A Critical Assessment of the State of the Art of Statistically-based Annual Cash-flow Prediction Models

by:

Kenneth S. Lorek*

and

Christine Fimia-Moe**

November 14, 2011

*Corresponding Author
The W. A. Franke College of Business
Northern Arizona University
PO Box 15066
Flagstaff, AZ 86011-5066
928-523-7406
Ken.Lorek@nau.edu

**Senior Audit Associate
KPMG LLP
701 West 8th Avenue, Suite 600
Anchorage, AK 99501
cfimia-moe@kpmg.com

A Critical Assessment of the State of the Art of Statistically-based Annual Cash-flow Prediction Models

ABSTRACT

We provide a synthesis and critical analysis of the methodological issues pertaining to the development of extant annual cash-flow prediction models as well as a set of specific recommendations pertaining to future research in this area. By so doing we explicate the linkage between accounting accruals and cash-flow data, a topic of considerable interest to standard-setting bodies, accounting researchers, and members of the financial community interested in deriving firm valuations. Specifically, we recommend that future empirical work on cash-flow prediction models should: (1) employ data consistent with SFAS No. 95 reporting standards; (2) consider the employment of time-series estimation procedures that capture firm-specific, contextual relationships; (3) assess predictive ability by generating cash-flow forecasts in an inter-temporal holdout period that is not used in model identification and estimation; and (4) assess predictive performance employing multi-step ahead projections.

Key Words: Statistically-based cash-flow prediction models, Time-series estimation, Cross-sectional estimation, SFAS No. 95.

A Critical Assessment of the State of the Art of Statistically-based Annual Cash-flow Prediction Models

I. Introduction

The purpose of this critical assessment of the state of the art of statistically-based cash-flow prediction models is to: (1) synthesize relatively recent empirical findings pertaining to annual statistically-based cash-flow prediction models; (2) assess the comparative advantages and disadvantages of using specific statistically-based cash-flow prediction models that have been employed in this developing literature; and (3) provide a set of specific recommendations regarding future academic research on this important topic. The timing of our work should be beneficial to the on-going efforts of academics to refine extant statistically-based cash flow prediction models such that predictive ability is enhanced.

Cash-flow prediction has been a primary consideration to standard-setting bodies in accounting. According to the Statement of Financial Accounting Concepts No. 1 [Financial Accounting Standards Board (FASB) 1978, par. 37] "...financial reporting should provide information to help investors, creditors, and others assess the amount, timing, and uncertainty of prospective net cash inflows to the related enterprise." More recently, the International Accounting Standards Board (IASB) and FASB issued a joint statement emphasizing that: "...an entity's investors and creditors (both present and potential) are directly interested in the amounts, timing, and uncertainty of their cash flows..." (IASB, FASB 2006 p. 18). Clearly, both

professional bodies have underscored the continued importance of the linkage between cash-flow predictions and financial reporting objectives.

The prediction of cash flows is germane to a wide set of economic decisions faced by analysts, managers, and researchers in accounting and finance. Bowen et al. (1986) suggest that cash-flow predictions serve as inputs in such diverse settings as: (1) distress prediction, (2) risk assessment, (3) determination of credit ratings, (4) valuing closely-held companies, and (5) assessing the differential security market effects of net earnings versus cash flows. Additionally, researchers have identified proxies for the market's expectation of cash flow from operations (CFO) in an attempt to separate firms' cash-flow "surprises" from accrual "surprises" (Defond and Hung 2003). The former are viewed as potentially value relevant while the latter are transitory shocks without value implications. We observe that one's ability to disentangle such permanent and transitory shocks is entirely dependent upon the descriptive validity and predictive power of the statistical proxy for the market's expectation of CFO. Therefore, refinements in the modeling of statistically-based CFO prediction models are needed to advance research in this area.

Statistically-based CFO forecasts are also particularly salient because they have the potential to be less biased than those provided by sell-side analysts. O'Brien (1988) and Dreman and Berry (1995), among others, have documented that analysts are overly optimistic (biased) in their earnings projections. To the extent that such biases also influence analysts' CFO projections, statistically-based forecasts of CFO provide a valuable alternative to researchers interested in the predictive ability of CFO data. In a related context, Barniv, Myring and Thomas (2005) provide relatively recent empirical evidence that analysts' CFO forecasts are only issued for 1% of their U. S. sample firms for which analysts' earnings forecasts are available. Therefore, it appears that statistically-based CFO forecasts are currently the *only* widespread source of CFO expectational data thereby increasing our interest in the enhancement of their predictive power.

The remainder of this article is organized as follows. The next section includes a summary of extant statistically-based models that have been employed in annual CFO

prediction studies. Next, we provide a critical assessment of the methodological issues that surround the prediction of annual CFO data. We then make specific recommendations for future research pertaining to the statistical modeling of CFO data and provide some concluding remarks.

II. CFO Expectation Models

Research assessing the efficacy of statistically-based annual cash-flow prediction models has been championed by Dechow et al. (1998), and Barth et al. (2001). These studies exhibit considerable analytical rigor in developing linkages between a set of independent variables and the dependent variable, CFO.[1] Both studies model the time-series properties of annual income statement and balance sheet components and develop lead/lag relationships between net earnings and/or accruals with respect to future realizations of CFO. Specifically, Dechow et al. (1998) employed firm-specific time-series regression models where future values of annual CFO were regressed upon current annual CFO and *aggregate* annual net earnings during the 1963–1992 test period. Since most of their database preceded the issuance of SFAS #95, Dechow et al. employed a proxy for the reported CFO series that was constructed via an algorithm that adjusted net earnings by adding/subtracting non-cash expenditures and changes in working capital. Overall, their empirical results suggest that aggregate earnings are a better predictor of future CFO than past values of CFO.

While Barth et al. (2001) adopted the overall modeling structure popularized by Dechow et al. (1998), they provided some important methodological refinements. First, Barth et al.'s period of analysis (1987–1996) allows them to employ annual CFO data reported in accordance with SFAS No. 95 as opposed to using an algorithm to obtain a proxy for the CFO series like Dechow et al. Second, Barth et al. disaggregate earnings into cash flows and six major accrual components (e.g., change in accounts receivable, change in accounts payable, change in inventory, depreciation, amortization, and other accruals) to determine whether employment of accruals in CFO prediction models enhance explanatory power. This process is consistent with the notion that the accruals which are imbedded in the aggregate net earnings figure

provide implicit signals of future cash inflows and outflows. Third, they estimate their CFO prediction models cross-sectionally as opposed to using the time-series approach employed by Dechow et al. Overall, their empirical results suggest that disaggregated accruals are a better predictor of future CFO than aggregate earnings. That is, they provide evidence that disaggregating net earnings into six major accrual components enhances the predictive ability of CFO.

More recent work like Yoder (2007), Cheng and Hollie (2008) and Luo (2008) serves to underscore the continued interest in developing more accurate CFO prediction models. Yoder extends the cross-sectional annual CFO prediction model developed by Barth et al. by including cash flow implications of growth in future sales rather than relying upon the random walk assumption employed by Barth et al. Cheng and Hollie investigate the persistence of core and non-core cash flow components and present evidence that disaggregating CFO improves in-sample descriptive goodness of fit. Finally, Luo decomposes CFO into individual sources (i.e., nonrecurring tax benefits of exercised nonqualified employee stock options, cash expenditures pertaining to investing activities, and cash proceeds from the sale or securitization of accounts receivable) and shows that in-sample prediction is enhanced. While these more recent efforts provide useful insights, they rely exclusively upon cross-sectional estimation procedures and employ in sample goodness-of-fit measures as proxies for predictive performance. We argue against the efficacy of such methodological choices in the next section.

The aforementioned studies have popularized the employment of the following annual statistically-based CFO prediction models: an aggregate earnings model [Equation (1)] which is relatively parsimonious and relies upon aggregate net earnings and a considerably more complex disaggregated accrual model [Equation (2)], both of which are stipulated below:

$$CFO_{i,t+1} = b_0 + \sum_{\tau=0}^{k} b_{t-\tau} EARN_{i,t-\tau} + u_{it} \qquad (1)$$

$$CFO_{i,\,t+1} = c_0 + c_1 CFO_{it} + c_2 \Delta AR_{it} + c_3 \Delta INV_{it} + c_4 \Delta AP_{it} + c_5 DEPR_{it} + c_6 AMORT_{it} + c_7 OTHER_{it} + u_{it} \quad (2)$$

where i and t denote firm and year, τ ranges between 0 and 2, EARN is net earnings before extraordinary items, Δ AR is a change in accounts receivable, Δ INV is a change in inventory, Δ AP is a change in accounts payable, DEPR is depreciation expense, AMORT is amortization expense, and OTHER is the aggregate of remaining accruals not specifically detailed above.

III. Methodological Issues

The statistical modeling of annual CFO data is replete with methodological issues that may affect the external validity of each study's findings. First, researchers need to acquire a database of CFO numbers on which they conduct their statistical modeling. Earlier work employs algorithmic proxies of CFO using relatively simplistic algorithms that rely upon financial statement subcomponents. More recent work, however, relies upon CFO data reported in accordance with SFAS No. 95. Second, statistical models are either estimated on a time-series basis or cross-sectionally. Each estimation procedure has comparative advantages and disadvantages. Third, predictive ability has been assessed via descriptive goodness-of-fit criteria (i.e., adjusted r^2) or, more preferably, through the employment of an inter-temporal holdout period on which the forecast accuracy of statistical models may be determined. Finally, predictive performance has been evaluated by generating either one-step ahead and/or multi-step ahead CFO predictions. Subsequent sections discuss the specific methodological issues pertaining to each of the aforementioned issues.

A. *Use of Algorithmic Proxies of CFO versus CFO Data Reported in Accordance with SFAS #95*

Earlier empirical work in this area employs relatively simplistic algorithms (Bernard and Stober 1989; Lorek, Schaefer, and Willinger 1993; Finger 1994; Lorek and Willinger 1996; Dechow et al. 1998; and Kim and Kross 2005, among others) that rely upon subcomponents of the financial statements to derive a proxy for CFO using the indirect method. Basically, non-cash expenses like depreciation and amortization as well as certain changes in net working capital are added/subtracted from operating income to develop a surrogate for CFO. We note, however, that SFAS No. 95 was promulgated in 1988 mandating that all listed firms report a cash-flow statement displaying CFO effective for fiscal years ending after July 15, 1988. Relatively recent empirical work like Barth et al. employs CFO data reported in accordance with SFAS No. 95.

Kim and Kross (2005) argue that CFO reported under the auspices of SFAS No. 95 guidelines are less subject to measurement error than algorithmic proxies utilized in previous studies. Hribar and Collins (2002) provide empirical evidence that there are systematic differences between SFAS No. 95 CFO data and proxies derived using simplistic algorithms. Therefore, the findings of earlier studies employing algorithmic proxies for CFO must be viewed with some skepticism (e.g., Bernard and Stober 1989; Lorek, Schaefer and Willinger 1993; Finger 1994; Dechow et al. 1998, and Lorek and Willinger 1996, among others) and future empirical work should only employ data generated via SFAS No. 95 guidelines. In essence, the external validity of CFO predictive ability work using algorithmic proxies for CFO is lessened given the possibility of measurement error being induced by the simplistic algorithm. Studies executed subsequent to SFAS No. 95 like Barth et al. (2001) avoid such measurement error problems by utilizing data reported in conformity with SFAS No. 95.

B. Cross-Sectional versus Time-Series Estimation

Lorek, Schaefer and Willinger (1993), Lorek and Willinger (1996), and Dechow et al. (1998) estimate their cash-flow prediction models on a time-series basis whereas Bernard and Stober (1989), Barth et al. (2001), and Kim and Kross (2005) employ cross-sectional estimation procedures. Each estimation approach provides comparative advantages and disadvantages. Studies invoking time-series estimation typically examine relatively smaller samples of firms given the need to construct time-series databases across years in order to estimate firm-specific parameters. For example, Lorek and Willinger (2009) obtain a time series of CFO data from 1989 to 2004 for a sample of 1,174 firms. While this approach benefits from the ability to capture firm-specific, contextual relationships between the independent variables and CFO, it may suffer from diminished generalizability. On the other hand, studies employing cross-sectional estimation typically have reduced data requirements thereby increasing sample size and potentially enhancing external validity. For example, Barth et al. (2001) acquire 10,164 firm/year observations in their study. Cross-sectional estimation, however, constrains parameters to be constant across firms and time, restricting parameter values from capturing any firm-specific information. Neill et al. (1991), among others, provide a particularly lucid discussion on the impact of restricting parameter values in this manner. They caution that cross-sectional estimation procedures may be inappropriate if the relationships between the independent variables and CFO are firm and/or industry specific. More recently, Lorek and Willinger (2009) provide descriptive evidence that the beta parameter in Equation (1) exhibits considerable firm-specific variability thus casting considerable doubt upon the propriety of cross-sectional estimation procedures that suppress such variability.[2] This evidence is particularly salient in that it underscores the comparative advantage that time-series estimation enjoys over cross-sectional estimation.

C. Descriptive Fit versus Predictive Ability

Kim and Kross (2005), among others, distinguish between "explanatory power" of statistically-based CFO models versus "out-of-sample" ex-ante predictive ability. The former

criterion may be subject to what is referred to as the "descriptive-predictive paradox" where descriptive fit may not imply predictive power due to potentially over-fitting the data. Watts and Leftwich (1977) and Zhang, Cao, and Schniederjans (2004) cite research that has consistently found that models that fit well descriptively may not perform well in forecasting settings. This casts considerable doubt upon the external validity of the findings of Bernard and Stober (1989) and Barth et al. (2001) who employ descriptive fit as the *singular* proxy for predictive ability. According to Kim and Kross (2005) caution must be exercised in imputing predictive characteristics to alternative CFO prediction models based solely on in-sample descriptive goodness-of-fit metrics. A more rigorous approach is to use data in an identification period to identify the structure of the prediction model and estimate its parameters. The predictive ability of the model is assessed, however, on data *not used* in model identification or estimation employing an inter-temporal holdout period. Kim and Kross (2005) and Lorek and Willinger (2009) are examples of this latter approach.

D. One-Step Ahead versus Multi-Step Ahead Predictions

A considerable number of cash-flow prediction studies limit the prediction of future CFOs to one step-ahead (e.g., Lorek, Schaefer and Willinger 1993, Lorek and Willinger 1996, and Kim and Kross, 2005, among others). While empirical evidence on predictive power derived from such studies is of importance, financial statement analysis texts (e.g., Palepu and Healy 2008 and Penman 2007, among others) underscore the need for multiple year-ahead CFO forecasts as inputs to firm valuation. In fact, Lundholm and Sloan (2007) employ valuation software that invokes n-step ahead forecasts (i.e., n=5, 10 or 20) for CFO predictions. Due to the crucial linkage between such multi-step ahead predictions of CFO and firm valuation methodologies, predictive ability studies should concentrate upon providing predictive evidence using such longer horizons. Consistent with this notion, Barth et al. (2001) assess predictive performance (i.e., using their proxy of descriptive goodness-of-fit as discussed above) across multiple years by assessing predictive ability across one-thru-four years-ahead. Since we are unaware of any publicly available source of multi-step ahead CFO predictions attributed to

sell-side analysts, further research refining extant statistically-based CFO prediction models that generate relatively accurate n-step ahead CFO forecasts is particularly relevant.

1V. Concluding Remarks

The academic literature devoted to specifying statistically-based, annual CFO prediction models has been growing due to the importance of generating relatively accurate CFO predictions for analysts, managers, standard setters and researchers in accounting and finance. While considerable strides have been made in the development of statistically-based cash-flow prediction models, additional research is necessary to refine extant models. Our review of this literature suggests that future improvements in CFO prediction models are most likely if time-series versus cross-sectional estimation procedures are invoked, inter-temporal holdout periods are employed vis-à-vis exclusive reliance on goodness-of-fit measures, and multi-step ahead rather than one-step ahead predictions are generated. Such refinements in CFO predictive ability studies will enhance methodological rigor, increase the confidence associated with predictive findings by enhancing external validity, and provide empirical results of direct interest to a host of users of CFO data.

REFERENCES

Barniv, R., M. Myring, and W. Thomas. 2005. "The Association Between the Legal and Financial Reporting Environments and Forecast Performance of Individual Analysts." *Contemporary Accounting Research* 22 No. 4: 727-758.

Barth, M. E., D. P. Cram, and K. Nelson. 2001. "Accruals and the Prediction of Future Cash Flows." *The Accounting Review* 76 (January): 27-58.

Bernard, V. L. and T. L. Stober. 1989. "The Nature and Amount of Information in Cash Flows and Accruals." *The Accounting Review* Vol. 64, No. 4: 624-652.

Bowen R., D. Burgstahler, and L. Daley. 1986. "Evidence on the Relationships Between Earnings and Various Measures of Cash Flow." *The Accounting Review* 61: 713-725.

Brown, S., K. Lo, and T. Lys. 1999. "Use of R^2 in Accounting Research: Measuring Changes in Value Relevance over the Last Four Decades." *Journal of Accounting and Economics* 28: 83-115.

Cheng, C. S. A. and D. Hollie. 2008. "Do Core and Non-core Cash Flows from Operations Persist Differentially in Predicting Future Cash Flows?" *Review of Quantitative Finance and Accounting* 31: 29-53.

Dechow, P. M., S. P. Kothari, and R. L. Watts. 1998. "The Relation Between Earnings and Cash Flows." *Journal of Accounting and Economics 25: 133-168.*

Defond, M. and Hung, M. (2003). "An Empirical Analysis of Analysts' Cash Flow Forecasts." *Journal of Accounting and Economics* 35 (1): 73-100.

Dreman, D. N. and M. A. Berry. 1995. "Forecasting Errors and Their Implications for Security Analysis." *The Financial Analysts Journal* (May-June): 30-41.

Financial Accounting Standards Board. 1978. Statement of Financial Accounting Concepts No. 1: *Objectives of Financial Reporting by Business Enterprises*. Stamford, Ct: FASB.

Financial Accounting Standards Board. 1987. Statement of Financial Accounting Standards No. 95: *Statement of cash flows*. Stamford, Ct. FASB.

Finger, C. 1994. "The Ability of Earnings to Predict Future Earnings and Cash Flow." *Journal of Accounting Research* Vol. 64, No. 4: 624-652.

Hribar, P. and D. W. Collins. 2002. "Errors in Estimating Accruals: Implications for Empirical Research." *Journal of Accounting Research* 40: 105-134.

International Accounting Standards Board and Financial Accounting Standards Board (IASB/ FASB). 2006. *Preliminary views on an improved conceptual framework for financial reporting: The objective of financial reporting and qualitative characteristics of decision-useful financial reporting information.* International Accounting Standards Committee. London.

Kim, M. and W. Kross. 2005. "The Ability of Earnings to Predict Future Operating Cash Flows Has Been Increasing – Not Decreasing." *Journal of Accounting Research* 43: 753-780.

Lorek, K. S., T. F. Schaefer, and G. L. Willinger. 1993. "Time-series Properties and Predictive Ability of Funds Flow Variables." *The Accounting Review* 68: 151-163.

Lorek, K. S. and G. L. Willinger. 1996. "A Multivariate Time-Series Prediction Model for Cash-Flow Data." *The Accounting Review* 71: 81-101.

Lorek, K. S. and G. L. Willinger. 2009. "New Evidence Pertaining to the Prediction of Operating Cash Flows." *Review of Quantitative Finance and Accounting* 32: 1-15.

Lundholm, R. and R. Sloan. 2007. *Equity Valuation and Analysis. 2nd ed. McGraw-Hill.*

Luo, M. 2008. "Unusual Operating Cash Flows and Stock Returns." *Journal of Accounting and Public Policy* 27: 420-429.

Neill, J. D., T. F. Schaefer, P. R. Bahnson, and M. E. Bradbury. 1991. "The Usefulness of Cash Flow Data: A Review and Synthesis." *Journal of Accounting Literature 10: 117-150.*

O'Brien, P. 1988. "Analysts' Forecasts as Earnings Expectations." *Journal of Accounting and Economics* (January): 53-83.

Palepu, K. G. and P. M. Healy. 2008. *Business Analysis & Valuation.* 4th ed. Southwestern College Publishing.

Penman, S. H. 2007. *Financial Statement Analysis and Security Valuation.* 3rd ed. McGraw-Hill.

Watts, R. L. and R. W. Leftwich. 1977. "The Time Series of Annual Accounting Earnings." *Journal of Accounting Research* (Autumn): 253-271.

Yoder, T. 2007. "The Incremental Predictive Ability of Accrual Models with Respect to Future Cash Flows." Working Paper, Mississippi State University.

Zhang, W., Q. Cao, and M. J. Schniederjans. 2004. "Neural Network Earnings Per Share Forecasting Models: A Comparative Analysis of Alternative Methods." *Decision Sciences* 35 (2): 205-237.

ENDNOTES

[1] Dechow et al. (1998) discuss the difficulties of developing quarterly CFO prediction models such as seasonality and the re-specification of lead/lag relationships between interim accrual data and quarterly CFO. See Lorek and Willinger (1996) for an example of such work.

[2] Whether the advantages of time-series estimation over cross-sectional estimation exist regardless of the specific nature of the CFO series (i.e., levels, per-share, or deflated by total assets) is an interesting empirical issue. For example, Brown Lo, and Lys (1999) provide evidence that the use of undeflated variables in regression models where scale effects are present may result in an omitted correlated variable problem with biased coefficients and forecast errors.