally, actively managed equity portfolios can have more constraints when it comes to implement cross-sectional models. We must seek to redefine the cross-sectional IC to accommodate different portfolio construction concerns. It remains to be seen whether this can be achieved for general case.

Acknowledgement

Edward Qian thanks members of Global Asset Allocation Group at Putnam Investments for stimulating this research. The authors benefited from discussion with Frederick Copper and Michael Abata at Putnam Investments. We are also grateful to Vladimir Liberson for his help in some of the calculations.

Appendix

This appendix provides mathematical details of mean-variance optimization, subject to both dollar and beta neutral constraints. And it also establishes a relationship between active alpha and appropriately defined cross-sectional IC.

The aim of the MV optimization is to maximize the objective function

\[
G(\hat{A}_t) = \hat{A}_t^\prime \hat{f}_t - \frac{1}{2} \lambda (\hat{A}_t^\prime S \hat{A}_t)
\]

where \( \hat{A}_t \) is the vector of active positions, \( \hat{f}_t \) is the vector of cross-sectional forecasts, \( S \) is the covariance matrix, and \( \lambda \) is the risk aversion parameter. We impose dollar neutral and beta neutral constraints on the optimization. We focus on these two since they are among the very basic constraints that most active equity portfolios require. Mathematically, the constraints are

\[
A_t^\prime \hat{f}_t = 0 \quad \text{and} \quad A_t^\prime \hat{\beta} = 0,
\]

where \( \hat{f}_t \) is a vector of ones and \( \hat{\beta} \) is the vector of betas.

The optimal solution for the active weights can be found by the method of Langrangian multipliers. The optimal active weights are

\[
\hat{A}_t = \hat{\lambda}^{-1} S^{-1} \left( \hat{f}_t - \lambda^{-1} I \hat{f}_t - \lambda^{-1} \hat{\beta} \right).
\]
The Langrangian multipliers are given by

\[
\begin{aligned}
l_1 &= \left( \beta' \cdot S^{-1} \cdot \beta \right) \left( \beta' \cdot S^{-1} \cdot \beta \right) - \left( \beta' \cdot S^{-1} \cdot \beta \right) \left( \beta' \cdot S^{-1} \cdot \beta \right) \\
l_2 &= \left( \beta' \cdot S^{-1} \cdot \beta \right) \left( \beta' \cdot S^{-1} \cdot \beta \right) - \left( \beta' \cdot S^{-1} \cdot \beta \right) \left( \beta' \cdot S^{-1} \cdot \beta \right)
\end{aligned}
\]  

\[(A4)\]

The term inside the parenthesis of equation (A3) is often referred to as the adjusted forecasts, which are obtained by subtracting the constant \( l_1 \) and \( l_2 \) multiplied by the beta from the forecasts. We denote the adjusted forecasts by

\[
\begin{aligned}
\gamma &= \left( \gamma - l_1 \cdot \beta - l_2 \cdot \beta \right).
\end{aligned}
\]  

\[(A5)\]

The product of actual returns and active positions produces the active alpha for the time period \( t \),

\[
\alpha_t = \beta \cdot \gamma, \quad \beta = \lambda^{-1} \cdot \beta \cdot S^{-1} \cdot \gamma,
\]  

\[(A6)\]

where \( \beta \) is the vector of actual returns. Since the forecasts are being adjusted by the constraints, we can adjust the actual return in the same manner as in equation (A5) without changing the value of alpha. Mathematically,

\[
\alpha_t = \lambda^{-1} \cdot \beta \cdot S^{-1} \cdot \gamma,
\]  

\[(A7)\]

where

\[
\gamma = \left( \gamma - k_1 \cdot \gamma - k_2 \cdot \beta \right).
\]  

\[(A8)\]

Even though we could choose the parameters \( k_1 \) and \( k_2 \) arbitrarily, it makes sense to choose \( k_2 \) as the market return \( r_M \) so the return (A8) is beta adjusted. The parameter \( k_1 \) adjusts the overall level of the actual return and its value will be determined later.

The choice of the covariance matrix \( S \) is made simple since our portfolio is market neutral. We can now select a diagonal covariance matrix with specific variances as the diagonal elements. Suppose the specific risks of stocks are \( (\sigma_1, \Lambda, \sigma_N) \), then we let

\[
S = \text{diag}(\sigma_1^2, \Lambda, \sigma_N^2).
\]  

\[(A9)\]

With equation (A9), the alpha given in equation (A7) simplifies to
\[ \alpha_i = \lambda^{-1} \sum_{i=1}^{N} \frac{\tilde{F}_i}{\sigma_i} \tilde{R}_i = \lambda^{-1} \sum_{i=1}^{N} F_i R_i. \]

Therefore, alpha is products of adjusted returns and adjusted forecasts, both then scaled by the specific risk. We shall call the final terms respectively the refined forecast and the refined return. We rewrite the sum of the products as covariance between the refined returns and forecasts and their means

\[ \alpha_i = \lambda^{-1} \{(N-1)\text{cov}(F_i, R_i) + N \cdot \text{avg}(F_i) \cdot \text{avg}(R_i)\}. \]

We now choose the free parameter \( k_i \) in equation (A8) such that the average of the refined returns vanishes. In financial terms, we require that the risk-adjusted specific returns for stocks in the portfolio have mean zero. As a result, the second terms in equation (A11) vanishes, and we have

\[ \alpha_i = \lambda^{-1} (N-1) \cdot \text{corr}(F_i, R_i) \cdot \text{std}(F_i) \cdot \text{std}(R_i). \]

We have rewritten the covariance as a product of correlation and dispersions. Equation (A12) shows that alpha consists of three terms: the correlation between the refined forecasts and returns, the dispersion of the refined forecasts, and the dispersion of the refined actual returns.
Cross-sectional Equity Models

Edward Qian & Ronald Hua
Putnam Investments
Northfield Conference, May 2002
Quantitative Models

- Time series models forecast asset returns over time
  - Tactical asset allocation, active currency management
  - Time series information coefficient (IC)

- Cross-sectional models forecast relative returns
  - Tactical asset allocation, active currency management
  - Active equity management
  - Cross-sectional IC
IC

- **Time series IC**
  - Single asset, over many time periods
  - Information ratio (IR) = IC
  - Extensive research

- **Cross-sectional IC**
  - Multiple assets, single period
  - Time series of cross-sectional IC
  - Little research
  - What is the IR of cross-sectional models?

\[ \rho_t = \text{corr}(\hat{f}_t, \hat{r}_t) \]
Naïve Model Construction

- Factor selection
  - Historical average cross-sectional IC
  - Valuation, momentum, earning quality, etc

- Back test
  - Active MV optimization versus benchmark
  - Risk models

- Performance
  - IR is not good, why?
Problems

- IC is not the whole story for realized alpha
  - IC determines model’s ranking ability
  - Dispersions in forecasts and actual returns determine magnitude of realized alpha

- MV optimization produces risk-adjusted optimal portfolios
  - Conventionally defined IC is not risk adjusted
  - Discount between IC and risk-adjusted alpha

- Other constraints
An Analytic Framework

- Relationship between realized alpha and IC and dispersions of forecasts and actual returns
- The role of IC when MV optimization is used to construct portfolio
  - Beta neutral optimal portfolio
  - Modified definition of IC
- The information ratio of cross-sectional models
Realized Alpha and IC

- Alpha = IC * opportunity * opportunity

\[ \alpha_t = \lambda^{-1}(N_t - 1) \text{corr}(\tilde{f}_t, \tilde{r}_t) \text{std}(\tilde{f}_t) \text{std}(\tilde{r}_t) \]

- Alpha = IC * opportunity

\[ \alpha_t = \lambda^{-1}(N_t - 1) \text{corr}(\tilde{f}_t, \tilde{r}_t) \text{std}(\tilde{r}_t) \]
Proof

- Math

\[ \alpha_t = \sum_{i=1}^{N_t} w_{i,t} \cdot r_{i,t} = \lambda^{-1} \sum_{i=1}^{N_t} (f_{i,t} - \bar{f}_t) \cdot r_{i,t} \]

\[ = \lambda^{-1} \sum_{i=1}^{N_t} (f_{i,t} - \bar{f}_t) \cdot (r_{i,t} - \bar{r}_t) \]

\[ = \lambda^{-1} (N_t - 1) \text{cov}(\bar{f}_t, \bar{r}_t) \]
Alpha = IC * Opportunity

- When equation is obvious, why try to be intuitive
- Automotive analogy
  - IC: gear
    - Forward or reverse
  - Opportunity in forecast: gas pedal
    - How hard do you apply acceleration
    - Cruise control
  - Opportunity in actual return: horse power
    - Is it a Yugo or BMW
    - The car is a rondomobile
IC in MV Optimization

- We usually don’t construction active portfolios so naively
- Many practitioners use MV optimization
- Active portfolios are often risk adjusted and beta adjusted
- Is IC still relevant?
Another Look

- When is these weights MV optimal
  \[ w_{i,t} = \lambda^{-1}(f_{i,t} - \bar{f}_t) \]
- Individual weight depends only on individual forecast. No correlation in covariance matrix
- No risk adjustment. Risks are the same for all stocks
- How to rescue IC
IC Is Saved

- Optimal portfolios
  - Dollar neutral, beta neutral, risk adjusted
- CAPM formulation

\[ r_i - r_f = \beta_i (r_M - r_f) + \epsilon_i, \]
\[ \epsilon_i \sim N(0, \sigma_i^2). \]
Realized Alpha and “IC”

- Alpha = “IC”* “opportunity”* “opportunity”
  \[ \alpha_t = \lambda^{-1}(N_t - 1) \text{corr}(\bar{F}_t, \bar{R}_t) \text{std}(\bar{F}_t) \text{std}(\bar{R}_t) \]

- Alpha = “IC”* “opportunity”
  \[ \alpha_t = \lambda^{-1}(N_t - 1) \text{corr}(\bar{F}_t, \bar{R}_t) \text{std}(\bar{R}_t) \]
Refined Forecast and Return

- $F$ and $R$ are dollar and beta adjusted and normalized by specific risk. We call them refined forecast and return.

$$F_{i,t} = \frac{f_{i,t} - l_1 - l_2 \beta_i}{\sigma_i} = \frac{\tilde{f}_{i,t}}{\sigma_i},$$

$$R_{i,t} = \frac{r_{i,t} - k_1 - r_{M,t} \beta_i}{\sigma_i} = \frac{\tilde{r}_{i,t}}{\sigma_i}.$$
New IC

- In this case the IC should be the correlation between the refined forecast and refined return
- The dispersions should those of refined forecast and refined return
- IC is saved
An Example

- How you calculate IC makes a big difference
- Three stocks

Table 1. An example of three stocks

<table>
<thead>
<tr>
<th>Stock</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$\sigma$</th>
<th>$r$</th>
<th>$\tilde{f}$</th>
<th>$\tilde{r}$</th>
<th>$F$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10%</td>
<td>1.25</td>
<td>50%</td>
<td>10%</td>
<td>-7.1%</td>
<td>-10%</td>
<td>-1.43</td>
<td>-0.20</td>
</tr>
<tr>
<td>2</td>
<td>20%</td>
<td>1.50</td>
<td>50%</td>
<td>30%</td>
<td>5.7%</td>
<td>20%</td>
<td>1.14</td>
<td>0.40</td>
</tr>
<tr>
<td>3</td>
<td>30%</td>
<td>0.25</td>
<td>50%</td>
<td>-10%</td>
<td>1.4%</td>
<td>-10%</td>
<td>0.29</td>
<td>-0.20</td>
</tr>
</tbody>
</table>

l1 31%
l2 -11%
k1 7%
rn 10%
Example

- Raw forecast has an IC of –0.50. It would overweight stock 3 and underweight stock 1.
- The refined forecast has an IC of +0.76. It would overweight stock 2 the most and underweight stock 1.
- The overweight in stock 3 is small.
The Long-term Performance

- Information ratio
  \[ IR = \frac{\text{avg}_t(\alpha_t)}{\text{std}_t(\alpha_t)} \]

- The goal of active investment management is to achieve positive excess return with minimum tracking error

- How do we access the IR of a cross-sectional model?

- Is “information ratio” of IC good enough?
The Long-term Performance

- It will depend on the interaction between the three terms:
  \[ \alpha_i \propto \text{corr}(F_t, \bar{R}_t) \text{std}(F_t) \text{std}(\bar{R}_t) \]

- Positive correlation among the three terms is desirable. Why?

- Intuition - IC is high, then dispersion is likely to be high, therefore alpha is high.

- If IC is negative, then dispersion is likely to be low, limiting the loss.
Examples

- Factor - free cash flow yield
  - Free cash flow to enterprise value (FCF2EV)
    \[
    \text{FCF2EV} = \frac{\text{CashFlowFromOperation} + \text{InterestExpense} - \text{CapitalSpending}}{\text{MarketCap} + \text{TotalDebt} + \text{PreferredStock} - \text{Cash}}
    \]

- Performance with raw factor
- Performance with refined factor (beta neutral)
- Trimmed Russell 3000 universe
- Monthly data from 01/1987 – 06/2001
Raw Factor IC

![Figure 1(b)](image_url)

- IC
- 12-month Average
Raw Factor Dispersion

![Figure 1(c)](image_url)

- std(f)
- 12-month Average

Northfield Conference, May 2002
Raw Return Dispersion

![Figure 1(d)](image)
Raw Factor

- **Statistics**
  
<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>std(f)</th>
<th>std(r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>avgₜ</td>
<td>0.05</td>
<td>7.0%</td>
<td>13.3%</td>
</tr>
<tr>
<td>stdₜ</td>
<td>0.09</td>
<td>1.7%</td>
<td>3.8%</td>
</tr>
<tr>
<td>IR</td>
<td>0.49</td>
<td>4.12</td>
<td>3.54</td>
</tr>
</tbody>
</table>

- **Correlation**
  
<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>std(f)</th>
<th>std(r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>1.00</td>
<td>-0.06</td>
<td>-0.23</td>
</tr>
<tr>
<td>std(f)</td>
<td>-0.06</td>
<td>1.00</td>
<td>0.06</td>
</tr>
<tr>
<td>std(r)</td>
<td>-0.23</td>
<td>0.06</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Figure 1(a)
Performance

- IR = 0.30 monthly
- If we set dispersion of factors as constant, IR = 0.30 monthly
- Both are much lower than IR if IC
- Realized tracking error for the last two years was significantly higher
Model With Refined Factor

- Recall

$$F_{i,t} = \frac{f_{i,t} - l_1 - l_2 \beta_i}{\sigma_i} = \frac{\tilde{f}_{i,t}}{\sigma_i},$$

$$R_{i,t} = \frac{r_{i,t} - k_1 - r_{M,t} \beta_i}{\sigma_i} = \frac{\tilde{r}_{i,t}}{\sigma_i}$$

- Beta and specific risk come from BARRA US equity model
Figure 2(b)

Refined Factor IC

-0.30
-0.20
-0.10
0.00
0.10
0.20
0.30
0.40

IC
12-month Average

Jan-87
Jan-88
Jan-89
Jan-90
Jan-91
Jan-92
Jan-93
Jan-94
Jan-95
Jan-96
Jan-97
Jan-98
Jan-99
Jan-00
Jan-01
Refined Factor Dispersion

Figure 2(c)

std(F) 12-month Average
Refined Return Dispersion

Figure 2(d)

std(R) 12-month Average
# Refined Factor

## Statistics

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>std(F)</th>
<th>std(R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{avg}_t)</td>
<td>0.03</td>
<td>0.7%</td>
<td>1.1%</td>
</tr>
<tr>
<td>(\text{std}_t)</td>
<td>0.06</td>
<td>0.2%</td>
<td>0.1%</td>
</tr>
<tr>
<td>IR</td>
<td>0.61</td>
<td>3.24</td>
<td>8.35</td>
</tr>
</tbody>
</table>

## Correlation

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>std(F)</th>
<th>std(R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>1.00</td>
<td>-0.04</td>
<td>-0.16</td>
</tr>
<tr>
<td>(\text{std}(F))</td>
<td>-0.04</td>
<td>1.00</td>
<td>0.24</td>
</tr>
<tr>
<td>(\text{std}(R))</td>
<td>-0.16</td>
<td>0.24</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Figure 2(a)

Alpha

12-month Average

Jan-87 Jan-88 Jan-89 Jan-90 Jan-91 Jan-92 Jan-93 Jan-94 Jan-95 Jan-96 Jan-97 Jan-98 Jan-99 Jan-00 Jan-01

-10.0% -7.5% -5.0% -2.5% 0.0% 2.5% 5.0% 7.5%

α
Performance

- IR = 0.68 monthly
- If we set the dispersion of refined factors constant, IR = 0.58 monthly. In this case it benefits from using dispersion to influence size of active positions. But we found for many factors, this is not true
- Both are close to IR of IC. The refined model is much better than the raw model
Summary

- Analytical framework of cross-sectional models
- \( \text{Alpha} = \text{IC} \times \text{opportunity} \times \text{opportunity} \)
- For realistic portfolio construction processes, forecast, return, and IC needed to be redefined
- The information ratio of the model depends on the strength of IC, as well as statistical relationship among three parties
- For free cash flow yield, the beta neutral cross sectional model is much superior