On the Use of Intercepts as Performance Measures

Jim Musumeci
Department of Finance, MOR 107
Bentley University
Waltham, MA 02452
jmusumeci@bentley.edu
781.891.2235

Current Draft: January, 2018

The author thanks Atul Gupta, Dhaval Dave, Diana Harrington, Joe Marks, Mindy Musumeci, Ralph Oliva, David Rakowski, Kartik Raman, Richard Sansing, Jeff Stark, and Jingyun Yang for their helpful comments. The usual disclaimer applies.

On the Use of Intercepts as Performance Measures

Abstract

Intercepts have long been used as measures of abnormal performance in finance. Jensen's alpha made its debut in 1968, but more recently researchers have used intercepts from the Fama-French Three-Factor model and others as *ex post* performance measures. A positive intercept is taken as evidence that the stock (or portfolio) exceeded expectations, and conversely for a negative intercept.

We find that even conditional on the premise that the model used is valid, the intercept's ability to detect abnormal performance is critically dependent on accurately estimating the slope(s). Unfortunately, the same unsystematic perturbation to stock returns that causes abnormal performance typically distorts this estimate of the slope, in which case the intercept can become misleading, in some cases so misleading that it will move in the direction *opposite* that of the perturbation. This paper quantifies the distortion in the intercept due to errors in estimating the slope(s). It also identifies necessary and sufficient conditions for the intercept to be so misleading as to move in the direction opposite that of the shock, estimates the frequency of these conditions, and provides hypothetical and actual examples. Finally, while buy-and-hold abnormal returns remain a promising alternative that are unaffected by this issue, for any who prefer to continue using intercepts we offer an alternative approach that alleviates the problem of distorted slopes.

JEL Codes: G10, G14, C13
Keywords: Performance Measurement

On the Use of Intercepts as Performance Measures

Consider the following conversation between a student and his professor:

> Student: I don't understand why I got a B+ in the course and my roommate got an A-. We scored exactly the same on every assignment except the first one, and I scored a 90 while he scored only an 80.
> Professor: You earned a lower grade *because* you received a better score on the first assignment. Had you also scored an 80, you would have received an A- as well. Had you scored worse—70 or less—you would have received an A.

This case will soon be heard by the department chair or associate dean, and it is difficult to imagine the assigned grade of B+ will be upheld. It is axiomatic that better performance on any single component of total performance should, *ceteris paribus*, be associated with a better assessment of total performance. Nevertheless, we find that a common measure of stock or portfolio performance in finance—the intercept from any model that regresses asset returns on other factors, such as the Fama-French Three-Factor model—often fails this simple requirement.

Jensen's alpha began a long history of using intercepts as measures of excess performance. The Capital Asset Pricing Model (CAPM) asserts $E[R_s] = R_f + \beta[E(R_m) - R_f]$, which Jensen [1968] expresses as $[\tilde{R}_{s,t} - R_{f,t}] = \beta[\tilde{R}_{m,t} - R_{f,t}] + \varepsilon_{s.t}$. If $[\tilde{R}_{m,t} - R_{f,t}]$ is taken as an independent variable and $[\tilde{R}_{s,t} - R_{f,t}]$ as a dependent variable, then estimating parameters of a linear regression $\left[\tilde{R}_{s,t} - R_{f,t}\right] = \hat{\alpha} + \hat{\beta}[\tilde{R}_{m,t} - R_{f,t}] + e_{s.t}$ should produce an intercept $\hat{\alpha}$ equal to zero if the stock has exactly met CAPM expectations. Positive values of $\hat{\alpha}$ suggest the stock has done better than expected; negative values indicate it has underperformed. Intercepts in subsequent models such as the Fama-French Three-Factor model have been used with similar inferences drawn from the sign of the intercept.[1]

Unfortunately, for these intercepts to work properly, it is important to have a good estimate of the slope(s), and the same unsystematic events that lead to abnormal performance will also distort our slope estimates. For ease of exposition, we begin with an analysis of Jensen's alpha. After developing intuition regarding the problem (and a solution for anyone who chooses to continue using intercepts), we proceed to extend the analysis to intercepts of the Fama-French model and find the problem is *more* frequent in that context than it is for Jensen's alpha. Indeed the problem (and solution) can be extended to virtually any other multifactor model as well. We emphasize that we are not in any way suggesting such models are flawed or

[1] Cremers et al [2012] find the Fama-French model produces significantly non-zero intercepts for common benchmark indices for reasons related to the weighting of the factor components. The problem we identify here is distinct from the one they analyze.

inappropriate. Instead, we focus on a much narrower issue, namely, that the use of such models' intercepts as performance measures is problematic.

For a simple example in the context of Jensen's alpha, consider only two pairs of observations of a stock's and the market's returns in excess of the riskless rate and suppose the stock's true (unobservable) β is one. Suppose also that for the first observation, the market's return in excess of the riskless rate is $RP_{m,1}$ = 1%, and for the second observation it is $RP_{m,2}$ = 3%. Finally, suppose that for each of the observations of stock return, the net effect of all unsystematic events is 0, so that $RP_{s,1}$ = 1% and $RP_{s,2}$ = 3% as well. With only two observations, the regression line passes exactly through both points, and so based on these observations our estimate of β is one, and Jensen's alpha is zero. The stock has performed exactly as expected.

Now suppose instead that for the second observation, the net effect of the unsystematic influences on share price was -1%, so that the actual observed value of $RP_{s,2}$ is 2%. The estimate $\hat{\beta}$ falls to 0.5, but Jensen's $\hat{\alpha}$ rises to 0.5%. Although the net effect of the unsystematic events was to cause this stock to perform worse, the resulting estimation error in $\hat{\beta}$ has caused it to have a positive Jensen's $\hat{\alpha}$, which should indicate superior performance. Similarly, if for the second observation the net effect of unsystematic effects on the stock were +1%, the estimate $\hat{\beta}$ would rise to 1.5, and Jensen's $\hat{\alpha}$ would fall to -.5%.

Note that this odd effect of a shock causing a movement in $\hat{\alpha}$ that is directly opposite the direction of the shock occurs only if it is the observation farther from zero[2] that experiences the shock. In our simple example, if the first observation's dependent variable experiences a shock of ±1%, Jensen's $\hat{\alpha}$ moves in the same direction (albeit by an exaggerated amount). As we demonstrate in Appendix A, because the regression line pivots around the point ($\overline{RP}_m, \overline{RP}_s$), only if the mean value of $\overline{RP}_m$ is between the $RP_{m,i}$ experiencing the perturbation and the origin can Jensen's $\hat{\alpha}$ move in the direction opposite that of the shock, but all errors in estimating β have implications for the estimate $\hat{\alpha}$.

We also show in Appendix A that our example with only two observations is not as limiting as it might seem. The same paradoxical change in Jensen's $\hat{\alpha}$ also occurs in larger samples, as such shocks alter $\hat{\beta}$ by tilting the the regression line in the direction of the shock.[3] Specifically, in Appendix A we find that for an unsystematic perturbation c to the stock return of the i^{th} observation, the change in $\hat{\alpha}$ expressed as a function of c and i is

---

[2] I.e., if both values of the independent variable were negative, it would now be the more negative one that would cause the aberrant result.

[3] There is not a problem when the shock is spread evenly over each observation, but such a case seems inconsistent with market efficiency. In any event, there are an infinite number of ways to spread abnormal performance across multiple observations; this paper focuses solely on the case in which the shock affects only one observation (or a small number of observations). This approach is consistent with the literature suggesting stock returns are a mixture of normal distributions, e.g., Kon [1984] and Harris [1986], and that they are characterized by skewness and kurtosis, e.g., Fama (1965), Fama et al (1969), and Brown and Warner (1980).

$$\Delta\hat{\alpha} = \hat{\alpha}(c,i) - \hat{\alpha}(0,i) = \frac{c}{n} - \frac{c}{n}\frac{(RP_{m,i} - \overline{RP}_m)}{\hat{\sigma}^2_{RP_m}} \cdot \overline{RP}_m. \quad [1]$$

The first term on the right side reflects the change in $\hat{\alpha}$ that is due to the increase in the mean stock return; the second term is the indirect effect on $\hat{\alpha}$ due to the change in the estimate of $\hat{\beta}$ that c causes. We note that if the perturbation occurs when the market risk premium is exactly at its mean, $\overline{RP}_m$, then the second term is zero and the change in $\hat{\alpha}$ is just the average change in the stock's risk premium, $\frac{c}{n}$. Moreover, the average value of the second term is zero for the sample, so the average unconditional distortion is zero. However, in any specific case, the farther away the market risk premium is from its mean when the perturbation c occurs, the more $\hat{\alpha}$ is affected. If we factor $\frac{c}{n}$ from the right side of [1], we get

$$\hat{\alpha}(c,i) - \hat{\alpha}(0,i) = \frac{c}{n}[1 - \frac{(RP_{m,i} - \overline{RP}_m)}{\hat{\sigma}^2_{RP_m}} \cdot \overline{RP}_m]. \quad [2]$$

The term in brackets represents one plus the proportion of distortion that occurs because the perturbation c affects $\hat{\alpha}$ directly and also indirectly through its effect on the estimate of the slope. It is when the magnitude of the second term in brackets is larger that we get greater distortion, which causes Jensen's $\hat{\alpha}$ to misrepresent the shock. We define the entire expression in brackets to be the sensitivity factor for month i, or SF_i. One perhaps surprising characteristic of [2] is that, except for the perturbation c itself, the change in $\hat{\alpha}$ is completely independent of stock returns. Thus we can estimate the frequency of the problem by focusing exclusively on market returns, specifically the distribution of the sensitivity factor SF_i. As suggested by the example in the introduction, there is no guarantee that the sensitivity factor will even be positive, which implies that a positive shock c can lead to a decrease in $\hat{\alpha}$. Finally, neither sample size nor the magnitude of c affects the sensitivity factor; thus if the sensitivity factor is negative, neither a sufficiently small shock nor a sufficiently large sample will keep c and the change in $\hat{\alpha}$ from having opposite signs, although both smaller c or larger n will cause the distortion to be smaller.

This potential distortion in Jensen's alpha is reminiscent of Blume's (1975) observation that while the expected value of $\hat{\beta}$ conditional on β is β, the expectation of β conditioned on $\hat{\beta}$ is not generally $\hat{\beta}$.[4] Blume's tests, however, found that this distortion generally didn't affect the final results very much. We find the opposite is true for Jensen's alpha. When the sample contains an observation that produces a negative sensitivity factor, the resulting estimates of

[4] An example explaining why (discussed on p. 788 of Blume) is that the population mean of β is presumably unimodal with a mean of one, so that if estimation error is symmetric around zero and also unimodal, when $\hat{\beta}$ exceeds one, there are more values of $\beta < \hat{\beta}$ that would produce that $\hat{\beta}$ than there are values of $\beta > \hat{\beta}$ that would do so. Thus, when $\hat{\beta} > 1$, E($\beta|\hat{\beta}$) < $\hat{\beta}$, and similarly, E($\beta|\hat{\beta}$) > $\hat{\beta}$ when $\hat{\beta} < 1$.

Jensen's alpha (and, for that matter, of beta) are significantly distorted for any firms happening to have unusually high or low performance that month. The estimate of Jensen's alpha given a shock that happens to occur in a month with a negative sensitivity factor is *unconditionally* unbiased, but is biased up for firms with negative shocks and biased down for firms with positive shocks.

This distortion is also similar to a violation of first-order stochastic dominance because values of two stocks' returns may be identical to each other for all but one observation, and yet in the presence of a negative sensitivity factor we would conclude that the stock with the higher return that month underperformed the stock with the lower return. We take it as self-evident that if actual performance is positively related to each component of performance, then any metric that *ceteris paribus* alleges a stock performed better when a component performed worse has a serious problem.

II. Frequency of Negative Sensitivity Factors for Jensen's Alpha

We first focus on the frequency of negative sensitivity factors; after all, if it is a once in a century event, it might be dismissed as only an academic curiosity with little practical significance. On the other hand, it merits greater attention from academics and practitioners if it is a fairly frequent occurrence.

For the 720-month interval from February, 1955 to January, 2015 we collected monthly 3-month T-Bill rates from the St. Louis Federal Reserve website and S&P 500 index values from the CRSP database. We then used these series to estimate periodic monthly market risk premia, and alternately partitioned the 720-month dataset into 24 non-overlapping 30-month and 20 non-overlapping 36-month subsets. For convenience, within each subset of data we order the data from smallest to largest risk premium on the market. Because [2] is monotonic in $RP_{m,i}$ within each subset, SF_1 is the minimum value for the factor and SF_n the maximum. Within each subset, we found SF_1, SF_n, and the standard deviation of SF_i.

Panels A and B of Table 1 report the results for these two partitions of the data. Recall that for any given window the mean value of the sensitivity factor must be 1.0, but that individual values of the sensitivity factor will be more distant from 1.0 when the observation under consideration has a value more distant from the sample mean. For the 30-month windows, the average standard deviation of the samples' sensitivity factors was about .212, suggesting that the change in Jensen's $\hat{a}$ associated with a shock c would be accurately measured within a factor of 1 ± .212 about 68% of the time. The 36-month windows of Panel B produce a similar result, with an average standard deviation of the sensitivity factor of about .199. These suggest that, on average, Jensen's $\hat{a}$ does a reasonably good job of measuring unsystematic shocks. Indeed, we found that the sensitivity factor was in the interval [.5, 1.5] just over 93% of the time.

However, the range of standard deviations of the sensitivity factors suggests that while the average standard deviations are in the ballpark of 20%, there is still significant dispersion among the different partitions. The 30-month windows, for example, have a minimum standard deviation of .003 and a maximum of .693. The 36-month windows have a similarly large range, from a minimum of .007 to a maximum of .461. It is when these standard deviations are large that we encounter significant problems. For example, the 30-month window with the standard deviation of .693 was 2/1/1995—7/1/1997, and the most extreme months in that window were June, 1996 (with a sensitivity factor of 2.627) and June, 1997 (with a sensitivity factor of -.255). Consider any stock during the 2/1/1995—7/1/1997 window. If that stock experienced an unsystematic perturbation of, say, +30% during June, 1996, the sensitivity factor indicates that instead of an increase in Jensen's $\hat{a}$ of $\frac{30\%}{30}$ = 1%/month, we would actually measure it to be 2.627%/month. Even more bizarrely, if a +30% shock occurred during June, 1997, instead of showing a 1% increase in Jensen's $\hat{a}$, estimating the regression parameters would give us a *decrease* of -.255%/month. The reason in both cases is that the shock of 30% changed our estimate $\hat{\beta}$ so that it completely dwarfed the shock's effect on $\hat{a}$.

It is fairly easy to see the main source of the dispersion in standard deviations of the sensitivity factor across the different windows by taking closer looks at equation [2] and at Table 1. Equation [2] shows the second term of the sensitivity factor to be $\frac{(RP_{m,i} - \overline{RP}_m)}{\hat{\sigma}^2_{RP_m}} \cdot \overline{RP}_m$. Because $\overline{RP}_m$ is a factor of this term, this term's standard deviation will be proportional to it, and windows where the mean market risk premium happens to be relatively close to zero will have low standard deviations and Jensen's $\hat{a}$ will have little estimation error. Similarly, when the mean market risk premium has a relatively large absolute value, Jensen's $\hat{a}$ will have significantly larger error. This is corroborated by both Panels of Table 1. In Panel A, the largest standard deviation of sensitivity factors is 2/1/1995—7/1/1997, which is also the one with the highest absolute value of average market risk premium (1.996%/month). Similarly, the window with the lowest standard deviation of sensitivity factors (.00342) is 2/1/1960—7/1/1962, and it is also the window featuring the lowest absolute value of market risk premium, 0.013%. The correlation between the absolute value of the average market risk premium and the standard deviation of the sensitivity factors is quite large at 0.889. Similar results are obtained for the 36-month windows.

It is also true that $\frac{1}{\hat{\sigma}^2_{RP_m}}$ is a factor of the second term of the sensitivity factor, but it varies much less than $\overline{RP}_m$. For this reason, although the correlation between $\frac{1}{\hat{\sigma}^2_{RP_m}}$ and the standard deviation of sensitivity factors is positive (.199), it is not anywhere near as large as the .889 correlation between $\overline{RP}_m$ and the standard deviation of sensitivity factors for 30-month windows.

Overall, with 720 months divided into 24 time periods of 30 months each, we found three months in two distinct time periods that had a negative sensitivity factor (October, 1962; November, 1996; and June, 1997). Depending on whether we count months or windows, our best estimate of the probability that a 30-month window will contain a month with a negative sensitivity factor is between 8.33% and 12.5%. Similarly, we found two months (March, 2003 and September, 2008) with negative sensitivity factors when we considered 36-month windows, so that our best estimate of the probability that a 36-month window will contain a month with a negative sensitivity factor is about 10.0%. We note that the 30-month windows and the 36-month windows had no common months with negative sensitivity factors. Thus we cannot simply identify the problem months in isolation; whether or not a month has a negative sensitivity factor depends on the window of which it is a part.

III. Simulations Showing the Effect of Negative Sensitivity Factors

To get an idea of the potential effects of the problem and explore potential solutions, we begin this section with a simulation, and then proceed in the next section to an empirical examination with actual data during a timeframe containing a negative sensitivity factor. The reason for proceeding in this order is that only with a simulation can we control the underlying parameters to be certain that our results are attributable to a negative sensitivity factor rather than actual excess performance.

The simulations are conducted in the following way. First, we simulate 29 normally distributed monthly market risk premia with mean of 1.5% and a standard deviation 5%. Next, for each of these 29 market risk premia, a stock risk premium was simulated as the market return plus a normally distributed noise term with mean zero and standard deviation 5%. Thus, for these 29 observations, the population α is zero and the population β is one.

Next, a 30th monthly market risk premium is generated to ensure the sensitivity factor is negative (provided the sample's average market risk premium is positive). Specifically, the 30th market risk premium is chosen to be

$$\overline{RP}_m + \frac{\hat{\sigma}^2_{RPm}}{\overline{RP}_m} + 3\%. \qquad [3]$$

This condition makes it extremely likely that, provided $\overline{RP}_m > 0$, then the sensitivity factor in Eq. [2] will be negative. Then to actually examine the results of Eq. [2], the stock's residual was simulated to be three standard deviations, or 15%, above or below (probability = 0.5 each) the market risk premium for the 30th observation.[5] For this set of 30 observations, because 29

[5] This may seem excessive, but even if the residuals are normally distributed, this should occur about a quarter of a percent of the time. However, Fama (1965), for example, finds in his Table 1 (p. 47) that DJIA stock returns are more than three standard deviations from their mean over four times as often as that, and hints that for smaller stocks the deviations from a normal are likely more pronounced than that.

stock returns are chosen with an average residual of zero and the 30th stock return has a residual of ±15% (with equal likelihoods of being positive or negative), the unconditional expectation of the 30th month's residual return is also zero, and so $\hat{\beta}$ is unbiased and the expected average Jensen's alpha is zero. However, conditioned on whether the stock's residual in the 30th month is positive or negative, the average excess return for all thirty months is either -0.5% or +0.5%.

This process was repeated 2277 more times, and then partitioned into subsamples with negative (N = 1193) and positive (N = 1185) residuals. While the problems with negative sensitivity factors occur whether the average market risk premium is positive or negative, in the latter case the last term of the 30th market risk premium in expression [3] would need to be -3%. For simplicity and without loss of generality, we trimmed the partitions by market risk premium to 1000 in each of the negative and positive 30th month residual subsamples to avoid negative (and unusually large positive) market excess returns.

Table 2 reports the entire sample's (N = 2000) results (1) of regressions using only the first 29 observations, (2) of using all 30 observations as is conventionally done, and finally (3) of using only the first 29 observations to estimate $\hat{\beta}$, and then using this estimate of $\hat{\beta}$ to estimate $\hat{\alpha}$ using all 30 observations. The purpose of this last (and seemingly unusual) procedure is to keep the 30th observation from distorting $\hat{\beta}$, and yet to consider the 30th observation's effect on abnormal performance. After all, by construction, that expected abnormal performance for all 30 observations is either +50 or -50 basis points, and so the 30th observation has the potential to distort $\hat{\beta}$ and consequently $\hat{\alpha}$. This procedure is in the spirit of trimming or Winsorizing, except that those two techniques are often applied indiscriminately, whereas here we make an adjustment only for months with a negative sensitivity factor, which a priori we know will distort Jensen's alpha. For Table 2 we should expect the unconditional effect of the last observation's perturbation to be zero, i.e., E[$\hat{\beta}$] = 1 and E[$\hat{\alpha}$] = 0.

The first three columns of Table 2 describe the average market risk premium, average maximum market risk premium, and average sensitivity factor for all 2000 simulations. By construction, the last month's average sensitivity factor of -0.0814 is negative so that the shock in the 30th month will result in a larger (smaller) estimate $\hat{\beta}$ and a smaller (larger) estimate of Jensen's alpha when the last month's residual is positive (negative). The next two columns show the average estimates for the first 29 observations in each sample. Because these were based on simulations with population parameters β = 1 and α = 0, it is no surprise that we do not reject either of the null hypotheses pertaining to these two.

The next two columns show the results when the 30th market risk premium produces a negative sensitivity factor and the actual stock risk premium for the 30th observation is ±3 standard deviations (or ±15%) from the market risk premium. In this table, we see no unconditional effect on our estimates $\hat{\beta}$ or $\hat{\alpha}$ because the distortions are offsetting. Similarly, in the last column, we find estimating $\hat{\beta}$ from only the first 29 observations, but then applying it to all 30 observations also produces virtually the same results for $\hat{\alpha}$. As expected, neither the unconditional value of the average $\hat{\beta}$ nor that of $\hat{\alpha}$ is significantly different from the population means of one and zero.

Things are quite different, however, when we partition the sample by the sign of the 30th month's residual and look at conditional values of $\hat{\beta}$ and $\hat{\alpha}$. In Table 3 we examine the results conditioned on the last month's residual = -3σ = -15%. Because this shock is applied only to the last month's stock return, $\hat{\beta}$ and $\hat{\alpha}$ for the first 29 observations should be close to the population values, and indeed they are at 0.994 and 0.0050%. However, the next two columns show the results for all 30 observations are quite different. As shown by Eq. [2], the 30th observation's negative shock will likely to bias $\hat{\beta}$ downwards and Jensen's alpha upwards, despite the negative shock. Indeed, the average $\hat{\beta}$ falls to 0.7516, allowing us to reject the (true) null that β = 1 with a t-value of -57.7691. In addition, Jensen's alpha rises to 0.0817%, with a t-value of 2.6703, this despite the fact that the simulated average excess monthly return for the stock is a negative 50 basis points.

Perhaps not surprisingly, of the three estimates of Jensen's alpha [(1) with $\hat{\beta}$ and $\hat{\alpha}$ both based on the first 29 observations, or (2) both based on all 30 observations, or with (3) $\hat{\beta}$ based on the first 29 observations, then $\hat{\alpha}$ based on this estimate of $\hat{\beta}$ applied to all 30 observations], the third choice provides the estimate of Jensen's alpha closest to the population value of -0.5%.[6] This method produces a negative estimate of Jensen's alpha, -0.4943%, with a t-value of -15.676, in sharp contrast to the standard method, which would produce a set of Jensen's alphas that has the wrong sign with a significant t-value of 2.6703

Finally, in Table 4 we examine the results when the 30th stock return's residual is +3σ = 15%. The results are analogous to those of Table 3. When only the first 29 observations are considered, average $\hat{\beta}$ is once again insignificantly different from one (1.0014, with a t-value of 0.2299), and similarly average $\hat{\alpha}$ is insignificantly different from zero (0.0052%, t = 0.1686). When all 30 observations are considered and the parameters estimated the standard way, the 30th month's positive stock residual causes average $\hat{\beta}$ to rise to 1.2498, making it significantly greater than 1 (t = 57.7975). Similarly, average $\hat{\alpha}$ falls to -0.0724%, making it significantly less than zero (t = -2.3503). Only in the last column ($\hat{\beta}$ estimated from the first 29 observations, average $\hat{\alpha}$ measured from this $\hat{\beta}$ applied to all 30 observations) do we find an estimate of average $\hat{\alpha}$ that at 0.5042% is close to the known parameter value of 0.5%, which is significantly different from zero and which, unlike the standard method's Jensen's alpha, has the right sign (t = 16.0396).

In the next section, we apply this last technique to a sample with a negative sensitivity factor and actual market prices, and compare it with the standard method.

[6] We also tried several variations of weighted least squares to estimate $\hat{\beta}$ and $\hat{\alpha}$ for the full set of 30 observations, but none were effective in detecting the average monthly abnormal performance of 50 basis points.

IV. An Examination of Jensen's Alpha for stocks during an interval that contains a negative sensitivity factor

The most perverse case and the one on which we focus our attention is when $\frac{(RP_{m,i} - \overline{RP}_m)}{\hat{\sigma}^2_{RP_m}} \cdot \overline{RP}_m > 1$, as this will produce a negative sensitivity factor so that the effect of a positive (negative) perturbation c in month i is to make Jensen's alpha smaller (larger). To see the practical extent of this problem, we first observe that this term for June, 1997 in the 30-month window [January, 1995; June, 1997] is 1.2547. Because this exceeds 1, the sensitivity factor is negative and so the effect of an abnormal shock on any stock in June, 1997 will move Jensen's α in the opposite direction, i.e., a positive shock will decrease Jensen's alpha and a negative shock will increase it. To see the effects of this, we began with the 125 best performing stocks and the 125 worst performing stocks during June, 1997. After deleting firms that did not have observations for all 30 months between January, 1995 and June, 78 firms with the smallest returns and 93 firms with the largest returns remained.[7] *Ceteris paribus*, the idea behind Jensen's alpha as traditionally calculated would lead us to expect the best (worst) performing stocks in June, 1997 to have, on average, better (worse) performance during the entire window, but the analysis presented so far suggests this month's abnormal performance will have the opposite effect on the estimate of Jensen's alpha.

We converted these raw returns to excess returns for each month by subtracting the periodic monthly T-Bill rate, obtained from the Federal Reserve Bank of St. Louis website. First we estimated $\hat{\beta}$ and $\hat{\alpha}$ for the set of the worst 78 June, 1997 performers, using only the first 29 observations and not those of June, 1997. We found an average $\hat{\beta}$ of 0.6360 and an average $\hat{\alpha}$ of -0.424%. We might expect that when we expand this set to include June, 1997, the month in which poor performance is the criterion for inclusion in the sample, we would find a lower Jensen's alpha if it is accurately measuring performance. As per the previous discussion, we found the opposite: the average Jensen's alpha *rose* from -0.424% to only -0.075%, or only about seven and a half basis points below zero. The reason for this is the poor performance in June, 1997 also flattened the regression lines, decreasing the average $\hat{\beta}$ from 0.6360 to -0.2257, and consequently increasing Jensen's alpha.

Because the aberrant $\hat{\alpha}$ in the set with all 30 observations is due to the effect of the 30th observation on $\hat{\beta}$ (decreasing it from 0.6360 to -0.2257), we also estimated $\hat{\beta}$ using only the first 29 observations, but then applied this estimate to the set of all 30 observations to estimate Jensen's alpha. Not surprisingly (because the stocks were selected based on their low returns during June, 1997), this gave us a substantially smaller value for Jensen's alpha, specifically

[7] We note that this is slightly different than the results of the previous section, which simulated shocks and combined them with the CAPM prediction. Here we are taking the most extreme observations, which might conceivably not be a shock at all, depending on the stock's true (unknown) beta.

-1.721%, and the only one of the three estimates that was significantly different from zero (t = -1.967). These and other statistics are summarized in Panel A of Table 5.

Finally, we replicated this procedure for the 93 stocks with the largest raw returns in June, 1997. Here if Jensen's alpha (as traditionally calculated) is accurately conveying abnormal performance, we should expect the estimate based on all thirty observations to be positive because exceptional performance in the 30th month was why each stock was selected. When we used only the first 29 observations, we found an average $\hat{\beta}$ of 1.390 and an average $\hat{\alpha}$ of -1.630%. When we expanded this set to include June, 1997, instead of increasing to reflect every stock's exceptional performance during that month, average Jensen's alpha decreased to -2.185%. The reason for the decrease in $\hat{\alpha}$ is that the June, 1997 outlier return made the regression lines steeper (average $\hat{\beta}$ rose from 1.390 to 2.761), and this made $\hat{\alpha}$ smaller. When we once again estimated $\hat{\beta}$ using the set of only 29 observations, but then applied this estimate of $\hat{\beta}$ to the entire set of 30 observations, average Jensen's alpha rose to 0.5516%. Thus the estimate of Jensen's alpha using all 30 observations (-2.185% per month) increased by 2.737%/month to 0.5516%/month when we follow the outlined procedure. These and other results are summarized in Panel B of Table 5.

V. Extension to the Fama-French Model

The previous analysis of Jensen's alpha can be extended to other models as well. In Appendix B we derive the expression for the sensitivity factor in the Fama-French Three-Factor model, and find it to be similar to equation [2], specifically,

$$\hat{a}(c,i) - \hat{a}(0,i) = \frac{c}{n}[1 - n(\Delta\beta_{RP_m} \cdot \overline{RP}_m - \Delta\beta_{SMB} \cdot \bar{R}_{SMB} - \Delta\beta_{HML} \cdot \bar{R}_{HML})]. \quad [4]$$

A priori, it is unclear whether we would expect negative sensitivity factors to become more frequent or less frequent when we add two factors to get the Fama-French Model. On the one hand, the presence of three variables might suggest there are two extra chances for a factor return to have a value sufficiently extreme as to cause a change in the sensitivity factor to be similarly extreme. On the other hand, perhaps the additional two variables might create a damping (or diversifying) effect so that even one extreme factor return is insufficient, in which case we might expect a lower frequency of negative sensitivity factors. In addition, negative sensitivity factors are related to the distance between the sample mean and zero, so R_{HML} (overall mean = 0.399%) and R_{SMB} (overall mean = 0.211%) are less likely to cause negative sensitivity factors than is R_M - R_F (overall mean = 0.654%). Nevertheless, the frequency of negative sensitivity factors is primarily an empirical question, and in fact we find them to be more common for the Fama-French model than for Jensen's alpha. Specifically, we find that for the Fama-French model, 0.9259% of months have a negative sensitivity factor, and eight out of our 36 thirty-month windows (or 22.22%) contain at least one such month. In comparison, the same numbers for Jensen's alpha were a little less than half as frequent: three of 720 months, or only 0.4167%, and two of 24 windows, or 8.33%.

For an example in a relatively recent window, we consider the period from January, 1, 1992 to June 30, 1994. This window features a month with a negative sensitivity factor, specifically, -0.1637 in December, 1992. We considered returns on Apple stock during this same period. As it was, Apple's actual return that month was -0.42%, and using the FF model, its intercept during the entire 30-month window was 0.87%. Suppose that Apple had earned a return that was 30.00% higher that month, increasing that month's return from -0.42% to +29.58%. This causes the average monthly return during the 30-month window to rise by 1.00%, and yet because the shock came in a month with a negative sensitivity factor, the intercept doesn't increase, but instead *falls* by 0.64% to 0.23%. Note that although the sensitivity factor given by equation [4] is a function of all three independent variables, the problem here doesn't seem to be that December, 1992's value of $R_m - R_f$ (0.93%) is much higher than its sample mean (0.38%), but rather that December's values of R_{SMB} (2.05%) and especially R_{HML} (5.89%) are much higher than their sample mean values of 0.02% and 1.29%. Using only the twenty-nine observations with positive sensitivity factors to estimate betas for all three factors, and then applying these betas to the entire set of thirty observations produces a much more reasonable intercept of 2.61%.

VI. Conclusions

Ceteris paribus, we would ordinarily expect an unsystematic shock to share price to cause the intercept to move in the same direction as the shock. After all, this is the basis for interpreting an intercept as a reasonable measure of excess performance. We find, however, that factor returns sufficiently extreme as to cause the intercept to move in the direction *opposite* that of the shock are not uncommon. Whether an observation does or doesn't have this property is independent of the stock's shock, and instead depends primarily on the characteristics of the factor return relative to its sample mean. For each month's observations, we define a "sensitivity factor" that quantifies the effect of unsystematic perturbations on their respective values for $\hat{\beta}$ and, consequently, on $\hat{\alpha}$, and find that while most sensitivity factors are positive, negative ones will cause this aberrant relationship between the stock's shock and the intercept. This aberrant effect is in turn due to the fact that a positive shock on a day when the factor return is much higher than its average (and its average is much greater than zero) will tilt the slope estimate upwards (and the intercept estimate downwards), and similarly for negative shocks or when the market return is much smaller than average.

We derive expressions that will tell us when the factor returns are sufficiently extreme as to have this property, and find that while it is not the norm, it did occur in 8.33% of 30-month windows and in 10% of 36-month windows between 1955 and 2015 for Jensen's alpha, and in 22.22% of the 30-month windows between 1927 and 2016 for the Fama-French Three-Factor Model. Moreover, we conduct simulations to estimate the effect of a single unsystematic shock in a month with a negative sensitivity factor and indeed find that an average Jensen's alpha that is insignificantly different from zero absent a shock becomes significantly *negative* when a *positive* shock is introduced (and conversely, significantly *positive* in the presence of a *negative* shock). This calls into question the widespread use of intercepts as performance measures. At a minimum, researchers should check for negative sensitivity factors and consider deleting those observations when estimating the slopes (but then including them when estimating the intercept); simulations suggest this method works well. Other than that, the only plausible alternative appears to be to measure performance without relying on an intercept from a factor model. The Buy-and-Hold-Abnormal-Return (BHAR) approach used by Ikenberry et al (1995), Lyon et al (1999), Mitchell and Stafford (2000), and others is such an alternative.

Appendix A—Derivation of Sensitivity Factor for Jensen's Alpha

The Capital Asset Pricing Model concludes that $E(R_s) = R_f + \beta[E(R_m) - R_f]$, or equivalently that $E(R_s) - R_f = \beta[E(R_m) - R_f]$. If we define RP to be the risk premium[8] (the amount by which the asset's returns exceed the riskless rate), this becomes

$$RP_s = \beta[RP_m]. \qquad [A1]$$

Given a set of data $\{(RP_{m,i}, RP_{s,i})\}_{i=1}^{n}$, we can estimate the parameters $\hat{\alpha}$ and $\hat{\beta}$ in

$$RP_{s,i} = \hat{\alpha} + \hat{\beta} RP_{m,i}. \qquad [A2]$$

Given equation [A1], we would expect the intercept $\hat{\alpha}$ to be zero when we estimate parameters of the regression. $\hat{\alpha}$ is Jensen's alpha; when positive, it suggests that the stock did better than the CAPM predicted, and conversely when it is negative. The problem is that deviations from the line $RP_s = \beta[RP_m]$ affect not only the estimate $\hat{\alpha}$, but also the estimate of slope, $\hat{\beta}$. Under certain conditions an unsystematic shock to stock returns can cause a change in $\hat{\beta}$ that overwhelms any effect on $\hat{\alpha}$. In other words, a single negative (positive) shock can make a stock seem to perform better (worse) than if the shock didn't occur. This appendix analyzes the conditions under which this can happen.

If we define $\hat{\sigma}^2_{RP_m}$ to be the maximum-likelihood estimator sample variance of the market risk premia (i.e., unadjusted for degrees of freedom), then the ordinary least squares estimate $\hat{\beta}$ is

$$\hat{\beta} = \frac{\sum(RP_{s,i} - \overline{RP}_s)(RP_{m,i} - \overline{RP}_m)}{\sum(RP_{m,i} - \overline{RP}_m)^2} = \frac{\sum(RP_{s,i} RP_{m,i}) - n\overline{RP}_s \overline{RP}_m}{n\hat{\sigma}^2_{RP_m}}. \qquad [A3]$$

Because a regression line passes through the sample means, the estimate of $\hat{\alpha}$ is

$$\hat{\alpha} = \overline{RP}_s - \hat{\beta} \cdot \overline{RP}_m. \qquad [A4]$$

Next we examine a shock c to the i^{th} observation of $RP_{s,i}$, and express the new parameters $\hat{a}(c,i)$ and $\hat{\beta}(c,i)$ as functions of c and i.

[8] Following Jensen, we use the term "risk premium" to refer to either the expected risk premium, $E(R) - R_f$, which must be positive provided beta is positive, or the observed risk premium, $R - R_f$, which can be negative, and indeed in a sufficiently large sample must occasionally be negative to satisfy the no-arbitrage condition. Because Jensen's alpha is a test of observed outcomes, all our references to "risk premium" refer to the observed risk premium."

Consider the second expression for $\hat{\beta}$ in [A3]. After the shock, the only summand in the first term of the numerator that will change is the i[th] term, which will change from $RP_{s,i}RP_{m,i}$ to $(RP_{s,i} + c)RP_{m,i}$. The new average risk premium on the stock will be the old risk premium, $\overline{RP_s}$, plus $\frac{c}{n}$. Thus the net change in the numerator will be $c(RP_{m,n} - \overline{RP_m})$. The denominator will be the same in both cases, so we have

$$\hat{\beta}(c,i) = \hat{\beta}(0,i) + \frac{c(RP_{m,i} - \overline{RP}_m)}{n\hat{\sigma}^2_{RP_m}}. \qquad \text{[A5]}$$

Whether the perturbation causes the slope to increase or decrease depends on the numerator of the second term on the right-hand side of [A5]. If c and $RP_{m,i}$ – $\overline{RP_m}$ have the same sign, then the slope will be steeper; if they have different signs, then the slope will be shallower.

We next consider $\hat{a}(c,i)$. As shown earlier,

$$\hat{a}(0,i) = \overline{RP_s} - \hat{\beta}(0,i) \cdot \overline{RP_m}. \qquad \text{[A6]}$$

Because the shock c causes the average risk premium on the stock to change by $\frac{c}{n}$, we have

$$\hat{a}(c,i) = [\,\overline{RP_s} + \frac{c}{n}] - \hat{\beta}(c,i) \cdot \overline{RP_m}. \qquad \text{[A7]}$$

Making use of [A5] and taking the difference between [A7] and [A6] gives us

$$\hat{a}(c,i) - \hat{a}(0,i) = \frac{c}{n} - \frac{c(RP_{m,i} - \overline{RP}_m)}{n\hat{\sigma}^2_{RP_m}} \cdot \overline{RP_m}, \text{ or} \qquad \text{[A8]}$$

$$\hat{a}(c,i) - \hat{a}(0,i) = \frac{c}{n}[1 - \frac{(RP_{m,i} - \overline{RP}_m)}{\hat{\sigma}^2_{RP_m}} \cdot \overline{RP_m}]. \qquad \text{[A9]}$$

The first term of the right-hand side of [A8], $\frac{c}{n}$, reflects the actual average change in performance due to the shock c. The second term, $-\frac{c(RP_{m,i} - \overline{RP}_m)}{n\hat{\sigma}^2_{RP_m}} \cdot \overline{RP_m}$, is a distortion due to the fact that the estimate of β changes. Because $\hat{a}(c,i)$ - $\hat{a}(0,i)$ will have the same sign as c if and only if the term in brackets on the right hand side, $1 - \frac{(RP_{m,i} - \overline{RP}_m)}{\hat{\sigma}^2_{RP_m}} \cdot \overline{RP_m}$, is positive, the most extreme distortions will occur if this term is negative, as then a positive (negative) shock c will produce a decrease (increase) in Jensen's $\hat{a}$.

Appendix B—Derivation of Sensitivity Factor for Fama-French Intercepts

In a multifactor model $Y = \hat{\alpha} + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \varepsilon$, OLS estimates are found by solving the three simultaneous equations

$$m_{11}\hat{\beta}_1 + m_{12}\hat{\beta}_2 + m_{13}\hat{\beta}_3 = m_{Y1},$$

$$m_{12}\hat{\beta}_1 + m_{22}\hat{\beta}_2 + m_{23}\hat{\beta}_3 = m_{Y2},$$

$$\text{and } m_{13}\hat{\beta}_1 + m_{23}\hat{\beta}_2 + m_{33}\hat{\beta}_3 = m_{Y3},$$

where

$$m_{ij} = \sum_{k=1}^{n} (X_{i,k} - \bar{X}_i)(X_{j,k} - \bar{X}_j)$$

and

$$m_{Yj} = \sum_{k=1}^{n} (Y_k - \bar{Y})(X_{j,k} - \bar{X}_j)$$

for i, j = 1, 2, 3, and n the number of observations in the sample. An application of Cramer's Rule gives the expression for $\hat{\beta}_1$, for example, to be

$$\hat{\beta}_1 = \frac{\begin{vmatrix} m_{Y1} & m_{21} & m_{31} \\ m_{Y2} & m_{22} & m_{32} \\ m_{Y3} & m_{23} & m_{33} \end{vmatrix}}{\begin{vmatrix} m_{11} & m_{21} & m_{31} \\ m_{12} & m_{22} & m_{32} \\ m_{13} & m_{23} & m_{33} \end{vmatrix}}$$

The expressions $\hat{\beta}_2$ and $\hat{\beta}_3$ are analogous. As always, $\hat{\alpha} = \bar{Y} - \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 + \hat{\beta}_3 \bar{X}_3$

As in Appendix A, suppose there is a shock to Y_n so that it is increased by c, the new $\hat{\beta}_1$ will equal

$$\hat{\beta}_1(c) = \frac{\begin{vmatrix} m_{Y1} + c(X_{1,n} - \bar{X}_1) & m_{21} & m_{31} \\ m_{Y2} + c(X_{2,n} - \bar{X}_2) & m_{22} & m_{32} \\ m_{Y3} + c(X_{3,n} - \bar{X}_3) & m_{23} & m_{33} \end{vmatrix}}{\begin{vmatrix} m_{11} & m_{21} & m_{31} \\ m_{12} & m_{22} & m_{32} \\ m_{13} & m_{23} & m_{33} \end{vmatrix}}$$

Thus the change in $\hat{\beta}_1$ due to the shock c will be $\Delta\hat{\beta}_1$ = $\hat{\beta}_1$(c) $- \hat{\beta}_1$ =

$$\frac{\begin{vmatrix} c(X_{1,n} - \bar{X}_1) & m_{21} & m_{31} \\ c(X_{2,n} - \bar{X}_2) & m_{22} & m_{32} \\ c(X_{3,n} - \bar{X}_3) & m_{23} & m_{33} \end{vmatrix}}{\begin{vmatrix} m_{11} & m_{21} & m_{31} \\ m_{12} & m_{22} & m_{32} \\ m_{13} & m_{23} & m_{33} \end{vmatrix}}$$

With similar expressions for $\Delta\hat{\beta}_2$ and $\Delta\hat{\beta}_3$. Because the new average value of Y will be the old average value plus $\frac{c}{n}$ and because the expression is linear in the shock c, the change in $\hat{\alpha}$ = $\hat{\alpha}(c, i)$ - $\hat{\alpha}(0, i)$ = $\frac{c}{n}[1 - n(\Delta\beta_{RP_m} \cdot \overline{RP}_m - \Delta\beta_{SMB} \cdot \bar{R}_{SMB} - \Delta\beta_{HML} \cdot \bar{R}_{HML})]$.

References

Blume, M., 1975, Betas and Their Regression Tendencies, *Journal of Finance*, 30(3): 785-795.

Brown, S., and J. Warner, 1980, Measuring Security Price Performance, *Journal of Financial Economics*, 8(3): 205-258.

Cermers, M., A. Petajisto, and E. Zitzewitz, 2012, Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation, *Critical Finance Review*, 2: 1—48.

DeBondt, W.F.M., and R. Thaler, 1987, Further Evidence on Investor Overreaction and Stock Market Seasonality, *Journal of Finance* 42(3): 557-581.

Fama, E., 1965, The Behavior of Stock-Market Prices, *Journal of Business* 38(1): 34-105.

Fama, E., L. Fisher, M. Jensen, and R. Roll, 1969, The adjustment of stock prices to new information, *International Economic Review*, 10(1): 1-21.

Harris, L., 1986, Cross-Sectional Tests of the Mixture of Distributions Hypothesis, *Journal of Financial and Quantitative Analysis* 21(1): 39-46.

Ikenberry, D., J. Lakonishok, and T. Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics*, 39: 181-208.

Jensen, M., 1968, The Performance of Mutual Funds in the Period 1945-1964, *Journal of Finance* 23(2): 389-416.

Kon, S., 1984, Models of Stock Return—A Comparison, *Journal of Finance* 39(1): 147-165.

Lyon, D., J. Barber, and C. Tsai, 1999, Improved methods for tests of long-run abnormal stock returns, *Journal of Finance* 54(1), 165-201.

Mitchell, M., and E. Stafford, 2000, Managerial Decisions and Long-Term Stock Price Performance, *Journal of Business* 73(3): 287-329.

Table 1—Distribution of Sensitivity Factor SF_i for Jensen's Alpha

Panel A: 30-month intervals

SUBSET DATE RANGE [BEGINNING DATE, ENDING DATE]	$\overline{RP_M}$	$\hat{\sigma}_{RM}$	$RP_{M,1}$ (MIN)	$RP_{M,30}$ (MAX)	SF_1 (MIN)	SF_{30} (MAX)	STANDARD DEVIATION OF SF_I
[2/1/1955, 7/1/1957]	0.007672	0.038488	-0.06791	0.081125	0.606464	1.404955	0.206203
[8/1/1957, 1/1/1960]	0.003247	0.034627	-0.0751	0.049695	0.869889	1.21946	0.096998
[2/1/1960, 7/1/1962]	0.00013	0.039331	-0.08823	0.061281	0.994686	1.007678	0.003418
[8/1/1962, 1/1/1965]	0.011334	0.028173	-0.05052	0.099207	-0.29798	1.913698	0.416152
[2/1/1965, 7/1/1967]	-0.00043	0.034095	-0.08188	0.074219	0.968754	1.028639	0.01308
[8/1/1967, 1/1/1970]	-0.00778	0.037772	-0.08306	0.076021	0.575564	1.47246	0.212963
[2/1/1970, 7/1/1972]	0.004562	0.041124	-0.09584	0.082813	0.781633	1.280187	0.114762
[8/1/1972, 1/1/1975]	-0.01517	0.058839	-0.12618	0.156964	0.496909	1.780014	0.266632
[2/1/1975, 7/1/1977]	0.004766	0.039664	-0.07273	0.114206	0.657025	1.242864	0.124304
[8/1/1977, 1/1/1980]	-0.00128	0.040663	-0.09818	0.080175	0.922657	1.065007	0.032454
[2/1/1980, 7/1/1982]	-0.01177	0.044628	-0.1145	0.090944	0.3719	1.627994	0.272854
[8/1/1982, 1/1/1985]	0.010839	0.042486	-0.06756	0.108711	0.392027	1.486989	0.26392
[2/1/1985, 7/1/1987]	0.014841	0.043778	-0.08978	0.127217	0.099793	1.838088	0.350694
[8/1/1987, 1/1/1990]	-0.00315	0.057572	-0.22281	0.081804	0.783864	1.083593	0.056648
[2/1/1990, 7/1/1992]	0.004592	0.043467	-0.10052	0.108188	0.739535	1.264283	0.109285
[8/1/1992, 1/1/1995]	0.000723	0.023366	-0.04866	0.033882	0.954579	1.067648	0.032006
[2/1/1995, 7/1/1997]	0.019958	0.029805	-0.05003	0.073946	-0.25472	2.62661	0.692701
[8/1/1997, 1/1/2000]	0.009978	0.050242	-0.14989	0.076953	0.726135	1.653704	0.205442
[2/1/2000, 7/1/2002]	-0.01608	0.050671	-0.09638	0.091978	0.479757	1.700053	0.328274
[8/1/2002, 1/1/2005]	0.008399	0.040158	-0.11139	0.085115	0.586672	1.645396	0.21636
[2/1/2005, 7/1/2007]	0.00369	0.021066	-0.036	0.039232	0.694294	1.341363	0.181191
[8/1/2007, 1/1/2010]	-0.00919	0.06193	-0.17	0.093783	0.601527	1.255146	0.153453
[2/1/2010, 7/1/2012]	0.009386	0.047458	-0.08211	0.107706	0.576123	1.394454	0.2046
[8/1/2012, 1/1/2015]	0.012618	0.02437	-0.03562	0.050361	0.170457	2.060321	0.535631
COLUMN MAXIMUM	0.019958	0.06193	-0.03562	0.156964	0.994686	2.62661	0.692701
COLUMN MINIMUM	-0.01608	0.021066	-0.22281	0.033882	-0.29798	1.007678	0.003418
COLUMN AVERAGE	0.002579	0.040574	-0.09229	0.08523	0.562398	1.477525	0.212084

This table displays parameters of 30-month non-overlapping windows starting in February, 1955 and ending in January, 2015. The first four columns report the mean, standard deviation, maximum, and minimum of market risk premia within the windows. The next three columns display the minimum, maximum, and standard deviation of sensitivity factors (SF) within the window. As discussed in the text, any unsystematic shock that occurs during a month with a negative sensitivity factor will cause Jensen's alpha to move in the direction opposite that of the shock.

Panel B: 36-month intervals

SUBSET DATE RANGE [BEGINNING DATE, ENDING DATE]	$\overline{RP_M}$	$\hat{\sigma}_{RM}$	$RP_{M,1}$ (MIN)	$RP_{M,30}$ (MAX)	SF_1 (MIN)	SF_{36} (MAX)	STANDARD DEVIATION OF SF_I
[2/1/1955, 1/1/1958]	0.00226	0.03088	-0.06791	0.08113	0.88802	1.09963	0.05741
[2/1/1958, 1/1/1961]	0.00921	0.039365	-0.07510	0.06128	0.48286	1.83720	0.30666
[2/1/1961, 1/1/1964]	0.00458	0.02858	-0.08823	0.09921	0.71241	1.28206	0.11963
[2/1/1964, 1/1/1967]	0.00020	0.037923	-0.08188	0.07422	0.98180	1.02018	0.00703
[2/1/1967, 1/1/1970]	-0.00442	0.038537	-0.08306	0.07602	0.75117	1.25457	0.12001
[2/1/1970, 1/1/1973]	0.00534	0.062018	-0.09584	0.08281	0.71324	1.37455	0.14265
[2/1/1973, 1/1/1976]	-0.00778	0.034928	-0.12618	0.15696	0.75363	1.34279	0.12904
[2/1/1976, 1/1/1979]	-0.00460	0.044248	-0.09818	0.08017	0.63701	1.32884	0.13548
[2/1/1979, 1/1/1982]	-0.00382	0.043465	-0.11450	0.09094	0.77778	1.19026	0.08884
[2/1/1982, 1/1/1985]	0.00415	0.060016	-0.07183	0.10871	0.76351	1.17187	0.09831
[2/1/1985, 1/1/1988]	0.00657	0.040706	-0.22281	0.12722	0.77373	1.43018	0.11256
[2/1/1988, 1/1/1991]	0.00273	0.030423	-0.10052	0.08554	0.85961	1.17505	0.06901
[2/1/1991, 1/1/1994]	0.00662	0.028811	-0.05254	0.10819	0.25231	1.43558	0.22397
[2/1/1994, 1/1/1997]	0.00997	0.04928	-0.05003	0.06918	0.26832	1.74146	0.35602
[2/1/1997, 1/1/2000]	0.01322	0.053857	-0.14989	0.07695	0.64314	1.91332	0.27594
[2/1/2000, 1/1/2003]	-0.01497	0.025764	-0.11139	0.09198	0.48822	1.56763	0.28585
[2/1/2003, 1/1/2006]	0.00996	0.044922	-0.03540	0.08010	-0.08256	1.70005	0.39763
[2/1/2006, 1/1/2009]	-0.01384	0.052535	-0.17000	0.04648	-0.10125	1.42533	0.31678
[2/1/2009, 1/1/2012]	0.01420	0.026802	-0.11018	0.10771	0.50521	1.65814	0.27799
[2/1/2012, 1/1/2015]	0.01200	0.03088	-0.06273	0.05036	0.34082	2.28407	0.46055
COLUMN MAXIMUM	0.0142	0.062018	-0.0354	0.15696	0.9818	2.28407	0.46055
COLUMN MINIMUM	-0.01497	0.025764	-0.22281	0.04648	-0.10125	1.02018	0.00703
COLUMN AVERAGE	0.002579	0.040675	-0.09841	0.087758	0.570449	1.461638	0.199068

This table displays parameters of 36-month non-overlapping windows starting in February, 1955 and ending in January, 2015. The first four columns report the mean, standard deviation, maximum, and minimum of market risk premia within the windows (all expressed in decimals, not percent). The next three columns display the minimum, maximum, and standard deviation of sensitivity factors (SF) within the window. As discussed in the text, any unsystematic shock that occurs during a month with a negative sensitivity factor will cause Jensen's alpha to move in the direction opposite that of the shock.

Table 2—Simulation Results

Full Sample with 2000 Firms (1000 with Stock Residual = -3σ = -15% in Last Month and Stock Residual = 3σ = 15% in Last Month)

	Average Market Risk Premium	Average Maximum Market Risk Premium	Average Sensitivity Factor for Last Observation	Average $\hat{\beta}$ first 29 observations	Average $\hat{\alpha}$ first 29 observations	Average $\hat{\beta}$ all 30 observations	Average $\hat{\alpha}$ all 30 observations	Average $\hat{\alpha}$ when only the first 29 observations are used to estimate $\hat{\beta}$, but then all 30 used to estimate $\hat{\alpha}$
Grand mean	2.388%	24.553%	-0.0814	1.0004	0.0051%	1.0007	0.0046%	0.0054%
Standard deviation of mean across 1000 simulations	0.461%	20.388%	0.4475	0.1940%	0.9713%	0.2839%	0.9737%	1.1124%
t-stat [H_0]				0.0889 [H_0: $\beta = 1$]	0.2346 [H_0: $\alpha = 0$]	0.1136 [H_0: $\beta = 1$]	0.2129 [H_0: $\alpha = 0$]	0.2183 [H_0: $\alpha = 0$]

Simulations with all 2000 observations, 1000 with last residual =-3σ and 1000 with last residual =3σ. The first three columns provide descriptive statistics for the simulations. The next two columns describe estimates of $\hat{\beta}$ and $\hat{\alpha}$ from using only the first 29 observations, and the two columns after that do the same when the 30th month is included. In no column is $\hat{\beta}$ significantly different from one nor $\hat{\alpha}$ from zero, suggesting that the estimates are unconditionally unbiased. However, as Tables 3 and 4 will show, they are conditionally biased.

Table 3—Simulation Results

Subsample with Stock Residual = -3σ = -15% in Last Month

	Average Market Risk Premium	Average Maximum Market Risk Premium	Average Sensitivity Factor	Average $\hat{\beta}$ first 29 observations	Average $\hat{\alpha}$ first 29 observations	Average $\hat{\beta}$ all 30 observations	Average $\hat{\alpha}$ all 30 observations	Average $\hat{\alpha}$ when only the first 29 observations are used to estimate $\hat{\beta}$, but then all 30 used to estimate $\hat{\alpha}$
Grand mean	2.386%	9.751%	-0.0780	0.9994	.0050%	0.7516	0.0817%	-0.4943%
Standard deviation of mean across 1000 simulations	0.453%	13.965%	0.4526	0.1949	0.9695%	0.1359	0.9669%	0.9946%
t-stat [H_0]				-0.1026 [H_0: $\beta = 1$]	0.1632 [H_0: $\alpha = 0$]	-57.7691 [H_0: $\beta = 1$]	2.6703 [H_0: $\alpha = 0$]	-15.676 [H_0: $\alpha = 0$]

The first three columns provide descriptive statistics for the simulations. The next two columns describe estimates of $\hat{\beta}$ and $\hat{\alpha}$ from using only the first 29 observations, and the two columns after that do the same when the 30th (extremely negative performance) month is included. Because the 30th observation decreases the slope, it increases the intercept to find a significant *positive* reaction when the 30th (exceptional negative performance month) is included. The last column provides statistics when $\hat{\beta}$ is estimated using only the first 29 observations, while $\hat{\alpha}$ is estimated by applying the 29-observation estimate $\hat{\beta}$ to all thirty observations. The penultimate column makes clear that Jensen's alpha does not work well if a stock has exceptional performance in a month with a negative sensitivity factor, while the suggested adjustment in the last column does work well (by construction, expected average abnormal return = -15%/30 = -0.5%).

Table 4— Simulation Results

Subsample with Stock Residual = +3σ = 15% in Last Month

	Average Market Risk Premium	Average Maximum Market Risk Premium	Average Sensitivity Factor	Average $\hat{\beta}$ first 29 observations	Average $\hat{\alpha}$ first 29 observations	Average $\hat{\beta}$ all 30 observations	Average $\hat{\alpha}$ all 30 observations	Average $\hat{\alpha}$ when only the first 29 observations are used to estimate $\hat{\beta}$, but then all 30 used to estimate $\hat{\alpha}$
Grand mean	2.389%	39.295%	-0.0849	1.0014	0.0052%	1.2498	-0.0724%	0.5042%
Standard deviation of mean across 1000 simulations	0.469%	14.124%	0.4421	0.1931	0.9726%	0.1366	0.9738%	0.9935%
t-stat [H_0]				0.2299 [H_0: β = 1]	0.1686 [H_0: α = 0]	57.7975 [H_0: β = 1]	-2.3503 [H_0: α = 0]	16.0396 [H_0: α = 0]

The first three columns provide descriptive statistics for the simulations. The next two columns describe estimates of $\hat{\beta}$ and $\hat{\alpha}$ from using only the first 29 observations, and the two columns after that do the same when the 30th (extremely positive performance) month is included. Because the 30th observation increases the slope, it decreases the intercept to find a significant *negative* reaction when the 30th (exceptional positive performance month) is included. The last column provides statistics when $\hat{\beta}$ is estimated using only the first 29 observations, while $\hat{\alpha}$ is estimated by applying the 29-observation estimate $\hat{\beta}$ to all thirty observations. The penultimate column makes clear that Jensen's alpha does not work well if a stock has exceptional performance in a month with a negative sensitivity factor, while the suggested adjustment in the last column does work well (by construction, expected average abnormal return = 15%/30 = 0.5%).

Table 5—Examination of Stocks with the Smallest and Largest Returns During June, 1997, a Month with a Negative Sensitivity Factor

Panel A: 78 Stocks with the Lowest Returns

	Median	Mean	T-statistic (H$_0$: β = 1 or H$_0$: α = 0, as appropriate)
$\hat{\beta}$ only 29 observations	0.4361	0.6360	-1.270
$\hat{\alpha}$ only 29 observations	-0.264%	-0.424%	-0.483
$\hat{\beta}$ all 30 observations	-0.3450	-0.2257	-4.777
$\hat{\alpha}$ all 30 observations	-0.125%	-.075%	-0.087
$\hat{\alpha}$ all 30 observations when $\hat{\beta}$ using only 29 observations	-2.322%	-1.721%	-1.967

Panel B: 93 Stocks with the Largest Returns

	Median	Mean	T-statistic (H$_0$: β = 1 or H$_0$: α = 0, as appropriate)
$\hat{\beta}$ only 29 observations	1.359	1.390	2.2883
$\hat{\alpha}$ only 29 observations	-2.307%	-1.630%	-3.0272
$\hat{\beta}$ all 30 observations	2.601	2.761	8.2150
$\hat{\alpha}$ all 30 observations	-2.809	-2.185%	-3.9310
$\hat{\alpha}$ all 30 observations when $\hat{\beta}$ using only 29 observations	0.102%	0.552%	1.0343

In a fashion similar to the simulations, the worst (best) performers appear to have performed better (worse) when the month of their extreme performance is included. The suggested adjustment to Jensen's alpha does not suffer this problem.

Table 6—Distribution of Sensitivity Factor SF_i for the Fama-French Model
30-month intervals

SUBSET DATE RANGE [BEGINNING DATE, ENDING DATE]	$\overline{R_M - R_F}$	$\overline{SMB}$	$\overline{HML}$	**MIN SF_I**	**MAX SF_I**
[1/1/1927, 6/30/1929]	0.023287	-0.00622	-0.00337	-0.10269	1.61838
[7/1/1929, 12/31/1931]	-0.03723	-0.00431	-0.00444	0.180059	1.619833
[1/1/1932, 6/30/1934]	0.019707	0.018733	0.017913	0.107722	1.488343
[7/1/1934, 12/31/1936]	0.027813	0.012393	0.01049	0.074815	2.475646
[1/1/1937, 6/30/1939]	-0.00446	-0.00371	-0.00934	0.022362	1.726773
[7/1/1939, 12/31/1941]	-0.00149	0.006253	0.009593	0.296228	4.463833
[1/1/1942, 6/30/1944]	0.01769	0.010167	0.01433	0.10809	9.300841
[7/1/1944, 12/31/1946]	0.012347	0.007617	0.00565	0.32839	5.839063
[1/1/1947, 6/30/1949]	0.00234	-0.00588	0.003907	0.262817	1.571122
[7/1/1949, 12/31/1951]	0.02042	-0.0001	0.00614	-0.06247	2.111119
[1/1/1952, 6/30/1954]	0.011127	-0.00284	0.000363	0.394562	1.911453
[7/1/1954, 12/31/1956]	0.013933	-0.00128	0.0054	-0.3995	1.746838
[1/1/1957, 6/30/1959]	0.01244	0.00243	0.00154	0.417782	1.96373
[7/1/1959, 12/31/1961]	0.005673	0.00038	0.000277	0.526314	1.332521
[1/1/1962, 6/30/1964]	0.006227	-0.00563	0.009507	0.275094	2.292669
[7/1/1964, 12/31/1966]	0.002307	0.01035	0.001847	0.340451	2.017358
[1/1/1967, 6/30/1969]	0.001857	0.011263	0.003743	0.032917	1.719114
[7/1/1969, 12/31/1971]	0.00348	0.000623	1.33E-05	0.874278	1.152201
[1/1/1972, 6/30/1974]	-0.01377	-0.01176	0.012157	0.030659	2.23991
[7/1/1974, 12/31/1976]	0.010603	0.006397	0.009823	-0.06603	2.828564
[1/1/1977, 6/30/1979]	0.00251	0.01479	0.003657	0.216626	2.698129
[7/1/1979, 12/31/1981]	0.001427	0.005767	-0.0011	0.470518	1.556957
[1/1/1982, 6/30/1984]	0.00442	0.00462	0.01084	-0.35425	1.806594
[7/1/1984, 12/31/1986]	0.017873	-0.00424	0.004243	0.007336	3.527036
[1/1/1987, 6/30/1989]	0.006113	-0.00202	0.003513	0.222677	1.681859
[7/1/1989, 12/31/1991]	0.00408	-0.0002	-0.00687	0.326189	1.916983
[1/1/1992, 6/30/1994]	0.00375	0.000203	0.012903	-0.16373	1.694864
[7/1/1994, 12/31/1996]	0.015367	-0.00303	-0.00014	0.050813	2.050985
[1/1/1997, 6/30/1999]	0.014737	-0.00715	0.00039	0.344328	1.866878
[7/1/1999, 12/31/2001]	-0.0062	0.010087	0.01096	-0.00072	1.750597
[1/1/2002, 6/30/2004]	0.00168	0.008	0.00511	0.357655	2.058335
[7/1/2004, 12/31/2006]	0.008563	0.001783	0.007073	-0.03813	1.852731
[1/1/2007, 6/30/2009]	-0.00997	0.00201	-0.00628	0.39678	1.460182
[7/1/2009, 12/31/2011]	0.012587	0.00364	-0.00285	0.335633	1.58255
[1/1/2012, 6/30/2014]	0.015237	-0.00173	0.004483	0.405374	2.260659
[7/1/2014, 12/31/2016]	0.00774	0.00114	0.000933	0.470761	1.451022
COLUMN MINIMUM	-0.03723	-0.01176	-0.00934	-0.3995	1.152201
COLUMN AVERAGE	0.006506	0.002182	0.003956	0.185826	2.295435
COLUMN MAXIMUM	0.027813	0.018733	0.017913	0.874278	9.300841

This table displays parameters of 30-month non-overlapping windows starting January 1, 1927 and ending in December 31, 2016. The first three columns report the mean values of $R_m - R_f$, SMB, and HML (all expressed as decimals, not percents) within the 30-month window. The next two columns display the minimum and maximum sensitivity factors (SF) within the window. As discussed in the text, any unsystematic shock that occurs during a month with a negative sensitivity factor will cause the intercept to move in the direction opposite that of the shock.