Using Key Performance Indicators and Risk Measures in Continuous Monitoring

Mark J. Nigrini  
*The College of New Jersey*

Arlo J. Johnson  
*Manager, Franchise Audit*

**ABSTRACT:** The paper describes a continuous monitoring methodology in an environment with multiple cues indicative of fraud and errors. A case study describes an application by a restaurant franchisor on the monthly sales numbers reported by franchisees. The system scores each restaurant based on variables judged to be possible indicators of fraud or errors in the monthly sales reports. The variables are weighted according to their perceived importance to obtain a final risk score for each restaurant. The output is a small set of audit targets believed to have a high risk of reporting errors. The variables can be categorized as being (1) key performance indicators, (2) statistical irregularities, (3) motivation or pressure related variables, and (4) variables related to compliance with the franchising agreement.

**Keywords:** continuous monitoring; fraud detection; risk assessment; detective controls.

**INTRODUCTION**

Continuous monitoring is present in many aspects of our personal life. For example, we expect our water supply, airplane flights, and prisons to be continuously monitored. An Internet search for “continuous monitoring” will return results such as monitoring storms, emissions, volcanoes, glucose for diabetics, and foreign broadcasts for intelligence purposes. With monitoring so pervasive it is puzzling that the concepts in Vasarhelyi and Halper (1991) have taken so long to be widely used by auditors and accountants.

PricewaterhouseCoopers (PWC) (2006) report that about half of chief audit executives had a system of continuous auditing or monitoring in place. The fully operational systems focused on fraud detection and key performance indicators to identify deteriorating business activities. PWC conclude that continuous auditing is still considered an emerging phenomenon, but that such methodologies are increasingly viewed by internal audit as a means to enhance their audit processes and to meet stakeholder needs and demands for faster and higher quality real-time assurance. PWC (2007a) paints a similar picture with about half
of audit executives reporting the use of some form of continuous monitoring, but only 11 percent of respondents describing the process as fully operational. PWC (2007b) predicts that over the next five years internal auditors will devote more time to risk management, fraud, internal controls, and process flows. They believe that auditors will need to optimize the use of technology and conduct audits on a more targeted basis in response to specific risk concerns. They rated continuous monitoring and auditing software applications as the most important technology related to internal audit over the next five years.

Alles et al. (2006) describe the approach taken and the lessons learned from the continuous monitoring of business process controls at Siemens. De Aquino et al. (2008) discuss the context and environment conducive to a continuous monitoring implementation. Alles et al. (2007) call for research on the role of the internal auditor in designing, developing, and implementing continuous auditing. They also call for research to address applications, best practices, and transitioning to a continuous audit environment. Given the few published papers on actual applications, this paper describes a continuous monitoring methodology and its application at a large corporation.

The next section of this paper describes a risk scoring method that could be adapted to various continuous monitoring and risk assessment applications. In the case study, the system used at a franchising company is described. Some preliminary results and future plans for the system are then discussed. The concluding sections include possible research avenues.

CONTINUOUS MONITORING METHODOLOGY

This section describes a continuous monitoring system, called the audit risk score method, based on an adaptation of the IT-monitoring framework of the International Federation of Accountants (IFAC) (2002). The adaptation, by the authors, gives the following series of steps in a continuous monitoring application:

- Determine the scope of the monitoring, and the methods and techniques to be applied.
- Determine the indicators that will be used.
- Design and document the system.
- Record the findings and prepare management reports.
- Update the system to improve the predictive ability of the system.

Methods and techniques are important because without methods and techniques (usually based on computer-assisted tools and data analysis techniques), the remainder of the steps in the framework cannot occur. The methodology uses a scoring system where the variables are seen to be indicators of some attribute or behavior of interest. Examples might include fraudulent or biased reporting by divisional controllers, fraudulent baggage claims for an airline, check kiting by an account holder, or fictitious vendors. The methodology is applicable where the user wants to score each business unit (e.g., a division, bank account holder, vendor, or securities dealer) with a risk score so that audit efforts can be directed at those with the highest scores. Given resource constraints, auditors need to direct their attention to the audit units that are believed to contain the most significant anomalies.

The risk score method combines scores from several variables to identify a behavior of interest. Each variable is weighted based on its perceived importance. The variables (similar to red flags) are chosen using professional judgment and industry knowledge. In the financial reporting and fraud applications developed to date, the variables or red flags have been grouped as follows.
Using Key Performance Indicators and Risk Measures in Continuous Monitoring

High Values
This occurs when an amount is high either in absolute terms or is higher than expected. Past examples included scrap variances, warranty variances, supplies usage, receivables, accruals, and materials variances. An audit unit should only be given a high-risk score if the audit unit’s numbers are higher than some carefully evaluated norm.

Erratic Behavior
This occurs where a set of numbers has a relatively high standard deviation over time as compared to some norm and the volatility is seen to be linked to risk. To compare entities of varying sizes, the standard deviation could be divided by the mean. Erratic behavior is also evidenced by a low correlation between the numbers for a specific audit unit and the average or budgeted trend.

Change in Balance
This occurs when there is a large change in the reported amount from one period to the next. To compare entities of varying sizes, the change could be divided by the mean.

Opposite to Expected
This occurs when an amount or balance moves in the opposite direction to what might be expected. Past examples include a decrease in the dollar amount of sales allowances in a period with an increase in sales, or an increase in manufacturing overhead in a period with declining manufacturing activity.

Other Special Situations
These require some imagination and creativity on the part of the auditors. In a check kiting application a customer that deposited checks every business day scored high on the frequency of deposits variable. In a travel agent application, a high score for the carrier void concentration variable occurred when a travel agent voided a high percentage of ticket sales against one carrier (airline) only. In a financial reporting application, the quarter-end profits pad variable tested whether the profits in March, June, September, and December were significantly above the normal monthly profit levels. There was also a quarter-end sales pad to score divisions with high sales at the end of the quarter. The smooth earnings variable had the divisions scoring high when the division had smooth earnings in a business where fluctuations were normal. The seldom-used accounts variable had the division scoring high when the divisional controller used any of three accounts that were meant for rare and unusual transactions and were also known to be used as cookie jar reserves.

The risk score combines scores from multiple sources into a final risk score. The indicators are based on various attributes including high values and special situations. Since the reporting entity would be aware of the possibility of an audit or review and would most likely avoid blatant fabrications, an effective risk scoring method would need to use some reasonably rigorous statistics and mathematics. The next section describes an environment within which a risk score was developed.

AUDIT ENVIRONMENT
The risk score methodology was used in a company that operates about 5,000 franchised restaurants. The franchisees are required to report their monthly sales numbers shortly after the end of the month. Based on the reported sales numbers, the franchisor
bills the franchisee for royalty and advertising fees. The sales reports are processed by the accounts receivable department and time is needed to follow up on missing values and obvious errors. By the end of the second week the sales file for the preceding month is finalized. There is, however, a continual reconciliation process that occurs to account for sales adjustments identified by the franchisees.

Sales reporting errors (intentional or unintentional) are usually in the direction of understated sales and result in a revenue loss for the franchisor. Using the Vasarhelyi (1983) taxonomy of errors, these errors could be (1) computational errors, (2) integrity errors (unauthorized deletion of transactions), (3) timing errors (incorrect time period), (4) irregularities (deliberate fraud), or (5) legal errors (transactions that violate legal clauses) such as omitting nonfood revenues. The cost of a revenue audit on location is high, and the full cost is not only borne by the franchisor, but also partially by the franchisee in terms of the costs related to providing data and other evidence. A system of identifying high-risk sales reports was important to minimize the costs of auditing compliant locations. The risk score approach also facilitated less expensive correspondence or desk audits that could be conducted when the questions were limited in scope.

The system scored each restaurant based on the perceived risk of underreported sales and was developed by the franchise audit section of internal audit. The first author’s role was that of an academic consultant who was solely responsible for the computer programming and initial data analysis, and jointly responsible for selecting and scoring the variables. The second author was solely responsible for downloading the data from the company’s systems and jointly responsible for selecting and scoring the variables.

This environment is similar to many others where audit units self-report dollar amounts and other statistics, and the recipient has to evaluate which of these might contain errors (e.g., individual tax returns, pollution reports, airline baggage, and insurance claims). The data used for the risk scoring system was downloaded from various financial and marketing systems. Microsoft Access was used for data analysis and reporting.

**FRANCHISE APPLICATION OF RISK SCORING SYSTEM**

Each restaurant was scored on ten variables that were seen to be indicators of, or proxy variables for, fraud and errors. For each variable a score of 0 suggested a low risk of underreported sales, while a score of 1 suggested a high risk of underreported sales.

The initial scoring objective for each variable was to score one-third of the restaurants with 0, one-third of the restaurants with 1, and the remaining restaurants with evenly distributed scores from 0 to 1. If this were achieved, the scores would be symmetrically distributed around the mean and the scores would tend toward having a relatively high dispersion (variance). However, the average restaurant would then be scored with the average risk level (0.50). Consequently, one-half of the restaurants would be scored so that they would be seen to have a more than 50 percent chance of underreported sales. This was far higher than the suspected level of underreporting. The initial scoring objective was discarded and then simply used as a rough guide. It was also not always possible to obtain a large variance for a single variable. For example, very few restaurants actually used excessive round numbers and so most restaurants would score a 0 for that variable.

Another scoring objective was to avoid the use of complex formulas since these are conducive to programming errors, and mathematical issues such as division by zero. For example, the food cost proportion is calculated by dividing the food costs for a month by the monthly sales. It is quite likely that a restaurant that opens on June 1 would buy some
Using Key Performance Indicators and Risk Measures in Continuous Monitoring

Figure 1 shows the initial scoring objective on the left. The achieved scores were, on average, closer to the step function shown on the right, with about one-half of the audit units scoring a 0.

Figure 2 shows a small group of restaurants with abnormally low scores and high scores. Noise occurs in the data when a franchisee purchases from a vendor for a specified location and then later redistributes some of the food and supplies from that location to another location. This, together with the fact that purchases in one month might not be

V1: High Food and Supplies Costs

Franchisees are required to buy food and supplies from a selection of approved vendors. The food cost proportion for May would be food cost ($xx) divided by zero ($0).

The variables (abbreviated “V”) and their weights were chosen based on the industry knowledge of the auditors, the prior experiences of the first author, and to a small extent the available data. The system used ten variables because these were all that the authors could think of during the planning phase.
used until the next month, supports the principle that no single variable by itself is a perfect indicator of reporting errors.

The data analysis showed that the average food cost as a proportion of sales was 0.305 and the standard deviation of these costs was 0.043. The median food cost proportion was 0.315. The V1 scoring formula was set as follows:

<table>
<thead>
<tr>
<th>Food Proportion</th>
<th>Score</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;= 0.31</td>
<td>0.0</td>
<td>Average or slightly lower than average</td>
</tr>
<tr>
<td>0.31 &lt; Proportion &lt;= 0.32</td>
<td>0.2</td>
<td>Slightly higher than average</td>
</tr>
<tr>
<td>0.32 &lt; Proportion &lt;= 0.33</td>
<td>0.4</td>
<td>Higher than average</td>
</tr>
<tr>
<td>0.33 &lt; Proportion &lt;= 0.34</td>
<td>0.6</td>
<td>Much higher than average</td>
</tr>
<tr>
<td>0.34 &lt; Proportion &lt;= 0.35</td>
<td>0.8</td>
<td>High</td>
</tr>
<tr>
<td>Above 0.35</td>
<td>1.0</td>
<td>Very high</td>
</tr>
</tbody>
</table>

The graph of the V1 scores across all locations is shown in Figure 2. About two-thirds of the restaurants score 0 because they have food cost proportions that are at or below average. The average score for this variable was 0.17, because most restaurants had food cost proportions near or below average.
V2: Very High Food and Supplies Costs

It was believed that V1 by itself did not do an adequate job of significantly raising the final scores of restaurants with high and very high food cost proportions. V2 was introduced to give an extra boost to the scores of high and very high food cost locations. The V2 scoring formula was as follows:

<table>
<thead>
<tr>
<th>Food Proportion</th>
<th>Score</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;= 0.325</td>
<td>0.0</td>
<td>Higher than average</td>
</tr>
<tr>
<td>0.325 &lt; Proportion &lt;= 0.333</td>
<td>0.5</td>
<td>Much higher than average</td>
</tr>
<tr>
<td>Above 0.333</td>
<td>1.0</td>
<td>Very high</td>
</tr>
</tbody>
</table>

A second variable, V2, was used to raise the final scores of the restaurants thought to be high-risk locations. The average score for V2 was 0.136, reflecting the fact that only a small proportion of restaurants had very high food costs.

V3: Sales Decreasing

The third indicator of risk was a decreasing sales trend. The belief was that as a franchisee underreported an ever-increasing percentage of sales, their sales trend would be below average. This variable forms part of the \textit{opposite to expected} group of variables. In a time of economic growth and inflation the expectation is that sales will increase over time. This variable is also not a perfect predictor of reporting errors because underreported sales could take place together with increasing sales.

Figure 3 shows the sales changes for the immediately past quarter against the same quarter one year earlier. The graph is truncated at −10 percent and +15 percent. About one-fifth of all restaurants had a sales decline, with the remaining locations showing a sales increase. The scoring system for V3 gave all locations that had a sales change that was worse than average a positive score. The largest scores were for the locations with the largest quarter-on-quarter decreases. The scoring system is shown below:

<table>
<thead>
<tr>
<th>Sales Change</th>
<th>Score</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than −4 percent</td>
<td>1.00</td>
<td>Worst 15 percent of changes</td>
</tr>
<tr>
<td>−0.04 &lt; Change &lt;= −0.02</td>
<td>0.8</td>
<td>Close to the largest declines</td>
</tr>
<tr>
<td>−0.02 &lt; Change &lt;= 0.00</td>
<td>0.6</td>
<td>Slightly negative</td>
</tr>
<tr>
<td>0.00 &lt; Proportion &lt;= 0.02</td>
<td>0.2</td>
<td>Positive change, worse than average</td>
</tr>
<tr>
<td>0.02 &lt; Proportion &lt;= 0.04</td>
<td>0.1</td>
<td>Positive change, slightly worse than average</td>
</tr>
<tr>
<td>Above 0.04</td>
<td>0</td>
<td>Better than average</td>
</tr>
</tbody>
</table>

Since −4.00 percent was the cutoff for the highest score and +4.00 percent was the cutoff for the lowest score, it would have been possible to use a formula for a continuous scoring function as opposed to a step function. A step function was preferred because it made it easier for management to visualize the link between a given sales change and a risk score.
V4: Food Costs Increasing as a Percentage of Sales

The objective for V4 was to give those locations a high score where the food cost percentage was increasing over time. An increase in this percentage over time was seen as a precursor of problems on the horizon. This variable forms part of the high values group of variables, and in this case high means higher than the location’s own historic averages. The monthly percentages were noisy because large purchases in the final week of a month could distort the food cost proportion for that month and for the following month. V4 was based on the slopes from linear regression equations run on several months of sales and food cost data.

The sales numbers were on average about three times as large as the food cost numbers and consequently the sales slope was usually about three times as large as the food cost slope. The approach taken was to compare the sales slope as a percentage of the intercept, to the food cost slope as a percentage of its intercept. To illustrate the calculations, the numbers for a hypothetical location might be that the sales are increasing, on average, by $1,000 per month and the food costs are increasing, on average, by $800 per month. The food cost percentage starts off at 50 percent and gradually increases to 55 percent at the end of the 18-month period. The regression equations would be:
Sales = 79,000 + 1,000 * period
Food costs = 39,200 + 800 * period

with the period starting at 1 and ending at 18 for the 18-month period.

The slopes would be converted to a percentage of the intercept as follows:

\[ \text{Slope}_c = \left(\frac{\text{Slope}}{\text{intercept}}\right) \times 100 \] (3)

The slope differences would be calculated as follows:

\[ \text{Slope}_{\text{Diff}} = \text{Slope}_{CS} - \text{Slope}_{CF} \] (4)

where the subscript Diff refers to difference, C refers to converted, and S and F refer to sales and food, respectively. The calculations for the hypothetical example give a slope difference of \(-0.775\). Negative slope differences mean that over the period the food costs are increasing as a percentage of sales. The slope differences were calculated for all restaurants and the ordered results are shown in Figure 4.

Figure 4 shows that about one-half of the restaurants had slope differences that were negative. This means that, to varying degrees, the food costs were increasing as a percentage

---

**FIGURE 4**
Analysis and Scoring of Food Cost Percentage Changes

---

\(^a\) Figure 4 shows the slope differences sorted in ascending order. The second graph shows the scores applied to the slope differences.
of sales over the 18-month period for one-half of the restaurants. Variable 4 was scored as follows:

<table>
<thead>
<tr>
<th>Slope Difference Range</th>
<th>Variable 4 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Slope_{Diff} &lt; -1.00$</td>
<td>1</td>
</tr>
<tr>
<td>$-1 \leq Slope_{Diff} \leq 0$</td>
<td>$</td>
</tr>
<tr>
<td>$Slope_{Diff} &gt; 1.00$</td>
<td>0</td>
</tr>
</tbody>
</table>

This scoring formula gave high scores for large slope differences (smaller than $-1$) and smaller scores for differences in the $[0, -1]$ range. Slope differences greater than or equal to 0 (which occurred when the food cost proportion was decreasing) resulted in a V4 score of 0. The V4 scores (sorted descending) are shown in the second graph in Figure 4.

**V5: Irregular Seasonal Pattern for Sales**

The logic behind V5 was that restaurants with sales patterns that deviated from the seasonal norms represented a higher risk of underreported sales. This variable forms part of the *erratic behavior* group of variables. Figure 5 shows the typical sales pattern for a calendar year. The months with seasonally high sales are July, August, and December. February usually has the lowest sales because of winter and because it usually only has 28 days.

Figure 5 also shows the sales pattern of a hypothetical restaurant where the sales decrease significantly in the last two months of the period. The correlation between the sales for the specific restaurant and the usual seasonal pattern in Figure 5 is 0.28. Correlation values range from $-1$ to $+1$, and a correlation of 0 means that there is no correlation and an increase in the seasonal pattern is matched with a 50 percent chance of an increase/decrease in the sales of a specific restaurant. A correlation of $+1$ means that there is a perfect, positive, linear relationship between the seasonal pattern and that of a specific restaurant. A correlation of $-1$ would occur when an *increase* in the seasonal pattern is matched with exactly the same absolute value *decrease* in the sales for a specific restaurant, or a *decrease* is matched with exactly the same absolute value *increase*. Correlations between $-1$ and $+1$ (other than 0) indicate the strength of the positive or negative relationship.

Correlations by themselves are imperfect indicators of risk in the reported numbers. For example, a low correlation matched with an above average increase in sales suggests a low risk of reporting errors. The trend in sales should be evaluated together with the correlation to the seasonal pattern. The risk scoring method includes V3 (sales trend) and V5 (correlation) to take both trend and correlation into account. The pattern shown on the right hand graph of Figure 5 is indicative of a high-risk situation because of both the low correlation and the sharply decreasing trend.

Figure 6 shows the calculated correlations sorted from smallest to largest. Variable 5 was scored as follows:
Using Key Performance Indicators and Risk Measures in Continuous Monitoring

Using Key Performance Indicators and Risk Measures in Continuous Monitoring

FIGURE 5
Analysis of Seasonal Pattern of Sales

Seasonal Pattern of Sales

Seasonal Pattern of Sales

Dollars

Month

$\text{Average sales}$

$\text{Average Sales}$

$\text{Regression line fitted to actual sales}$

$\text{Actual sales for a restaurant}$

*Figure 5 shows the seasonal pattern of the sales over the course of a year. The second graph shows the same seasonal pattern and the sales of a specific restaurant together with a fitted regression line.*

<table>
<thead>
<tr>
<th>Calculated Correlation</th>
<th>Variable 5 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Correlation} &lt; -0.4$</td>
<td>1</td>
</tr>
<tr>
<td>$-0.4 \leq \text{Correlation} \leq 0.6$</td>
<td>$(\text{Correlation} \times -1) + 0.6$</td>
</tr>
<tr>
<td>$\text{Correlation} &gt; 0.6$</td>
<td>0</td>
</tr>
</tbody>
</table>

The result of the scoring formula is that correlations of 0.6 (and higher) were given a 0 score, and correlations of −0.4 (or lower) were scored at 1. A correlation midway between −0.4 and 0.6 would be scored at 0.5. The V5 scores that are linked with the correlations are shown in the second graph of Figure 6.

**V6: Round Numbers Reported as Sales Numbers**

The use of rounded numbers was first introduced to the auditing literature in Nigrini and Mittermaier (1997). The belief was that restaurants that reported round numbers as sales amounts represented a higher risk for underreported sales. This variable forms part of the other special situations group of variables.
To score V6 required a judgment call as to what constituted a round number and what constituted an abnormally high count of round numbers. The position was taken that a round number would be a number with 0 in the units position and no cents shown after the decimal point. For example, $17,030.00 would be a round number, but $17,035.00 would not be a round number. Under Benford’s Law the probabilities of the digits tend toward being uniformly distributed when moving from the first (leftmost) to the last (rightmost) digits. Most franchisees reported in whole dollars. The tendency toward uniformity meant that the ending two-digit combinations were equally likely for all practical purposes. Round numbers were deemed to be numbers ending in 00, 10, 20, ... , 90. These digit combinations were one-tenth of all the possible last-two-digit combinations (00, 01, 02, ... , 99). The expectation was that one-tenth of all reported numbers would be round numbers due purely to chance alone.

An analysis of the round numbers counts showed that about three-quarters of all restaurants had either 0, 1, or 2 round numbers for the 18 monthly sales reports. The expectation was that each restaurant would have 1.8 (one-tenth of 18) round numbers in an 18 month period. The V6 scoring chart is shown below:
A count of either two or three round numbers exceeded the expected count, but not by a large margin. Counts of 4 and higher were abnormally high.

**V7: Repeating Numbers Reported as Sales Numbers**

The use of repeated numbers was introduced to the auditing literature in Nigrini and Mittermaier (1997) as the number duplication test. Because of the seasonal nature of the business, restaurants were unlikely to report exactly the same dollar amount more than once in an 18-month period. This variable is a part of the *other special situations* group of variables. The average range of the reported numbers was $16,250 over the 18-month period giving a low probability of duplicate numbers. A duplicate number was seen to have a high probability of being an error. The duplicate number analysis showed that 106 restaurants repeated a sales number in the 18-month period. An extract from the table showed the following at the low and the high ends of the listing of the duplicate amounts.

<table>
<thead>
<tr>
<th>Restaurant #</th>
<th>Amounts Duplicated</th>
</tr>
</thead>
<tbody>
<tr>
<td>omitted</td>
<td>$1.00</td>
</tr>
<tr>
<td>omitted</td>
<td>$12,239.00</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>omitted</td>
<td>$344,986.00</td>
</tr>
</tbody>
</table>

V7 was scored so that the restaurant was given a score of 1.00 if any amount was duplicated during the 18-month period and 0 otherwise. The average score for this variable was very low at 0.02.

**V8: Inspection Rankings**

The franchisor regularly inspected the franchised facilities and rated a number of factors related to customer service, hygiene, and operating procedures. The belief was that if a franchisee was conscientious in following the operating procedures, they were probably also following the reporting requirements. Similarly, tardiness in operations was seen to have a high likelihood of spilling over into tardiness in reporting. A high inspection ranking was a sign of a positive attitude towards the franchisor and a desire to have a pleasant relationship. This variable forms part of the *other special situations* group of variables.

The scores for V8 were based on a weighting of the restaurant’s score for the most recent month and the score for the year to date. Restaurants with poor inspection results were awarded a V8 score of 1.00, and those with inspection results only slightly worse than average were awarded a V8 score of 0.50. Restaurants with inspection scores that were average or better than average were awarded a V8 score of 0. The average V8 score...
across all restaurants was 0.30 because only one-half of all restaurants scored worse than average.

**V9: High Receivable Balance**

The logic behind this variable was that a high receivables balance was a sign that the franchisee had cash flow issues. This variable forms part of the other special situations group of variables and fits in with the pressure aspect of the fraud triangle (pressure, opportunity, and rationalization). Cash flow issues were linked to motivation and might translate into underreported sales. The V9 scores are shown below:

<table>
<thead>
<tr>
<th>Description</th>
<th>Score for V9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant balance over 60 days</td>
<td>1.00</td>
</tr>
<tr>
<td>Moderate balance over 60 days</td>
<td>0.75</td>
</tr>
<tr>
<td>Small balance over 60 days</td>
<td>0.50</td>
</tr>
<tr>
<td>Zero amount owing over 60 days</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**V10: Use of Automated Reporting Procedures**

The company had an Internet-based sales reporting system that requested significant additional information in addition to the monthly sales total, such as a breakdown of other income sources, sales-related statistics, and permitted deductions. Full compliance was encouraged but was not required by the franchising agreement. The logic behind V10 was that if a franchisee voluntarily used the system and reported all the minute details requested, then this was a sign of voluntary cooperation and that there was a reduced chance that the franchisee was engaging in willful sales underreporting. This variable formed part of the other special situations group of variables. The scores awarded for V10 took into account whether (1) the franchisee used the system in the preceding month, (2) the franchisee had used the system for an extended period, and (3) the franchisee reported all or only some of the line items requested.

**Final Scores**

The ten variables were each weighted from 0.05 to 0.20. Variables that had low weights (e.g., round numbers and repeating numbers) were variables that only occurred in a small proportion of the population, or were judged to have less predictive power than the variables with relatively large weights. The final scores were calculated and sorted from largest to smallest. The results showed a small set of about 140 scores that exceeded 0.50, which in probabilistic terms suggests loosely that these restaurants have a higher than 50 percent chance of some level of sales underreporting. These restaurants will be the focus of the company’s audit efforts in 2008. About 270 restaurants had scores of 0, which seems plausible. This means that about 5 percent of all restaurants did not display a single sign believed to be linked to sales underreporting.

The final scores were compared to those of an earlier pilot study. The correlation between the pilot study scores and the 2007 scores was 0.15, indicating that there was virtually no relationship between the past and the present scores. The reasons were (1) different weights, (2) the addition of new variables and the deletion of some of the old
variables, and (3) changed conditions. The low correlation between the old and new scores suggests that the system needs to be regularly updated.

Overview of the Reporting System and Future Plans

The results were communicated as Access reports. Several reports were available, including:

- The final scores for all restaurants sorted by score descending.
- A report for each of the ten variables listing the restaurants that scored high on that variable only.
- A report where the user could input a restaurant reference and the final score for that restaurant would be shown together with the scores for that restaurant for each of the ten variables.

The plans are for the risk scoring system to be updated semiannually with current data. Future system upgrades might be based on (1) changes in the weights of the variables, (2) changes to the scores associated with the values of the variables, (3) deletion or addition of variables, and (4) the inclusion of prior scores as an input to the new scores. The inclusion of prior scores in the value of the current score would mean that the system has some memory. Prior high scores would linger in the calculation of the current score.

Initial Findings

Given internal audit’s preoccupation with the legal requirements of Sarbanes-Oxley, at the time of writing the system had not yet been used to schedule franchisee audits. The reported numbers of the highest scoring franchisees (which usually own more than one location) were scrutinized as a preliminary evaluation of the system. Internal audit noticed that six of the top 50 franchisees, as ranked by average restaurant score, had ceased to operate their locations. This closure rate was significantly above average and showed that the risk scoring system could also assist as a management tool to identify weak operators and allow the franchisor company to offer assistance before the situation was irreversible.

The highest overall score was 0.897. This restaurant scored 1.00 on all variables except for the round numbers and repeated numbers variables, and marginally less than 1.00 on the sales correlation variable. The restaurant was located near a college which explains the weak correlation. Colleges have vacations and have fewer people around in July, August, and December. The usual slowdown because of the cold weather in February does not affect college locations. While this may explain the weak correlation it does not explain the scores of 1.00 on the other seven variables.

DISCUSSION AND FUTURE WORK

The paper describes a continuous monitoring methodology and its application to the monthly sales reports in a franchisor environment. The system scored each restaurant on each of ten variables believed to be linked to sales reporting issues. Final scores at or near 1 suggested an increased risk of fraud or errors in the reported monthly sales numbers. The logic and methods could be adapted to other continuous monitoring environments in which auditors or management want a formal system to evaluate the risk of fraud or errors.

Alles et al. (2007) note that the implementation of continuous monitoring requires five basic concepts. These are (1) metrics (the measurements of a system, which in this case are the sales reports and other indicators), (2) standards or models (the norms obtained from the data analysis phase), (3) analytics (the scoring methods), (4) alarms (a relatively
high score would be an alarm), and (5) methods of measurement (the reports generated by the system). The risk scoring system conforms in all respects to the five concepts of continuous business activity monitoring.

If a risk scoring system were developed in which the weights of the variables had no relationship to their predictive ability, and the variables themselves were not really linked to fraud or errors, then the system would be the same as a random selection of audit units. To better understand the link between using multiple variables, and a simple additive formula in an audit selection context, the authors call for experimental research using auditor subjects in a realistic audit setting where subjects are asked to rank reports based on the perceived likelihood of fraudulent reporting.

REFERENCES


