

Fraud Detection using Autobox's™ Automatic Intervention Detection

David P. Reilly, Vice President

Automatic Forecasting Systems

<http://www.autobox.com>

Paul Sheldon Foote, Ph.D.

California State University, Fullerton

<http://business.fullerton.edu/pfoote>

David H. Lindsay, Ph.D., C.F.E., C.P.A., C.I.S.A.

California State University, Stanislaus

Annhenrie Campbell, Ph.D., C.P.A., C.M.A., C.G.A.

California State University, Stanislaus

International Symposium on Forecasting 2002, Dublin, Ireland

June 23 – 26, 2002

Introduction

Current auditing standards require independent auditors to evaluate the risk that financial statements may have been fraudulently altered. Internal auditors are usually expected to search for possible frauds within the company. Universities have responded to the enhanced demand for fraud detection skills with the addition of forensic accounting and fraud auditing courses to the curriculum [Davia, 2000]. Audit tools and technologies for fraud detection have developed along with the evolution of financial reporting from a periodic, manual activity to an ongoing, technology-intensive process.

Among the almost endless variety of possible frauds, many are perpetrated by altering a company's financial records. These direct modifications are termed "interventions" in the terminology of the Box-Jenkins time series analysis examined here. This study evaluates whether automatic intervention detection(AID) can be effectively used to distinguish companies with fraudulent reported data from those with no indication of fraudulent reports.

Fraud is costly. The Association of Certified Fraud Examiners' current estimate of the annual direct cost of fraud exceeds \$600 billion [Albrecht and Albrecht, 2002]. The cost in lost investor and customer confidence is incalculable. Auditors face direct costs from fraud whenever shareholders sue them upon discovery of ongoing fraud [Wells, 2002].

Therefore, auditors must assess the risk of fraud in each audit engagement. Most audit tools have been developed from the perspective of this overall evaluation rather than primarily as aids in the search for specific frauds. The best audit tool for fraud discovery may be the experience, background and expertise of the individuals on each audit engagement [Moyes and Hasan, 1996].

"Red flag questionnaires" have long been successful decision aids to direct both novice and skilled auditors' attention to specific fraud indicators [Pincus, 1989]. The red flags are both incentives for employees of an audit client to commit fraud as well as opportunities that may allow fraud to occur [Apostolou et al, 2001].

Auditors also have statistical tools available, most of which can be applied simply by using a personal computer. Discovery sampling, long used by auditors, is used to test for the existence of errors and inaccuracies in randomly sampled sets of transactions. Signs of fraud deserving further investigation may be unearthed through this simple, direct process.

Another statistical approach compares the frequency of digits in reported data to the Benford Distribution, the expected frequency of the digits in numbers naturally issued from some underlying process such as financial record-keeping. Purposefully altered numbers rarely conform to the Benford statistic so a deviation may indicate the need for further examination [Busta and Wienberg, 1998].

Other digital analyses of reported amounts use computers to search through data files for specific items such as even dollar quantities or amounts just under approved limits. A kind of ratio analysis is available which compares the relative incidence of specific amounts, for example the ratio of the highest to the lowest value in the data [Coderre, 1999].

Advanced computer-based tools are being developed for reviewing very large quantities of data. Data mining software can be used to sift through entire databases and sort the information along various parameters to locate anomalous patterns that may require further investigation. Database programming expertise is costly, but less expensive off-the-shelf data mining software can be used in the audits of smaller companies. In all environments, outcomes tend to be better when a fraud expert is available to review the program output [Albrecht and Albrecht, 2002].

An audit variant of data mining is data extraction in which auditors use software to collect quantities of specific information for review. Extensive training is also needed for effective use of data extraction tools. Such tools are frequently applied when a fraud is already suspected rather than as a routine screening process [Glover, 2000].

Continuous auditing processes are being developed to support ongoing real-time financial reporting. Such auditing processes may use warehoused data generated by mirroring live data for audit purposes [Rezaee et al, 2002]. Automatic, real-time testing for fraud using neural network technology is already successfully identifying fraudulent credit card and cell phone transactions [Harvey, 2002]. Computerized neural network

simulations are applied to differentiate between patterns inherent in a set of "training" data and sets of test or live data. Neural networks can find discrepancies in data patterns in situations with high quantities of similar and repetitive transactions.

Fraud detection tools are no longer limited to aids for audit specialists to use to review dated information. Newer techniques are more likely to operate on information as it is being produced or released. Automatic, unsupervised detection methodologies will integrate the detection of fraud and the review of financial information to enhance its value and reliability to the user [Pathria, 1999].

Intervention Detection

If someone wants to commit fraud by modifying the accounting information system, such modifications would be termed interventions in the terminology of Box-Jenkins time series analysis. The purpose of this study was to evaluate whether intervention detection could be used successfully to distinguish between real companies with reported fraudulent data and real companies with no reports of fraudulent data.

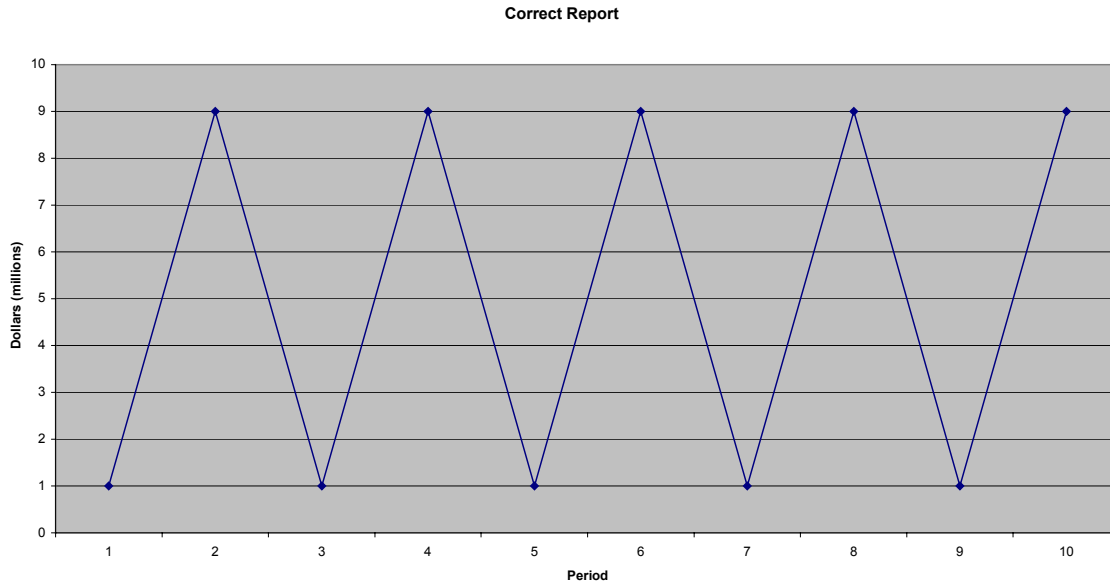
This study proceeds as follows:

1. For a hypothetical company, show what a spreadsheet user would see in the way of charts and statistics for correct data and for data containing a single fraudulent intervention.
2. For the hypothetical company, explain how to analyze the results from using a software system designed for intervention detection.
3. Describe an experiment of using the data from real firms. Some of the firms in the experiment had reported cases of fraud. Other firms have had no reported cases of fraud.
4. Conclude whether or not intervention detection can detect fraud in the more difficult cases of real firms with data containing one or more fraudulent interventions.

A Hypothetical Cyclical Company

Assume that a hypothetical cyclical company should have reported the following results in millions of dollars for the last 10 periods: 1, 9, 1, 9, 1, 9, 1, 9, 1, 9.

A Microsoft Excel user should have seen this chart:



Using Microsoft Excel and an 80% confidence level, the summary statistics for a correct report would have been:

Correct Statistics

Mean	5
Standard Error	1.333333333
Median	5
Mode	1
Standard Deviation	4.216370214
Sample Variance	17.77777778
Kurtosis	-2.571428571
Skewness	0
Range	8
Minimum	1

Maximum	9
Sum	50
Count	10
<u>Confidence Level(80.0%)</u>	<u>1.844038403</u>

Confidence Level, for Microsoft Excel users, is based upon:

= CONFIDENCE (alpha, standard_dev, size)

Estimate the confidence interval for a population mean by a range on either side of a sample mean. Alpha is the confidence level used to compute the confidence level. For this study, we used an alpha of 0.20, indicating an 80% confidence level. Standard_dev is the population standard deviation for the data range (assumed to be known). Size is the sample size (an integer).

For the hypothetical cyclical company, we used:
 Alpha = 0.20
 Standard_dev = 4.216370214
 Size = 10

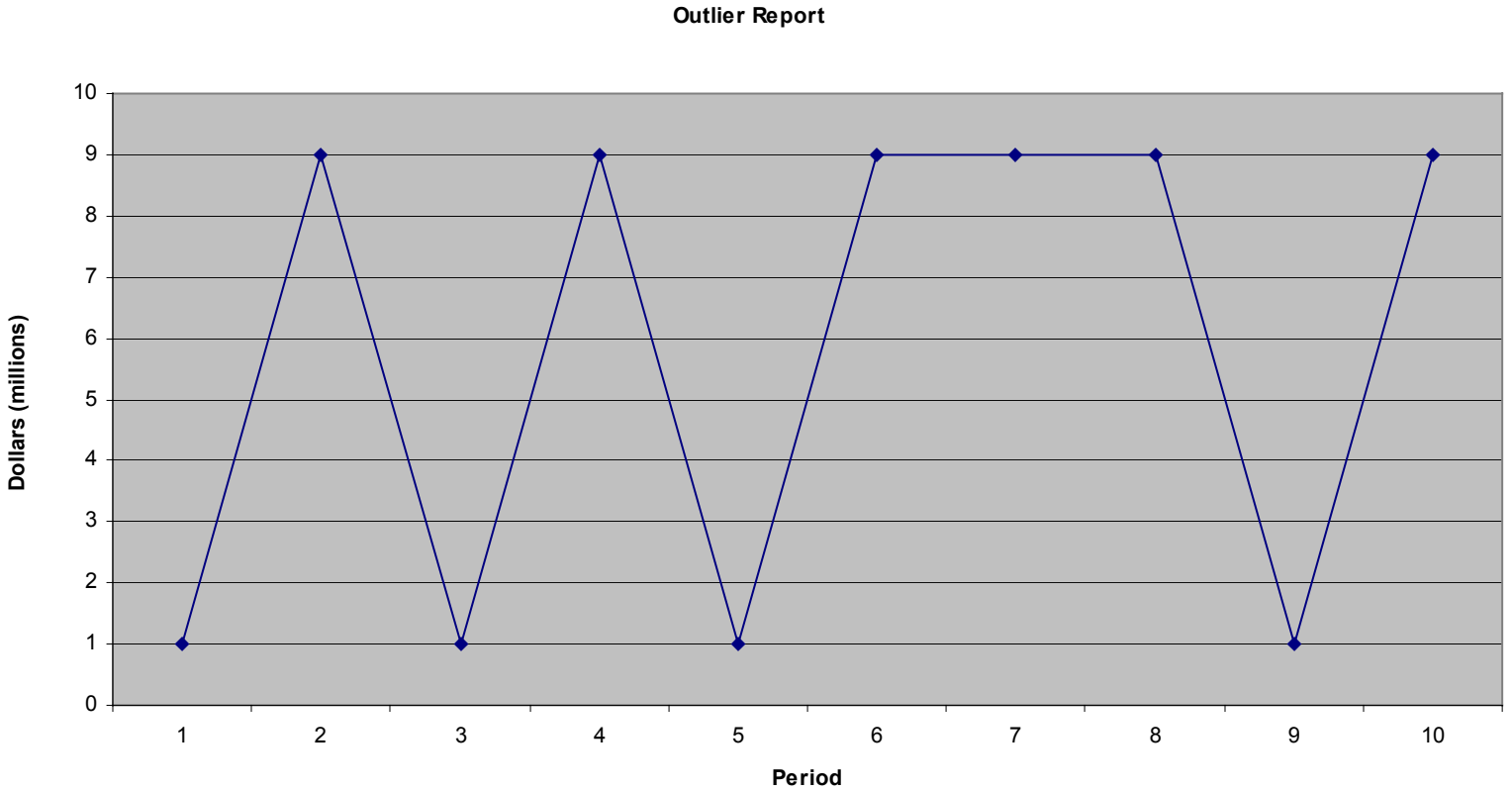
The resulting confidence level of 1.844038403 indicates a range of:

$$5 + 1.844038403 = 6.844038403$$

$$5 - 1.844038403 = 3.155961597$$

A Hypothetical Cyclical Company with a Single Outlier

Suppose, instead, the hypothetical cyclical company issued only a single, fraudulent report in Period 7 by reporting \$9 million instead of \$1 million. The resulting Microsoft Excel chart of the reports for 10 years would be:

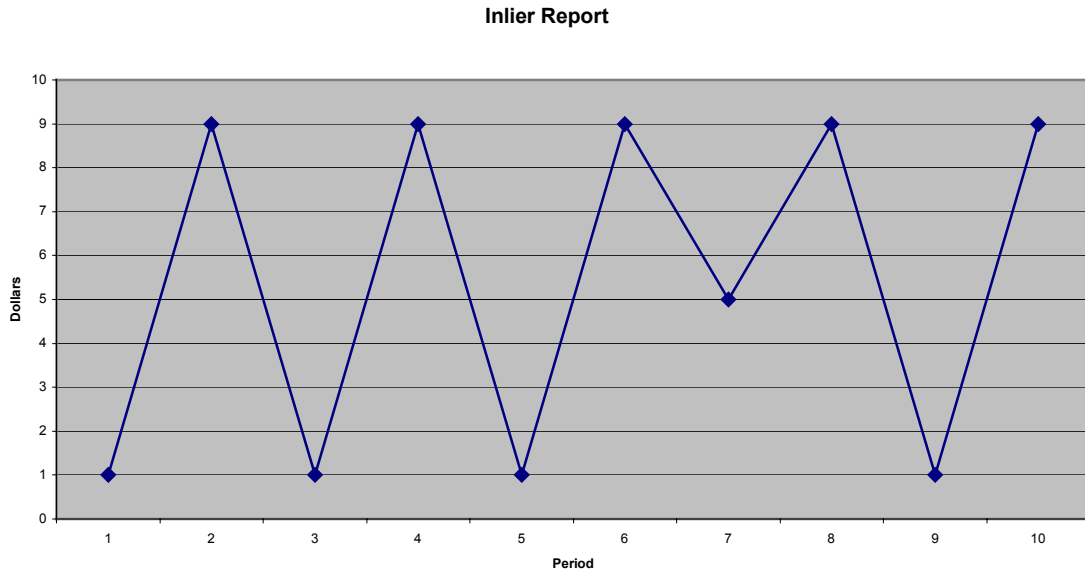


Using Microsoft Excel and an 80% confidence level, the summary statistics for an outlier report would have been:

<i>Outlier Statistics</i>	
Mean	5.8
Standard Error	1.306394529
Median	9
Mode	9
Standard Deviation	4.131182236
Sample Variance	17.06666667
Kurtosis	-2.276785714
Skewness	-0.484122918
Range	8
Minimum	1
Maximum	9
Sum	58
Count	10
Confidence Level(80.0%)	1.806781262

A Hypothetical Cyclical Company with a Single Inlier

Suppose, instead, the hypothetical cyclical company issued only a single, fraudulent report in Period 7 by reporting \$5 million instead of \$1 million. The resulting Microsoft Excel chart would have been:



Using Microsoft Excel and an 80% confidence level, the summary statistics for an inlier report would have been:

Inlier Statistics

Mean	5.4
Standard Error	1.257864151
Median	7
Mode	9
Standard Deviation	3.977715704
Sample Variance	15.82222222
Kurtosis	-2.300034267
Skewness	-0.237276944
Range	8
Minimum	1
Maximum	9
Sum	54
Count	10
Confidence Level(80.0%)	1.739662351

From the two cases (outlier and inlier example), we can see that the confidence limit has been unchanged along with the standard deviation. Thus standard procedures are useless in detecting fraud.

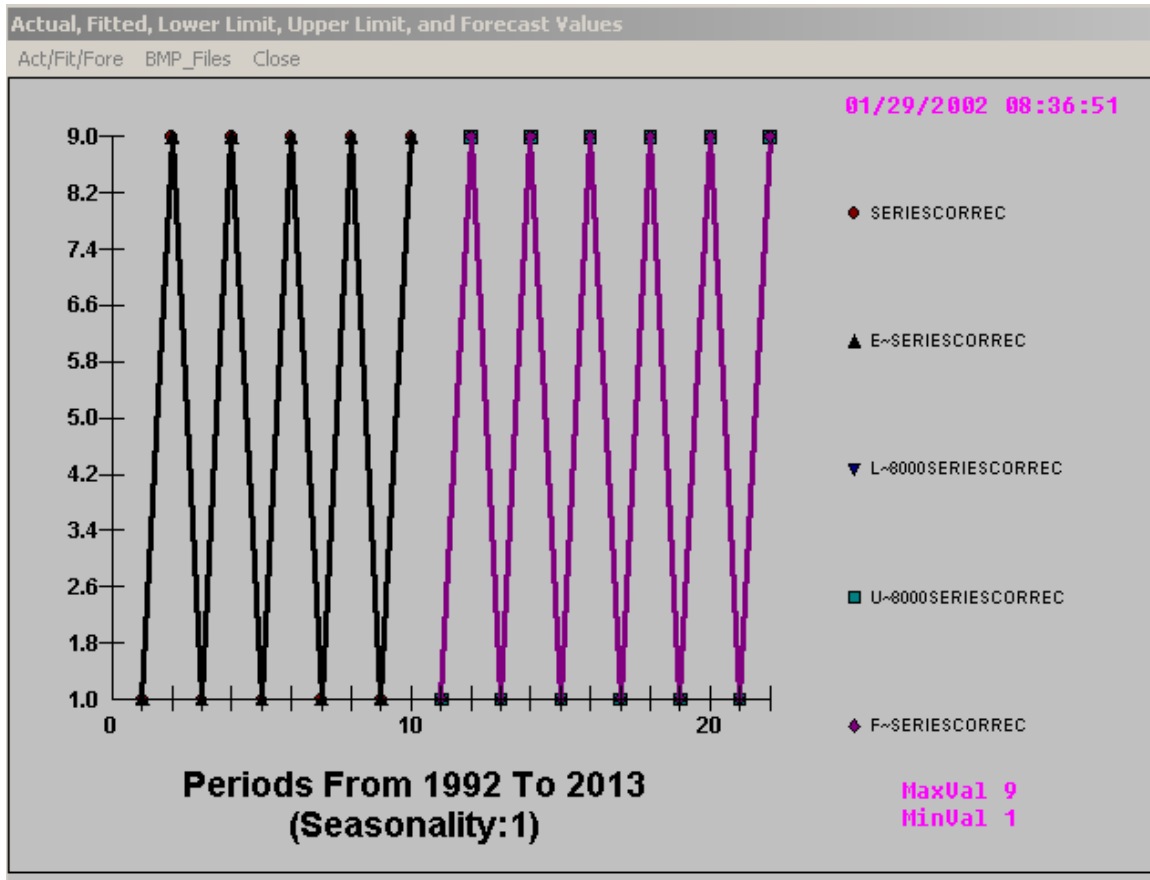
Automatic Fraud Detection for the Hypothetical Cyclical Company

If a company had strong cyclical patterns and a single, large, fraudulent intervention, then it would be easy to spot the suspicious interventions simply by generating charts of the data series. However, for real firms, there can be multiple interventions over time and no strong cyclical or seasonal patterns in the data. Fraud auditors need tools for identifying suspicious interventions in the accounting data of any company.

Automatic Forecasting Systems' Autobox™ (<http://www.autobox.com>) and FreeFore (<http://www.autobox.xom/freef.exe>) are able to detect many types of interventions, including: (1) outliers (2) inliers (3) steps (4) level changes (5) local time trends. In this study, we tested the ability of AID to distinguish companies known to have had fraud cases from those companies with no published indications of fraud.

Automatic Forecasts for the Correct Series

If the hypothetical cyclical company had issued correct reports each year for the period 1992–2001, Autobox™ would have generated the following forecasts:



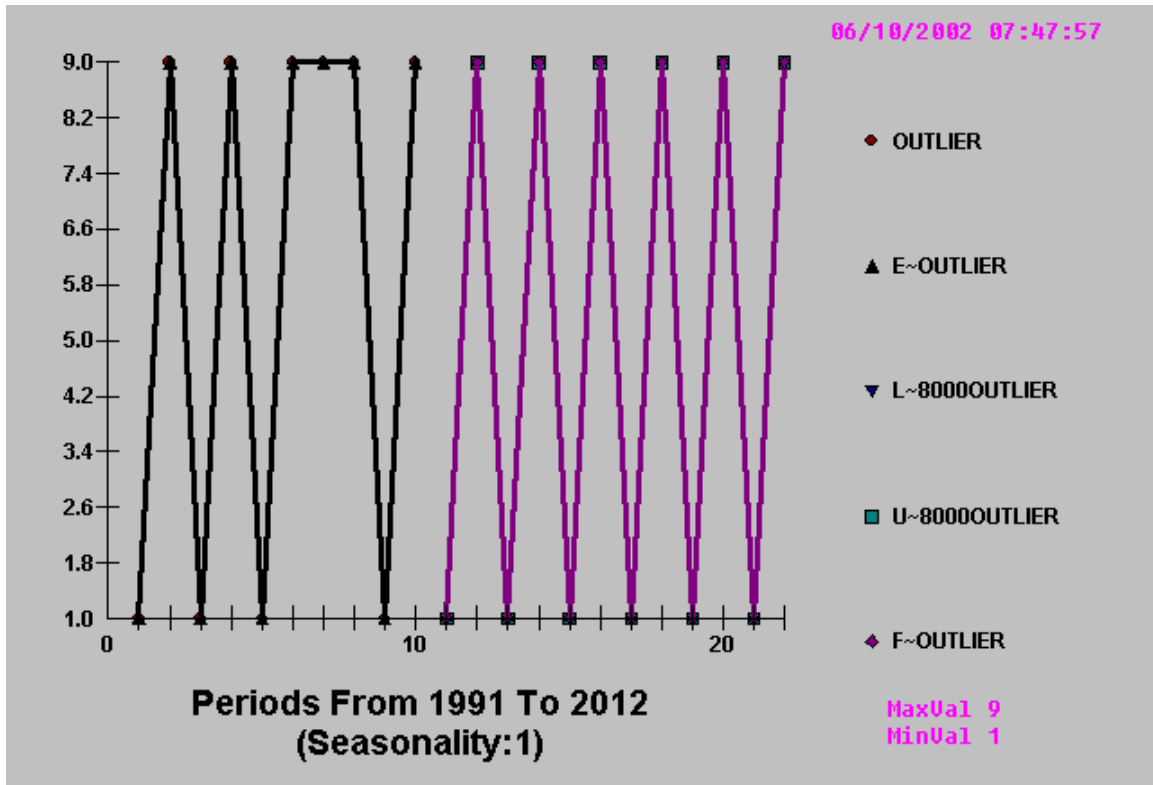
Note that Autobox™ identified the strong cyclical pattern of the data and continued that pattern in forecasts for future periods. So, all future forecasts are either 1 or 9. For those of you who are curious, the model is

$$Y(t)=10-Y(t-1) +a(t) .$$

This is also called “lumpy demand” and is quite prevalent in logistical studies.

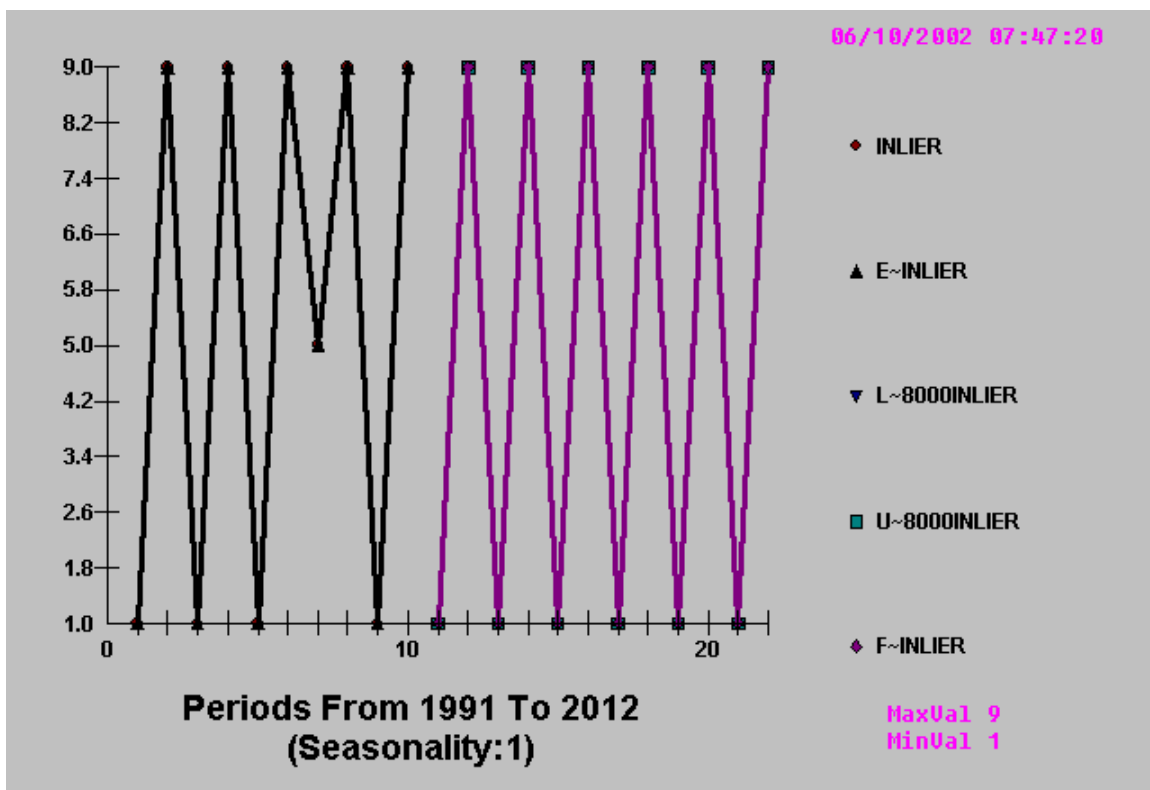
Automatic Forecasts for the Outlier Series

The insertion of one fraudulent report in a series of 10 years of reports has no major impact upon the automatic forecasts. The unusual value (“9”) at time period 8 (1998) is identified and replaced with the normative value of “1”. Forecasts are unaffected by this anomaly and are thus robust.



Automatic Forecasts for the Inlier Series

The fraudulent reporting of one inlier value over a 10-year historical period at time period 7 (1997) of the value “5” creates an interesting problem. If the overall mean is approximately “5”, how can the value of “5” be identified as unusual? The answer is simple as the Expected Value is not always equal to the mean and is based upon patterns in the series. If no pattern exists then the Expected Value is equal to the mean, but not otherwise. Autobox™ identifies the anomaly, replaces it with the Expected Value and proceeds.



Methodology

Eight companies, identified in the general press as having engaged in financial fraud, were selected. Newspaper articles outlining the specifics of each fraud were downloaded from NEXIS. Table 1 lists the eight fraud firms and the type of fraud each engaged in.

**Table 1
Company and Type of Fraud**

Company	Nature of the Fraud
Cendant	Inflated earnings and improper use of reserves
Con Agra	Improper revenue recognition of fictitious sales
Enron	Failure to disclose liabilities and improper recognition of revenue
Grace	Improper use of reserves to facilitate income smoothing
McKesson	Premature recognition of sales revenue
Rite Aid	Recognition of fictitious vender credits
Sunbeam	Fictitious sales and improper use of reserves
Waste Management	Improper revenue recognition

Each fraud firm was pair-matched with a non-fraud firm classified within the same SIC code. If data availability permitted the identification of multiple non-fraud firms, two such pair-match firms(if available) were randomly selected. The fraud firms and the pair-matched non-fraud firms are listed in Table 2.

**Table 2
Fraud Firms and Pair-Matches**

	Fraud Firm	Pair-Match 1	Pair-Match 2
1	Cendant	Advance Tobacco Products	Competitive Technologies
2	Con Agra	Sara Lee	Classica
3	Enron	Mercury Air Group	World Fuel Service
4	Grace	Great Lakes Chemical	None
5	McKesson	Bergen Brunswig	None
6	Rite Aid	Drug Emporium	None
7	Sunbeam	Decorator Industries	None
8	Waste Management	Rich Coast	Wastemasters

Financial Statements were obtained from the COMPUSTAT Annual Industrial File. Since this is an exploratory study, data was downloaded for all available data items. However, examination of the data on a firm by firm basis revealed that missing data. These fields were eliminated, since their data couldn't be compared across firms. The 45 items are defined in Table 3.

Table 3
COMPUSTAT DATA ITEMS TESTED

N1	CASH AND SHORT TERM INVEST.
N2	RECEIVABLES
N3	TOTAL CURRENT ASSETS
N4	TOTAL CURRENT LIABILITIES
N5	TOTAL ASSETS
N6	NET PLANT AND EQUIPMENT
N7	TOTAL LONG TERM DEBT
N8	TANGIBLE COMMON EQUITY
N9	NET SALES
N10	OPERATING INCOME BEFORE DEPRECIATION.
N11	DEPRECIATION AND AMORTIZATION EXPENSE
N12	INTEREST EXPENSE
N13	TOTAL INCOME TAXES
N14	SPECIAL ITEMS
N15	INCOME BEFORE SPECIAL ITEMS
N16	AVAIL FOR COMMON AFTER ADJ.
N17	COMMON SHARES OUTSTANDING
N18	CUMULATIVE ADJUSTMENT FACTOR
N19	CAPITAL EXPENDITURES
N20	INVESTMENTS IN OTHERS
N21	DEBT IN CURRENT LIABILITIES
N22	DEF. TAXES & INV. CREDIT(BS)
N23	RETAINED EARNINGS
N24	TOTAL INVESTED CAPITAL
N25	COST OF GOODS SOLD
N26	DEBT DUE IN ONE YEAR
N27	PRI EPS INCL. EXTRORD ITEMS
N28	SHARES USED TO COMPUTE EPS
N29	DILUTED EPS EX. EXTRAORDINARY ITEMS
N30	PRIMARY EPS EX. EXTRAORDINARY ITEMS
N31	COMMON EQUITY AS REPORTED
N32	NON-OPERATING INCOME/EXPENSE
N33	OTHER CURRENT ASSETS
N34	OTHER ASSETS
N35	ACCOUNTS PAYABLE
N36	OTHER CURRENT LIABILITIES
N37	DEFERRED TAXES (BAL. SHEET)
N38	OTHER LIABILITIES
N39	DEBT (CONVERTIBLE)
N40	DEBT (SUBORDINATED)
N41	DEBT (NOTES)
N42	DEBT (DEBENTURES)
N43	DEBT (OTHER LONG-TERM)
N44	CAPITALIZED LEASE OBLIGATION
N45	COMMON STOCK

For each fraud firm and its pair-match, ten years of data were downloaded. The ten-year period ended one year prior to the discovery of the fraud.

We counted the number of Interventions at the last point in time and used that to rank the companies within each of the eight sets. AUTOBOX correctly identified 6 of the 8 fraudulent firms.

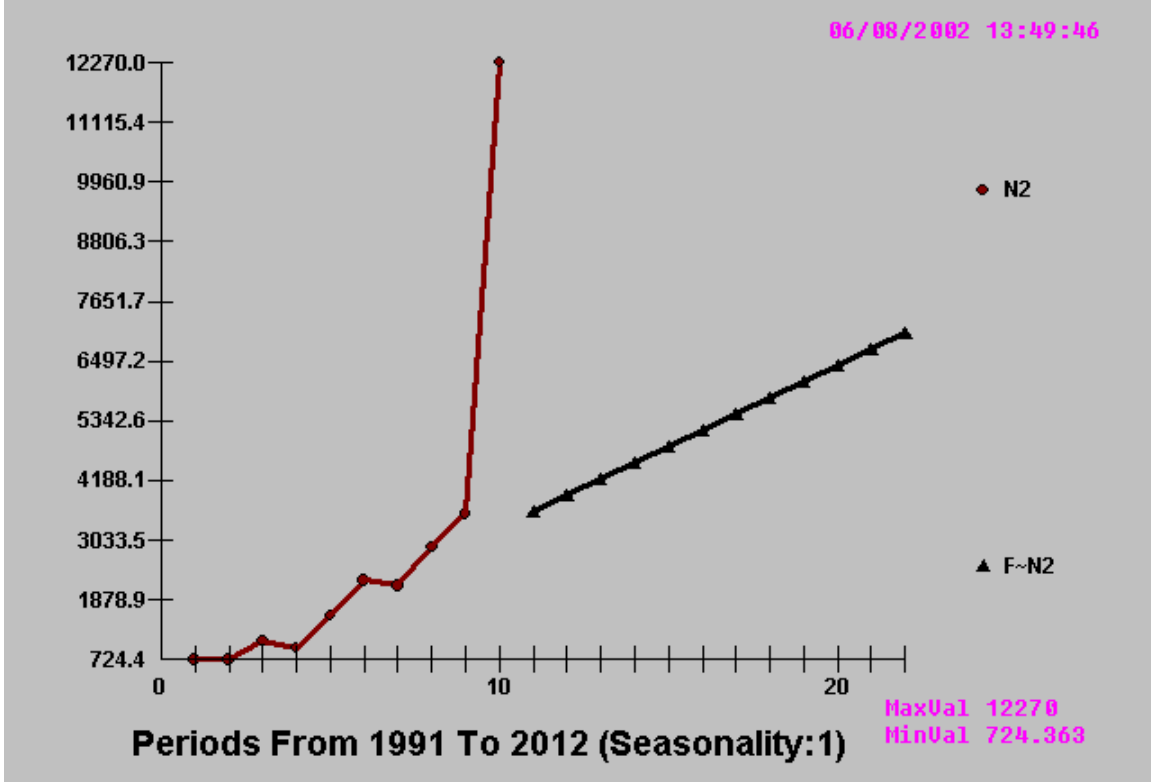
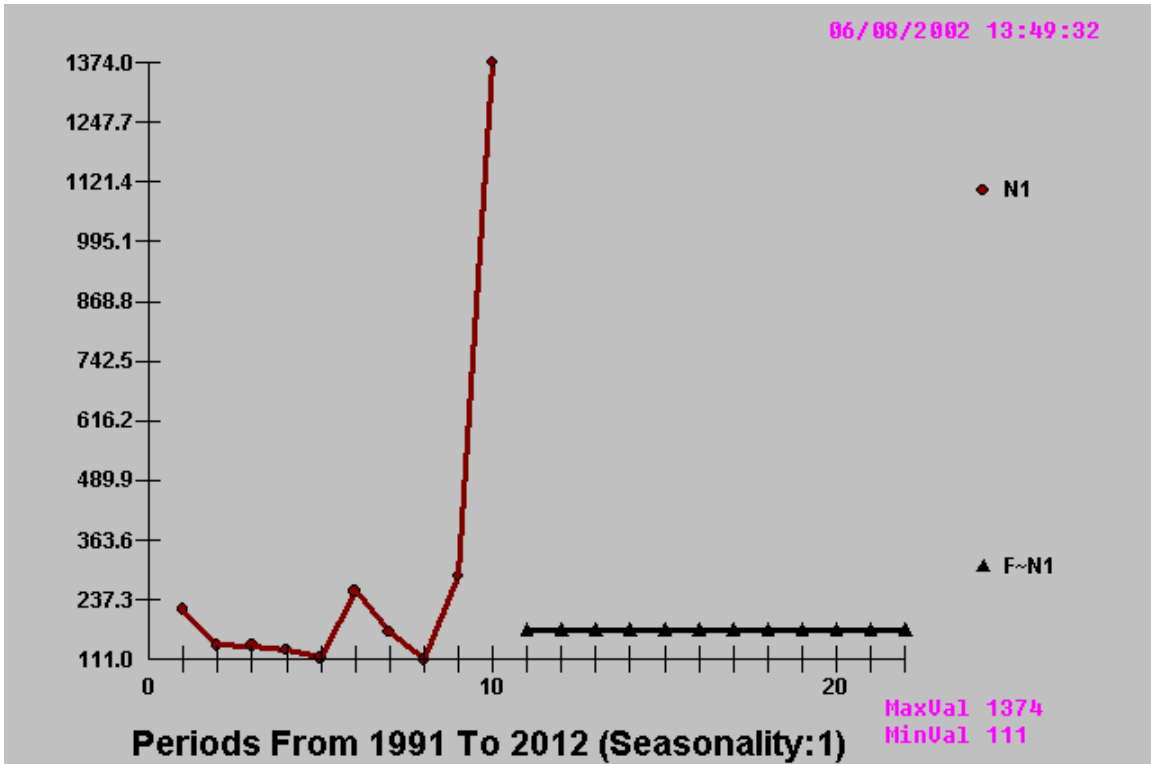
A tolerance threshold of 22 or more seems indicate a good benchmark of when companies seem to be in a “red flag” zone.

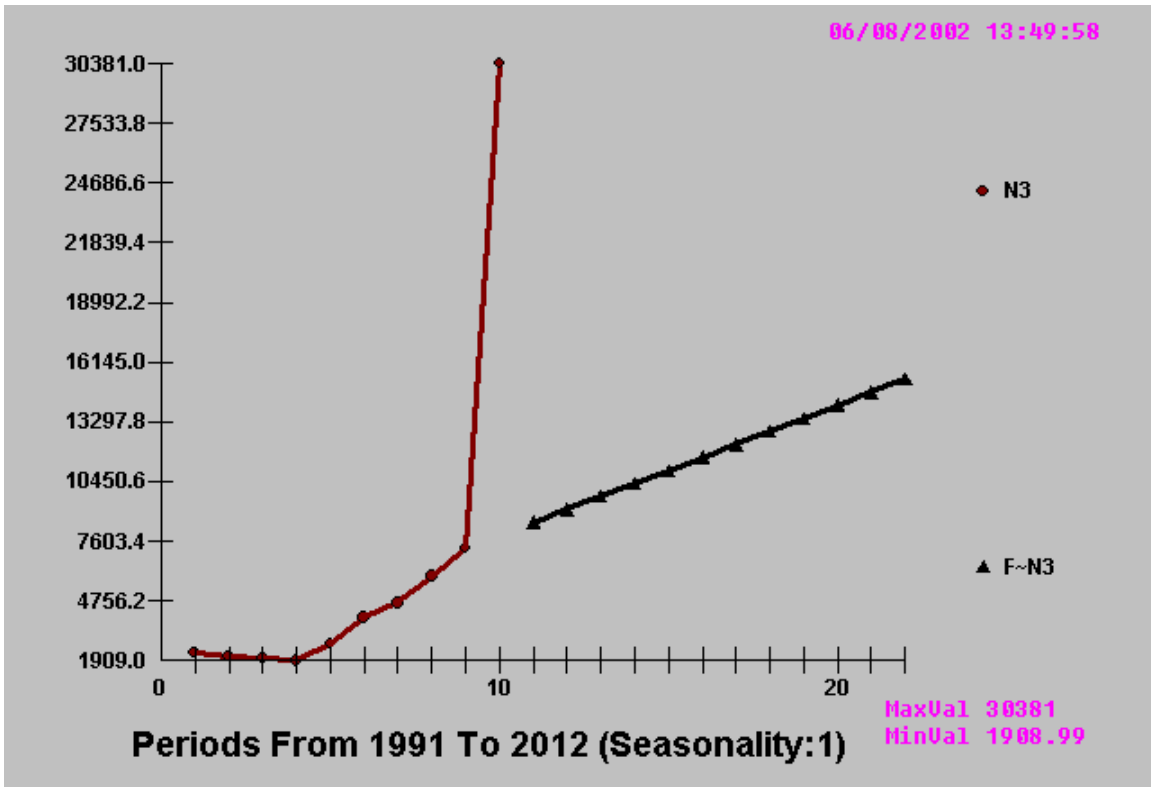
Number of 45 B/S Items found as interventions
in the year before fraud was publicly identified

Fraud Firm	Pair Match 1	Pair Match 2	
Cendant	26 Advance Tobacco Products	3 Competitive Technologies	12
Con Agra	9 Sara Lee	15 Classica	10
Enron	22 Mercury Air Group	8 World Fuel Service	15
Grace	21 Great Lakes Chemical	11	
McKesson	29 Bergen Brunswig	16	
Rite Aid	29 Drug Emporium	10	
Sunbeam	2 Decorator Industries	7	
Waste Mgt	38 Rich Coast	4 Wastemasters	22

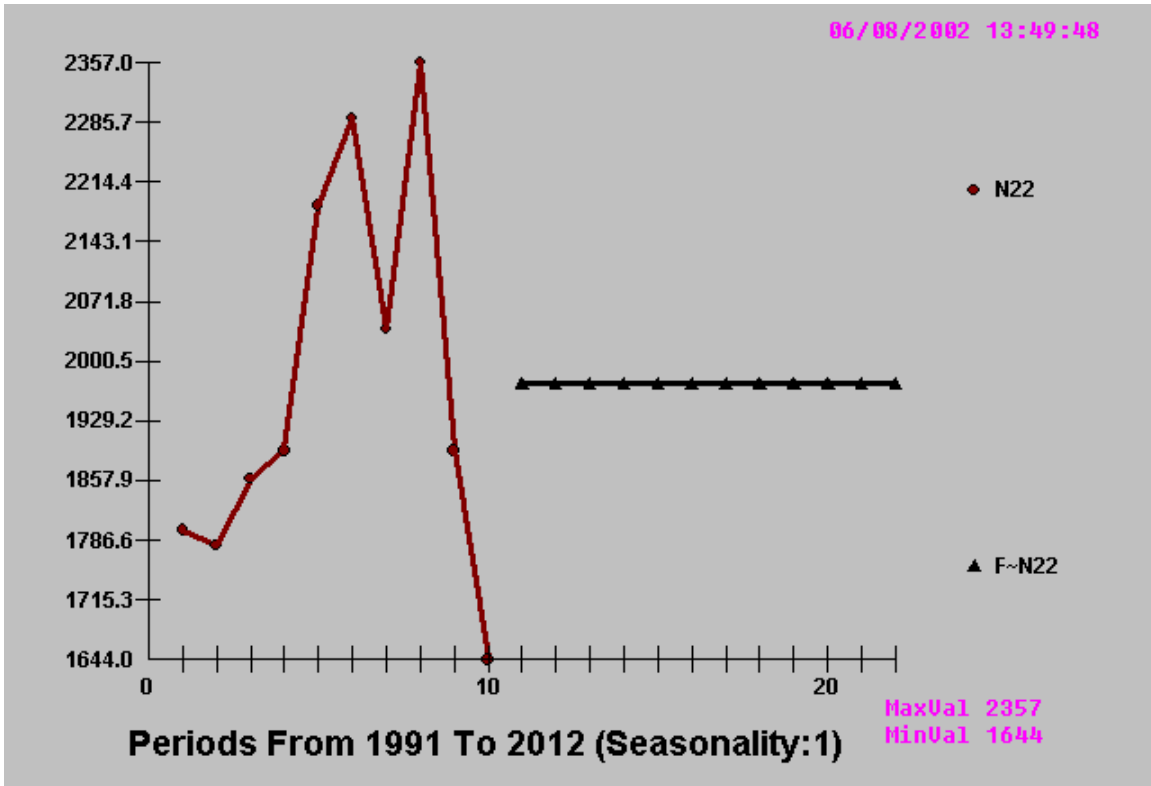
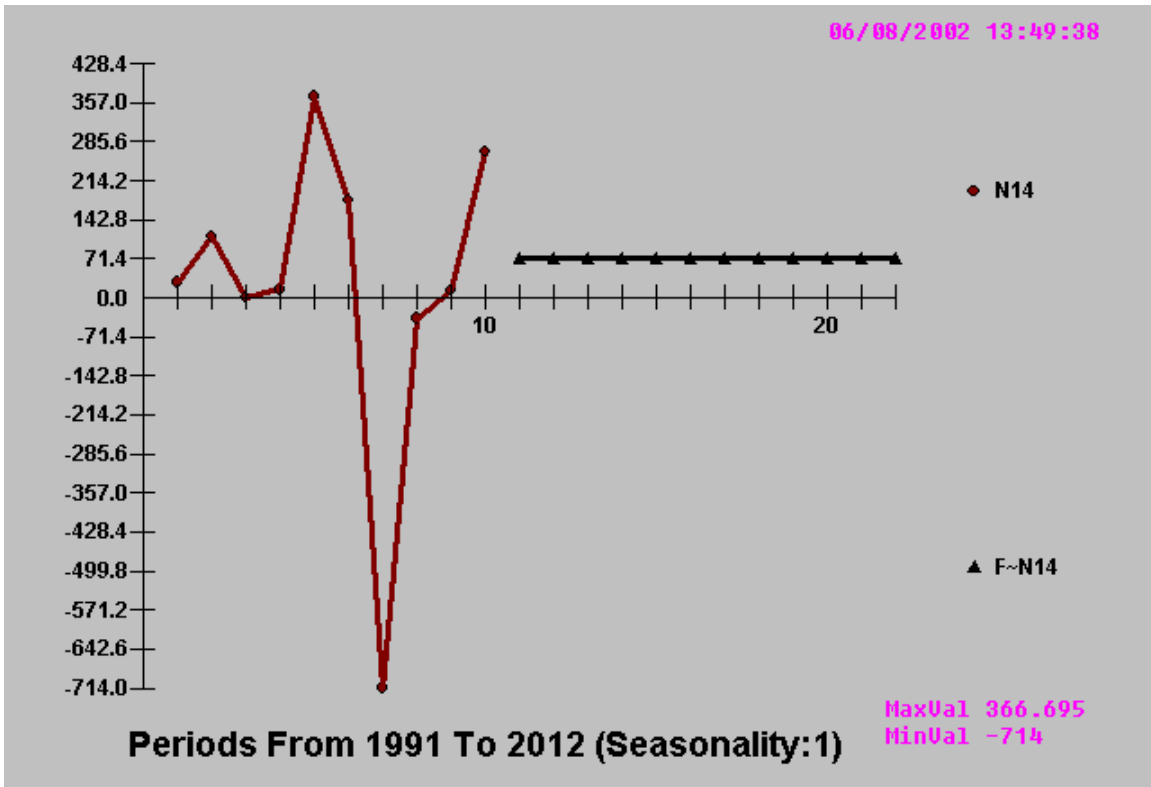
¹ The frauds of Con Agra and Sunbeam were not detected. Sunbeam would have been detected, if we had looked at the next to last point in time instead of the last point in time. It appears that “Chainsaw” cleaned up his act before departing.

To illustrate the identified interventions we now present three of the items identified to be anomalous for Enron.





and two of the items that were not identified as anomalous.



OUTLIER DETECTION vs. PLUS/MINUS 3-SIGMA CHARTS

The problem is that you can't catch an **outlier** without a model (at least a mild one) for your data. Else how would you know that a point violated that model? In fact, the process of growing understanding and finding and examining outliers must be iterative. This isn't a new thought. Bacon, writing in *Novum Organum* about 400 years ago said: "Errors of Nature, Sports and Monsters correct the understanding in regard to ordinary things, and reveal general forms. For whoever knows the ways of Nature will more easily notice her deviations; and, on the other hand, whoever knows her deviations will more accurately describe her ways." [Sir Francis Bacon]

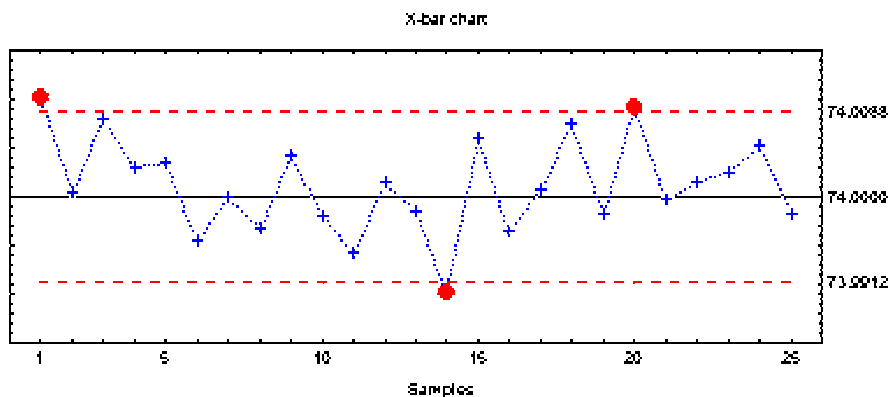
Some analysts think that they can remove outliers based on abnormal residuals to a simple fitted model sometimes even "eye models". If the outlier is outside of a particular probability limit (95 or 99), they then attempt to locate if there is something missing from model. If not, it's gone. This deletion or adjustment of the value so that there is no outlier effect is equivalent to augmenting the model with a 0/1 variable where a 1 is used to denote the time point and 0's elsewhere. This manual adjustment is normally supported by visual or graphical analysis which as we will see below often fails. Additionally, this approach begs the question of "inliers" whose effect is just as serious as "outliers". Inliers are "too normal or too close to the mean" and if ignored will bias the identification of the model and its parameters. Consider the time series 1,9,1,9,1,9,5,9 and how a simple model might find nothing exceptional whereas a slightly less simple model would focus the attention on the exceptional value of 5 at time period seven.

To evaluate each and every unusual value separately is inefficient and misses the point of intervention detection or data-scrubbing. A sequence of values may individually be within "bounds", but collectively might represent a level shift that may or may not be permanent. A sequence of "unusual" values may arise at a fixed interval representing the need for a seasonal pulse "scrubbing". Individual values may be within bounds, but collectively they may be indicative of non-randomness. To complicate things a little bit more, there may be a local trend in the values. In summary, there are four types of "unusual values" ; 1. Pulse, 2. Seasonal Pulse, 3. Level Shift and 4. Time Trends .

In order to assess an unusual value one needs to have a prediction. A prediction requires a model. Hopefully the model utilized is not too simple, just simple enough.

The Original Data May Be Contaminated With Outliers

An original time series plot, example shown below, is a chronological or sequential representation of the readings. The mean is computed and the standard deviation is then used to place (+ / -) 3 standard deviation limit lines. These are then super-imposed on the actual data in order to assess what a reasonable spread or variation should be. Outlier points (points above or below 3 standard deviations) are immediately identified and thus may be deleted from the next stage of the analysis.



The flaw in the above logic is obvious. Outliers will distort the computation of the standard deviation thus inflating the standard deviation and masking the exceptions. Thus we need to simultaneously determine the process standard deviation and the outliers. This problem is exacerbated when you have autocorrelated data as this has an effect on the standard deviation.

If the data is negatively autocorrelated, i.e. high then low then high etc., the standard deviation is overstated. Similarly, if the data is positively autocorrelated i.e. slow drifts on either side of the mean the standard deviation is understated.

Some would argue that the outliers can be identified via an "influential observation approach" or "Cook's Distance approach".

Essentially this detection scheme focuses on the effect of its deletion on the residual sum of squares. But this approach usually fails because the outlier is an "unusual value" to its prediction and that prediction requires a model.

Autobox™ solves this problem by running computer-based experiments where an initial model is either identified or not before intervention detection is pursued. This is compounded by controlling for time trends and/or level shifts in distinctly different trials leading to a global optimization.

Formal Presentation of How Outliers Are Identified

Outliers can be represented as intervention variables of the forms: pulse, level shifts and seasonal pulses and local time trends. The procedure for detecting the outlier variables is as follows. Develop the appropriate ARIMA model for the series. Test the hypothesis that there is an outlier via a series of regressions at each time period. Modify the residuals for any potential outlier and repeat the search until all possible outliers are discovered. These outliers can then be included as intervention variables in a multiple input Box-Jenkins model.

The noise model can be identified from the original series modified for the outliers. This option provides a more complete method for the development of a model to forecast a univariate time series. The basic premise is that a univariate time series may not be homogeneous and therefore the modeling procedure should account for this. By homogeneous, we mean that the underlying noise process of a univariate time series is random about a constant mean. If a series is not homogeneous, then the process driving the series has undergone a change in structure and an ARIMA model is not sufficient.

The Autobox™ heuristic that is in place checks the series for homogeneity and modifies the model if it finds any such changes in structure. The point is that it is necessary for the mean of the residuals to be close enough to zero so that it can be assumed to be zero for all intents and purposes. That requirement is necessary, but it is not sufficient. The mean of the errors (residuals) must be near zero for all time slices or sections. This is a more stringent requirement for model adequacy and is at the heart of intervention detection. Note that some inferior forecasting programs use

standardized residuals as the vehicle for identifying outliers. This is inadequate when the ARIMA model is non-null.

We now present a summary of the mathematical properties underlying this procedure. For purposes of this discussion, we present the following equation, which is the general ARIMA model:

$$\nabla \Phi_p(B) (N_t - \mu) = \theta_0 + \theta_q(B) A_t \quad (\text{equation 1})$$

where N_t = the discrete time series,

μ = the mean of the stationary series,

∇ = the differencing factor(s),

Φ_p = the autoregressive factor(s),

θ_0 = the deterministic trend,

θ_q = the moving average factor(s),

A_t = the noise series,

and B = the backshift operator.

Outliers can occur in many ways. They may be the result of an error (i.e. a recording error). They may also occur by the effect of some exogenous intervention. These can be described by two different, but related, generating models discussed by Chang and Tiao (1983) and by Tsay (1986). They are termed the innovational outlier (IO) and additive outlier (AO) models.

An additive defined as, t_0

$$Y_t = N_t + W E_t \quad (\text{equation 2})$$

An innovational outlier defined as, t_0

$$Y_t = N_t + \frac{\theta(B)}{\Phi(B)} W E_t \quad (\text{equation 3})$$

where Y = the observed time series, t in length

W = the magnitude of the outlier,

t_0

$E_t = 1$ if $t = t_0$,

$= 0$ if $t \neq t_0$,

that is, E_t is a time indicator signifying the time occurrence of the outlier, and N_t is an unobservable outlier free time series that follows the model

given by equation 1. Expressing equation 2 in terms of white noise series A_t in equation 1, we find that for the AO model

$$Y_t = \frac{\Theta(B)}{\Phi(B)} A_t + W E_t^{t_0} \quad (\text{equation 4})$$

while for the the IO model

$$Y_t = \frac{\Theta(B)}{\Phi(B)} (A_t + W E_t^{t_0}) \quad (\text{equation 5})$$

Equation 4 indicates that the additive outlier appears as simply a level change in the t_0^{th} observation and is described as a o "gross error" model by Tiao (1985). The innovational outlier represents an extraordinary shock at time period t_0 since it influences observations $Y_{t_0}, Y_{t_0+1}, \dots$ through the memory of the system described by:

$$\frac{\Theta(B)}{\Phi(B)}$$

The reader should note that the residual outlier analysis as conducted in the course of diagnostic checking is an AO type. Also note that AO and IO models are related. In other words, a single IO model is equivalent to a potentially infinite AO model and vice versa. To demonstrate this, we expand equation 5 to

$$Y_t = \frac{\Theta(B)}{\Phi(B)} A_t + \frac{\Theta(B)}{\Phi(B)} W E_t^{t_0} \quad (\text{equation 6})$$

and then express equation 6 in terms of equation 4

$$Y_t = \frac{\Theta(B)}{\Phi(B)} A_t + W^* E_t^{t_0} \quad (\text{equation 7})$$

$$W^* = \frac{\Theta(B)}{\Phi(B)} W$$

Due to estimation considerations, the following discussion will be concerned with the additive outlier case only. Those interested in the estimation, testing, and subsequent adjustment for innovative outliers should read Tsay (1986). Note that while the above models indicate a single outlier, in practice several outliers may be present.

The estimation of the AO can be obtained by forming:

$$\Pi(B) = \frac{\Theta(B)}{\Phi(B)} = (1 - \Pi_1 B - \Pi_2 B - \dots) \quad (\text{equation 8})$$

and calculating the residuals E_t by

$$\begin{aligned} E_t &= \Pi(B) Y_t && (\text{equation 9}) \\ &= \Pi(B) \frac{(\Theta B)}{\Phi B} A_t + W E_t^{to} \\ &= A_t + W \Pi(B) E_t^{to} \end{aligned}$$

By least squares theory, the magnitude W of the additive outlier can be estimated by

$$\begin{aligned} \hat{W}_{t_0} &= n^2 \Pi(F) E_{t_0} && (\text{equation 10}) \\ &= n^2 (1 - \Pi_1 F^2 - \Pi_2 F^2 - \dots - \Pi_{n-t_0} F^{n-t_0}) E_{t_0} \end{aligned}$$

$$n^2 = (1 + \Pi_1^2 + \Pi_2^2 + \dots + \Pi_{n-t_0}^2) - 1$$

and F is the forward shift operator such that $F e_t = e_{t+1}$. The variance of \hat{W} is given by:

$$\begin{aligned} \text{Var}(\hat{W}_{t_0}) &= n^2 \sigma^2 && (\text{equation 11}) \\ \text{where } \sigma^2 &\text{ is the variance of the white noise (random) process } A_t. \end{aligned}$$

Based on the above results, Chang and Tiao (1983) proposed the following test statistic for outlier detection:

$$Y_{t_0} = \hat{W}_{t_0} / n \sigma. \quad (\text{equation 12})$$

If the null hypothesis of no outlier is true, then Y_{t_0} has the to standard normal distribution. Usually, in practice the true parameters Π and σ^2 are unknown, but consistent estimates exist. Even more important is the fact that to, the time of the outlier, is unknown, but every time point may be checked. In this case one uses the statistic:

$$Y = \max_{\{t_0: 1 \leq t_0 \leq n\}} |Y_{t_0}| \quad (\text{equation 13})$$

and declares an outlier at time t_0 if the maximum occurs at t_0 and is greater than some critical value C . Chang and Tiao (1983) suggest values of 3.0, 3.5 and 4.0 for C . The outlier model given by Equation 4 indicates a pulse change in the series at time t_0 . A step change can also be modeled simply by replacing E_t^{to} with S_t^{to} where:

$$S_t^{t_0} = \begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if not} \end{cases} \quad (\text{equation 14})$$

We note that $(1-B)S_t^{t_0} = E^{t_0}$. Using S^{t_0} one can apply least squares to estimate the step change and perform the same tests of hypothesis reflected in Equations 12 and 13. In this way, significant pulse and/or step changes in the time series can be detected.

Summary and Conclusions

This study examined whether automatic intervention detection can be effectively used to distinguish companies with fraudulent reported data from those with no indication of fraud. Eight companies, identified in the general press as having engaged in financial statement fraud, were pair-matched with firms within the same SIC code that had not been identified by press as having engaged in financial statement fraud¹. Financial statement data were obtained for the firms for a ten-year period ending one year prior to the discovery of the fraud. Hence, during the test period, the media had not yet publicized the fact that the firm was engaged in fraud. The number of interventions at the last point in time was used to rank the companies within the eight sets.

Intervention detection correctly identified 6 or 7 of the fraudulent firms². All fraud firms had a large number of interventions. These results are consistent with the supposition that automatic intervention detection can be effectively used to detect fraud firms.

Suggestions for Future Research

A critical next step is the application of intervention detection to test samples that reflect realistic base rates in which most firms are not engaged in fraudulent reporting. This study used 45 COMPUSTAT data fields. Future studies may attempt to determine which of these fields are relevant to

¹ Wastemasters, a pair-matched firm to Waste Management, had not been identified in the general press as having engaged in financial statement fraud. However, Wastemasters subsequently was sued for activities related to its debentures.

² Sunbeam would have been detected, if we had looked at the next to last point in time instead of the last point in time.

the detection of fraud and which are not. If the number of fields can be reduced, it will be possible to identify additional pair-matched firms.

Depending upon whether or not Sunbeam is included as one of the frauds detected, intervention detection was able to detect the fraud firm between 75 and 87.5 percent of the cases. This is true, even though only ten years of data was examined. This suggests that the integrity of data is paramount. Future studies may fine tune the methodology by seeking to determine the incremental benefit of adding additional years of data.

References

- Albrecht, W.S. and C.C. Albrecht. 2002. Root Out Financial Deception, *Journal of Accountancy*, (193)4, pp. 30-34.
- Apostolou, B.A., J.M. Hassell, S.A. Webber and G.E. Sumners. 2001 "The relative importance of management fraud risk factors, *Behavioral Research in Accounting*. (13)pp. 1-24.
- Box, G.E.P., and Tiao, G. (1975). "Intervention Analysis with Applications to Economic and Environmental Problems," *Journal of the American Statistical Association*, Vol 70,70-79
- Busta, B., and R. Weinberg, 1998. Using Benford's law and neural networks as a review procedure. *Managerial Auditing Journal*. (13)6, pp. 356-366.
- Chang, I., and Tiao, G.C. (1983). "Estimation of Time Series Parameters in the Presence of Outliers," Technical Report #8, Statistics Research Center, Graduate School of Business, University of Chicago, Chicago.
- Coderre, D. 1999 Computer assisted techniques for fraud detection. *The CPA Journal* (69)8, pp. 57-59.
- Davia, H.R. 2001. *Fraud 101: Techniques and Strategies for Detection*. New York. John Wiley & Sons, Inc.
- Downing, D.J., and McLaughlin, S.B. (1986). "Detecting Shifts in Radical Growth Rates by the Use of Intervention Detection," Engineering Physics and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge.
- Glover, S.M., D. Prawitt, M.B. Romney. 2000. The software scene. *The Internal Auditor*. (57)4 pp. 49-57.
- Harvey, F. January 12, 2002. A key role in detecting fraud patterns: neural networks. *Financial Times*. London. p. 3.
- Moyes, G.D. and I. Hasan. 1996. "An empirical analysis of fraud detection likelihood." *Managerial Auditing Journal*. (11)3, pp. 41-46.
- Pincus, K. 1989. The efficacy of a red flags questionnaire for assessing the possibility of fraud, *Accounting, Organizations and Society*, pp. 153-63.

Reilly, D.P. (1980). "Experiences with an Automatic Box-Jenkins Modeling Algorithm," in *Time Series Analysis*, ed. O.D. Anderson. (Amsterdam: North-Holland), pp. 493-508.

Reilly, D.P. (1987). "Experiences with an Automatic Transfer Function Algorithm," in *Computer Science and Statistics Proceedings of the 19th Symposium on the Interface*, ed. R.M. Heiberger, (Alexandria, VI: American Statistical Association), pp. 128-135.

Rezaee, Z., A. Sharbatoghlie, R. Elam, P.L. McMickle. 2002. "Continuous auditing: Building automated auditing capability." *Auditing*. (21)1, pp. 147-163.

Tsay, R.S. (1986). "Time Series Model Specification in the Presence of Outliers," *Journal of the American Statistical Society*, Vol. 81, pp. 132-141.

Wells, J.T. 2001. ". . . And nothing but the truth: uncovering fraudulent disclosures." *Journal of Accountancy*. (192)7, pp. 47-52.