THE ACCOUNTING REVIEW
Vol. 83, No. 4
2008
pp. 1101–1124

American Accounting Association
DOI: 10.2308/accr.2008.83.4.1101

Extracting Forward-Looking Information from Security Prices: A New Approach

Dan Weiss
Tel Aviv University

Prasad A. Naik
University of California, Davis

Chih-Ling Tsai
University of California, Davis

ABSTRACT: This paper proposes a new index to extract forward-looking information from security prices and infer market participants' expectations of future earnings. The index, called market-adapted earnings (MAE), utilizes stock returns and fundamental accounting signals to estimate market expectations of future earnings at the firm level. MAE outperforms time-series models (e.g., random-walk) in predicting future earnings. Results demonstrate the usefulness of MAE for firms that have no analyst following.

Keywords: *single-index-model; market-adapted earnings*.

I. INTRODUCTION

The prices-lead-earnings stream of research suggests that security prices reflect a richer information set than historical earnings because prices reflect all information available to market participants (Beaver et al. 1980, hereafter BLM). Several studies examine the information content of security prices with respect to future earnings (e.g., BLM; Collins et al. 1987; Ayers and Freeman 2000). However, as Kothari (2001, 148) argues, "researchers have found it difficult to harness the information in prices at the firm level to make an economically important improvement." This challenge motivates us to develop an improved proxy for the market expectations of future earnings.

To construct an improved proxy, we employ a new method to extract information from security prices at the firm level. Similarly to BLM, we utilize information in security prices to infer expected future earnings. However, our model differs from BLM in two ways: we allow for (1) a nonlinear returns-earnings relationship and (2) multiple accounting signals. Because of the

The authors are grateful for valuable discussions with and constructive suggestions from Gil Aharoni, Joseph Aharony, Yakov Amihud, Eli Amir, Brad Barber, Uri BenZion, Amihud Dotan, Haim Falk, Dan Givoly, Paul Griffin, Itay Kama, Lakshmanan Shivakumar, and Avi Wohl. The authors thank the editor, Patricia Dechow, and two anonymous referees for their insightful comments that enriched the paper. Comments from the participants of the Tel Aviv Conference in Accounting are highly appreciated. The authors thank Michael Bowers for outstanding research assistance.

Editor's note: This paper was accepted by Patricia N. Dechow.

Submitted: February 2006
Accepted: December 2007
Published Online: July 2008

nonlinearity and the multivariate setting, neither BLM's grouping technique nor reverse regression (e.g., Beaver et al. 1987) can be used. As a result, we apply the recently developed Single-Index Model (SIM), which is a semiparametric model (see, e.g., Horowitz 1998) that provides an approach for efficient estimation in a multi-dimensional environment without assuming a linear returns-earnings relationship.

Using SIM, we construct an index, called *market-adapted earnings* (MAE), which incorporates information from stock returns together with multiple accounting signals to estimate market expectations of future earnings change. Unlike the equal-weight index proposed by Lev and Thiagarajan (1993, 210), MAE places unequal market-based weights on fundamental accounting signals. Hence, we expect MAE to provide better predictions, especially because it also incorporates information from both stock returns and fundamental accounting signals through a nonlinear link function. To this end, we use MAE to forecast one-year-ahead earnings change and examine whether it improves the prediction of future earnings.

Based on a sample of 636 firms from 1992–2001, we find that MAE results in more accurate earnings forecasts than either the random-walk model or an accounting-based model (Abarbanell and Bushee 1997; Lev and Thiagarajan 1993). Specifically, the median absolute forecast error of MAE is 20.9 percent (17.0 percent) lower than the median absolute forecast error of the random-walk (accounting-based) model. In addition, the prediction accuracy of MAE over the time-series models increases as the information environment becomes richer (i.e., the number of analysts covering a firm increases).

How do MAE predictions compare to analysts' forecasts? We find that the median absolute forecast error of analysts' forecasts is 21.0 percent lower than that of MAE, which is expected because analysts utilize a larger information set than MAE (e.g., Fried and Givoly 1982; O'Brien 1988). Furthermore, the analysts' advantage over MAE increases when more information is available. However, financial analysts do not supply one-year-ahead earnings forecasts for all firms; only less than half of the Compustat universe is covered by the Institutional Brokers' Estimate System (I/B/E/S).

For firms with no analyst coverage, the median absolute forecast error for MAE is 14.4 percent (12.0 percent) lower than that for random-walk (accounting-based) model. Furthermore, MAE predicts 55.2 percent (53.6 percent) observations more accurately than the random-walk (accounting-based) model. Thus, MAE outperforms both the random-walk and accounting-based models. The above findings demonstrate the usefulness of MAE for an accurate earnings prediction when no analyst following is available.

In sum, this study contributes to the literature in three ways. First, we show that researchers can benefit from improved estimation of market expectations of future earnings for the sizable class of firms without analyst coverage by extracting forward-looking information from stock prices. Second, we introduce the SIM methodology that is shown to be a useful tool for extracting information from security prices. Third, we demonstrate that the accuracy of earnings prediction increases as the information set is more comprehensive.

The paper continues as follows. Section II reviews the relevant literature. Section III presents our research design, describes the SIM methodology, and displays results from SIM estimations. Section IV compares the prediction accuracy of market-adapted earnings with analysts' forecast as well as both the random-walk and accounting-based models. Section V demonstrates the benefit of MAE for firms without analyst coverage. Section VI concludes.

II. LITERATURE REVIEW

Prices-Lead-Earnings

The concept of prices leading earnings was introduced by BLM, which is based on the premise that security prices reflect a richer information set than do the time-series of earnings.[1] Insofar as current price is the capitalized present value of a firm's expected future earnings, security prices contain useful information for the prediction of future earnings growth. BLM assume that earnings follow a compound stochastic process, and they infer market participants' expectations regarding future earnings from observed changes in security prices. They assess the distribution of future earnings conditionally on known prices and earnings, and they expand the information set underlying inferred earnings expectations from known earnings to known prices and earnings. Furthermore, BLM provide evidence that price-based forecasts are slightly more accurate than the random-walk model with a drift.[2]

BLM use a grouping method, which diversifies away the measurement error in earnings. Because grouping entails loss of estimation efficiency, Beaver et al. (1987) introduce a reverse regression to exploit information contained in prices. Notably, the reverse regression places the measurement error in the disturbance term, rather than in the explanatory variable.[3] Using another approach involving simultaneous equations, Beaver et al. (1997) provide evidence that earnings and prices are endogenously determined.

To examine the availability of rich information as a driver for prices leading earnings, Freeman (1987) shows that the security prices of large firms anticipate accounting earnings earlier than those of small firms. In the same line, Collins et al. (1987) use firm size as a proxy for the richness of the information environment and examine its relationship to the predictive power of price-based earnings forecasts. They find that price-based earnings forecasts reduce forecast error when compared with time-series models for large firms, but increase forecast error for small firms. Later, Ayers and Freeman (2000, 2003) indicate that large firms' prices also have a longer lead than small firms' prices with respect to industry-wide earnings, and they provide evidence that price leads increase as analysts' coverage and institutional ownership increase.[4] Moreover, Kothari and Sloan (1992) provide evidence that price-based earnings forecasts are informative in long-horizon settings.

Table 1 summarizes pertinent dimensions of the above studies. In sum, this line of research indicates that using the prices-lead-earnings framework results in a marginal predictive advantage over random-walk (with or without drift) for predicting one-year-ahead earnings, which raises questions regarding our ability to exploit information incorporated in prices. To address this problem, we propose a new methodology for utilizing information incorporated in security prices and focus on retrieving forward-looking information with respect to one-year-ahead earnings.

Fundamental Analysis

The second line of research linked to this study is the fundamental analysis literature, which uses financial ratios as leading indicators of future performance. This line of research exploits

---

1 See Brown (1993) for a literature review on time-series models for earnings forecasting.

2 BLM report that the mean absolute error of price-based forecasts is 1.8 percent lower than the mean absolute error of a random-walk with drift forecasting model (see BLM, Table 4).

3 An immediate benefit of reverse regression is an unbiased slope coefficient. Cready et al. (2000) further analyze the biases in estimated coefficients when reverse regression is used in a multi-interacted variable setting. See also Kothari (1992).

4 Ayers and Freeman (2003) and Jiambalvo et al. (2002) report that institutional ownership enhances prices leading earnings. Institutional ownership is characterized as professional and sophisticated, but does not necessarily imply a rich information environment.

TABLE 1

Comparative Summary of Literature Review

	Prices-Lead-Earnings			Fundamental Analysis			
Paper	Beaver et al. (1980)	Beaver et al. (1987)	Collins et al. (1987)	Ou and Penman (1989b)	Lev and Thiagarajan (1993)	Abarbanell and Bushee (1997)	This paper
Goal	"We infer market participants' expectations regarding future earnings from observed changes in security prices"	Suggest reverse regression "to assess the information content of security prices with respect to accounting earnings."	"focusing on firm size and its relation to the predictive accuracy of price-based earnings forecasts."	Compare the ability of security prices and accounting signals to predict future earnings.	"identify…a set of financial variables (fundamentals) claimed by analysts to be useful in security valuation."	"We investigate how detailed financial statement data …enter the decisions of market participants by examining whether… signals are informative about subsequent earnings changes."	We examine whether information incorporated in stock returns and accounting signals, when taken together, can improve the prediction of future earnings (over existing models and analysts' earnings forecasts).
Measure used for earnings prediction[a]	$Exp[E_{t+j} \mid E_t, P_t]$	$Exp[E_{t+j} \mid P_t]$	$Exp[E_{t+j} \mid P_t]$	$Exp[Sign\ E_{t+j} \mid A_t]$, $Exp[E_{t+j} \mid E_t, P_t]$	$Exp[Sign\ E_{t+j} \mid A_t]$	$Exp[E_{t+j} \mid A_t]$ and $Exp[SignE_{t+j} \mid A_t]$	$Exp[MAE_t \mid P_t, A_t]$

(continued on next page)

TABLE 1 (continued)

	Prices-Lead-Earnings			Fundamental Analysis			
Method	OLS + Grouping data	Reverse Regression	Reverse Regression	Logit model used to generate an index (called Pr)	Score based on good or bad news in accounting signals	OLS	Single-index model (Semiparametric)
Model	Linear	Linear	Linear	Linear	Linear	Linear	Nonlinear
Independent signals	Earnings only	Security prices	Security prices	18 fundamental signals	Earnings and 12 fundamental signals	Earnings and nine fundamental signals	Stock returns, earnings, and four fundamental signals
Earnings prediction	"This approach can potentially lead to earnings forecasting models that are more accurate than the random-walk with a drift."	"Provides the basis for forecasting earnings based upon current and past values of price changes."	"Price-based earnings will outperform univariate time-series forecasts by a greater margin for larger firms."	Information in prices that leads future earnings is contained in financial statements.	"relating the fundamental-based quality scores to subsequent earnings changes."	"We find that fundamental signals have incremental explanatory power relative to current-year earnings."	SIM outperforms random-walk and accounting-based models in accurately predicting earnings for firms with no analyst coverage.

[a] *P* = security prices; *E* = reported earnings; *A* = accounting signals reported in financial statements.

accounting information reported in financial statements to predict future earnings and identify mispriced securities. Ou and Penman (1989a) combine a large set of accounting signals into an index that expresses the probability of a one-year-ahead earnings increase based on the signals. Ou (1990) employs this index to provide evidence that investors use non-earnings accounting variables to revise their expectations of future earnings. Ou and Penman (1989b) also compare the ability of security prices and accounting signals to predict future earnings, arguing that information in prices that leads future earnings is contained in financial statements. In this study, we examine a related aspect of Ou and Penman's claim using a new methodology that utilizes information used by market participants in forming their expectations on future earnings.

In addition to the above studies, Lev and Thiagarajan (1993) report evidence on the value-relevance of 12 accounting signals and demonstrate that an equally weighted index of the ratios is informative in assessing earnings persistence and change. Later, Abarbanell and Bushee (1997, 5) use accounting signals suggested by Lev and Thiagarajan (1993) in a linear earnings forecasting model and show an incremental improvement in explanatory power over the current-year reported earnings change.

Although the fundamental analysis literature (see Table 1) generally supports the ability of fundamental ratios to forecast future earnings, the lure of above-normal returns based on fundamental analysis demonstrates that the information contained in these signals is not fully incorporated into prices (e.g., Ou and Penman 1989a; Abarbanell and Bushee 1998; Rajgopal et al. 2003). Hence, several studies employ ratios that combine information from both security prices and earnings; for instance, Penman (1998), and Fama and French (2000) use price-earnings or price-to-book ratios to predict earnings. In this paper, we examine whether the information incorporated in stock returns and accounting signals, when taken together, can improve the prediction of future earnings.

III. RESEARCH DESIGN

This study utilizes information incorporated in stock returns and accounting signals for estimating earnings components that are expected to persist. The logic of our model stems from BLM (1980, 4), who characterize future earnings as resulting from a combination of two processes:

(1) Value-relevant earnings: a process reflecting the impact on earnings of events that affect security prices, i.e., have a valuation implication.
(2) Noise in earnings: a process capturing value-irrelevant events resulting in an earnings component with no effect on security prices, i.e., uncorrelated with stock prices or returns. This process includes earnings components that are not expected to persist.

To show that reported earnings do not follow a random-walk process, BLM use the information incorporated in security prices. They introduce a first-order moving average of the first differences in earnings and predict a linear relationship between the percentage change in price and the percentage change in earnings. Our basic motivation is similar to BLM (1980, 23): "viewing earnings as a compound process … provides a basis for forecasting earnings." However, our prediction model differs from BLM since we not only relax the linearity assumption and allow for a nonlinear returns-earnings relationship, but also incorporate multiple accounting signals.

In the context of nonlinearity, Freeman and Tse (1992), Cheng et al. (1992), Das and Lev (1994), and Beneish and Harvey (1998) report on the S-shaped returns-earnings relationship. The economic rationale underlying the S-shape is that the market does not expect extreme earnings changes to be permanent, and the S-shape is interpreted as reflecting small abnormal returns to extreme earnings changes. Das and Lev (1994) adopt a nonparametric estimation procedure, locally weighted regression, to demonstrate the superior prediction power of out-of-sample stock returns over a linear specification. Following earlier research, they suggest differential earnings persistence as a source of nonlinearity.

In the context of multiple accounting signals, Lev and Thiagarajan (1993) report significant associations of multiple fundamental signals with (1) stock returns, and (2) the expected direction of future earnings change. Keeping prediction in mind, these associations may impede a linear regression of future earnings changes on a set of current fundamental accounting variables and stock returns due to multicollinearity.

In sum, Das and Lev (1994) use earnings as the only explanatory variable (at the exclusion of supplementary signals[5]) in a nonparametric model to study nonlinearity, whereas the fundamental analysis studies formulate linear parametric models to incorporate multiple signals (at the exclusion of nonlinearity). To combine both the nonlinearity and multiple signals in a common framework, we propose semiparametric SIM to predict future earnings.

Single-Index Model (SIM)

SIM is a semiparametric model for estimating a conditional mean function in a multi-signal setting.[6] Let R_{it} be stock return of firm i at year t, and let S_{it} be a $1 \times k$ random vector of realized accounting signals ($k = 5$, reported earnings change and four supplementary fundamental accounting signals). In SIM, the conditional mean function is $E[R_{it} \mid S_{it}] = G(S_{it}\beta_t)$, where β_t is an unknown $k \times 1$ constant vector, $\beta_t = (\beta_{0t}, \beta_{1t}, \beta_{2t}, \beta_{3t}, \beta_{4t})'$ in our case, and $G(.)$ is an unknown link function, whose specific functional form is determined by the data, rather than being assumed by researchers. For the sake of identification, we let $\beta_{0t} = 1$ (see Horowitz 1998, 15), which does not affect the prediction (more on this issue in the following subsection). In addition, $G(.)$ is assumed to be a differentiable function on the support of $S\beta$, where the scalar index $S\beta$ is the market-adapted earnings, *MAE*. Thus the single-index model is given by the cross-sectional regression model:

$$R_{it} = G_t(MAE_{it}) + \varepsilon_{it}, \tag{1a}$$

where MAE_{it} is given by:

$$MAE_{it} = \Delta E_{it} + \Sigma_j S_{jit}\beta_{jt}, \quad j = 1,\ldots,4. \tag{1b}$$

Using stock returns, earnings change and four supplementary accounting signals, we estimate MAE_{it} to obtain the relative weights β_{jt}. Consequently, the estimated MAE_{it} combines the earnings change and the four supplementary accounting signals with their corresponding estimated $\hat{\beta}_{jt}$ as follows:

$$M\hat{A}E_{it} = \Delta E_{it} + \sum_{j=1}^{4} S_{jit}\hat{\beta}_{jt}. \tag{2}$$

The form of MAE_{it} is similar to Abarbanell and Bushee's (1997, 5) earnings forecast model in its linear structure. Linear indexes for earnings predictions have been widely employed in the fundamental analysis literature (e.g., Lev and Thiagarajan 1993; Penman 1998). The function $G(.)$, as well as the vector β, are estimated endogenously by using information on both stock returns and

[5] Das and Lev (1994, 361) use a random sample of 622 observations for a nonparametric investigation of nonlinearity in the returns-earnings relation because "it is very computer intensive to estimate a total of 6,220 firm-year observations pooled across all ten years."

[6] See Horowitz (1998) for an in-depth discussion of single-index models. Except for Ait-Sahalia and Brandt (2001), who applied it in the finance literature, we believe this study marks the first application of SIM in the accounting literature.

fundamental accounting signals. In other words, the information content of stock returns is used to estimate the link function $G(.)$ and the market-based β-coefficients.

We achieve an efficient estimation of the index MAE_{it} *without* assuming a linear (or even pre-specifying a nonlinear) relationship between earnings and stock returns. The estimated $\hat{\beta}_{jt}$ are obtained from Li's (1991) Sliced Inverse Regression method as described in Naik and Tsai (2001). It enables us to extract linear combinations of explanatory variables without a pre-specified parametric model. The Appendix presents a detailed description of the estimation procedure.

SIM retains the desirable features of linear models, such as simplicity and interpretability of results, yet incorporates nonlinearity via the link function $G(.)$. Furthermore, SIM possesses the following four advantages:

(1) SIM uses information in stock returns to set the relative weights among the multiple signals. Lev and Thiagarajan (1993, 210) mention that the rather coarse nature of their index is due to the equal weights of the fundamental accounting signals, and Abarbanell and Bushee (1997, 8) argue that equally weighting all the signals may induce inaccuracy. The market-based β-coefficients estimate the relative weights of the fundamental signals as perceived by market participants.

(2) SIM allows us to test the hypothesis that the information in fundamental accounting signals is irrelevant with respect to market expectations of future earnings: H: $\beta_{1t} = \cdots = \beta_{4t} = 0$. If some of the fundamental signals are irrelevant with respect to market expectations of future earnings, then there will be no significant difference between the specifications with or without these signals.

(3) SIM extends Lev and Thiagarajan's (1993) model to a nonlinear setting. The nonlinear function $G(.)$ allows for asymmetric reactions to differential changes.[7]

(4) SIM avoids the curse of dimensionality because the index combines the variables in S. Therefore, SIM allows for the estimation of $G(.)$ with the same degree of precision that it would have if a one-dimensional index were observed (Horowitz 1998, 6). As a result, SIM uses a reasonable sample size to achieve efficient computations.

Market-Adapted Earnings to Forecast Future Earnings

To estimate market expectations of future earnings, we construct $M\hat{A}E^*_{it}$, which transforms $M\hat{A}E_{it}$ in Equation (2) as follows:

$$
\begin{aligned}
M\hat{A}E^*_{it} &= M\hat{A}E_{it} \times \frac{\overline{\Delta E_{it}}}{\overline{M\hat{A}E_{it}}} \\
&= \left(\frac{M\hat{A}E_{it}}{\overline{M\hat{A}E_{it}}} \right) \times \overline{\Delta E_{it}}, \qquad (3)
\end{aligned}
$$

where $\overline{\Delta E_{it}}$ and $\overline{M\hat{A}E_{it}}$ are the mean current earnings change and the mean $M\hat{A}E_{it}$, respectively. This transformation ensures that the mean value of the scaled $M\hat{A}E^*_{it}$ equals the mean value of reported earnings changes, $\overline{\Delta E_{it}}$. In other words, the sum of noise components in ΔE_{it} across all firms is zero.

[7] When the true link function is linear, i.e., $G(S\beta) = S\beta$, then changes in S generate proportional changes in stock returns.

This transformation achieves two goals. First, the ratio in the parenthesis of Equation (3) is unit-free (i.e., dimensionless). Consequently, we can set any nonzero value for the coefficient of ΔE_{it} in Equation (2) (i.e., β_{0t} other than 1) and obtain the same values for the scaled $M\hat{A}E_{it}^{*}$. Hence the assumption $\beta_{0t} = 1$ does not affect $M\hat{A}E_{it}^{*}$. Second, to forecast future earnings change, we multiply that ratio by $\overline{\Delta E_{it}}$. Note that both $M\hat{A}E_{it}^{*}$ and $M\hat{A}E_{it}$ contain information not only from accounting signals, but also from stock returns. Thus, the prediction procedure is well identified. We will employ $M\hat{A}E_{it}^{*}$ as our proxy for predicting future earnings change $\Delta E_{i,t+1}$.

Another property of $M\hat{A}E_{it}^{*}$ is that it is equivalent to a random-walk when the link function $G(.)$ is linear with no supplementary accounting signals, i.e., Equation (1a) becomes $R_{it} = \alpha_t + \beta_{0t}\Delta E_{it} + \varepsilon_{it}$. To see this point, we express Equation (3) as follows:

$$\begin{aligned} M\hat{A}E_{it}^{*} &= M\hat{A}E_{it} \times \frac{\overline{\Delta E_{it}}}{\overline{M\hat{A}E_{it}}} \\ &= (\hat{\beta}_{0t}^{SIM}\Delta E_{it}) \times \frac{\overline{\Delta E_{it}}}{\hat{\beta}_{0t}^{SIM}\overline{\Delta E_{it}}} \\ &= \Delta E_{it} \end{aligned}$$

Hence, the resulting $M\hat{A}E_{it}^{*}$ in this case is reduced to ΔE_{it}.

To gain insight into the potential advantage of our proposed earnings prediction model, we compare incremental predictive accuracy of the proposed model with the three benchmarks: (1) the classical random-walk; (2) the earnings forecast model of Abarbanell and Bushee (1997), which is based on Lev and Thiagarajan's (1993) fundamental accounting signals and reports significant relationships between most accounting signals and future earnings change; (3) the analysts' earnings forecasts for firms with analyst following.

Data

In the context of fundamental information analysis, Lev and Thiagarajan (1993) select 12 accounting signals of which five (gross margin, inventory, sales and administrative expenses, accounts receivable, and capital expenditures) are significant at the 5 percent confidence level (see their Table 2). Our preliminary analysis shows that gross margin, inventory, sales, and administrative expenses are significant, accounts receivable is marginally significant, and capital expenditures is insignificant. As a result, we chose the first four signals, which focus on the four key firm activities: profitability of sales, logistic operations, marketing and administration, and management of clientele. An advantage of excluding insignificant variables is that it enhances the precision of estimated parameters (Greene 2003, 151).

To measure the four accounting signals, we follow Lev and Thiagarajan (1993), who attach "good" or "bad" signs to accounting indicators of future economic performance based on analysts' interpretations, such that larger values indicate bad news. For instance, we compute the annual percentage change in inventory minus the percentage change in sales for each firm-year in the sample. Table 2 presents the definitions for the four signals.[8]

In addition to the above accounting signals, we use earnings per share before extraordinary items and discontinued operations (Compustat #58), as a measure of the reported earnings, E_t, deflated by share price at the beginning of the fiscal year. Changes in reported earnings, ΔE_t, are

[8] Additional discussion and references are in Lev and Thiagarajan (1993, 192).

TABLE 2
Definitions of Four Supplementary Accounting Variables

Accounting	Variables	Measurement	Good News	Bad News
ΔINV	Change in inventory[a]	Δinventory (#78 or #3) − Δsales (#12)	<0	>0
ΔGM	Change in gross margin	Δsales (#12) − Δgross margin (#12-#41)	<0	>0
ΔSGA	Change in sales and administrative expenses	Δsales and administrative expenses (#189) − Δsales (#12)	<0	>0
ΔREC	Change in accounts receivable	Δaccounts receivable (#2) − Δsales (#12)	<0	>0

[a] The inventory variable used is finished goods (when available) or total inventories (otherwise).

Following Lev and Thiagarajan (1993) and Abarbanell and Bushee (1997), the variables are defined such that good (bad) news is based on its negative (positive) expected relation with future benefits. Annual percentage changes are computed for each variable. Compustat items are in parentheses.

measured by the annual change in reported earnings deflated by share price at the beginning of fiscal year.[9] We also note that good news in each of the supplementary non-earnings signals is expressed by negative values and so the coefficients $\beta_{1t},\ldots,\beta_{4t}$ are expected to have negative values.

To measure the stock returns, we obtain information from Compustat annual files and the Center for Research in Security Prices (CRSP) monthly files. Annual abnormal return, R_{it}, is measured by the accumulated raw stock returns over the 12 months (from three months after the fiscal year begins until three months after the fiscal year-ends) minus the CRSP equally weighted monthly returns accumulated for the same period (see Das and Lev 1994).

To allow for intertemporal analysis, we follow Das and Lev's (1994, 360) approach to construct a constant-firm sample. As a result, the sample consists of firms that have complete data on Compustat and CRSP files for returns, earnings and the four accounting signals described in Table 2 for the 11-year period from 1990 to 2000, resulting in 636 firms.[10] The variables are adjusted for splits and dividends; high and low 1 percent of the variables are winsorized.

Estimating SIM

We estimate the following cross-sectional SIM model:

$$R_{it} = G(MAE_{it}) + \varepsilon_{it},$$

where $MAE_{it} = \Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it}$.

Table 3 presents the estimates of the regression coefficients and t-statistics. The average estimated effects across years for changes in inventory valuation, changes in gross margins, changes in SG&A, and changes in accounts receivable are −0.361 (−2.30), −1.062 (−3.20),

[9] Kothari (1992) demonstrates that the error-in-variables problem in a price-earnings regression is mitigated when earnings are deflated by share price and prices are assumed to lead earnings.
[10] This restriction may introduce a survivorship bias in our analysis (see also Das and Lev 1994).

TABLE 3

Single-Index Model Estimates of Regression Coefficients

SIM: $R_{it} = G(MAE_{it}) + \varepsilon_{it}$,
where $MAE_{it} = \Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it}$.

(1)	ΔINV (2)	ΔGM (3)	ΔSGA (4)	ΔREC (5)
Means across	−0.361	−1.062	−1.167	−0.271
years	(−2.30)[a]	(−3.20)	(−3.49)	(−1.34)
Pooled	−0.344	−0.670	−0.939	−0.278
sample	(−7.59)	(−8.64)	(−10.76)	(−4.65)

[a] t-statistics are in parentheses.

Regression coefficients, function $G(.)$ and market-adapted earnings are endogenously estimated from stock returns, reported earnings changes and four supplementary accounting signals: inventory valuation, gross margin, SG&A, and accounts receivable for 636 firms from 1990–2000.

Variable Definitions:

R_{it} = annual abnormal return, calculated by accumulating raw stock returns over the 12 months (from three months after the fiscal year begins until three months after the fiscal year ends minus the CRSP equally weighted monthly return accumulated for the same period); and

ΔE_{it} = annual change in earnings deflated by beginning of fiscal year share price.

ΔINV, ΔGM, ΔSGA, ΔREC are defined in Table 1.

−1.167 (−3.49), and −0.271 (−1.34), respectively, where t-statistics are given in parentheses. In addition, the estimated coefficients (t-values) for the pooled sample are −0.344 (−7.59), −0.670 (−8.64), −0.939 (−10.76), and −0.278 (−4.65), respectively, and these effects are all significant.[11]

The above significant and negative β coefficients demonstrate that market participants consider the first three fundamental accounting signals in forming their perceptions on future earnings, while results on accounts receivable are mixed. The findings also shed light on the relative weight of each signal as assigned by market participants in forming their expectations on future earnings. Specifically, the estimated weight assigned to news, as signaled through changes in gross margins and SG&A, is about three times greater than the weight assigned to news signaled through changes in the value of inventory. Consistent with the earlier studies, the estimated function $G(.)$ plotted in Figure 1 shows an S-shaped relationship between stock returns and MAE_{it}.

To gain further insights into both the nonlinearity and the supplementary accounting signals, we compare the SIM model with two models previously explored in the literature: the LT model and the DL model.

The LT Model—fundamental accounting signals (Lev and Thiagarajan 1993):

$$R_{it} = \alpha_t + \beta_{0t}\Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it} + \varepsilon_{it}.$$

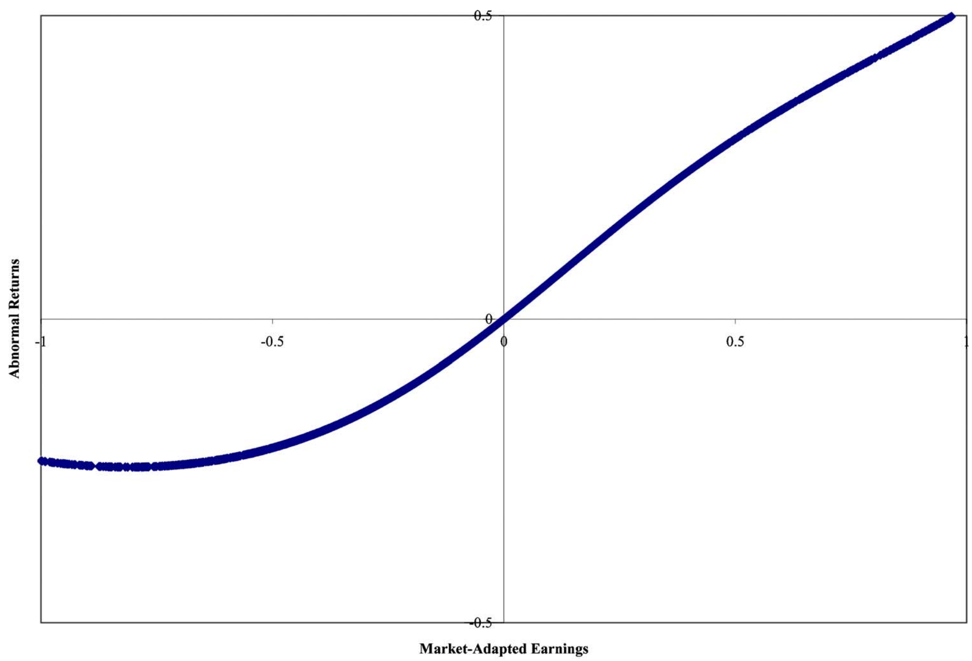
The DL Model—nonlinearity (Das and Lev 1994):

$$R_{it} = G(\Delta E_{it}) + \varepsilon_{it}.$$

[11] The t-statistics in Table 3 are not corrected for serial correlation or clustering by firms because such corrections are not available for single-index models (see Li and Racine 2007, Chapters 8 and 19).

FIGURE 1
Relationship between Abnormal Returns and Market-Adapted Earnings

Model: $R_{it} = G(MAE_{it})$, where $MAE_{it} = \Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it}$.[a]
The figure plots function $G(.)$ and shows an S-shaped relationship between stock returns and MAE, 6360 firm-year observations, 1990–2000.



MAE

[a] The upper and lower 1 percent of the observations are truncated.

Because the LT model follows the linear specification with supplementary accounting signals, we compare it with SIM to infer the net contribution of nonlinearity in improving the Lev and Thiagarajan (1993) specification. Similarly, because the DL model emphasizes the nonlinear link between returns and earnings, but ignores supplementary accounting signals, we compare it with SIM to infer the net contribution from introducing supplementary accounting signals.

Table 4 presents the results of these comparisons. Specifically, on average, the explanatory power of the SIM model measured by adjusted R^2 is 40.5 percent greater than that of the DL model, and 58.9 percent greater than that of the LT model. In addition, the SIM model has a greater explanatory power for 9 of the 11 years compared with the DL model and for all 11 years compared with the LT model.

This improvement in explanatory power due to the SIM model ranges between 2.9 percent and 156.1 percent for the LT comparison and between −29.7 percent and 264.1 percent for the DL comparison. These results are consistent with Lev and Thiagarajan (1993, Tables 2 and 3) and Das and Lev (1994, Table 1), who also report considerable intertemporal variation.

TABLE 4

A Comparative Analysis of Adjusted R^{2a}

DL Model: $R_{it} = G_t(\Delta E_{it}) + \varepsilon_{it}$

LT Model: $R_{it} = \alpha_t + \beta_{0t}\Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it} + \varepsilon_{it}$

SIM: $R_{it} = G_t(MAE_{it}) + \varepsilon_{it}$,

where $MAE_{it} = \Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it}$.

				Improvement (%)		
Year (1)	**DL Model Adjusted R^2 %** (2)	**LT Model Adjusted R^2 %** (3)	**SIM Adjusted R^2 %** (4)	**Model SIM over Model DL = [(4) − (2)] / (2)** (5)	**Model SIM over Model LT = [(4) − (3)] / (3)** (6)	**n** (7)
1990	11.66	11.16	13.50	15.8	21.0	636
1991	6.60	8.13	8.67	31.4	6.6	636
1992	14.85	9.12	16.42	10.6	80.0	636
1993	19.85	18.77	32.98	66.1	75.7	636
1994	10.62	19.84	38.67	264.1	94.9	636
1995	11.10	8.71	8.96	(19.3)	2.9	636
1996	26.67	11.71	18.74	(29.7)	60.0	636
1997	11.68	5.17	13.24	13.4	156.1	636
1998	5.54	3.57	6.89	24.4	93.0	636
1999	9.21	7.44	10.24	11.2	37.6	636
2000	8.08	10.63	12.75	57.8	19.9	636
Average improvement				40.5	58.9	

[a] The sample includes 636 constant firms, 1990-2000.

Variable Definitions:

R_{it} = annual abnormal return, calculated by accumulating raw stock returns over the 12 months (from three months after the fiscal year begins until three months after the fiscal year ends minus the CRSP equally weighted monthly return accumulated for the same period); and

ΔE_{it} = annual change in earnings, deflated by beginning of fiscal year share price.

ΔINV, ΔGM, ΔSGA, ΔREC are defined in Table 1.

We next test the economic significance of the incremental explanatory power of supplementary accounting signals. To this end, we compare the nonlinear DL model with the SIM model using an F*-test.[12] The second column of Table 5 reports F*-statistics for testing the null hypothesis (the DL model is the appropriate specification) against the alternative hypothesis (SIM is better specified). The F*-statistics for the 9 out of 11 years reject the null at the $\alpha = 0.05$ level. In addition, the third column of Table 5 reports Smirnov statistics for testing the null hypothesis (the

[12] For simplicity, we followed the classical methodology to approximate an F-statistic and used $F^* = [\ (SSE_1 - SSE) / k)] / [SSE / (N - k - 1)$, where N = number of observations, SSE (SSE_1) = residual sum of squares for the five-variable model (earnings-only sample), and $k = 4$, the number of additional variables (Kmenta 1986, 417). We also computed the degrees of freedom for the appropriate matrices and used Cleveland and Devlin's (1988) approximation for an F-statistic. Similar results hold for both tests.

TABLE 5

Test of Alternative Specifications of the Returns-Earnings Relationship for 636 Constant Firms 1990–2000

DL Model: $R_{it} = G_t(\Delta E_{it}) + \varepsilon_{it}$

LT Model: $R_{it} = \alpha_t + \beta_{0t}\Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it} + \varepsilon_{it}$

SIM: $R_{it} = G_t(\text{MAE}_{it}) + \varepsilon_{it}$,
where $MAE_{it} = \Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it}$.

	H_0: DL is the Appropriate Specification H_1: SIM is the Appropriate Specification		H_0: LT is the Appropriate Specification H_1: SIM is the Appropriate Specification	
Year (1)	**F*-statistic[a] (2)**	**Smirnov Statistic[b] (3)**	**F-statistic[a] (4)**	**Smirnov Statistic[b] (5)**
1990	4.336 (0.0007)	0.057 (0.0181)	3.5750 (0.0015)	0.064 (0.0050)
1991	13.412 (0.0000)	0.047 (0.0741)	1.4191 (0.1990)	0.069 (0.0020)
1992	3.922 (0.0017)	0.085 (0.0001)	9.8048 (0.0000)	0.131 (0.0000)
1993	31.331 (0.0000)	0.039 (0.2151)	22.7074 (0.0000)	0.027 (0.7100)
1994	73.080 (0.0000)	0.074 (0.0006)	31.4839 (0.0000)	0.052 (0.0400)
1995	NA[c]	0.050 (0.0478)	1.1180 (0.3501)	0.053 (0.0350)
1996	NA[c]	0.060 (0.0111)	4.6068 (0.0000)	0.108 (0.0000)
1997	3.672 (0.0028)	0.127 (0.0000)	10.5024 (0.0000)	0.131 (0.0000)
1998	3.488 (0.0041)	0.045 (0.0975)	3.9167 (0.0003)	0.083 (0.0001)
1999	2.927 (0.0127)	0.131 (0.0000)	3.9286 (0.0005)	0.182 (0.0000)
2000	9.453 (0.0000)	0.189 (0.0000)	3.6413 (0.0022)	0.134 (0.0000)

[a] For simplicity, we followed the classical methodology to approximate an F*-statistic and used $F^* = [(SSE_1 - SSE) / (k)SSE / (N - k - 1)]$, where N = number of observations, $SSE(SSE_1)$ = residual sum of squares for the five-variable model (earnings-only sample), and $k = 4$, the number of additional variables (Kmenta 1986, 417). We also computed the degrees of freedom for the appropriate matrices and used Cleveland and Devlin's (1988) approximation for an F-statistic. Similar results hold for both tests.

[b] The two sample Smirnov tests (also called the Kolmogorov-Smirnov test) follow Lindgren (1968, 494).

[c] The residual sum of squares for the SIM is higher than for the DL Model, suggesting that the DL Model has better explanatory power than the SIM for 1995 and 1996.

(continued on next page)

TABLE 5 (continued)

Cell entries in column 2 (4) report an F* (F)-statistic that tests for the null hypothesis that the DL (LT) model is the appropriate specification, against the alternative that the SIM is better specified. Numbers in parentheses represent the corresponding p-values. A significant F-statistic implies a rejection of the null.

Cell entries in column 3 (5) report a Smirnov statistic that tests for the null hypothesis that the specification of the DL (LT) model and the SIM are equivalent estimations, against the alternative that they are different estimations. Numbers in parentheses represent the corresponding p-values. A significant Smirnov statistic implies a rejection of the null.

specification of the DL model and SIM are equivalent) against the alternative hypothesis (they are different). Smirnov statistics for the 8 out of 11 years are significant at the $\alpha = 0.05$ level, thus rejecting the null hypothesis. The above evidence demonstrates that, beyond nonlinearity, the supplementary accounting signals contain meaningful information.

We further test the incremental explanatory power of nonlinearity given multiple accounting variables. We use F-statistics to test for the null hypothesis (i.e., the LT model is the appropriate specification) against the alternative hypothesis (i.e., SIM is better specified). The fourth column of Table 5 reports F-statistics, which are significant for 9 of the 11 years, suggesting that we can reject the null at the $\alpha = 0.05$ level. The fifth column reports a Smirnov statistic for testing the null hypothesis (the specifications of the LT model and SIM are equivalent) against the alternative (i.e., they are different). Smirnov statistics for 10 of the 11 years are significant at the $\alpha = 0.05$ level and. Hence. we reject the null hypothesis. Consequently, this evidence indicates that, beyond multiple accounting signals, nonlinearity provides meaningful information.

To assess robustness, we replicate the above analyses by using eight supplementary signals in Lev and Thiagarajan (1993)—that is, including four additional signals (changes in capital expenditures, effective tax rate, order backlog, and labor force)—and by using the sampling criteria used by Abarbanell and Bushee (1997). We find that three of the four additional signals are insignificant (only order backlog was significant). In addition, across 11 years, we find that on average the explanatory power of the SIM model as measured by adjusted R^2 is 30.3 percent greater than that of the DL model, and 41.4 percent greater than that of the LT model. In sum, the evidence suggests that both nonlinearity and supplementary accounting signals contribute to the improved specification of SIM. We next examine the predictive power of market-adapted earnings.

IV. OUT-OF-SAMPLE EARNINGS PREDICTION

In this section, we compare one-year-ahead forecasts from the market-adapted earnings, *MAE**, with those obtained from the random-walk model and the accounting-based model (ABM) proposed by Abarbanell and Bushee (1997).

The ABM Model—accounting-based model (Abarbanell and Bushee 1997)

$$\Delta E_{i,t+1} = \alpha_t + \beta_{0t}\Delta E_{it} + \beta_{1t}\Delta INV_{it} + \beta_{2t}\Delta GM_{it} + \beta_{3t}\Delta SGA_{it} + \beta_{4t}\Delta REC_{it} + \varepsilon_{i,t}.$$

To study out-of-sample earnings prediction, we first estimate the regression coefficients based on the accounting information reported for fiscal years t and $t-1$. Next, we use the estimated model coefficients to forecast earnings change in $t+1$ for each firm i, based on its accounting signals reported for the fiscal year t. Accordingly, we need two years of data for out-of-sample forecasts.

In ABM, the average estimated effects across years for changes in inventory valuation, changes in gross margins, changes in SG&A, and changes in accounts receivable are -0.040 (-1.880), -0.016 (-2.240), -0.045 (-2.051), -0.068 (-0.283), respectively, where t-statistics are given in parentheses. Note that these four estimated effects in ABM differ significantly from the corresponding estimated coefficients in SIM (see the first row of Table 3), leading to different forecasting performance.

To compare prediction accuracy across MAE^*, random-walk and ABM models, we analyze the sample from 1992 to 2001. Using MAE^* to predict out-of-sample earnings change, we employ information observed in year t to forecast earnings change in year $t+1$. The incremental out-of-sample earnings prediction power of MAE^* over the two benchmark models is evaluated based on the potential increase in predictive accuracy. The absolute percentage forecast error is defined as the percentage difference between actual and forecasted earnings divided by the actual earnings (firm subscript suppressed):

$$\text{Earnings forecast error} = \left|\frac{\text{Actual} - \text{Forecasted}}{\text{Actual}}\right| = \left|\frac{E_{t+1} - \text{Forecasted } E_{t+1}}{E_{t+1}}\right| = \left|\frac{E_{t+1} - (E_t + \text{Forecasted } \Delta E_{t+1})}{E_{t+1}}\right| = \left|\frac{\Delta E_{t+1} - \text{Forecasted } \Delta E_{t+1}}{E_{t+1}}\right|. \quad (4)$$

This measure of earnings forecast error has been used in earlier studies (e.g., BLM; Collins et al. 1987; Elgers and Murray 1992).

Table 6 presents the absolute earnings forecast errors obtained from the three models. Because percentage forecast errors can become quite large when scaled by relatively small actual earnings values, as in Collins et al. (1987), we report the median rather than the mean. Table 6 shows that the median MAE^* forecast error is smaller than the median random-walk forecast errors in nine out of the ten years. For the pooled sample, the median MAE^* forecast error is 0.273, which is 20.9 percent lower than the median random-walk forecast error, 0.345. Furthermore, 56.1 percent of the pooled sample observations have a lower absolute earnings forecast error for MAE^* than that for the random-walk model (t-statistic = 9.781).[13] Hence MAE^* outperforms the random-walk model.

Similarly, Table 6 shows that MAE^* yields lower forecast errors than the ABM. The median MAE^* forecast error is smaller than the median forecast error of the ABM in nine out of the ten years. For the pooled sample, the median MAE^* forecast error is 0.273, which is 17.0 percent lower than the median ABM forecast error, 0.329. In addition, 57.7 percent of the pooled sample observations have a lower absolute earnings forecast error for MAE^* than that for the ABM (t-statistic = 12.328). Hence MAE^* outperforms the accounting-based model.

Across the three models, the evidence indicates that employing the SIM model to extract forward-looking information from stock returns results in a higher level of accuracy than either the random-walk or the ABM models. In other words, the accuracy of earnings forecasts enhances when we utilize information contained in stock returns.[14]

In the rest of this section, we compare the performance of MAE^* with analysts' earnings forecasts; in the subsequent section, we present results when no analyst coverage is available. To this end, we obtain analysts' consensus earnings forecasts from I/B/E/S. Specifically, we use the consensus forecasts of annual earnings announced in the fourth month after the start of the fiscal year to keep a contemporaneous analysis with no timing advantage. Firm-years observations with

---

[13] We use the normal approximation of the binomial probabilities to test the hypothesis that the probability of a lower error is 0.5 (see BLM, 24).

[14] When stock returns are positive (i.e., good news), further analyses reveal that the accuracy of MAE^* increases (median absolute forecast error = 0.206); when stock returns are negative (i.e., bad news), MAE^* accuracy decreases (median absolute forecast error = 0.429). One reason for larger forecast errors under bad news is that earnings realizations exhibit large extreme values, partly due to write-offs and nonrecurring items.

TABLE 6

Out-of-Sample Earnings Prediction
*MAE** versus Random-Walk and Accounting-Based Model

	Random-Walk (*RW*)			Market-Adapted Earnings (*MAE**)			Accounting-Based Model (*ABM*)		
Predicted Year	Median Absolute Error[a]	Percent times *MAE** has Lower Error[c]	Median Difference[b] *RW* – *MAE**	Median Improvement Over *RW*	Median Absolute Error[a]	Median Improvement Over the *ABM*	Median Absolute Error[a]	Percent times MAE* has Lower Error[c]	Median Difference[b] *ABM* – *MAE**
(1)	(2)	(3)	(4)	(5) = [(2) – (6)] / (2)	(6)	(7) = [(8) – (6)] / (8)	(8)	(9)	(10)
1992	0.299	0.524	0.017	16.8	0.249	14.1	0.290	0.561	0.025
		(1.190)						(3.093)	
1993	0.415	0.601	0.103	33.4	0.277	31.8	0.406	0.613	0.076
		(5.076)						(5.710)	
1994	0.395	0.616	0.121	42.6	0.227	−15.8	0.196	0.403	−0.038
		(5.869)						(−4.917)	
1995	0.374	0.590	0.079	43.2	0.213	20.2	0.267	0.555	0.022
		(4.520)						(2.776)	
1996	0.298	0.571	0.048	31.1	0.205	26.0	0.277	0.623	0.065
		(3.569)						(6.186)	
1997	0.325	0.552	0.037	20.8	0.258	22.5	0.333	0.619	0.051
		(2.617)						(6.027)	
1998	0.335	0.550	0.025	18.3	0.274	18.5	0.336	0.583	0.034
		(2.538)						(4.203)	
1999	0.371	0.560	0.044	14.0	0.319	15.8	0.379	0.588	0.052
		(3.014)						(4.441)	
2000	0.345	0.577	0.058	22.3	0.268	26.8	0.366	0.626	0.068

(continued on next page)

TABLE 6 (continued)

	Random-Walk (*RW*)			Market-Adapted Earnings (*MAE**)			Accounting-Based Model (*ABM*)		
Predicted Year	Median Absolute Error[a]	Percent times *MAE** has Lower Error[c]	Median Difference[b] *RW* − *MAE**	Median Improvement Over *RW*	Median Absolute Error[a]	Median Improvement Over the *ABM*	Median Absolute Error[a]	Percent times MAE* has Lower Error[c]	Median Difference[b] *ABM* − *MAE**
(1)	(2)	(3)	(4)	(5) = [(2) − (6)] / (2)	(6)	(7) = [(8) − (6)] / (8)	(8)	(9)	(10)
		(3.886)						(6.334)	
2001	0.327	0.473	−0.032	−14.6	0.374	14.4	0.437	0.601	0.072
		(−1.348)						(5.076)	
All	0.345	0.561	0.049	20.9	0.273	17.0	0.329	0.577	0.038
		(9.781)						(12.328)	

[a] Absolute earnings forecast error is the absolute value of the percentage difference between actual and forecasted earnings divided by actual earnings.

[b] The difference in absolute earnings forecast errors is computed for all observations in each year and the median is reported.

[c] Percent of firm-year observations with absolute earnings forecast error lower for market-adapted earnings than for the alternative model. Numbers in parentheses are t-statistics for testing the null hypothesis that the probability of a lower error is 0.5 using a normal approximation of the binomial probabilities.

analyst coverage comprise 66.2 percent of our sample (4213 out of 6360). The analyst earnings forecast error is computed using Equation (4), where firms' actual earnings are obtained from I/B/E/S for comparability with the forecast (see Richardson et al. 2004, 896).

Table 7 presents the comparative results for observations with analyst coverage—see the first row. The median absolute forecast error of *MAE** is 21.0 percent higher than that for the analysts' forecasts. Moreover, analysts forecasts have smaller error than *MAE** on 53.2 percent (=100 − 46.8) of the observations (t-statistic = −4.128). This finding indicates that analysts provide investors with more accurate earnings forecasts than *MAE**, and it is consistent with the extant literature (e.g., Fried and Givoly 1982; Brown et al. 1987; Banker and Chen 2006).

Furthermore, we examine the analysts' advantage over *MAE** in low versus high information environments as measured by the number of analysts covering each firm (see the second and third rows in Table 7). In high (low) coverage, analysts' forecasts outperform *MAE** by 22.2 percent (16.6 percent). Analysts do better than *MAE** because their forecasts are based on all available information, not just financial variables. In contrast, in low information environment, the performance of both analysts and *MAE** degrades because fundamentals are likely to be more volatile. Overall, analysts' incremental usefulness over *MAE** is more pronounced in rich-information environments.

TABLE 7

Out-of-Sample Earnings Prediction:
***MAE** versus Analyst Earnings Forecasts**

	Analyst Earnings Forecasts (*AF*)			Market-Adapted Earnings (*MAE**)	
Group	**Median Absolute Error[a]**	**Percent times *MAE** Has Lower Error[c]**	**Median Difference[b] *AF* − *MAE****	**Median Improvement Over *AF***	**Median Absolute Error[a]**
(1)	**(2)**	**(3)**	**(4)**	**(5) = [(2) − (6)] / (2)**	**(6)**
ALL firm-year observations with analyst coverage	0.203	0.468 (−4.128)	−0.028	−21.0	0.246
Low coverage	0.239	0.495 (−0.437)	−0.004	−16.6	0.278
High[d] coverage	0.174	0.442 (−5.371)	−0.041	−22.2	0.213

[a] Absolute earnings forecast error is the absolute value of the percentage difference between actual and forecasted earnings divided by actual earnings.

[b] The difference in absolute earnings forecast errors is computed for all observations in each year and the median is reported.

[c] Percent of firm-year observations with absolute earnings forecast error lower for market-adapted earnings than that for the alternative model. Numbers in parentheses are t-statistics for testing the null hypothesis that the probability of a lower error is 0.5 using a normal approximation of the binomial probabilities.

[d] Analyst coverage is high when the number of analysts is equal to or above its median (=8).

Finally, we assess the relative advantage of *MAE** over the random-walk and accounting-based earnings forecasts as the richness of the information environment varies. We expect a higher relative advantage of *MAE** over the random-walk and the accounting-based earnings forecasts when more information is available. Table 8 shows that *MAE** improves the random-walk model (ABM) by 26.1 percent (20.6 percent) under high coverage and by 10.4 percent (10.7 percent) under low coverage. Thus the incremental advantage of *MAE** over both the random-walk and ABM models is higher in richer information environments (i.e., high coverage) because it extracts information from stock prices, which incorporate more information in richer information environments. The next section examines whether *MAE** offers better predictions for firms with no analyst coverage.

V. FIRMS WITH NO ANALYST COVERAGE

When analysts' forecasts are not available, researchers use potentially inferior proxies for market expectations of future earnings. As in the previous section, we compare the predictive performance of *MAE** with that of the random-walk and ABM models for firms with no analyst coverage. Results reported in the first row of Table 8 indicate that *MAE** outperforms both the random-walk and ABM models. Specifically, the median *MAE** absolute forecast error is 14.4 percent (12.0 percent) lower than the median random-walk (ABM) absolute forecast error. In addition, *MAE** outperforms the random-walk model (ABM) on 55.2 percent (53.6 percent) of the observations (t-statistic is 2.796 (3.324)).[15]

The reason for the superiority of *MAE** over accounting-based earnings forecasts for the firms with no analyst coverage is due to its ability to exploit information contained in stock prices. In other words, *MAE** yields improved price-based forecasts over accounting-based forecasts because it utilizes a richer information set. Thus *MAE** is of most useful when analysts' earnings forecasts are not available, because investors and researchers cannot benefit from alternative earnings forecasts.

VI. SUMMARY

Understanding how market participants form earnings expectations has long been a subject of interest to accounting researchers. This paper introduces a new estimation approach, SIM regression, for extracting forward-looking information from stock returns. We propose and estimate market-adapted earnings, which estimates market expectations of future earnings conditioned on stock returns and multiple accounting signals. We demonstrate an economically meaningful predictive advantage of *MAE** over both the random-walk model and ABM. In particular, *MAE** is shown to be useful for a class of firms with no analyst following. Hence, researchers can employ market-adapted earnings as a proxy for market expectations of future earnings for firms with no analyst coverage.

[15] As a robustness check, we replicate the analysis using a less restrictive sample. Specifically, we utilize all firm-year observations with available data for our variables from Compustat and CRSP. To avoid the small deflator problem, we consider stock prices to be at least five dollars (e.g., Lim 2001), obtaining 29,498 observations for 5,033 firms. The less restrictive sample consists of 15,195 observations with no analyst coverage (51.5 percent of the sample observations in contrast with 33.8 percent in the constant-firms sample). Results (not reported here) indicate that *MAE** is more accurate than the random-walk model (*ABM*) by 8.6 percent (8.0 percent) under no analyst coverage. Furthermore, the absolute forecast error for *MAE** is lower among 54.2 percent (54.0 percent) of the observations than that for the random-walk model (*ABM*) with t-statistic = 10.354 (9.861). Thus, in the less restrictive sample, these findings confirm the predictive advantage of *MAE** over both the random-walk model and *ABM*. We thank an anonymous referee for suggesting this analysis.

TABLE 8

Out-of-Sample Earnings Prediction

***MAE** versus Random-Walk and Accounting-Based Model for Firms with No, Low, and High Analyst Coverage**

	Random Walk (*RW*)			Market-Adapted Earnings (*MAE**)			Accounting Based Model (*ABM*)			
Group	Median Absolute Error[a]	Percent times *MAE** Has Lower Error[c]	Median Difference[b] *RW* – *MAE**	Median Improvement Over *RW*	Median Absolute Error[a]	Median Improvement Over the *ABM*	Median Absolute Error[a]	Percent times *MAE** Has Lower Error[c]	Median Difference[b] *ABM* – *MAE**	n
(1)	(2)	(3)	(4)	(5) = [(2) – (6)] / (2)	(6)	(7) = [(8) – (6)] / (8)	(8)	(9)	(10)	(11)
No coverage	0.396	0.552 (2.796)	0.014	14.4	0.339	12.0	0.386	0.536 (3.324)	0.013	2147
Low coverage[c]	0.310	0.569 (3.624)	0.021	10.4	0.278	10.7	0.311	0.569 (6.331)	0.038	2098
High coverage[d]	0.288	0.627 (6.759)	0.038	26.1	0.213	20.6	0.335	0.581 (11.720)	0.038	2115

[a] Absolute earnings forecast error is the absolute value of the percentage difference between actual and forecasted earnings divided by actual earnings.

[b] The difference in absolute earnings forecast errors is computed for all observations in each year and the median is reported.

[c] Percent of firm-year observations with absolute earnings forecast error lower for market-adapted earnings than that for the alternative model. Numbers in parentheses are t-statistics for testing the null hypothesis that the probability of a lower error is 0.5 using a normal approximation of the binomial probabilities.

[d] Analyst coverage is high when the number of analysts is equal to or above its median (=8).

APPENDIX
SINGLE-INDEX MODEL ESTIMATION

Here, for a given year t, we describe the estimation of the single-index model:

$$R_i = G(MAE_i) + \varepsilon_i,$$

where $MAE_i = \Delta E_i + \beta_1 \Delta INV_i + \beta_2 \Delta GM_i + \beta_3 \Delta SGA_i + \beta_4 \Delta REC_i$, and R_i denotes the dependent variable and ε_i is the normal error term with zero mean and constant variance.

Let the vector X_i contain the five regressors ΔE_i, ΔINV_i, ΔGM_i, ΔSGA_i, and ΔREC_i. We first compute the usual covariance matrix:

$$\hat{\Sigma}_x = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})(X_i - \bar{X})', \tag{A1}$$

where X_i' denotes a row for the ith firm, and $\bar{X}$ contains the means of the five variables across all N firms.

Next, we create a matrix (R, X) and sort it by the increasing values of R_i. We partition the *sorted* (R, X) matrix into H groups or slices, and we denote each sub-matrix of regressors in slice h by X_h, where $h = 1,\ldots,H$. Slicing captures the nonlinearity of the link function $G(.)$ without the need to know its exact shape, thus mitigating the risk of specification error (see Duan and Li 1991).

We compute the sample means of the regressors for each group X_h across the respective firms in the slice h and denote them by $\bar{X}_h$. The weighted average across all H slices yields the covariance matrix:

$$\hat{\Sigma}_\eta = \sum_{h=1}^{H}\hat{p}_h(\bar{X}_h - \bar{X})(\bar{X}_h - \bar{X})', \tag{A2}$$

where $\hat{p}_h$ is the proportion of firms in slice h.

Then, to estimate β, we extract the principal eigenvector γ of the generalized eigenvalue decomposition problem:

$$\hat{\Sigma}_\eta \gamma = \lambda \hat{\Sigma}_x \gamma, \tag{A3}$$

where λ is the largest eigenvalue. We also obtain the standard errors of $\hat{\beta}$ (see Chen and Li 1998, 297) by taking the squared root of the diagonal of the matrix:

$$\frac{1 - \hat{\lambda}}{\hat{\lambda}} N^{-1} \hat{\Sigma}_x^{-1}. \tag{A4}$$

Using these estimated parameters and their standard errors, we can thus compute the t-values to determine the significant effects at the 95 percent confidence level.

Finally, we normalize the first element of $\hat{\beta}$ to unity and construct the Single Index $z_i = X_i'\hat{\beta}$ to estimate the unknown function $\hat{G}$ by applying the standard local polynomial regression to the model $R_i = G(z_i) + \varepsilon_i$ (see Simonoff 1996; Das and Lev 1994). The local polynomial regression requires a selection of a bandwidth parameter, which is set to unity. For software, see note at: http://recanati.tau.ac.il/_Uploads/20SIM.pdf.

REFERENCES

Abarbanell, J. S., and B. J. Bushee. 1997. Fundamental analysis, future earnings, and stock returns. *Journal of Accounting Research* 35 (1): 1–24.

———, and ———. 1998. Abnormal returns to a fundamental analysis strategy. *The Accounting Review* 73 (1): 19–45.

Ait-Sahalia, Y., and M. W. Brandt. 2001. Variable selection for portfolio choice. *The Journal of Finance* 56: 1297–1351.

Ayers, B. C., and R. N. Freeman. 2000. Why do large firms' prices anticipate earnings earlier than small firms' prices? *Contemporary Accounting Research* 17 (2): 191–212.

———, and ———. 2003. Evidence that analyst following and institutional ownership accelerates the pricing of future earnings. *Review of Accounting Studies* 8: 47–67.

Banker, R., and L. Chen. 2006. Predicting earnings using a model based on cost variability and cost stickiness. *The Accounting Review* 81: 285–307.

Beaver, W. H., R. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting and Economics* 2: 3–28.

———, ———, and S. G. Ryan. 1987. The information content of security prices—A second look. *Journal of Accounting and Economics* 9: 139–157.

———, M. L. McAnally, and C. H. Stinson. 1997. The information content of earnings and prices: A simultaneous equations approach. *Journal of Accounting and Economics* 23: 53–81.

Beneish, M. D., and C. R. Harvey. 1998. Measurement error and nonlinearity in the earnings-returns relation. *Review of Quantitative Finance and Accounting* 11: 219–247.

Brown, L. D. 1993. Earnings forecasting research: Its implications for capital market research. *International Journal of Forecasting* 9: 295–320.

———, R. L. Hagerman, P. A. Griffin, and M. E. Zmijewski. 1987. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics* 9: 61–87.

Chen, C. H., and K. C. Li. 1998. Can SIR be as popular as multiple linear regression? *Statistica Sinica* 8: 289–316.

Cheng, C. S. A., W. S. Hopewood, and J. C. McKeown. 1992. Non-linearity and specification problems in unexpected earnings response regression models. *The Accounting Review* 67 (3): 579–598.

Cleveland, W. S., and S. J. Devlin. 1988. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association* 83 (403): 596–610.

Collins, D. W., S. P. Kothari, and J. D. Rayburn. 1987. Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics* 9: 111–138.

Cready, W. M., D. H. Hurtt, and J. A. Seida. 2000. Applying reverse regression techniques in earnings-return analyses. *Journal of Accounting and Economics* 30: 227–240.

Das, S., and B. Lev. 1994. Nonlinearities in the returns-earnings relation: Tests of alternative specifications and explanations. *Contemporary Accounting Research* 11 (1): 353–379.

Duan, N., and K-C. Li. 1991. Slicing regression: A link-free regression method. *Annals of Statistics* 19: 505–530.

Elgers, P. T., and D. Murray. 1992. The relative and complementary performance of analyst and security-price-based measures of expected earnings. *Journal of Accounting and Economics* 15 (2): 303–316.

Fama, E. F., and K. R. French. 2000. Forecasting profitability and earnings. *Journal of Business* 73 (2): 161–175.

Freeman, R. N. 1987. The association between accounting earnings and security returns for large and small firms. *Journal of Accounting and Economics* 9 (2): 195–228.

———, and S. Y. Tse. 1992. A nonlinear model of security price responses to unexpected earnings. *Journal of Accounting Research* 30 (2): 185–209.

Fried, D., and D. Givoly. 1982. Financial analysts' forecasts of earnings: A better surrogate for market expectations. *Journal of Accounting and Economics* 11: 1–31.

Greene, W. H. 2003. *Econometric Analysis*. Englewood Cliffs, NJ: Prentice Hall.

Horowitz, J. L. 1998. *Semiparametric Methods in Econometrics*. New York, NY: Springer.

Jiambalvo, J., S. Rajgopal, and M. Venkatachalam. 2002. Institutional ownership and the extent to which stock prices reflect future earnings. *Contemporary Accounting Research* 19: 117–145.

Kmenta, J. 1986. *Elements of Econometrics*. 2nd edition. New York, NY: Macmillan Publishing Co.

Kothari, S. P. 1992. Price-earnings regressions in the presence of prices leading earnings. *Journal of Accounting and Economics* 15 (2–3): 173–202.

———, and R. G. Sloan. 1992. Information in prices about future earnings. *Journal of Accounting and Economics* 15 (2–3): 143–171.

———. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31: 105–231.

Lev, B., and R. Thiagarajan. 1993. Fundamental information analysis. *Journal of Accounting Research* 31 (2): 190–215.

Li, K. C. 1991. Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association* 86: 316–342.

Li, Q., and J. S. Racine. 2007. *Nonparametric Econometrics*. Princeton, NJ: Princeton University Press.

Lim, T. 2001. Rationality and analyst forecast bias. *The Journal of Finance* 56: 369–385.

Lindgren, B. W. 1968. *Statistical Theory*. New York, NY: Macmillan Publishing Co.

Naik, P. A., and C. L. Tsai. 2001. Single-index model selections. *Biometrika* 88 (3): 821–832.

O'Brien, P. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10: 53–83.

Ou, J. A., and S. H. Penman. 1989a. Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics* 11: 295–329.

———, and ———. 1989b. Accounting measurement, price-earnings ratio, and the information content of security prices. *Journal of Accounting Research* 27 (Supplement): 111–144.

———. 1990. The information content of nonearnings accounting numbers as earnings predictors. *Journal of Accounting Research* 28 (1): 144–163.

Penman, S. H. 1998. Combining earnings and book value in equity valuation. *Contemporary Accounting Research* 15 (3): 291–324.

Rajgopal, S., T. Shevlin, and M. Venkatachalam. 2003. Does the stock market fully appreciate the implications of leading indicators for future earnings? Evidence from order backlog. *Review of Accounting Studies* 8: 461–492.

Richardson, S., S. H. Teoh, and P. D. Wysocki. 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21 (4): 885–924.

Simonoff, J. S. 1996. *Smoothing Methods in Statistics*. New York, NY: Springer.