WHITE PAPER #8

Characteristics of Estimates from a Fair Value Methodology; Comparing Alternative Models

Peter Ciampi, Interactive Data Pricing and Reference Data, Inc.
March 11, 2004

We propose in this white paper that a properly constructed fair value methodology should produce fair value estimates that are both closer to the next day's opening price than to the current day's local closing price and not inappropriately biased. This white paper (1) discusses these characteristics, (2) outlines tests for measuring them, (3) applies those tests to the results of Interactive Data Pricing and Reference Data's Fair Value Information Service model for a hypothetical portfolio of the 100 largest European and Asian companies by capitalization in 2003 (referred to as the "Sample"), and (4) discusses an objective way to compare models with similar test results.

Closeness

Fair value estimates (or estimates) should be "closer" to the next day's opening price (the open) than the previous day's closing price (the close) is. The following three tests can be used to measure "closeness".

Counting Test: This test counts how many times the estimate was closer to the open than the close was. In the Sample, evaluations produced by our model were closer to the open on 74% of the days. However, the utility of the counting test is limited because it doesn't measure the degree of closeness. For instance, if the close is $10.00 and the open is $11.00, then a $10.01 estimate is equivalent to an $11.01 estimate using the counting test.

Reduction in Average Absolute Distance Test: To measure the degree of closeness, you could use a reduction in average absolute distance test. In the preceding example, the distance from the open–to-close price was $1.00; so the $10.01 estimate, with a fair value estimate-to-open distance of $0.99, was 1% closer whereas the $11.01 estimate, with a fair value estimate-to-open distance of $0.01, was 99% closer. (You would use absolute value because being one cent above the open, i.e., $11.01, is the same distance as one cent below, i.e., $10.99.) In our Sample, the average close-to-open distance was .62% whereas the average fair value estimate-to-open distance was .33%. This tells us that the fair value evaluations produced by our model were 43% closer than the closes for the corresponding securities.

Variance Reduction Test: A variance reduction test is similar to the reduction in average absolute distance test, except that it places more weight on avoiding large differences.[1] The variance reduction test computes how much smaller the variance of fair value estimate-to-open is than the variance of the close-to-open. In the Sample, the variance reduction of the evaluations produced by our model was 58%.

[1] Specifically, the variance reduction test measures the decrease in the average absolute distance *squared*. A fair value methodology performs well on this test if it approximates the *mean* expected open, whereas it performs best on the distance reduction test if it approximates the *median*. In practice, these are often similar. But since value is usually linked to mean expectations, the variance reduction test is a better test of the performance of a model.

Controlling for post-4:00 p.m. Eastern Time (ET) events

While the open can be used to measure the effectiveness of a fair value estimate, there is a lag between the fair value estimate at 4:00 p.m. (ET) and the open. For example, there is a 12 hour lag in Europe, where trading begins on the exchanges at 4:00 a.m. (ET). A fair value estimate may be appropriate at 4:00 p.m. (ET), but intervening events, such as surprise earnings from Intel or the start of unannounced stock purchases by the Bank of Japan, could cause measurement errors.

One way to examine events occurring during this gap is to use an objective source such as the S&P futures market, which is actively traded twenty-four hours a day. (The alternative of looking for news events could be criticized as subjective "data mining.")

In the Sample counting test, the evaluations produced by our model were closer on 74% of the days, but what happened on the 26% of days when our evaluations weren't closer? To understand this, you need to look at post-4:00 p.m. (ET) events. The Fair Value Information Service is designed to provide information that can be used to estimate a price for a foreign exchange-traded equity security in view of market information available at the close of the US markets, typically at 4:00 p.m. (ET). To examine the effects of post-4:00 p.m. (ET) events, you could consider S&P "reversals", meaning when the 4:00 p.m. to 4:00 a.m. (ET) change in the S&P futures is opposite the change in the S&P futures from 11:30 a.m. to 4:00 p.m. (ET). To be meaningful, the reversal must be more than minimal. For purposes of the Sample, we disregarded reversals of less than 20%. For example, if the S&P 500 Futures Index was up by 1.0%, then a post 4:00 p.m. (ET) move in the S&P 500 Futures by 0.2% or greater would be considered a reversal. With this criterion, there were reversals on 17% of the days in the Sample. When we superimpose that fact on the counting test, our Fair Value Information Service model was actually closer on 91% (74% + 17%) of the days.

Rather than counting S&P reversals, a more systematic way of controlling for post 4:00 p.m. (ET) events is through regression. Variance reduction is essentially the R-squared (a statistical measure of how well the regression line approximates the data points) from a regression of close-to-open on fair value estimate-to-open. By regressing close-to-open on both fair value estimate-to-open and on post-4:00 p.m. (ET) changes in the S&P future, we systematically control for these changes. With this more refined analysis, variance reduction for our Fair Value Information Service model increases from 58% to 71%.

Test	**%**	**Comments**
Closeness 1: Percent of days that Fair Value was closer to the open than the close was	74%	On 185 of 251 days, fair value estimate was closer
Closeness 1: Percent of days that fair value estimate was closer to the open than the close was, considering post-4:00 p.m. (ET) reversals	91%	On 228 of 251 days, fair value estimate was closer
Closeness 2: Reduction in absolute average distance	43%	Daily average absolute distance of fair value estimate-to-open was .33%. Close-to-open was .62%
Closeness 3: Variance reduction	58%	
Closeness 3: Variance reduction, considering post-4:00 p.m. (ET) reversals	71%	

Lack of bias

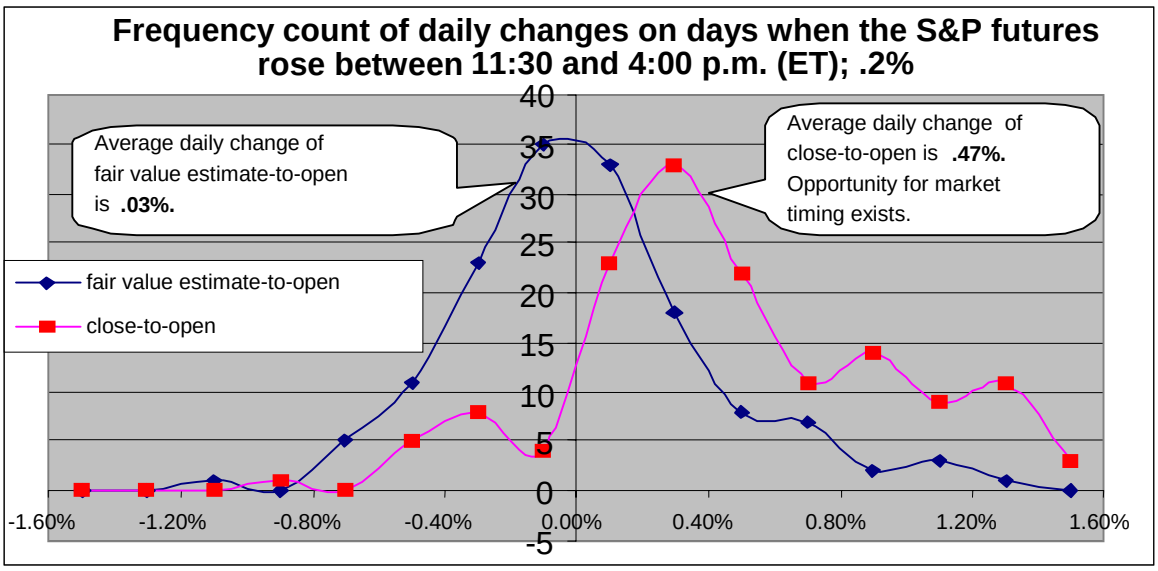
Fair value estimates from a properly constructed model should not be "biased." If the fair value estimate is consistently higher (or lower) than the next day's open, it would systematically increase (or decrease) a fund's net asset value. This is an undesirable feature of a model, which an arbitrageur could exploit. To simplify the discussion of bias, we restricted the Sample to the 145 days during the 251 day period when the S&P Futures Index rose. One test for bias is counting the number of days that the fair value estimate is above and below the open. If a model is unbiased, days above should be about equal to days below. In the Sample, the evaluations produced by our model were above the open 72 days and below the open on 71 days.

This counting test is helpful but, again, it does not measure the size of the distance above and below the open. For instance, if the fair value estimate was above the open by one dollar on 50 days and below it by one penny on 50 days, the system has a bias even though it passes the counting test. A test without this weakness is: the average of fair value estimate-to-open changes (not absolute value) should be about zero.

In the Sample, the average daily change of evaluations produced by our model was three basis points. This is not statistically different from zero, that is, it's within two standard errors of zero. The average for close-to-open was 47 basis points.

In White Paper #3, we outlined how fair value estimates should approximate the values that would prevail in a liquid market; they should not allow arbitrage profits. "Lack of Bias" is an alternative, but more direct, test of this quality. In the Sample, if local closing prices were used, an arbitrageur who purchased the fund on the 145 days when the S&P Futures Index increased would have profits on 127 (or 87%) of the days, totaling 68% (.47% * 145). Alternatively, based on evaluations produced by our model, the arbitrageur would have profits on 72 days but losses on 71 days. Total profits would have been 5% (.03% * 145).

If fair value estimates have no predictable future changes, fair value estimate-to-open changes should be distributed with an average of zero. The frequency plot of daily changes (below) shows that fair value estimate-to-open changes have this feature. It also shows that close-to-open changes do not have this feature. Instead, the close-to-open changes are both shifted and skewed to the positive.



Comparing the performance of alternative models

The tests discussed above can be used to determine the efficiency of fair value estimates. What happens, however, when competing models yield similar results for these tests? We propose the following method for comparing such models. Collect three sets of data:

- Actual Data—Close-to-open changes for a number of days
- Model 1 Data—Fair value estimate-to-open changes from Model 1
- Model 2 Data—Fair value estimate-to-open changes from Model 2

A regression of close-to-open on fair value estimate-to-open from Model 1 and on fair value estimate-to-open from Model 2 produces coefficients for the adjustment for each model. The coefficients define the weight on the two models that would have produced the best results for the time period being analyzed. For example, if the coefficients are 0.5 and 0.5, then a 50-50 weighting would have produced the best fair value prices. If they are 1.0 and 0.0, then the best approach is to use Model 1 and to ignore Model 2.

Summary

A properly constructed fair value methodology should produce estimates with two characteristics: (1) they should be closer than the local close to the next trade and (2) they should have no systematic bias.

We outlined tests to measure these characteristics and showed that some tests are more powerful than others. In performing these tests, it's important to control for significant events occurring in the gap between the time the fair value estimate is constructed and the time of the next exchange trade. We propose using the S&P future in this control.

In cases where two models appear to be similar, a regression of actual changes on data from each model can objectively differentiate their relative performance.