Regression Analysis and Strengthening the SEC's Efforts to Regulate Mutual Funds

by Efstathia Bura and Joseph L. Gastwirth*

Portions of this article are based on our previous article that appeared in Law, Probability and Risk in 2009.

Early in 2004 the Securities and Exchange Commission (SEC) proposed new conditions that Mutual Funds would need to satisfy in order to engage in otherwise prohibited transactions, i.e. those that the fund's advisor would also benefit.[1] In particular, the fund would be required to have a Chairman and at least 75% of its Directors independent of its advisors and sponsors.[2] The Commission adopted this and other new conditions in light of recent abuses in the mutual fund industry and felt that more independent directors would provide greater oversight of activities that involve inherent conflicts of interest between the funds and their managers in order to protect shareholders. The industry challenged the new proposal. In *Chamber of Commerce of the United States of America v. Securities and Exchange Commission* [3] the Court of Appeals of the D.C. Circuit held that while the Commission had the legal authority to establish additional conditions on the funds, its deliberations did not fully comply with the Administrative Procedures Act. In particular, the Commission had not considered the potential costs of more independent directors versus those of a proposed alternative involving fuller disclosure by investment companies. Within days after this decision was issued, the SEC published a Response Release[4] stating that it was unnecessary to reopen the rulemaking record for additional comments as the information in the existing record together with other publically available information was sufficient to remedy the deficiencies noted by the court. The Chamber of Commerce again challenged this procedure, in part because when an agency determines that additional fact gathering is needed, public notice and comments are typically required. In *Chamber of Commerce of the United States of America v. Securities and Exchange Commission II* [5] the court vacated the rule and gave the SEC time to reopen the record. The decision noted that the SEC estimated the cost of requiring an independent chair on the increased compensation for the chair and the cost of extra staff using information obtained from "extra record" summaries of two surveys, one of which was not even available during the previous comment period.[6]

In the first case, the Chamber referred to a study[7] supported by Fidelity Investments that had not found a statistically significant difference

* Efstathia Bura is Associate Professor and Joseph L. Gastwirth is Professor in the Department of Statistics at the George Washington University, Washington, DC 20052. They can be reached electronically at ebura@gwu.edu and jlgast@gwu.edu, respectively.

in the performance of funds with a high proportion of independent directors compared with other mutual funds. The Commission discounted that study because it had not

included other potential predictors.[8] The study encompassed all retail-oriented fund complexes with $10 billion or more in long-term assets, which account for 83% of industry long-term fund assets. The data used in the study indicated that the distributions of assets managed by the two types of funds are different. For example, the average assets of the management chaired funds were 74,180 ($ millions) with a standard deviation of 126,519 ($ millions), while the average assets of for the independent chaired funds was 38,490 ($ millions) with a standard deviation of 35,091.89 ($ millions).[9] Apparently, the SEC did not suggest that due to economies of scale fund groups managing more assets should have reduced expenses per share and consequently better performance. This effect should be accounted for in a study of differential returns of the two types of funds.

SEC Power Study

In anticipation of the further proceedings mandated by the second appellate decision, the SEC staff carried out a literature review[10] and a power study which will be referred to as the OEA/SEC memorandum [11]. The purpose of this power study was to demonstrate that statistical tests based on regression models that include relevant predictors would have a low probability of detecting a meaningful impact of the independent directors and chairs on the profitability of funds. The SEC staff examined the performance of independent versus management chaired mutual funds by considering the asset-weighted average annual returns for every family in the CRSP mutual fund database[12]. From the data, the standard deviation of returns for 2002 was calculated as approximately 12%. In the memo, the SEC staff formulated the problem in the hypothesis testing framework:

H_0: Mutual funds with an independent chair produce returns no higher than those with a management chair.

versus

H_1: Mutual funds with an independent chair produce higher returns than those with a management chair.

Statistical power is the probability that a statistical test will identify an effect of a pre-specified magnitude when it in fact exists. The SEC memorandum states that "economic analyses, including those published in peer-reviewed scholarly journals, place strong emphasis on the statistical significance of results. Discussions of economic significance and the power of statistical methods employed are often given a back seat to this pursuit of statistical significance." In order to address how empirical analyses suffer as a result of this imbalance, the OEA/SEC carried out power calculations as a function of sample size (up to 500).[13] The analysis assumed that 1 in 5 fund families are chaired by an independent director or trustee. The sample size is the total number of funds in the two groups with a 1:4 allocation between the independent and management chaired funds. The power of the two sample *t*-test to detect differences in returns in the range of 1% to 5% (100 to 500 basis points) at significance level 5% was computed. The results showed that the power of the test, applied to the total sample of 448 funds**,** to detect a 3% difference between independent chaired and management chaired mutual funds would only be about 55%.

This means that the probability of **not** detecting a 3% difference in performance between the two types of funds would be a substantial, 45%, even when the data from all 448 funds are utilized. Typically in statistical applications one desires studies to have a power of at least 80% to detect a meaningful difference. Thus, the sample size used in the study submitted by the Chamber was insufficient.

Noting that a careful study of the difference in performance of the two groups of funds would control for other variables that affect returns, the SEC stated that an appropriate method of analysis would be a regression that incorporated major predictors. In order to calculate the power of a regression model, with a coefficient of determination[14] R^2, the SEC multiplied the original variance of the returns in each group by $(1-R^2)$ and then used the power formula for the two-sample *t*-test assuming the variance in returns in both groups equaled this smaller variance. In particular, when the standard deviation of returns equals 12 and R^2=.50 (.75) the resulting standard deviation of the error term in the regression becomes 8.5 (6.0). [15]

Comments:

1. The OEA/SEC memorandum did not describe the explanatory variables that a regression model should contain nor did it present a regression model relating a fund's returns and the predictors. The SEC used the overall explanatory ability of the model, as measured by R^2, to carry out their power calculations for studies based on samples of the size of the available data and on the sample standard deviation (12%) of fund returns.
2. It is well known from statistical theory that the power and hence the required sample size to achieve a pre-specified power, depends on the joint distribution of the predictors or independent variables of the regression model[16].

Our First Power Study

Previously[17], we showed that the approach the SEC adopted is statistically correct only when the predictors have the same distribution in both types of funds. To illustrate the flaw in the logic of the OEA/SEC memorandum the case of a regression model with a single predictor variable that is normally distributed (bell-shaped) was considered.

Assume that a regression model is used to predict a response *y* from information on a single independent variable *x* and membership in one of two groups. In the context of the SEC memorandum, the response (*y*) is fund return and the two groups are the independent chaired and management chaired mutual funds. A reasonable independent variable might be the amount of assets (*x*) a fund manages. Informally the model can be expressed as

fund return = a +b(asset size) + random error

In the context of this regression model, power assesses the likelihood of a statistical test to correctly deduce the difference in performance as measured by the response *y* between the two groups. The power is a value between 0 and 1 and the closer it is to 1 the better the test is at detecting a difference in performance. To be meaningful in this context, a

test needs sufficiently high power to identify a modest 2% or 3% difference in the returns of the two fund groups.

We demonstrated[17] that when the normal distribution of the predictor variable in the funds with independent directors has a different mean than in the management chaired, the power results given in the OEA/SEC memorandum are not correct. More precisely, we showed:

1. In the OEA/SEC memorandum power curves as a function of both sample size and magnitude of the difference in returns of the two groups [18] were computed for the two-sided test even though the stated alternative was (correctly) one-sided that the returns are higher in mutual funds with an independent chair.
2. When the distribution of the predictor variable is different in the two groups of funds, the power of the test for detecting difference in returns is significantly smaller than the power of the same test when the predictor distribution is identical in the two groups. We exhibited this via a simulation where the predictor had a normal distribution with different means (same variances) in the two fund groups. The power loss increased noticeably as the difference in the two means increased and exceeded 30% at the maximum difference in means we considered.

Here the effect of the difference in the distributions of the predictors in the two groups of funds on the estimated power of a regression analysis is explored in greater detail **in** more realistic situations**.**

What is realistic? It would be quite rare that one can identify a unique predictor of fund returns. Most likely fund returns depend upon several factors that should be used in a reliable prediction model. Here we just consider two predictors and assess how even this simple extension of the regression model can decrease the power of evaluating differential performance in the two types of funds when the two predictors behave differently in the two groups.

Furthermore, most predictors used to model fund profit are not normally distributed but follow a skewed distribution, which is non-negative and has large variance. Income and the size of funds are skewed. Often such economic data can be modeled by a log-normal distribution[19] or a mixture of log-normal distributions. In the next section we show using real fund data that the distribution of fund assets not only differs in the two fund groups but is also a log-normal in the independent chaired funds and a mixture of two log-normal distributions in the management chaired funds.

To assess the effect on power using a model with two predictors has on the performance of the types of funds, we carry out a small simulation study with a normally distributed factor, as in our first study, and another that has a distribution as the one estimated for the fund assets. We will see that even this fairly simple, yet more realistic model can lead to significant loss in power for detecting differences in fund performance.

The distribution of assets

Bobrock and Mack[7] used the Strategic Insight Simfund mutual fund database as their principal source for identifying and categorizing fund families and funds. All fund complexes with $10 billion or more in long-term, open-end fund assets as of December 31, 2003 included in this database were identified - 68 in all. Among these complexes, the 55 which have a significant retail (sales force and/or direct marketed), active management fund business and offer industry-standard retail load or no-load pricing were selected for the analysis. Two of these complexes (Fidelity and Dreyfus) were separated into sales force and direct market product lines, so there are a total of 57 fund families included. This excluded purely institutional funds but allowed funds that have both institutional and retail classes. A total of 2,437 funds with $3,728 billion in portfolio-level assets were eligible by these criteria, or 83% of all industry long-term open end fund assets. Each of the 57 families was identified as either an "independent chair" or "management chair" family, based on all or substantially all of its funds having an independent board chair, or not. On this basis, 14 of the 57 fund families have independent chairs, and 43 have management chairs.

Exhibit A-1[7] reports the number of funds and assets and the funds and assets included in their performance calculations. These funds have a Morningstar ranking or rating calculation. Overall, 86% of the eligible funds, representing 97% of the eligible fund assets were included in 3-year rankings and Star rating calculations, both of which require a minimum 3-year track record. The most common reason for funds not being included in Exhibit A-1 was lack of a sufficient track record.

The asset data from Exhibit A-1 are used to estimate the distribution of fund assets for the independent and management chair funds. The back-to-back histograms of the management (left) and independent (right) chair fund assets are plotted in Figure 1.

Figure 1. Back-to-back histograms of the management chaired fund assets (black) and the independent chaired fund assets (grey). The data come from Exhibit A-1[7].

The histograms of the assets for both groups of funds indicate their distributions differ. Moreover, for the management chaired funds, the asset distribution appears to be bimodal. Both are right skewed and between the exponential and log-normal distributions, the log-normal best fits the assets of the independent chair funds. We also determined that the assets of management chaired fund follow a mixture of two log-normals. Interestingly the first component of the mixture is almost the same as the independent chair fund asset distribution. That is, the distribution of assets of the management chair funds F_0 is given by

$$F_0 = w\mathrm{LN}(\mu_1, \sigma_1) + (1-w)\mathrm{LN}(\mu_2, \sigma_2)$$

where $\mathrm{LN}(\mu_i, \sigma_i)$ indicates a log-normal distribution with parameters μ_i and σ_i, $i = 1, 2$, and w is the mixing proportion for the two distributions. The parameters μ_i and σ_i are the mean and standard deviation of the respective log-normals. The logarithm of these distributions is normally distributed (hence the name log-normal distribution[20]).

Most management chair fund assets follow the first lognormal distribution, $\mathrm{LN}(\mu_1, \sigma_1)$ except for three whose asset size by far exceeds the rest (see histogram in Figure 2). Also, the assets of the independent chair funds follow $\mathrm{LN}(\mu_1, \sigma_1)$.

So, if it were not for the three outliers, the distributions of the two groups of assets would be about the same.

Fitting the data reported in Exhibit A-1, the following estimates of the parameters of the two log-normal distributions are obtained:

$$F_0 = wLN(40{,}896.51,\ 36{,}159.48) + (1-w)LN(514{,}653.9, 25{,}748.79)$$
$$F_1 = LN(37{,}752.24,\ 33{,}379.41)$$

The mixing proportion w is 40/43=0.93. Notice that the asset distribution of the independent chair funds and the *first* component of the mixture of the asset distribution of the management chair funds are very close. Also, the difference between the means of the two distributions is about 475,000 with standard deviation of the independent chair fund assets and the *first* component of the mixture of the asset distribution of the management chair funds being about 34,000 and 25,749 for the *second* component in the mixture. We use these values in our power calculations in Tables 1, 2, 3 and Figure 2 below.

Several predictors with different predictor distributions result in a greater loss in power

In our calculations we use the model

fund return = a + b×Factor1 + c×Factor2 + random error

to predict a fund's returns. We can think of Factor 1 as "Number of positions held," for example, which measures experience and diversification of interest and for which the normal would be a potentially good fit. The normal distributions in the two groups of funds differ only in their means and have the same standard deviation. We consider asset size to be Factor2. It has distributions close to the estimated ones from Exhibit A-1. That is, we fix the mixing proportion to .90, the difference in the two means to 475,000 and the standard deviations of the two log-normals to 30000 and 20000, respectively. To comply with the OEA/SEC power study[11], we also assumed the variance of y to be 144 and the error variance to be 72, 20% (n/5) of the funds to have independent chairs, and the significance level to be $\alpha = 0.05$. To achieve the target 50% in R^2, the difference in the two normal means of Factor1 is set to 25 when all other parameters are fixed.

We compare the model where the two predictors have the same distributions (normal and log-normal, respectively) in the two fund groups with the model where the number of positions held is normal with different means and assets have two different distributions ($LN(\mu_1, \sigma_1)$ and the mixture $w\mathrm{LN}(\mu_1, \sigma_1) + (1-w)\mathrm{LN}(\mu_2, \sigma_2)$) in the independent chaired fund group and in the management chair fund group, respectively, with $w = 0.90$. The results of these calculations are reported in Tables 1, 2 and 3. In Table 1 we report the average power and corresponding standard deviation for detecting 1%, 2%, 3%, 4% and 5% difference in fund returns for sample sizes ranging from 10 to 500.

Table 1: Comparison of the simulated power of the t-test approximation used by the SEC with the true power of the two-predictor model described in the text: average and standard deviation (in parentheses) of the simulated power of the *t*-test used in regression when the difference in average returns is 1%, 2%, 3%, 4% and 5% for n=10, …,500.

	1% difference	**2% difference**	**3% difference**	**4% difference**	**5% difference**
Same distribution (SEC)	0.181 (0.058)	0.424 (0.17)	0.661 (0.238)	0.802 (0.231)	0.879 (0.202)
Different Distributions	0.136 (0.047)	0.284 (0.135)	0.437 (0.229)	0.579 (0.284)	0.671 (0.293)

From Table 1 we see that if one were to use number of positions held and assets as predictors of fund return for mutual funds as those of Exhibit A-1, the average loss in power for detecting a 3, 4 or a 5 point difference in performance between the two types of funds would be at least 20% for sample sizes ranging from 10 to 500 in the more realistic situation where the predictors have different distributions in the two types of funds. For a 2% difference we would lose 14% power and even for 1% difference, which is very hard to detect in the available sample, one would lose 5% power.

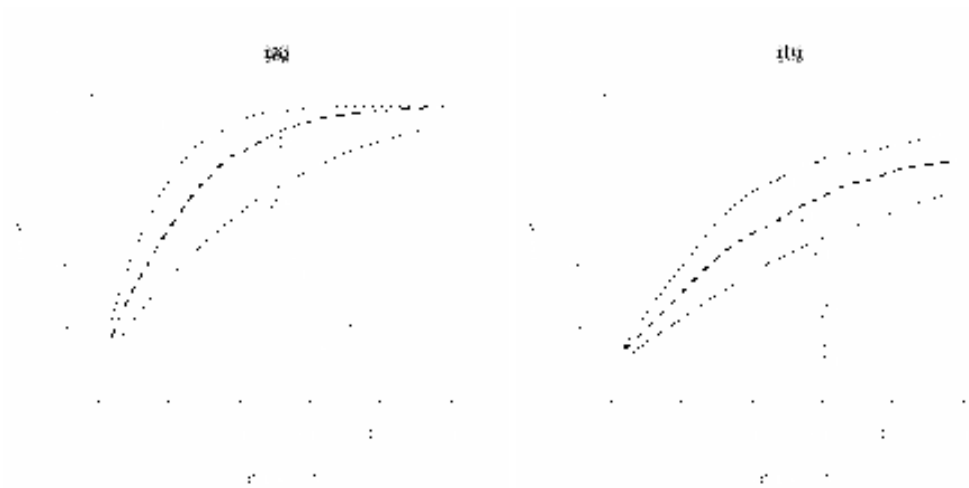

Figure 2. Smoothed power curves for testing difference in fund returns. Friedman's[21] super smoother was used to smooth the simulated power curves. *Panel (a):* The predictor distribution is the same in the two groups. The % difference in returns is indicated in the plot. *Panel (b):* The predictor distribution is different in the two groups. The % difference in returns is indicated in the plot. The flat solid line on the y-axis indicates a power of .80.

From Figure 2 we see that the power for detecting 1, 2, 3, 4, or 5 point difference in fund returns for the two fund groups is uniformly much lower when the predictors' distributions are different in the two fund groups. As the difference in the rate of returns increases, the power decreases much more drastically when the predictor is differently distributed compared to the case when the predictor has the same distribution that the SEC considered. Moreover, a sample of 488 funds would be required to detect a 3 point difference in returns with 80% power when the distributions of predictors are different, while only a sample of 303 is required if the distribution of assets was the same for the two types of funds. A sample of 342 (265) is required to detect a 4 (5) point difference in returns when the distributions of predictors are different, while samples of only 162 and 105, respectively, would be required when they are the same in the two fund types. Thus, much larger samples are required to detect the same difference in performance when the predictors have different distributions in the two types of funds.

In the OEA/SEC memorandum, approximate power of 0.80 is achieved for a difference of 5% at n=143.86 and for a difference of 4% at n=223.54. To make the comparison with the OEA/SEC study clearer, we approximate the power over a narrow window of sample sizes centered at 144 and 224. In Table 2 we report the average power along with its standard deviation over sample size ranging from 135-150 and in Table 3 for samples of size 215-230 when the two predictors have the same distributions (normal and log-normal, respectively) in the two fund groups and when the number of positions held is normal with different means and assets have two different distributions ($LN(\mu_1, \sigma_1)$ and the mixture $w\mathrm{LN}(\mu_1, \sigma_1) + (1-w)\mathrm{LN}(\mu_2, \sigma_2)$) in the independent chaired fund group and in the management chair fund group, respectively, with $w = 0.90$.

Table 2: Comparison of the simulated power of the t-test approximation used by the SEC with the true power of the two-predictor model described in the text: average and standard deviation (in parentheses) of the simulated power of the t-test for n=135, …,150, for testing 1%, 2%, 3%, 4% and 5% difference in fund returns.

	1% difference	**2% difference**	**3% difference**	**4% difference**	**5% difference**
Same distribution (SEC)	0.137 (0.007)	0.297 (0.026)	0.522 (0.039)	0.733 (0.042)	0.882(0.030)
Different Distributions	0.111 (0.016)	0.192 (0.056)	0.286 (0.113)	0.441 (0.151)	0.577 (0.142)

Table 3: Comparison of the simulated power of the t-test approximation used by the SEC with the true power of the two-predictor model described in the text: average and standard deviation of the simulated power of the *t*-test for *n*=215, …,230 for testing 1%, 2%, 3%, 4% and 5% difference in fund returns.

	1% difference	**2% difference**	**3% difference**	**4% difference**	**5% difference**
Same distribution (SEC)	0.171 (0.008)	0.416 (0.028)	0.684 (0.043)	0.883 (0.026)	0.968 (0.013)
Different Distributions	0.129 (0.02)	0.272 (0.058)	0.474 (0.042)	0.585 (0.201)	0.712 (0.174)

The first row in Tables 2 and 3 reports the power when the distributions of the two predictors are the *same* for the two types of funds: the case SEC considered. In the second rows the power to detect 1%-5% difference in fund returns when the distributions of the predictors in the two fund groups are *different* is reported. Here we see that for 1 point difference the average drop in power is about 5%, for a 2 point difference the drop is already 14% and for 3-5% can be as drastic as 30% at either sample size. Note also that it is very difficult to detect a difference of only 1% or 2% in the returns of the two types of funds, so even a modest increase in power would be helpful.

In particular, for the case which the SEC focused on, where detection of a 3% difference in performance is desired, the power would drop from 68.4% to 47.4% when the predictors have different distributions in the two types of funds. This is over 20% loss of power. It amounts to instead of having a relatively high probability of 68.4% of detecting a 3% difference with data from 224 funds (the SEC calculation) to having a 47.4% chance, less than 50%, of detecting a 3% difference in performance. In other words, despite the fact a model is fitted to 224 data points, the effectiveness of the model to detect a 3% difference is worse than simply tossing a coin.

Discussion

This paper illustrates the effect an incorrect statistical calculation can have on the strength of the conclusions derived from it. The calculation made by the SEC is only correct in the unrealistic situation where the distribution of the relevant predictor variables is the same in both groups. The data on mutual fund returns relevant to the issue of concern in the regulation did not satisfy this special condition; indeed, the average assets under management of the two groups differed by a factor of two. Previously we demonstrated[17] that even if the only difference between the two types of funds is in the means of the distributions of a single predictor, the calculation made by the SEC *over-estimates* the statistical power of the test. The correct calculation shows that the power of a study comparing the returns of the management chaired funds to those of the independently chaired funds would be significantly less than that found by the SEC. When one considers the more realistic scenario that a regression model for fund returns

would be based on more than one predictor, it is even less likely that the joint distribution of several predictors will be identical in both groups of fund managers. The results reported here show that once more factors are incorporated into the model, the power of the test for detecting difference in performance between the two types of funds will drop even more than when only one factor is taken into account. This means that such a study would likely have very low power of detecting a difference of 2%, or even 3%, in the returns of the two groups with the available data. In view of the large difference at least in the amounts (assets) managed by the two types of funds, this problem is likely to occur if the parties develop a full regression model in future proceedings. Additional results[22, 23] concerning the power of tests in multiple linear regressions further support our conclusion.

The importance of the power of a statistical test is often not fully appreciated in the legal and regulatory setting. Indeed, in a jury discrimination case, Berghuis v. Smith currently before the U.S. Supreme court, neither party informed the Court of the power of the standard method to detect a level of under-representation corresponding to the "four-fifths" differential in success rates used in employment discrimination cases. Yet, if the analysis was based on the jury composition data for the six months up to the time of the trial, its power would be slightly less than 40% of detecting this meaningful difference[24]. While it was appropriate for the SEC to consider the power of a statistical method to detect a meaningful difference in fund performance; the correct calculation would have strengthened the agency's argument that the available sample was *too small* to detect a meaningful differential in performance of the two types of funds.

Recent studies in the financial and business literature have been concerned with factors, e.g. the cost of acquiring information about a firm, that are related to the effectiveness of outside directors**.** Duchin et al.[25] explore this topic and provide references to literature. The studies cited by the Chamber of Commerce and submitted to the SEC, which compared the returns of funds with a high fraction of independent directors to those with the current legally specified minimum number; do not seem especially pertinent to the purpose of the proposed regulation. The regulation concerns those funds whose sponsors will make trades where there is a potential conflict of interest, e.g. an investment bank is managing an IPO (initial public offering) and also purchasing some of those shares for the funds it manages. One expects such trades to be a *relatively small* proportion of the trades made by a mutual fund. Hence, studies focusing on the profitability of this special type of trade relative to other trades made by the fund or to similar "special" trades made by funds with independent chairs and boards would be more germane to the problem the SEC regulation is addressing. This problem of choosing a statistically sound "control" group arises in the choice of comparison groups (or individual comparators) in equal employment cases where courts have expressed concern with studies that are over-inclusive, e.g. include all employees rather than just those qualified for the job or not subject to the same employment practice being scrutinized, or under-inclusive, e.g. comparing a plaintiff to a subset of all comparable applicants or employees rather than the full set.[26]

[1] *See* Investment Company Governance, Final Rule, 69 Fed. Reg. 46,378 (Aug. 2, 2004).
[2] *See* Role of Independent Directors of Investment Companies, Final Rule, 66 Fed. Reg. 3734 (Jan. 16, 2001) for the criteria a director needs to satisfy in order to be considered independent.
[3] 412 F.3d 133 (D.C. Cir. 2005)
[4] See Investment Company Governance, Release No. 26,985, 70 Fed. Reg. 39390-39398 (July 7, 2005).
[5] 443 F.3d 890 (D.C. Cir. 2006).
[6] Ibid at 907.
[7] Bobroff, G. & Mack, T. 2004 Assessing the Significance of Mutual Fund Board Independent Chairs, submitted as part of Fidelity's comment letter to the Commission for consideration of Investment Company Governance rules.
http://www.sec.gov/rules/proposed/s70304/fidelity031804.htm

[8] *See* n. 5 at 143 where the opinion notes that the study did not rule out "other important differences [than independence of the chairman] that may have impacted performance results," 69 Fed.Reg. at 46,383 n.52 (quoting study), and because it did not use a reliable method of calculating fund expenses.
[9] *See* Exhibit 4 in Bobroff and Mack, 2004.
[10] Spatt, C. S. 2006a *Literature review on independent mutual fund chairs and directors*. OEA/SEC MEMORANDUM.
www.**sec**.gov/rules/proposed/s70304/**oea**memo12**29**06-litreview.pdf

[11] Spatt, C. S. 2006b *Power Study as Related to Independent Mutual Fund Chairs*. OEA/SEC MEMORANDUM.
http://www.sec.gov/rules/proposed/s70304/oeamemo122906-powerstudy.pdf

[12] This is the CRSP Survivor-Free U.S. Mutual Fund Database produced by the Center for Research in Securities Prices at the University of Chicago Graduate School of Business.
[13] Figure 1, p. 9 of the OEA/SEC Memorandum (Spatt, 2006b).
[14] The R^2 statistic measures the proportion of the variance in the response (returns) that is explained by the predictors used in the regression.
[15] The total variation in returns is the sum of the variation due to the predictors plus that due to the error. Since the total standard deviation assumed by the SEC equals 12, the error variance changes with R^2. The numbers in the text are obtained from $\sigma_\varepsilon^2 = Var(y)(1-R^2) = 12^2(1-.50) = 72$ and the error standard deviation is $\sqrt{72} = 8.485 \approx 8.5$.
[16] See Kleinbaum, D. G., Kupper, L. L, Nizam, A. & Muller, K. 2008 *Applied Regression Analysis and Other Multivariate Methods* (4th ed.). North Scituate, MA: Duxburg Press, and Hsieh, F. Y., Bloch, D. A. & Larsen, M. D. 1998 A simple method of sample size calculation for linear and logistic regression. *Statistics in Medicine*, **17**, 1623-1634.
[17] Bura, E. and Gastwirth, J. L. (2009) How accurate are the power calculations relied upon the SEC? *Law, Probability and Risk,* **8**, 277-288.
[18] See Figures 2 and 3 on p. 11 of the OEA/SEC memorandum.
[19] Aitchison, J. and Brown, J.A.C. 1957.The Log-normal Distribution.Cambridge (UK): Cambridge University Press; Crow EL, Shimizu K, eds. 1988. Log-normal Distributions: Theory and Application. New York: Dekker; Johnson NL, Kotz S, Balakrishan N. 1994. Continuous Univariate Distributions. New York:Wiley.
[20] *See* Aitchison and Brown (1957)
[21] Friedman, J. H. 1984 A variable span scatterplot smoother. Laboratory for Computational Statistics, Technical Report 5, Stanford University, Department of Statistics.
[22] Andrews, D. W. K. 1989 Power in econometric applications. *Econometrica*, **57**, 1059-1090.
[23] Gatsonis, C. & Sampson, A. R. 1989 Multiple correlation: exact Power and sample size calculations. *Psychological Bulletin*, **106**, 516-524.
[24] Gastwirth, J.L & Pan, Q. (2009). Statistical Measures and Methods for Assessing the Representativeness of Juries: A Reanalysis of the data in *Berghuis v. Smith.*(Technical Report available on SSRN).

[25] Duchin, R., Matsusaka, J. & Ozbas, O. 2008 When Are Outside Directors Effective? University of Southern California Law and Economics Working Paper Series. Working Paper 71. http://law.bepress.com/usclwps/lewps/art71

[26] For example, in Blair v. Henry Filters, Inc., 101 FEP Cases 1345 (6th Cir. 2007) the statistical analyses of both the plaintiff (Id. at fn. 13) and defendant (at fn. 14) were faulty. In particular, the defendant submitted data on the ages of employees in all three of its companies rather than the one in which the plaintiff was employed. The opinion referred to Bender vs. Hecht's Department Stores, 455 F.3d 612 (6th Cir. 2006) where criteria for determining the relevant comparison population in RIF cases is described. This issue was also discussed by Judge Posner in Crawford v. Indiana Harbor Belt Railroad Co., 461 F.3d 844, 846 (7th Cir. 2006) where he notes that the comparators need not be "nearly identical" to the plaintiff but should be "sufficiently comparable" to suggest that the plaintiff "was singled out for worse treatment."