EARNINGS ESTIMATES, PREDICTOR SPECIFICATION, AND MEASUREMENT ERROR

BRUCE I. JACOBS

is a principal of Jacobs Levy Equity Management. He was formerly senior managing director of a quantitative equity management affiliate of the Prudential Asset Management Company. He has also taught at the University of Pennsylvania's Wharton School. He holds a B.A. from Columbia College, an M.S. from Columbia University's School of Engineering, an M.S.I.A. from Carnegie-Mellon University's Graduate School of Industrial Administration, and an M.A. and Ph.D. in finance from the Wharton School.

KENNETH N. LEVY

is a principal of Jacobs Levy Equity Management. He was formerly managing director of a quantitative equity management affiliate of the Prudential Asset Management Company. He holds a B.A. from Cornell University, and an M.B.A. and M.A. from the University of Pennsylvania's Wharton School. He has completed all requirements short of the dissertation for a Ph.D. in finance at Wharton.

MITCHELL C. KRASK

is a financial econometrician at Jacobs Levy Equity Management. His previous experience includes work for Prudential Asset Management Company and Mellon Equity Associates, as well as consulting to IBES and Bank of America. He has also taught at the Katz Graduate School of Business of the University of Pittsburgh, where he received an M.S. in statistics. He holds a B.S. in fluids engineering, M.S. degrees in economics and engineering, and a Ph.D. in economics, all from the University of Illinois at Chicago.

Securities researchers today are able to draw upon a wider array of data from a broader universe of companies and a more extensive time horizon than ever before. This new wealth of information offers new ways to fine-tune and improve investment decision-making — but it also offers greater leeway for error. While the right choice of data can enhance investment performance, the wrong choice may introduce measurement error that detracts from performance.

This article explores some of the crucial decisions that arise when expectational data are used to construct explanatory variables for predicting returns. We show how these decisions can lead to measurement error when variables are misspecified, and how treatment of incomplete data sets can affect empirical analyses. We focus on expectational earnings data and their use in constructing earnings predictors for portfolio screening and for quantitative modeling (forecast E/P and earnings trend, in particular). The findings are generalizable to a wide range of data, predictors, and investment approaches.

We begin with a brief exploratory analysis of the issues that arise in predictor specification. We then present some evidence on how predictor specification can affect the results of screening and modeling processes. The findings suggest that the importance of the specification problem varies, depending upon the predictor and the use to which it is being put.

We also discuss some issues that further complicate the specification problem. When data are unavailable, one must decide whether to exclude from the analysis stocks lacking the desired data, or to fill in the gaps using substitute data. We suggest a method that can be used to arrive at the best available data set when observations are missing. Further, we examine whether the importance of predictor specification varies, not only across predictors, but also across different types of stocks. As an illustration, we stratify stocks by extent of analyst coverage.

PREDICTOR SPECIFICATION AND MEASUREMENT ERROR

In screening stocks for portfolio selection or in modeling stock behavior, one typically considers a number of variables as potential predictors of return. These include forecast E/P; forecast earnings trend (changes in estimates); earnings surprise; forecast earnings controversy (dispersion of earnings estimates); growth rates in expected earnings; measures related to analyst coverage or neglect; and analyst participation rates in earnings changes (number of revisions). Beyond the problem of selecting the variable or set of variables that will provide the best estimate of future return, one faces the problem of selecting the data that will provide the best estimate of the variable.[1]

Consider forecast E/P. It is typically defined as the mean earnings forecast divided by stock price. But which "mean" does this mean? A mean based on all available estimates — that is, the consensus mean? Or a mean based on some but not all available estimates? Which estimates should such a subset include? Should inclusion be based upon timeliness? If so, how does one measure timeliness? On the basis of some fixed horizon — say, a six-week "flash" estimate using only those estimates revised over the last six weeks? Or should one include all the latest estimates available for a given sample of stocks, whenever they were made?[2]

Nor is the mean the only possible measure of the central tendency of analyst forecasts. Other candidates are the median, the trimmed mean, or the midpoint between the high and low estimates. All of these will provide an estimate that can be used to calculate forecast E/P.[3]

Expectational data may also cover many different fiscal periods. Expected earnings, for example, are often provided not only for the current fiscal year (i.e., fiscal year 1), but also for the following year and the year after that. Expectational estimates are also provided for quarterly earnings and for a long-term (five-year) growth rate. Should one use expectations for fiscal year 1 only? Or should information from other periods be used as well? If so, should one construct a separate predictor for each fiscal period, or combine periods, using a composite indicator?

Different choices of expectational earnings can lead to different estimates of E/P for the same company. That is to say, the various possible specifications of the predictor will produce a distribution of E/P estimates. Some estimates may be different enough to result in different relative valuations for the same company.[4] Predictive power may also differ across alternative specifications.

Use of less than the best available data set can reduce the accuracy of a given predictor, leading to measurement error. A predictor based upon a particular specification may be inferior because the data are less available, less timely, or more error-prone than alternative specifications. A mean based on consensus earnings data, for example, may be less accurate than a mean based on earnings revisions made in the past four weeks, because the consensus data are likely to include stale estimates. If this is the case, then use of consensus data to construct forecast E/P when more timely analysts' revisions are available will result in measurement error.

Measurement error can in turn affect the empirical analyses associated with quantitative modeling. In a simple linear regression, for example, measurement error in the forecast E/P will bias the estimated positive relationship between forecast E/P and subsequent return downward. In general, the greater a predictor's measurement error, the greater the bias toward zero. Intuitively, measurement error dilutes the information content associated with a given predictor. (See the appendix.)

In real life, it is difficult to know which specification is best. Furthermore, the best specification may differ both over time and across different types of stocks. In other words, the degrees of measurement error associated with alternative specifications may change over time with changes in the economy, the industry, or the firm, or with changes in data technology. At a given time, they may also differ across industries or sectors or market capitalizations. Predictor specification may also be sensitive to the investment horizon, with the specification best suited to predicting monthly returns not optimal for a daily or quarterly horizon. Finally, the best specification may depend on the investment strategy, and on related criteria such as portfolio turnover and risk.

Given the complexities involved in choosing among alternative predictor specifications, it may be wise to question whether attempts to improve specification are worth the effort required. Just how important is predictor specification? Is it more important for some predictors than others? For some investment approaches than others?

We consider these questions in the context of two predictors — forecast E/P and forecast earnings trend — for two alternative specifications — consensus versus flash data; two investment approaches — portfolio screening and return modeling; and two investment universes — one of 30 and the other of 3,000 stocks.

Alternative Specifications of E/P and Earnings Trend for Screening

Exhibit 1 presents the thirty Dow Jones industrial stocks as of December 1996, together with their prices and fiscal year 1 consensus and flash earnings estimates. These data are used to calculate two alternative specifications of forecast E/P for each stock. The table provides the calculated values and each stock's ranking by each specification.

The last four columns of Exhibit 1 can be used to compare the consensus forecast E/P (fiscal year 1 consensus earnings mean divided by price) with the flash forecast E/P (fiscal year 1 six-week flash earnings mean divided by price). There is little difference between the two specifications. The two sets of E/P values are highly correlated, with a Pearson correlation of 0.9990 and a Spearman rank correlation of 0.9996.[5]

As Exhibit 1 shows, the two E/P specifications result in identical rankings except in the case of two stocks — Philip Morris and Sears Roebuck. Sears is ranked ninth by the consensus forecast E/P but tenth by the flash forecast E/P, while Philip Morris is ranked tenth by consensus forecast E/P but ninth by flash forecast E/P.

The similarities in rankings by the two E/P specifications lead to similarities in portfolio composition. For the top five, bottom five, top ten, and bottom ten stock portfolios constructed on the basis of E/P, compositions would be the same whether E/P were specified with consensus earnings data or flash earnings data. These results suggest that the precise specification of E/P may matter little, at least in terms of screening a small universe of stocks for potential portfolio inclusion.

To extend the analysis, we consider the same E/P specifications applied to a larger universe of stocks and over a longer time period. Results for a 3,000-stock universe are similar to those for the 30-stock universe.[6] In particular, the consensus and flash forecast E/Ps are highly correlated, with a Pearson correlation of 0.9863 and a Spearman rank correlation of 0.9921.

Furthermore, portfolio compositions across the two specifications, although not identical, as is the case with the 30-stock universe, are quite close. Top 100, 300, and 500 portfolios selected from consensus and flash E/P forecasts have 90, 282, and 479 stocks in common, respectively, while bottom 100, 300, and 500 portfolios constructed from the alternative specifications have in common 96, 288, and 478 stocks.

The similarities between the alternative specifications of E/P also hold over a longer time frame — from April 1990 through December 1996. Over this period, Pearson correlations between the two specifications range from 0.8500 to 0.9975; Spearman rank correlations are even higher, ranging from 0.9725 to 0.9950. The proportions of stocks common to portfolios selected by the alternative specifications remain similar to those found for the 3,000-stock universe in December 1996.[7]

What holds true for E/P, however, may not hold true for other predictors. Exhibit 2 shows alternative specifications of earnings trend for fiscal year 1, defined as follows:

$$\text{Consensus Trend} = (\text{Current Consensus Mean} - \text{One Month Ago Consensus Mean})/\text{Price} \quad (1)$$

$$\text{Flash Trend} = (\text{Current Six-Week Flash Mean} - \text{One Month Ago Six-Week Flash Mean})/\text{Price} \quad (2)$$

The exhibit provides the calculation of these predictors

EXHIBIT 1
CONSENSUS VERSUS FLASH FORECAST E/P FOR THIRTY-STOCK UNIVERSE

TICKER	COMPANY	NO. OF FY1 CONSENSUS ESTIMATES	NO. OF FY1 FLASH ESTIMATES	CONSENSUS FY1 MEAN DEC. 1996	FLASH FY1 MEAN DEC. 1996	PRICE DEC. 1996	CONSENSUS FY1 E/P	FLASH FY1 E/P	RANK CONSENSUS FY1 E/P	RANK FLASH FY1 E/P
AA	Aluminum Co. of America	22	2	3.59	3.80	60.88	0.0590	0.0624	14	14
ALD	Allied Signal	19	3	3.60	3.62	68.25	0.0527	0.0530	18.5	18.5
AXP	American Express	18	4	3.57	3.60	53.63	0.0666	0.0671	11	11
BA	Boeing	25	9	2.96	2.94	101.50	0.0292	0.0290	28	28
BS	Bethlehem Steel	15	4	0.32	0.29	9.00	0.0356	0.0322	27	27
CAT	Caterpillar Tractor	28	4	6.92	6.94	74.25	0.0932	0.0935	3	3
CHV	Chevron	31	15	4.03	4.04	64.13	0.0628	0.0630	13	13
DD	Du Pont	23	7	6.61	6.66	93.00	0.0711	0.0716	7	7
DIS	Walt Disney	34	20	2.66	2.65	71.50	0.0372	0.0371	26	26
EK	Eastman Kodak	19	3	4.43	4.41	79.63	0.0556	0.0554	15	15
GE	General Electric	24	1	4.39	4.40	99.25	0.0442	0.0443	23	23
GM	General Motors	25	16	5.75	5.65	54.50	0.1055	0.1037	1	1
GT	Goodyear Tire & Rubber	14	6	4.37	4.35	49.50	0.0883	0.0879	5	5
IBM	IBM	23	6	11.01	11.09	158.63	0.0694	0.0699	8	8
IP	International Paper	23	9	1.53	1.48	39.63	0.0386	0.0373	25	25
JPM	J.P. Morgan	24	3	7.47	7.35	96.88	0.0771	0.0759	6	6
KO	Coca-Cola	26	2	1.40	1.40	48.38	0.0289	0.0289	29	29
MCD	McDonalds	33	9	2.22	2.22	45.50	0.0488	0.0488	20	20
MMM	Minnesota Mining & Mfg.	19	4	3.61	3.61	81.13	0.0445	0.0445	22	22
MO	Philip Morris	24	3	7.66	7.70	111.63	0.0686	0.0690	10	9
MRK	Merck	37	16	3.17	3.18	76.88	0.0412	0.0414	24	24
PG	Procter & Gamble	24	3	4.81	4.81	104.13	0.0462	0.0462	21	21
S	Sears Roebuck	33	11	3.06	3.06	44.38	0.0689	0.0689	9	10
T	AT&T	34	6	3.48	3.48	38.81	0.0897	0.0897	4	4
TX	Texaco	26	12	6.30	6.33	97.63	0.0645	0.0648	12	12
UK	Union Carbide	16	16	3.86	3.86	39.00	0.0990	0.0990	2	2
UTX	United Technologies	19	3	3.39	3.41	64.38	0.0527	0.0530	18.5	18.5
WX	Westinghouse	11	5	–0.05	–0.08	17.88	–0.0028	–0.0045	30	30
XON	Exxon	35	14	5.35	5.38	97.25	0.0550	0.0553	16	16
Z	Woolworth	9	6	1.17	1.17	22.00	0.0532	0.0532	17	17

Source: IBES data.

EXHIBIT 2
CONSENSUS VERSUS FLASH FORECAST TREND FOR THIRTY-STOCK UNIVERSE*

TICKER	COMPANY	NO. OF FY1 CONSENSUS ESTIMATES	NO. OF FY1 FLASH ESTIMATES	CONSENSUS FYI MEAN DEC. 1996	CONSENSUS FYI MEAN NOV. 1996	FLASH FYI MEAN DEC. 1996	FLASH FYI MEAN NOV. 1996	PRICE DEC. 1996	CONSENSUS FYI TREND	FLASH FYI TREND	RANK CONSENSUS FYI TREND	RANK FLASH FYI TREND
AA	Aluminum Co. of America	22	2	3.59	3.64	3.80	3.46	60.88	–0.000821	0.005585	25	1
ALD	Allied Signal	19	3	3.60	3.60	3.62	3.62	68.25	0.000000	0.000000	18	15
AXP	American Express	18	4	3.57	3.55	3.60	3.60	53.63	0.000373	0.000000	5	15
BA	Boeing	25	9	2.96	2.96	2.94	3.05	101.50	0.000000	–0.001084	18	26
BS	Bethlehem Steel	15	4	0.32	0.36	0.29	0.31	9.00	–0.004444	–0.002222	28	27
CAT	Caterpillar Tractor	28	4	6.92	6.90	6.94	6.92	74.25	0.000269	0.000269	6	6
CHV	Chevron	31	15	4.03	3.96	4.04	4.01	64.13	0.001092	0.000468	2	3
DD	DuPont	23	7	6.61	6.59	6.66	6.65	93.00	0.000215	0.000108	9	11
DIS	Walt Disney	34	20	2.66	2.65	2.65	2.65	71.50	0.000140	–0.000000	11	15
EK	Eastman Kodak	19	3	4.43	4.42	4.41	4.44	79.63	0.000126	–0.000377	13	21
GE	General Electric	24	1	4.39	4.39	4.40	4.41	99.25	0.000000	–0.000101	18	19
GM	General Motors	25	16	5.75	6.08	5.65	6.10	54.50	–0.006055	–0.008257	29	29
GT	Goodyear Tire & Rubber	14	6	4.37	4.38	4.35	4.38	49.50	–0.000202	–0.000606	23	24
IBM	IBM	23	6	11.01	11.02	11.09	10.98	158.63	–0.000063	0.000693	22	2
IP	International Paper	23	9	1.53	1.57	1.48	1.57	39.63	–0.001009	–0.002271	26	28
JPM	J.P.Morgan	24	3	7.47	7.46	7.35	7.42	96.88	0.000103	–0.000723	14	25
KO	Coca–Cola	26	2	1.40	1.39	1.40	1.40	48.38	0.000207	0.000000	10	15
MCD	McDonalds	33	9	2.22	2.23	2.22	2.22	45.50	–0.000220	0.000000	24	15
MMM	Minnesota Mining & Mfg.	19	4	3.61	3.61	3.61	3.60	81.13	0.000000	0.000123	18	10
MO	Philip Morris	24	3	7.66	7.66	7.70	7.68	111.63	0.000000	0.000179	18	9
MRK	Merck	37	16	3.17	3.16	3.18	3.16	76.88	0.000130	0.000260	12	7
PG	Procter & Gamble	24	3	4.81	4.81	4.81	4.81	104.13	0.000000	0.000000	18	15
S	Sears Roebuck	33	11	3.06	3.05	3.06	3.07	44.38	0.000225	–0.000225	8	20
T	AT&T	34	6	3.48	3.47	3.48	3.48	38.81	0.000258	0.000000	7	15
TX	Texaco	26	12	6.30	6.22	6.33	6.31	97.63	0.000819	0.000205	3	8
UK	Union Carbide	16	16	3.86	4.19	3.86	4.19	39.00	–0.008462	–0.008462	30	30
UTX	United Technologies	19	3	3.39	3.39	3.41	3.39	64.38	0.000000	0.000311	18	4
WX	Westinghouse	11	5	–0.05	–0.02	–0.08	–0.07	17.88	–0.001678	–0.000559	27	23
XON	Exxon	35	14	5.35	5.31	5.38	5.35	97.25	0.000411	0.000308	4	5
Z	Woolworth	9	6	1.17	1.09	1.17	1.18	22.00	0.003636	–0.000455	1	22

*Note: Walt Disney had a fiscal year change; thus we use FY2 mean data for November 1996 in trend calculations.

Source: IBES data.

and rankings for each stock in the thirty-stock universe.

Unlike the forecast E/P specifications, the alternative specifications of earnings trend do not lead to a similarity of results. In fact, for a number of companies, the forecast trends differ not only in magnitude but also in direction. Sears, for example, has an increasing earnings trend based on consensus data, but a decreasing trend based on six-week flash data. Aluminum Company of America has a decreasing consensus trend but an increasing flash trend.

The differences between the two specifications are reflected in their correlations. The Pearson correlation of 0.7688 and Spearman rank correlation of only 0.4270 are much lower than the correlations between the E/P specifications. The differences are also reflected in the two trend specifications' rankings of the thirty stocks. These have substantial implications for stock selection and portfolio composition.

As Exhibit 3 shows, although the bottom five portfolios selected by the two trend specifications hold four stocks in common, the top five portfolios hold only two of the same stocks. Only four stocks are common to both top ten portfolios, while seven are common to the bottom ten portfolios. In fact, three stocks in the top five portfolio based on the flash specification, including its top-ranked stock, Aluminum Company of America, are placed in the bottom ten portfolio ranked by consensus data. Conversely, Woolworth, the top-rated stock on the basis of consensus data, is ranked twenty-second on the basis of flash data. Over time, such dissimilarities between rankings by alternative specifications affect portfolio composition and lead to differences in performance.[8]

Similar results hold when the earnings trend specifications are applied to the larger 3,000-stock universe over the longer April 1990 through December 1996 period. The correlations between the six-week flash trend and the consensus trend are much lower than the correlations observed for the forecast E/P specifications, ranging from 0.450 to 0.750 for the Spearman rank and 0.200 to 0.950 for the Pearson. The proportions of stocks common to various-sized portfolios constructed on the basis of the two trend specifications are also lower, on the order of 70% for the top 100, 300, and 500 stock portfolios and 67% for the bottom 100, 300, and 500 stock portfolios.

These findings suggest that, in screening, the precise specification of the E/P predictor (at least as between six-week flash and consensus data) may not have much effect on portfolio results, especially when the investment universe consists of well-known, widely followed stocks. For forecast earnings trend, however, different specifications of the predictor may lead to very different portfolios and very different investment results.

EXHIBIT 3
CONSENSUS VERSUS FLASH TREND PORTFOLIOS

CONSENSUS TREND	FLASH TREND
TOP FIVE STOCKS	
Z	AA
CHV	IBM
TX	CHV
XON	UTX
AXP	XON
BOTTOM FIVE STOCKS	
UK	UK
GM	GM
BS	IP
WX	BS
IP	BA
TOP TEN STOCKS	
Z	AA
CHV	IBM
TX	CHV
XON	UTX
AXP	XON
CAT	CAT
T	MRK
S	TX
DD	MO
KO	MMM
BOTTOM TEN STOCKS	
UK	UK
GM	GM
BS	IP
WX	BS
IP	BA
AA	JPM
MCD	GT
GT	WX
IBM	Z
7 Tied for Tenth	EK

Source: December 1996 IBES data.

Alternative Specifications of E/P and Trend for Modeling Returns

Does the relationship between stock returns and their possible predictors depend on the specification of the predictors? To examine this, we fit the model:

$$\text{Return} = a + b(\text{Consensus Predictor}) + c(\text{Flash} - \text{Consensus Predictor}) + d(\text{Controversy}) + e(\text{Neglect}) \quad (3)$$

Here Return is the excess return for the subsequent month (relative to the Treasury bill rate). The Consensus and Flash Predictors used are based on fiscal year 1 earnings estimates. Controversy is defined as the standard deviation of fiscal year 1 earnings estimates, where the estimates are based on flash data if available or otherwise on consensus data. Neglect is defined as:

$$-\text{Log}(1 + \text{Number of Fiscal Year 1 Analysts})$$

Controversy and neglect are included to control for some important expectational-related return effects.[9]

We estimate two separate models — one for the forecast E/P predictor, and the other for the forecast earnings trend predictor — using the 3,000-stock universe and monthly data from April 1990 through December 1996.[10] The analysis includes only those stocks for which at least consensus data are available. All explanatory variables are standardized with winsorization set at plus or minus 5 standard deviations from the mean in order to truncate outliers. Methods of estimation include equal-weighted least squares regression and monotone regression.[11]

If the inclusion of flash data has explanatory power beyond that provided by consensus data, the coefficient c in Equation (3) will be significantly different from zero. This would indicate that the relationship between returns and predictors based upon flash data differs from that between returns and predictors based on consensus data. Furthermore, a positive and significant coefficient would suggest that companies with a positive flash-consensus differential would be expected to have higher excess returns, on average, than companies with a flash mean below the consensus mean.

Such a finding would imply not only that the relationship between returns and predictor is sensitive to specification, but also that the relationship between returns and flash data is stronger than the relationship between returns and consensus data. A priori, one might expect this to be the case, since flash data are more timely, and hence likely have higher information content than consensus data.

Exhibits 4 and 5 present the results from the estimated models. The evidence in Exhibit 4 pertaining to forecast E/P suggests that the return-predictor relationship is sensitive to specification. With

EXHIBIT 4
INCREMENTAL EFFECT OF THE FLASH FORECAST E/P
REGRESSION RESULTS FOR 3,000-STOCK UNIVERSE — APRIL 1990-DECEMBER 1996

	CONSENSUS FORECAST E/P	INCREMENTAL FLASH FORECAST E/P	CONTROVERSY	NEGLECT
		LEAST SQUARES REGRESSION		
Mean	0.3353	0.2210	−0.0448	−0.1540
Standard Error Mean	0.1009	0.0428	0.0924	0.0962
T-Statistic	3.3230	5.1618	−0.4845	−1.6006
P-Value	0.0014	0.0000	0.6294	0.1134
		MONOTONE RANK REGRESSION		
Mean	0.0412	0.0235	−0.0155	−0.0258
Standard Error Mean	0.0090	0.0040	0.0068	0.0091
T-Statistic	4.5710	5.8112	−2.2622	−2.8224
P-Value	0.0000	0.0000	0.0264	0.0060

EXHIBIT 5
INCREMENTAL EFFECT OF THE FLASH FORECAST TREND
REGRESSION RESULTS FOR 3,000-STOCK UNIVERSE — APRIL 1990-DECEMBER 1996

	CONSENSUS TREND	INCREMENTAL FLASH TREND	CONTROVERSY	NEGLECT
		LEAST SQUARES REGRESSION		
Mean	0.3859	0.0134	–0.1289	–0.1704
Standard Error Mean	0.0606	0.0341	0.0872	0.1001
T-Statistic	6.3696	0.3922	–1.4785	–1.7025
P-Value	0.0000	0.6960	0.1432	0.0925
		MONOTONE RANK REGRESSION		
Mean	0.0456	0.0031	–0.0115	–0.0284
Standard Error Mean	0.0062	0.0022	0.0065	0.0094
T-Statistic	7.4019	1.4320	–1.7692	–3.0297
P-Value	0.0000	0.1560	0.0807	0.0033

least squares estimation, the coefficient c for the flash-consensus differential is positive, with a p-value of less than 0.0001.[12]

This suggests that the incremental effect of flash data is highly significant, and that one may expect differences between the flash and consensus forecast E/Ps for a given company to lead to differences in return. Other things equal, those companies with flash forecast E/Ps higher than their consensus E/Ps will tend to enjoy higher returns than those companies with flash E/Ps lower than their consensus E/Ps. The regression estimate of 0.2210 for the flash-consensus forecast E/P differential suggests that, with a 1 standard deviation increase in the differential, average excess return can be expected to increase by around 22 basis points, other things being equal.

Our results suggest that use of consensus E/P can capture the positive relationship between returns and forecast E/P. Other things equal, average excess return increases by around 34 basis points with a 1 standard deviation increase in exposure to forecast E/P. But use of flash E/P can lead to even higher returns. The other predictors included in the model — controversy and neglect — have p-values of 0.6294 and 0.1134, respectively, indicating that both predictors are not significantly different from zero, even at the 10% level.

The results from the monotone regression in Exhibit 4 are somewhat different from those of the least squares regression, however.[13] All the predictors are significant at the 5% level. That is, all predictors are monotonically related to return, with consensus forecast E/P and the incremental flash forecast E/P positively related to returns, and controversy and neglect inversely related.

The average estimates from the monotone regressions may be interpreted as the marginal effect on stock return rank of an increase in the rank of each predictor, other things equal. While least squares coefficients represent the partial estimated return between standardized predictors and subsequent monthly returns, monotone regression coefficients estimate the relationship in terms of rank. Thus, over the period of study, an increase of 100 in the rank for the consensus E/P predictor is associated with an increase in return rank of 4.12. Similarly, an increase of 100 in the incremental flash forecast E/P rank is associated with an increase of 2.35 in the return rank.

Exhibit 5 reports the results from estimating the model using the earnings trend predictor rather than the E/P predictor. Here the evidence for an incremental effect from the use of flash data is much less conclusive. With least squares regression, the estimated incremental effect is positive, but very small, with an average value of 0.0134; that is, on average, excess return increases by only about 1 basis point with a 1 standard deviation increase in the flash-consensus trend differential. The p-value of 0.6960 also indicates that the incremental effect is not significant; one can conclude that the relationship between earnings trend and stock returns, at least over this period, does not differ between a quantitative model using consensus data and one using six-week flash data.

Clearly the most significant predictor in the least squares estimation is the fiscal year 1 consensus earnings trend. This predictor is positive and highly significant, with a p-value of less than 0.0001. Other things equal, our results suggest that, for every 1 standard deviation increase in consensus earnings trend, excess return increases by about 39 basis points.

The incremental flash effect is somewhat stronger in the monotone regression model, although still statistically insignificant. The consensus earnings trend predictor remains highly significant. As with the E/P predictor, however, the monotone regression gives stronger support than the least squares regression for significant relationships between returns and the controversy and neglect predictors. Neglect is significant at the 1% level, while controversy is significant at the 10% level.

The results from modeling stock returns using E/P and trend predictors provide evidence that the return-predictor relationship can be sensitive to predictor specification. Companies with flash E/Ps higher than their consensus E/Ps experience higher excess returns than companies having flash E/Ps lower than their consensus E/Ps. Specification of the earnings trend predictor, at least when the choice is between the two specifications considered here, does not seem to matter, however.[14]

The results also demonstrate that, while the general conclusions reached about consensus forecast E/P and trend and incremental flash effects are largely the same across different methods of estimation, the significance of controversy and neglect are sensitive to the estimation procedure employed. Unlike the least squares regressions, the monotone regressions provide support for significant inverse relationships between these predictors and subsequent return. That is, companies whose expected earnings estimates are more dispersed, and those with less analyst coverage, tended in this period to have lower excess returns.[15]

To summarize, specification of earnings predictor variables appears to matter, in the sense that alternative specifications of the same predictor can result in quite different investment decisions. Specification is not of the same importance to all predictors, however, nor to all types of analyses or estimation procedures. In particular, when we use the forecast E/P predictor to screen stocks for portfolio selection, it gives roughly equivalent results whether specified with consensus or flash data. The forecast trend predictor, however, can yield substantially different portfolios when specified with consensus rather than flash data.

In contrast, when we use the forecast E/P predictor to model returns, its relationship to subsequent returns differs markedly, depending on whether the predictor is constructed with consensus or with flash data. This holds true whether the relationship is estimated with least squares or monotone regression. Specification of the trend predictor is relatively less crucial in modeling returns, especially when limited to the two specifications analyzed here.[16]

PREDICTOR SPECIFICATION WITH MISSING VALUES

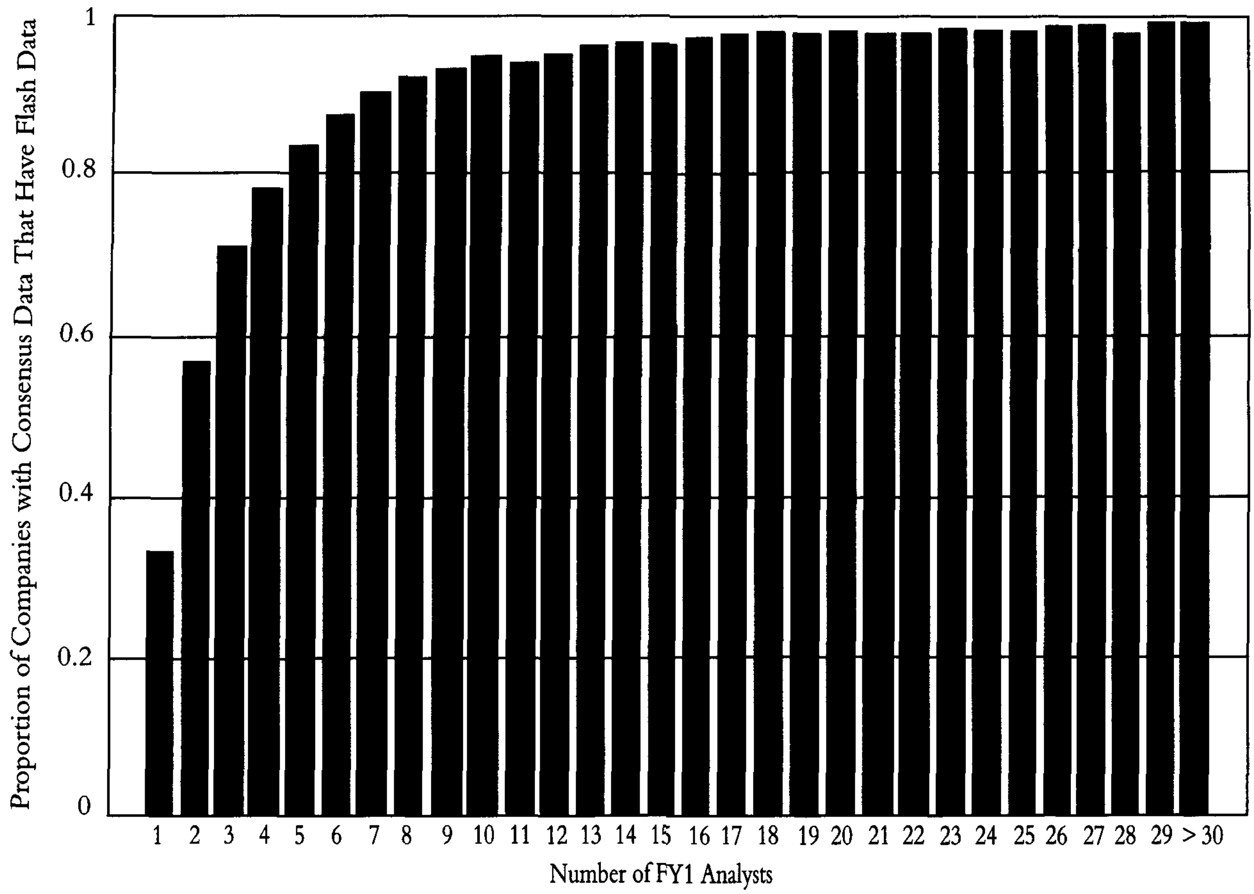
Besides having to choose among alternative specifications of a given predictor, one may face the problem of how to deal with missing data values. In the case of a universe of large-capitalization, widely followed stocks, this problem may not arise. For example, both consensus and flash data are available for all thirty of the Dow Jones industrials in Exhibits 1 and 2. For a broader universe of stocks and a greater number of predictor specifications, however, all information may not be available for every stock.

Exhibit 6 illustrates that the availability of flash earnings data may be limited, especially in the case of companies covered by only a few analysts. For the great majority of companies covered by nine or more analysts (over 90%), both consensus and flash data are available. As the number of analysts covering a stock declines, however, the percentage of companies with both consensus and flash data declines.[17]

What does one do when data are unavailable? One possible solution is to exclude companies with missing observations from the analysis. This could result, however, in a substantially reduced sample of companies for parameter estimation, especially if the model includes several variables with missing observations for different companies. In this situation, it may be worthwhile to consider other options.

One alternative is to impute estimated values to missing observations. One could, for example, assign some average value (e.g., the sector or industry average), or use the values from a comparable company or group of companies for which the data are available. In choosing among alternative treatments for missing values, the aim should be to arrive at the best possible estimates. The poorer the estimates, the greater the measurement error and the

EXHIBIT 6
ANALYST COVERAGE AND FLASH DATA AVAILABILITY
3,000-STOCK UNIVERSE — APRIL 1990-DECEMBER 1996



resulting bias in regression coefficients.

To get some idea of the impact on estimated returns of the treatment of missing observations, we examine the relationship between a six-week trend predictor and subsequent one-month returns, using all stocks in the 3,000-stock universe for which consensus data are available over the April 1990 through December 1996 period. We calculate flash predictors, using two methods of substituting for flash data when the data are unavailable or of questionable integrity.[18]

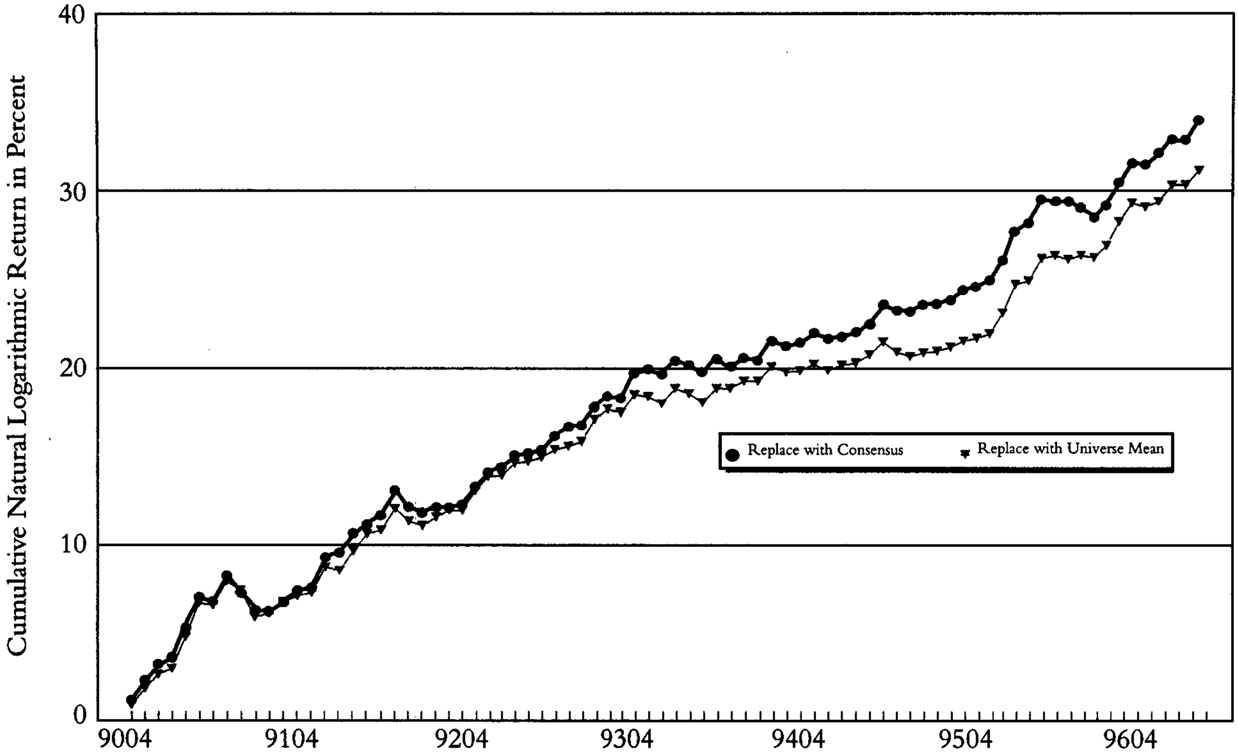
The first method uses the company's consensus data as a proxy for flash data. The second method uses the universe average flash. (The data are standardized, as before, with winsorization set at plus or minus 5 standard deviations from the mean.)

Use of consensus data when flash data are not available may result in less measurement error than use of the universe flash mean. One might expect to find a stronger relationship between returns and the first method of specification than between returns and the second method. Exhibit 7 illustrates the differences between the two methods, according to equal-weighted least squares regressions run over the period from April 1990 through December 1996.

Cumulative return under method one (flash/consensus data) is greater than cumulative return under method two (flash/universe average data), and the differential in favor of method one tends to grow over time. Under method one, cumulative return grows to nearly 34% over the period, compared with 30.5% under method two.[19] This finding is consistent with the notion that treatment of missing values via method two entails more measurement error than treatment via method one. Method two, in other words, ignores useful data that method one incorporates. As a result, the return-trend relationship is biased downward when trend is specified according to method two.

To help insure inclusion of the best estimates in computing predictors, one can employ a stepwise process that relies on the best data available for a given

EXHIBIT 7
TREATMENT OF MISSING FLASH TREND OBSERVATIONS AFFECTS RETURN*



*Simple least squares regression of one-month returns on fiscal year 1 forecast trend.

company at a given time. On the basis of theory, empirical evidence, experience with data, intuition, or other relevant considerations, one first determines the preferred ranking of all available data items. Specification of predictors then relies on this sequence.

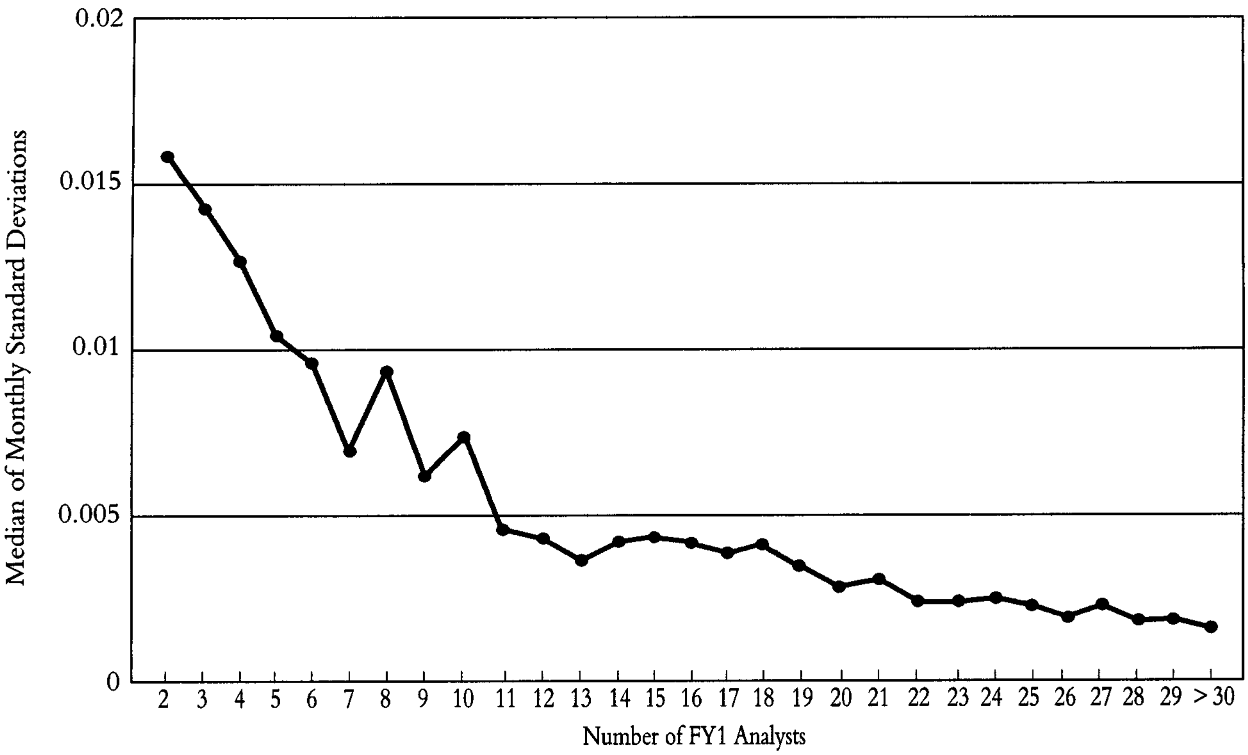
For example, if the most recent estimate is believed to be the most accurate, followed by the six-week flash mean, the consensus mean, the industry mean, and the universe mean, the value for a given predictor for each company would be calculated on the basis of this sequence, according to data availability. As data availability will vary across different companies, computed predictor values will be based upon different data items. But they will constitute the most accurate specifications available for each company at a given time.

PREDICTOR SPECIFICATION AND ANALYST COVERAGE

The issues that arise in earnings predictor specification become more complex as the number of analysts covering a stock increases. Consider a company followed by only one analyst. The consensus mean, the flash mean, and the most recent estimate collapse to the same value, and choice of predictor specification becomes a trivial issue. For companies with more than one analyst, however, the consensus mean will likely differ from the flash mean and the most recent estimate. As a result, different specifications will result in different predictor values. One is then faced with the problem of choosing among different specifications, or calculating the grand mean and dispersion across a number of specifications (and even here, one would need to determine whether all possible specifications are included).

Analyst coverage may also affect specification choice because it affects data availability. Exhibit 6, for example, shows that the availability of six-week flash data declines noticeably as the number of analysts falls below nine. The greater availability of data items for more widely followed companies opens the door to greater choice in predictor specification. There is some evidence to suggest, however, that predictor specification may be relatively less critical for widely followed companies. Exhibits 8 and 9

EXHIBIT 8
DIFFERENCE BETWEEN FLASH AND CONSENSUS FORECAST E/P DECLINES AS NUMBER OF ANALYSTS COVERING STOCK INCREASES — 3,000-STOCK UNIVERSE — APRIL 1990-DECEMBER 1996



illustrate why this may be the case.

Exhibit 8 shows the standard deviations of the differences between fiscal year 1 flash E/P and consensus E/P predictor values for the 3,000-stock universe over the April 1990-December 1996 period, stratified by the number of analysts following each stock. Exhibit 9 provides the standard deviations for the differences between fiscal year 1 flash and consensus trend predictor values, stratified by number of analysts. For both E/P and trend predictors, the difference between flash and consensus specifications declines noticeably as the number of analysts increases.

The decreased difference between the alternative specifications as analyst following increases reflects two factors. First, the number of analysts following a company tends to increase as the company's stock price increases; relatively higher prices in the denominators of the predictors for widely followed stocks tend to dilute the difference across earnings estimators in the numerators. Second, the range in the differences between six-week flash means and consensus means narrows as analyst coverage increases.[20]

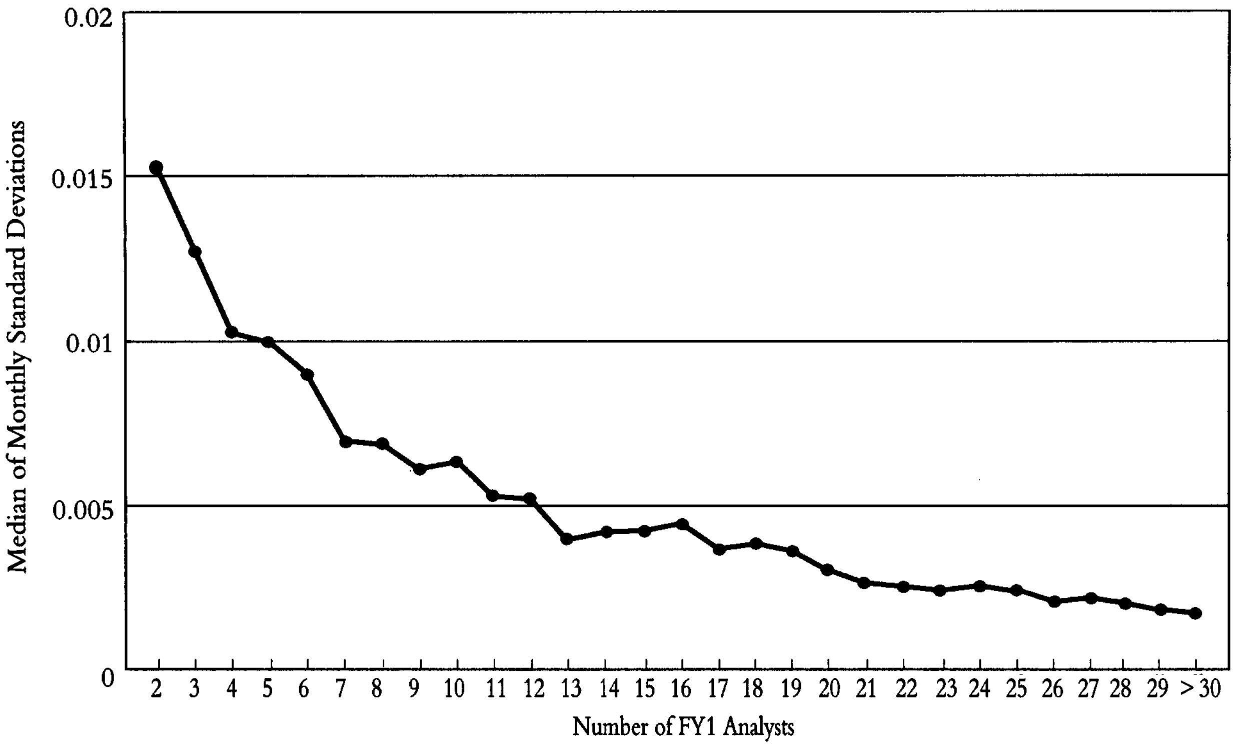
The implication is that the valuation of stocks with less analyst coverage may be more sensitive to predictor specification than the valuation of widely followed stocks. This is confirmed by an examination of differences in the rank orderings of the stocks, stratified by the number of analysts, between the two alternative E/P specifications.

For stocks followed by two to four analysts, rank order may change by as much as plus or minus 1,000 depending upon predictor specification. For stocks followed by twenty or more analysts, rank order changes tend to be much smaller, on the order of plus or minus 100.

THE RETURN-PREDICTOR RELATIONSHIP AND ANALYST COVERAGE

There are several reasons to believe that the relationship between an expectational predictor and returns is distributed differentially across a universe of stocks by the degree of analyst coverage. For one, investors may differentiate between widely followed

EXHIBIT 9
DIFFERENCE BETWEEN FLASH AND CONSENSUS FORECAST TREND DECLINES AS NUMBER OF ANALYSTS COVERING STOCK INCREASES — 3,000-STOCK UNIVERSE — APRIL 1990-DECEMBER 1996



and less widely followed companies when considering expectational earnings data. The more analysts covering a given stock, the greater may be their tendency to "herd" — that is, to tailor their earnings estimates so that they are in line with those of other analysts.

If such herding tendencies exist, then differences in estimates may tend to be small, and changes in earnings forecasts, rather than signaling an informative change in expected earnings distributions, may be more a reflection of analyst repositionings within the distribution of earnings estimates. In this case, investors may perceive earnings forecasts for widely followed companies as less meaningful than earnings forecasts for companies with a smaller analyst following, and the return-predictor relationship may be stronger for the latter than for the former.

There may, of course, be other reasons why the predictor-return relationship differs across companies with varying levels of analyst coverage. More widely followed companies, for example, may be priced more efficiently to begin with, or they may be less risky. Alternatively, the finding of such differences could reflect model or predictor misspecification.

To examine whether such distributed effects are present in our 3,000-stock universe, we fit a model over the period from April 1990 through December 1996:[21]

Return =

a + b (Predictor) +

c (Distributed Predictor for Large Coverage Stocks) +

d (Distributed Predictor for Small Coverage Stocks) +

e (Controversy) + f (Neglect) (4)

Here Return is excess return in the subsequent month. Predictors are calculated using consensus data for fiscal year 1. Controversy and Neglect are as defined previously.

The Distributed Predictor for Large Coverage Stocks is the marginal E/P or trend effect for companies having more than ten fiscal year 1 analysts. The Distributed Predictor for Small Coverage

Stocks is the marginal E/P or trend for companies having from one to four fiscal year 1 analysts.

The coefficient b represents the estimated predictor effect for companies covered by five to ten analysts. The estimated predictor effect for companies with more than ten analysts is computed by adding the coefficient c (the distributed effect for large coverage companies) to coefficient b. The predictor effect for companies with one to four analysts is computed by adding b and d (the distributed effect for small coverage companies). Thus c and d represent marginal effects relative to the base case group of companies followed by five to ten analysts. In addition, the t-values for c and d indicate whether the marginal effects are significantly different from the base case.

The break points for analyst coverage are chosen because they split the universe of stocks approximately into thirds. They also consistently partition the universe into thirds over the entire period of study, so modification of break points over time was not necessary. Other models were examined, including one based on an interaction effect between the number of analysts and earnings trend. The results are consistent with those reported for the model based upon partition into thirds.

If the distribution of the forecast predictors differs across the three stock groups, we would expect to see c and d coefficients significantly different from zero. In particular, if investors tend to view expectational data for widely followed stocks as less meaningful than the data for less widely followed stocks, we would expect the coefficient c to be negative and significant.

Exhibit 10 reports the least squares and monotone rank regression results for the fiscal year 1 consensus trend predictor. The evidence here, with either estimation procedure, suggests that the relationship between trend and subsequent monthly returns is distributed differentially, depending on the level of analyst coverage. Specifically, the relationship is weaker (i.e., less positive) for stocks with more than ten analysts than it is for stocks with five to ten analysts.

The least squares regression shows that, for those companies followed by five to ten analysts, a 1 standard deviation increase in the trend predictor is associated with an increase in return of about 44 basis points. This value is highly significant, with a p-value of zero.[22]

For companies followed by one to four analysts, a 1 standard deviation increase in the trend predictor is associated with a slightly smaller increase in return — around 41 basis points (0.4438 – 0.0340) — but the level of significance is such that we may conclude that the trend-return relationship does not differ between stocks followed by one to four analysts and those followed by five to ten analysts. For companies followed by more than ten analysts, however, a 1 standard deviation increase in the trend predictor results in a return increase of about 20 basis

EXHIBIT 10
EFFECTS OF FORECAST TREND BY NUMBER OF ANALYSTS
REGRESSION RESULTS USING CONSENSUS DATA FOR 3,000-STOCK UNIVERSE — APRIL 1990-DECEMBER 1996

	OVERALL TREND	DISTRIBUTED TREND EFFECT (> 10 ANALYSTS)	DISTRIBUTED TREND EFFECT (1-4 ANALYSTS)	CONTROVERSY	NEGLECT
		DISTRIBUTED LEAST SQUARES REGRESSION			
Mean	0.4438	–0.2423	–0.0340	–0.1396	–0.1741
Standard Error Mean	0.0990	0.1086	0.0863	0.0847	0.0997
T-Statistic	4.4821	–2.2316	–0.3939	–1.6480	–1.7459
P-Value	0.0000	0.0284	0.6947	0.1033	0.0847
		DISTRIBUTED MONOTONE REGRESSION			
Mean	0.0513	–0.0203	–0.0011	–0.0144	–0.0359
Standard Error Mean	0.0070	0.0068	0.0042	0.0062	0.0093
T-Statistic	7.2833	–2.9687	–0.2710	–2.3217	–3.8634
P-Value	0.0000	0.0039	0.7871	0.0228	0.0002

points (0.4438 − 0.2423) — significantly less than the increase observed for the other groups.

As in the previous least squares regression, we find that the controversy and neglect predictors are inversely related to subsequent returns, but only marginally (with p-values near or slightly below 0.10).

The monotone regression results in Exhibit 10 also indicate a significantly different and less positive trend-return relationship for widely followed stocks. The estimated effect for these companies is significantly smaller than the effects for the small coverage and medium coverage groups; furthermore, the monotone regression assigns even higher significance to the difference than the least squares regression does, with a p-value of 0.0039. Again, the trend effect does not appear to differ significantly between stocks with five to ten analysts covering them and stocks with one to four analysts.

Once again, the monotone regression gives stronger support than the least squares regression for the controversy and neglect predictors. Both have p-values well below 0.05, suggesting a significant inverse monotonic relationship between these predictors and subsequent returns.[23]

Exhibit 11 reports the least squares and monotone regression results for forecast E/P. The evidence here is somewhat mixed.

The least squares estimation yields no strong evidence of a significant distributed effect. The monotone regression, however, indicates a significant (p-value of 0.0241) distributed effect of −0.0106 for widely followed stocks. The E/P-return relationship for this group is significantly less positive than the E/P-return relationship for the median coverage group. The distributed effect for the small coverage group remains insignificant. And, once again, the monotone regression results in a finding of significance for the controversy and neglect predictors, with increases in either associated with decreases in subsequent return.

The overall results are thus unclear. Distributed effects appear to exist for the trend predictor specified with consensus data, regardless of the estimation procedure. For the E/P predictor, however, distributed effects show up only in monotone estimation. Given that the least squares estimators may be more sensitive to leverage points and outliers, one may want to place more reliance on the monotone results.[24]

Might one find distributed effects across other possible E/P and trend specifications? Across other expectational earnings predictors — say, forecast growth rates or earnings surprise? Might inclusion in the return model of different explanatory variables affect the results? Are distributed effects robust across other statistical paradigms? Do they appear in other investment strategies with different investment horizons? These are all important questions in the search for return opportunities.

EXHIBIT 11
EFFECTS OF FORECAST E/P BY NUMBER OF ANALYSTS
REGRESSION RESULTS USING CONSENSUS DATA FOR 3,000-STOCK UNIVERSE — APRIL 1990-DECEMBER 1996

	OVERALL TREND	DISTRIBUTED E/P EFFECT (> 10 ANALYSTS)	DISTRIBUTED E/P EFFECT (1-4 ANALYSTS)	CONTROVERSY	NEGLECT
		DISTRIBUTED LEAST SQUARES REGRESSION			
Mean	0.3720	−0.1191	−0.0157	−0.1231	−0.1573
Standard Error Mean	0.1457	0.1091	0.1024	0.0947	0.0961
T-Statistic	2.5540	−1.0917	−0.1535	−1.2995	−1.6365
P-Value	0.0126	0.2782	0.8784	0.1975	0.1057
		DISTRIBUTED MONOTONE REGRESSION			
Mean	0.0446	−0.0106	−0.0001	−0.0209	−0.0285
Standard Error Mean	0.0098	0.0046	0.0051	0.0071	0.0090
T-Statistic	4.5619	−2.2996	−0.0254	−2.9603	−3.1791
P-Value	0.0000	0.0241	0.9798	0.0040	0.0021

SUMMARY

Our examination of predictor specification indicates that specification can play an important role in model building. When a number of alternative specifications are possible, different specifications of the same predictor may not result in the same portfolio compositions or be related to stock returns in the same way. The choice of expectational data with which to specify a given predictor (including the selection of data to fill in gaps in data availability) thus has the potential to introduce noise and measurement error into investment decision-making.

The importance of specification choice may vary depending upon the predictor, the investment strategy, the estimation procedure used, and the number of analysts following a stock. In general, however, decisions regarding predictor specification have the potential to influence the results of empirical analyses. This is true not only for research based upon traditional methods of statistical analysis, but also for new wave techniques such as genetic algorithms and neural nets, which also require the specification of predictors or inputs. Finally, although we have focused on expectational earnings data for individual firms, our findings also have relevance for predictors based upon fundamental and technical data, as well as aggregate data for industries, sectors, and the overall market.

APPENDIX
THE EFFECT OF MEASUREMENT ERROR ON REGRESSION COEFFICIENTS

Theoretically, the effect of measurement error on the estimated coefficient in the simple linear regression model may be seen as follows. Suppose the appropriate model is:

$$\text{Return} = a + b\ (\text{Flash Predictor}) + e$$

Instead, however, we use the model:

$$\text{Return} = a + b\ (\text{Consensus Predictor}) + e$$

with Consensus Predictor = Flash Predictor + u. That is, the consensus predictor is an imperfect proxy for the appropriate flash predictor.

Linear regression calculates the coefficient as:

$$\hat{b} = \frac{\text{Covariance(Return, Consensus Predictor)}}{\text{Variance(Consensus Predictor)}}$$

or

$$\hat{b} = \frac{\text{Covariance(Return, Flash Predictor} + \text{u)}}{\text{Variance(Flash Predictor} + \text{u)}}$$

or

$$\hat{b} = \frac{\text{Covariance[a} + \text{b(Flash Predictor)} + \text{e, Flash Predictor} + \text{u]}}{\text{Variance(Flash Predictor} + \text{u)}}$$

or

$$\hat{b} = \frac{\text{b} \times \text{Variance(Flash Predictor)}}{\text{Variance(Flash Predictor)} + \text{Variance(u)}}$$

assuming that the measurement error, u, and residual model error, e, are independent and that measurement error is not a function of the flash predictor.

Alternatively, one can write the above as:

$$\hat{b} = \frac{\text{b}}{1 + [\text{Variance(u)/Variance(Flash Predictor)}]}$$

This suggests that the estimated coefficient $\hat{b}$ has a bias toward zero that depends upon the variance in the measurement error relative to the variance in the flash predictor. Other things equal, the greater the variance in the measurement error, the more biased the estimate of $\hat{b}$. Why? Because the consensus provides a more noisy estimate of flash earnings as measurement error increases.

For models with more than one explanatory variable, the effect of measurement error becomes a bit more complex. It will depend on, among other things, the number of predictors with measurement error, the correlations between predictors, the correlations between measurement errors, and the signs of the regression coefficients. For the special case where predictors are not correlated and not related to measurement errors, and measurement errors across predictors are independent, the regression coefficients will be biased toward zero. As predictors become more highly correlated, however, any bias will depend on the signs of the regression coefficients and the variance of the measurement errors relative to the variance of the predictors, other things equal.[25]

ENDNOTES

The authors thank Judith Kimball for her editorial assistance.

[1]One can choose from a variety of data vendors as well. For this study we use IBES data.

[2]There are a variety of options regarding the estimates to include, as well as their weights. A dynamic weighting strategy, for example, would weight more recent estimates more heavily than older estimates.

[3]Trimmed means remove a proportion of the most extreme observations from a data set and compute the mean of the remaining observations. This procedure reduces the influence of outliers.

[4]The different possible estimates represent proxies for a company's earnings. For any given company, one would like to use the proxies providing the best estimate of current and future earnings. Changes in technology and in the way analysts update estimates may also influence the estimates selected.

[5]The Pearson correlation estimates the linear association between the actual values of the predictors. The Spearman rank cor-

relation measures the association between the ranks, not the actual values. Because variables may be monotonically related, but in a highly non-linear way, the Spearman rank can capture information the Pearson cannot.

[6]The 3,000-stock universe consists of the approximately 3,000 most liquid U.S. stocks having IBES coverage.

[7]We find larger differences in the tails of the rankings, especially for the most highly ranked companies. The companies with the greatest differences tended to have limited analyst following.

[8]The effects of these differences may also depend on investment strategy. For example, a long-only manager using flash data would hold Aluminum Company of America and not Woolworth, while a long-only manager using consensus data would hold Woolworth and not Aluminum. If the two were long-short managers, however, one would hold Aluminum long and Woolworth short, while the other held Woolworth long and Aluminum short.

[9]See Jacobs and Levy [1988] for discussion of the benefits of disentangling related effects.

[10]The universe is updated regularly in order to reflect changes over the period of study. Eighty-one monthly cross-sectional regressions are run for each model. Parameter estimates for each month are unrestricted and allowed to vary from month to month.

[11]Monotone regression is based upon ranks. See Conover [1980] and Iman and Conover [1979]. We also ran robust regressions based on an iterative weighting procedure. Both Huber [1964, 1981] and Beaton-Tukey [1974] bi-weights are used. In general, these procedures reduce the influence of outliers on regression estimates. The results from these procedures are not presented here because they do not affect the conclusions of our analysis. For a more general discussion of alternative robust regression methods, see Rousseeuw and Leroy [1987].

[12]The p-value is the smallest level of significance for which the null hypothesis can be rejected. Reporting the p-value gives others the opportunity to determine how sensitive a hypothesis test is to changes in the significance level. For example, two test statistics, one with a p-value of 0.045 and another with a p-value of 0.0001, are both significant at the 5% level, but the conclusions based upon the former would be much more sensitive to changes in the level of significance.

[13]Monotone regression uses the ranks of both the dependent and explanatory variables. The inverse rank transform may be used to determine actual values for predictors stated in terms of ranks. In essence, monotone rank regression is to linear regression as Spearman rank correlation is to Pearson correlation.

[14]We also undertook tests with specifications using shorter than six-week flash horizons, which show significant incremental effects (and differences in portfolio composition) between flash and consensus specifications.

[15]One might expect firms about which there is more controversy to have lower returns, in the absence of short-selling, because the wider range of earnings forecasts tends to lead to higher prices, and lower subsequent returns. With regard to neglect, however, one might expect to find, a priori, a positive rather than a negative relationship (i.e., a small-firm effect). Over the period of study, however, large-cap stocks tended to outperform small-cap stocks. As analyst coverage is positively correlated with market capitalization, it is likely that our neglect predictor is capturing this return differential between large- and small-cap stocks.

[16]The finding of a significant incremental flash effect for E/P may seem surprising, given the high correlation between the two specifications. Results from a semiparametric model (relaxing the linearity assumption for the consensus and incremental flash effects) suggest that the return to the flash predictor is significantly higher in the positive tail, other things equal.

[17]The proportion of individual analysts revising forecasts appears to be independent of the level of coverage. This is true for stocks covered by one or numerous analysts. On average, each analyst tends to revise estimates about one-third of the time.

[18]Data were run through a set of integrity checks. If data looked questionable, for whatever reason, they were not used.

[19]Both trend predictors are positive and significant. We tested for an incremental effect of the difference between consensus and flash universe average controlling for the availability of flash data, and found statistical significance at the 1% level in both least squares and monotone regressions.

[20]Note that this phenomenon is not due to analysts making more frequent revisions for well-followed stocks.

[21]One might be concerned about possible collinearity and its impact on estimator precision for the neglect (analyst coverage) and the distributed effects predictors. We examined the degree of collinearity present in our models using the singular value decomposition and condition indexes proposed by Belsley, Kuh, and Welsch [1980]. In general, we find no evidence to suggest that collinearity is seriously degrading our estimates. Nor do we find that collinearity is changing (increasing) over time.

[22]A significance level of less than 0.05% would be needed not to reject the null hypothesis that the trend predictor, on average, is not significantly different from zero. This is one-twentieth the level set for conservative tests (where the significance level is set at 1%). Thus there is strong evidence suggesting that trend and returns are directly related.

[23]As earlier regression results show, these predictors tend to be negative but highly insignificant in the least squares estimation. They appear to be monotonically related to returns but not linearly related on the basis of the raw data. A primary reason for this finding is the existence of influential leverage points for these two predictors in the least squares model. Leverage points exert undue influence on unbounded influence estimators, such as least squares, and hence have a significant effect on regression coefficients. Use of robust regression procedures, such as least median squares or robust regression with Beaton-Tukey bi-weights, reduces the influence of such observations on regression coefficients. Application of these procedures to our data results in findings similar to those for the monotone regression.

[24]Use of other robust methods also results in contradictory findings regarding the significance of distributed effects in the E/P model, however. The iterative Beaton-Tukey procedure, for example, finds no distributed effects, while L1 (least absolute value) regression finds a significant negative distributed effect for large coverage stocks. Interestingly, the distributed-effects E/P model is the only one where alternative estimation procedures give conflicting results; in all other cases, the results from the alternative robust estimation procedures are consistent with those from monotone rank regressions.

[25]See Maddala [1977], Levi [1973], and Theil [1961, 1971].

REFERENCES

Beaton, A.W., and J.W. Tukey. "The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data." *Technometrics*, May 1974.

Belsley, D.A., E. Kuh, and R.E. Welsch. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley, 1980.

Conover, W.J. *Practical Nonparametric Statistics*. New York: John Wiley, 1980.

Huber, P.J. "Robust Estimation of a Location Parameter." *Annals of Mathematical Statistics*, March 1964.

——. *Robust Statistics*. New York: John Wiley, 1981.

Iman, R.L., and W.J. Conover. "The Use of the Rank Transform in Regression." *Technometrics*, November 1979.

Jacobs, B.J., and K.N. Levy. "Disentangling Equity Return Regularities: New Insights and Investment Opportunities." *Financial Analysts Journal*, May/June 1988.

Levi, M.D. "Errors in the Variables Bias in the Presence of Correctly Measured Variables." *Econometrics*, September 1973.

Maddala, G.S. *Econometrics*. New York: McGraw-Hill, 1977.

Rousseeuw, P.J., and A.M. Leroy. *Robust Regression and Outlier Detection*. New York: John Wiley, 1987.

Theil, H. *Economic Forecasts and Policy*, 2nd ed. Amsterdam: North Holland, 1961.

——. *Principles of Econometrics*. New York: John Wiley, 1971.