Communications of the Association for Information Systems

Volume 11 Article 34

June 2003

Outcome Based Systems Evaluation to Assess Information Technology Impact Using Arima Methods

Gerald E. Evans
University of Montana, jerry.evans@business.umt.edu

Bruce A. Costa
University of Montana, bruce.costa@business.montana.edu

Communications of the Association for Information Systems

OUTCOME-BASED SYSTEMS EVALUATION TO ASSESS INFORMATION TECHNOLOGY IMPACT USING ARIMA METHODS

GERALD E. EVANS
BRUCE A. COSTA
School of Business Administration
The University of Montana
jerry.evans@business.umt.edu

ABSTRACT

A new method of system evaluation that focuses on the impact the system has on a data series that served as the rationale for systems implementation was designed and modeled by the authors. Called outcome-based evaluation, this method is founded on the concept of intervention analysis and employs interrupted time series designs to determine the impact of an information system on specific organizational goals. Based on a review of the literature on the evolution of systems evaluation methods from focusing on user goals to user satisfaction and system usage, we conclude that user satisfaction and system usage are necessary but not sufficient criteria to establish system effectiveness or success. Thus, we establish a need for the proposed new method of system evaluation.

Three business case studies are presented in this article that demonstrate and validate an evaluation method using ARIMA models for the analysis. The value of this tool for managers is its means of assessing IT effectiveness and payoff contextually, thereby enabling businesses to clarify both their IT needs and their outcome expectations *a priori*.

Keywords: Information systems evaluation, systems outcomes, intervention analysis, outcomes based systems evaluation

I. INTRODUCTION AND REVIEW OF LITERATURE

Evaluating information systems is an essential part of the IS literature [e.g., Keen and Scott-Morton. 1977; Shannon and Weaver, 1949; Sprague and Carlson, 1982]. Boehm and Bell [1977] concluded that both the cost of a system and the information needs of the users must be considered in any information systems evaluation. This idea was later discussed and supported by Chandler [1982] who favored a multiple criteria approach. Chandler held that assessing the ability of a system to meet the user's goals must be a critical element in evaluating that system

[DeSanctis and Gallupe, 1987; Dennis et al, 1988; George et al. 1988]. At the core of evaluating information systems lies the user's cost and utilization, an approach detailed by Borovits and Galadi [1993].

Unfortunately, over time this initial notion of system evaluation employing user goals and system constraints became distorted in the contemporary application until the two primary measures of system success evolved merely into user satisfaction and system usage. User satisfaction, then, was measured [Bailey & Pearson, 1983;Ives, Olson, & Baroudi,1983] in relation to user involvement in system design [Ammoako-Gyampah & White [1983], Baroudi, Olson, & Ives, 1986; McKeen, Guimaraes, & Wetherby, 1994]. On rare occasions researchers went even further and examined user satisfaction as a means of determining system effectiveness [e.g., Iivari and Ervasti, 1994]. Following this argument to its [il]logical extreme, one enterprising study even demonstrated a connection between user satisfaction and system effectiveness [Gaitian, 1994]. In contrast, the position developed in this paper is that while user satisfaction and system usage are certainly necessary for system success, they are not sufficient in and of themselves nor are they measures of effectiveness or payoff. This paper focuses on a method of evaluating management systems by examining an organization's goals via time series data analysis associated with the process in question.

In their review of the existing research on IS success, DeLone and McLean [1992] noted that the problem in conceptualizing measures of IS success was not the absence of measures in the literature but the large number of different dependent variables. In fact, they suggested that there are almost as many different measures of IS success as studies. Perhaps this multitude of measures of systems success exists because each system was implemented for different reasons specific to its organization.

The outcome-based approach developed and applied in the present work is capable of recognizing that each system was conceived for specific reasons. Furthermore, the rationale of a systems design and implementation must form the basis for its evaluation. For example, most systems, particularly those in small businesses, are designed and implemented because something in a specific data stream (e.g. inventory levels, manufacturing costs, percent overdue accounts) indicates the existence of an opportunity or problem. Technology is then applied to take advantage of the opportunity or to solve the problem. This, view [Evans, 1999] is the concept of information systems as interventions:

> *Systems are implemented as an intervention to take advantage of an opportunity or to solve a problem.*

Viewing interventions in this manner is basic to understanding this methodology in other areas. Wichern and Jones [1977] used these methods to assess the impact of marketing interventions on the sale of Crest toothpaste. This initial application of ARIMA-based interrupted time-series analysis provides a foundation for broader applications to assess the impact on a time series of any treatment variable. Known also as interrupted time-series analysis, it examines a time series seeking a significant change concomitant with the application of an intervention or treatment. A complete treatise of intervention analysis and interrupted time series design is available in Chapter 10 of the SPSS Trends manual [1993] and in books by Cook and Campbell [1979] and McDowall, McCleary, Meidinger, and Hay [1980].

The application of these methods to systems evaluation follows the same logic. Since most systems are implemented to address a problem or opportunity discovered in the regularly-kept data records of a business enterprise, it makes sense to evaluate the effectiveness of the system based upon the impact the system had on the data series it was designed to influence. This approach is termed "outcome-base evaluation." Systems are designed and implemented to produce certain outcomes; therefore each system should be evaluated in terms of their impact upon those outcomes.

For the case for outcome-based evaluation to be compelling, four issues must be addressed:

1. a complete description of the tool must be provided. Although much has been written about intervention analysis, its application to IT evaluation must be set forth in a clear and convincing manner.

2. Understanding that the impact of IT is multi-faceted, outcome-based evaluation must be capable of demonstrating the impact of specific IT on one set of outcomes as distinguished from another set of outcomes. The possible effects of certain intermediary or moderating variables must also be able to be identified and isolated.

3. For outcome-based evaluation to be a universal assessment tool, it must be demonstrated to be effective on a wide range of outcomes. These outcome variables must include, but are not limited, to economic variables, performance variables, attitude and opinions, and enterprise-wide objectives.

4. Outcome-based evaluation must be viable in complex organizations where interventions such as IT produce both intended and unintended outcomes. These outcomes usually occur in unanticipated places in the organization and after some unpredictable time delay.

As Senge [1990] observed, outcomes from interventions occur somewhere else and at some other time in complex organizations, thereby making learning difficult. Because outcome-based evaluation of IT is based on program evaluation and evaluation research methodologies [e.g., Posavac and Carey, 1980], it is appropriate in complex organizations where multiple outcomes, both intended and unintended, result from the implementation of specific IT.

IT AS AN INTERVENTION

The concept of information technology as an intervention is straightforward. The application of technology to specific problems or opportunities represents managerial decisions to intervene in an organization to achieve a particular outcome. As noted earlier, both the cause and effect of the specific IT implementation are typically a routinely collected series of data representing some important performance measure of the organization. These measures could be financial ratios, cost data, or income. They could also be non-financial performance measures such as quality variables, turn-around times, or length of waiting lines. Attitude and opinion data such as those generated from customer or employee satisfaction surveys might represent further measures. These measures can be either enterprise-wide data or applicable only to specific business units or operations.

The necessary component for outcome-based evaluation to work is a time-series before and after implementing the IT. In the case of outcome-based evaluation, the time series is interrupted by the implementation of the IT. Although ordinary least squares regression analysis could be used to identify the nature and magnitude of the impact of the IT [Evans, 1999] ARIMA models are preferred since least-squares methods are often unsuitable for time-series data. The robustness of this approach lies in the power of ARIMA methods. For example, other time-series assumed to be unaffected by the IT can be used as a control group to protect the analysis from threats to internal validity that may stem from history or maturation. Furthermore, moderating or intervening variables can be included as co-variates, thus demonstrating the impact of the IT over and above the effect of other variables. Finally, phased implementation of the IT can be assessed by coding the intervention variable as a series of increasing pulses (e.g. 000…111…222…333…) or as a set of dichotomous dummy variables which are entered in sequence as independent variables in the ARIMA model. The latter approach makes it possible to identify the impact of each phase of the implementation separately and then cumulatively. Cook and Campbell [1979] provide a thorough discussion of interrupted time-series design and their analysis using these methods.

II. METHODS

Three case studies are used to demonstrate the application and usefulness of outcome-based evaluation. The cases were not been selected randomly. Rather, they are instances of applying outcome-based evaluation that best serve to validate and authenticate the evaluation methodology. The cases demonstrate different types of outcomes and IT implementations. Furthermore the cases go from simpler to increasingly complex organizations and IT to demonstrate the validity of the methodology.

CASE STUDY 1

A lumberyard and home center used a manual method for pricing invoices. Items sold to contractors were sold on 30-day accounts that were recorded on invoices without prices because prices were previously agreed to on quotes and bids. Some hardware items were recorded with prices but all lumber items did not show them. Invoices were then "priced out" by two or three individuals who specialized in this work. Such practice is not uncommon among small lumberyards. At some point, the owner began to investigate the accuracy of the pricing process and discovered that the average pricing error was approximately $220 a day. These errors were not normally distributed because a customer would surely appeal the error if the price were higher than agreed to but would tend to avoid mentioning cases where the prices were lower than quoted. Figure 1 shows the time series data for pricing errors for 90 days prior to the implementation of a system at this business.

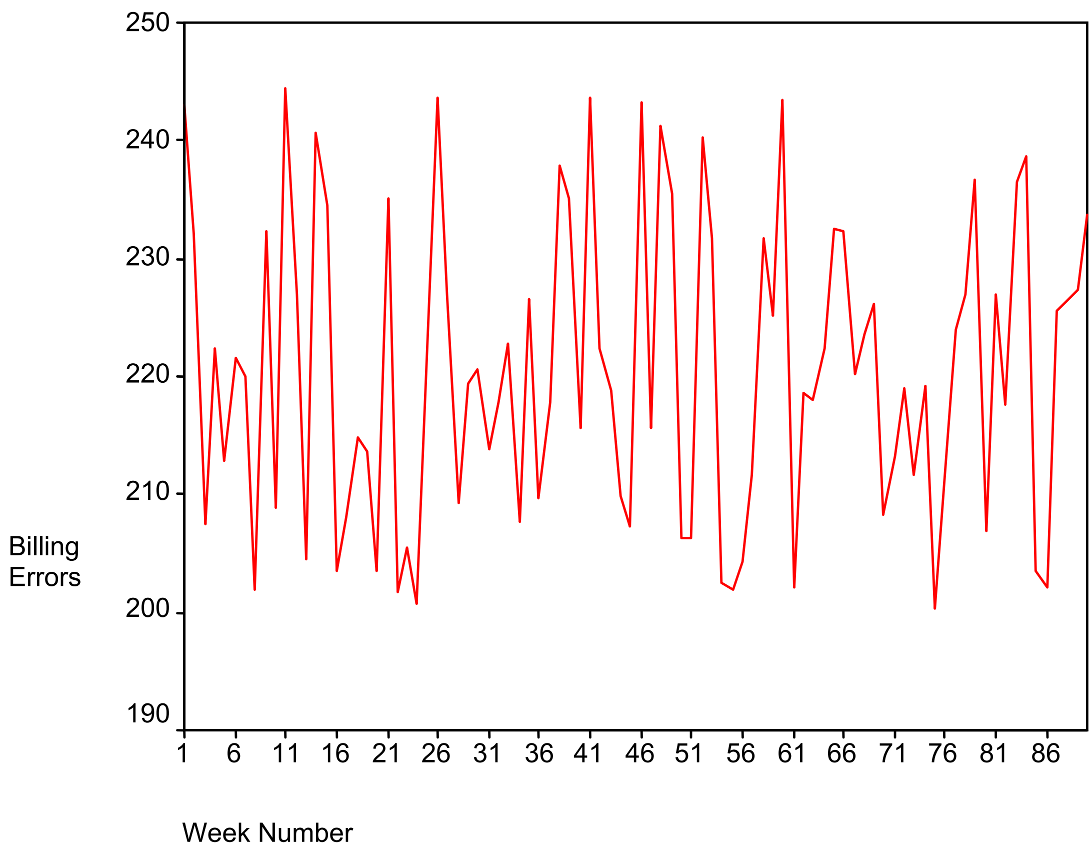


Figure 1. Billing Errors for 90 Days Prior to Intervention

As an intervention, a point-of-sale system (POS) was designed and installed to “fix” the problem. Figure 2 shows the time series for both 90 days prior to the systems implementation and 90 days following it. The effectiveness of the system is clearly evident

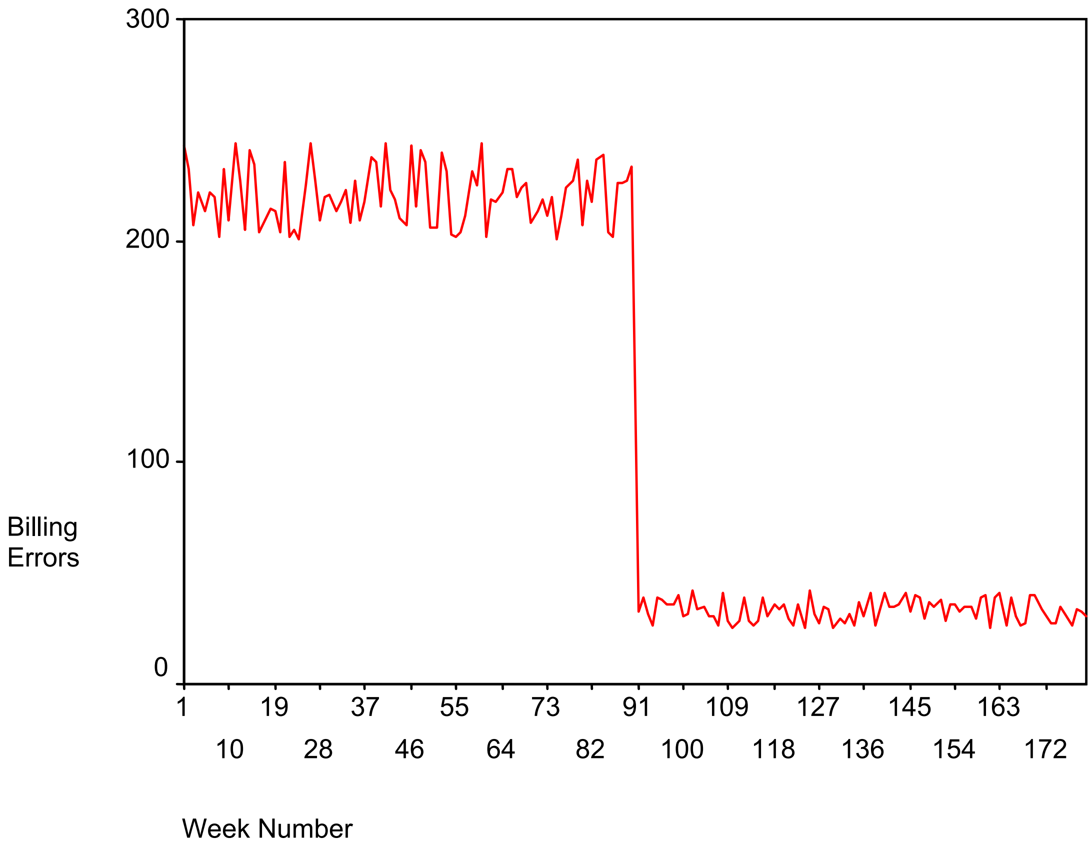
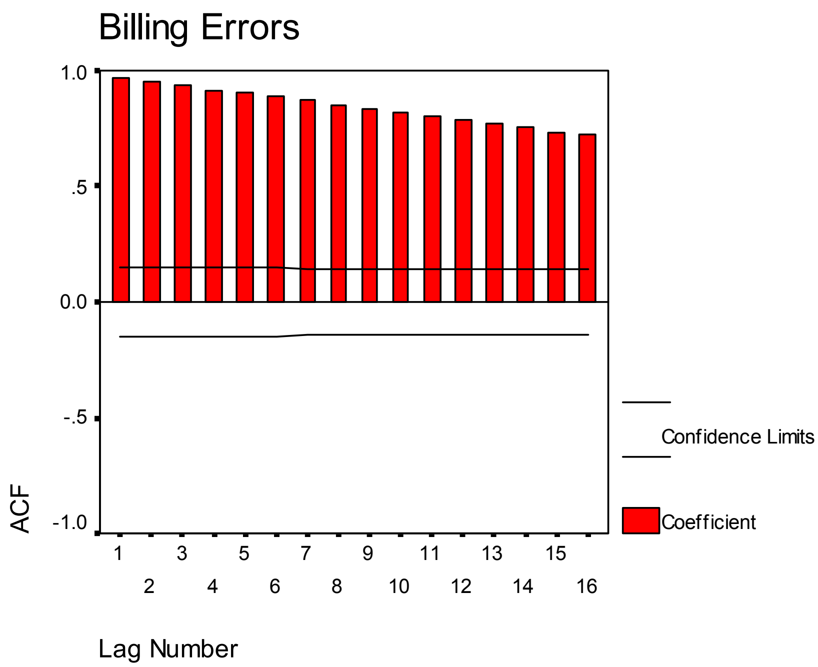
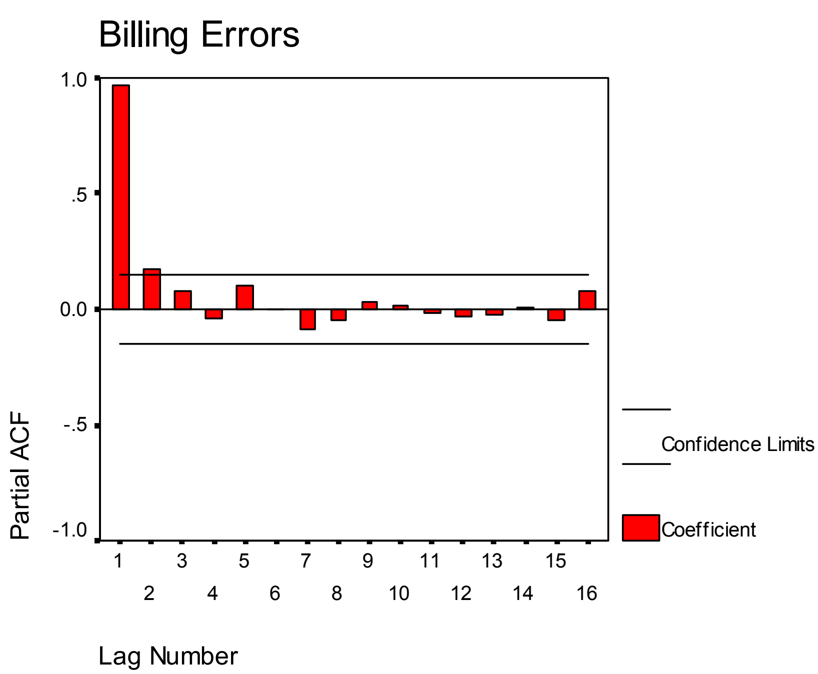


Figure 2. Billing Errors for 90 Days Prior to and After Intervention

To begin the analysis, autocorrelation functions (ACF) and partial autocorrelation function (PACF) are plotted in Figure 3 for the time series with a 95% confidence interval.



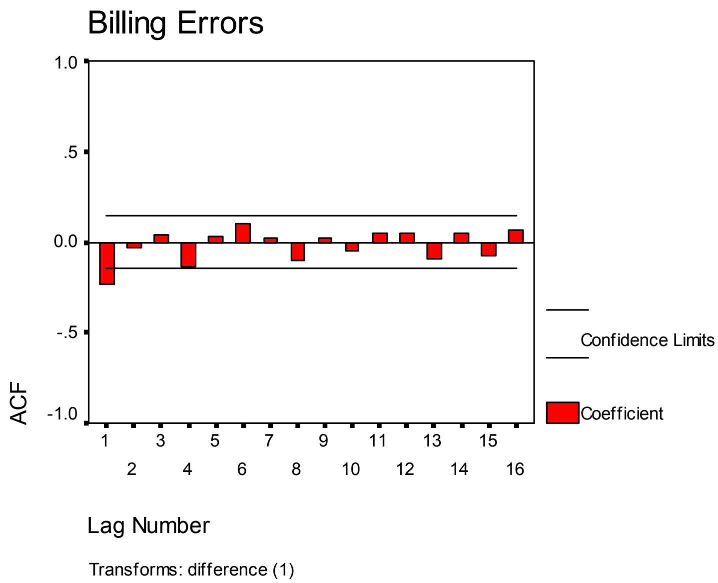
a. ACF Plot



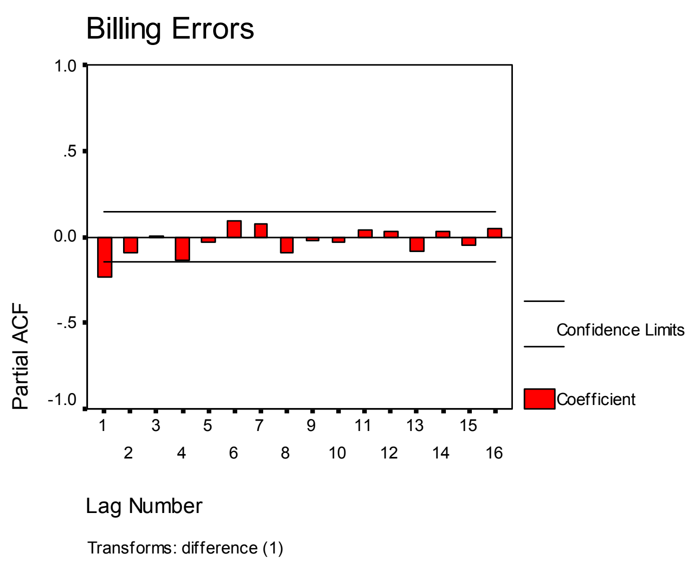
b. PACF Plot

Figure 3. ACF and PACF Plots

The ACF plot in Figure 3 indicates a series with a significant trend, often termed non-stationary in the ARIMA literature. Differencing the series creating an ARIMA (0,1,0) model usually produces a stationary series. The ACF and PACF plot of the differenced series is shown in Figure 4



a. ACF Plot



b. PACF Plot

Figure 4. Differenced Series

The pattern observed in Figure 4 of a rapidly declining ACF and PACF suggests a Moving Average model. Thus, an ARIMA (0,1,1) model appears to be a good starting point. The introduction of the POS system at period 91 was coded as a step variable with a value of "0" in periods 1-90 and of “1” in the remaining periods. The dummy variable for the intervention was entered into the ARIMA (0,1,1) model as an independent variable similar to the process in least squares regression analysis. The results are seen in Figure 5.

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	176	17384.291	95.961067

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
MA1	.99988	3.7976773	.263288	.79263628
DUMMY	-187.74582	2.9205662	-64.284049	.00000000
CONSTANT	.00614	.0281037	.218624	.82719663

Figure 5. Intervention Model for Billing

Figure 5 indicates that the impact of our intervention, coded as the dummy variable, is highly significant. Curiously, the moving average component is not. The analysis was repeated eliminating the moving average component and including only the integrated or difference component, an ARIMA (0,1,0). The results still produced a significant intervention variable. However, the test of a good ARIMA model is in the analysis of the residual. The plot of the ACF and PACF for the residuals for the ARIMA (0,1,0) model revealed several significant ACF’s and PACF’s indicating a poor fit. The residuals analysis for the ARIMA (0,1,1) showed only white noise in the residuals (Figure 6). The conclusion based on Figure 6 is that the moving average component is important for the overall model although not statistically significant when included with the dummy-coded intervention variable. This conclusion is affirmed when examining the total residual variance.

The residual variance for the ARIMA (0,1,1) model plotted in Figure 6 is 95.96. Residual variance increases to 184.26 for the ARIMA (0,1,0) model. Clearly, the ARIMA (0,1,1) model is a better fit for the data. More important, we can conclude that the introduction of the POS system impacted billing errors significantly at this firm. A further confirmation of the superiority of the ARIMA (0,1,1) lies in two statistics that test the fit of an ARIMA model: Akaike Information Criterion (AIC) and Schwartz Bayesian Criterion (SBC).[1] Both statistics are smaller for the ARIMA (0,1,1) model (AIC = 1333, SBC = 1342) compared to the ARIMA (0,1,0) model (AIC = 1409, SBC = 1419) indicating a better fit.

This first case study demonstrates the simplest form of outcome-based evaluation. A problem observed in a single performance variable (pricing errors) was the primary reason for the system’s implementation and subsequently became the primary outcome evaluated. The IT implemented

[1] The smaller the number, the better the fit for both AIC and SBC. Both AIC and BIC have solid theoretical foundations in information theory (for AIC) and integrated likelihood in Bayesian theory (for BIC). If the complexity of the true model does not increase with the size of the data set BIC is the preferred criterion, otherwise AIC is preferred.

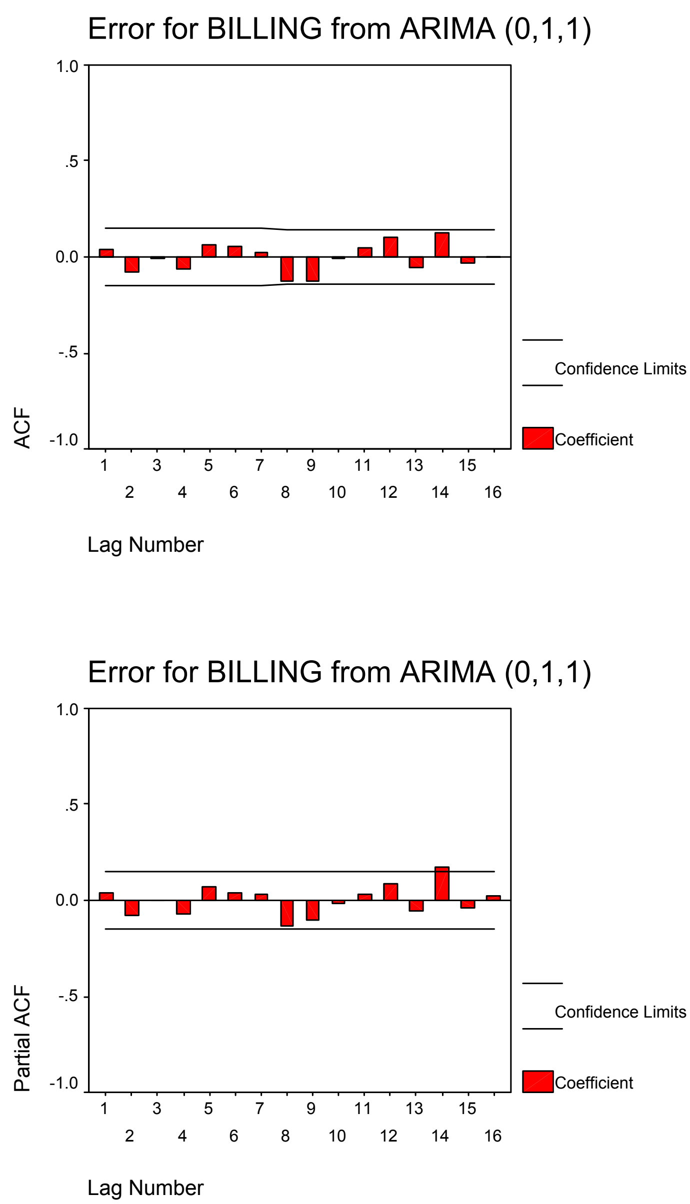


Figure 6. Comparison of ARIMA Errors for Billing

in this case was a simple point-of-sale system with obvious impacts on the outcome variable. The more interesting aspect of this first case is the application of the intervention analysis methodology that can be used in other fields of research for the evaluation of information systems. The next two cases demonstrate different applications.

The residual variance for the ARIMA (0,1,1) model plotted in Figure 6 is 95.96. Residual variance increases to 184.26 for the ARIMA (0,1,0) model. Clearly, the ARIMA (0,1,1) model is a better fit for the data. More important, we can conclude that the introduction of the POS system impacted billing errors significantly at this firm. A further confirmation of the superiority of the ARIMA (0,1,1) lies in two statistics that test the fit of an ARIMA model: Akaike Information

Criterion (AIC) and Schwartz Bayesian Criterion (SBC).[2] Both statistics are smaller for the ARIMA (0,1,1) model (AIC = 1333, SBC = 1342) compared to the ARIMA (0,1,0) model (AIC = 1409, SBC = 1419) indicating a better fit.

This first case study demonstrates the simplest form of outcome-based evaluation. A problem observed in a single performance variable (pricing errors) was the primary reason for the system's implementation and subsequently became the primary outcome evaluated. The IT implemented in this case was a simple point-of-sale system with obvious impacts on the outcome variable. The more interesting aspect of this first case is the application of the intervention analysis methodology that can be used in other fields of research for the evaluation of information systems. The next two cases demonstrate different applications.

CASE 2

Case 2 uses a second time series as a control group for the primary intervention analysis. The business was a 19-member law firm that had implemented integrated case management software. Features of the software included cataloguing general information about each case, determining which attorney's billable hours were assigned to each case, developing witness and exhibit lists, identifying critical points in the adjudication of each case, and creating a calendar to track filing deadlines and court dates. The primary reason for the implementation of the software was the rising costs associated with the clerical time devoted to doing each of these tasks manually. In addition, it was hoped that the software would assign billable hours more accurately and meet both deadlines and court dates.

The time series used for the intervention analysis was total clerical hours devoted to case management weekly. Operationally defined, this variable was the weekly hours logged by two clerical employees whose responsibility it had been to assign billable hours, compile witness and exhibit lists, and track deadlines and court appearance dates. The evaluation used 25 weeks of data prior to the implementation of the case management software and 17 weeks after the implementation of the software; the intervention point was the beginning of week 26. Figure 7 shows the time series for the 42 weeks of data used in this analysis. A "step" dummy code was used in which the value of the intervention variable was "0" for the first 25 weeks and "1" for the next 17 weeks.

The ARIMA analysis proceeds exactly as in the first case. First, the ACF and PACF are plotted for the original series. The initial plot revealed a non-stationary series that need to be differenced. The ACF and PACF plots of the differenced series indicated a Moving Average model: ARIMA (0,1,1). The intervention analysis using this model is shown in Figure 8.

---

[2] The smaller the number, the better the fit for both AIC and SBC. Both AIC and BIC have solid theoretical foundations in information theory (for AIC) and integrated likelihood in Bayesian theory (for BIC). If the complexity of the true model does not increase with the size of the data set BIC is the preferred criterion, otherwise AIC is preferred.

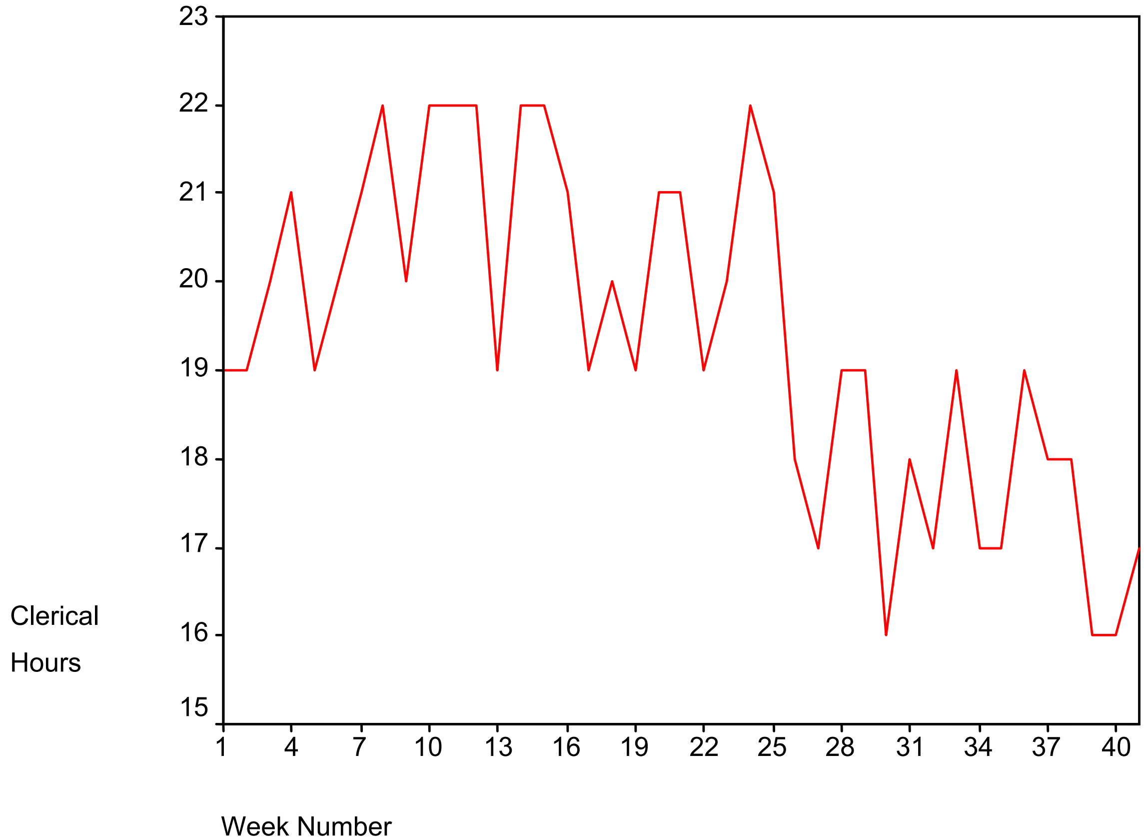


Figure 7. Clerical Hours

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	37	68.709564	1.8041860

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
MA1	.8275089	.10011259	8.2657829	.00000000
DUMMY	-1.9474949	.91164944	-2.1362322	.03934141
CONSTANT	-.0262977	.05074233	-.5182590	.60736313

Figure 8. Intervention Model

In this analysis that the dummy variable for the case management system impacted clerical hours significantly. This model was tested against other models using the same criterion as in Case 1: residual variance, residual analysis, and AIC and SBC statistics. Each comparison indicated that the ARIMA (0,1,1) model was the best fit.

The need for a control group stems from the fact that hours logged to case management could easily depend on the overall workload in the firm during the 42 weeks of the investigation. The obvious control variable would be total weekly billed hours during the 42 weeks of the investigation. This control variable is compelling because weekly billed hours is the single best operational definition of workload in a law firm. Figure 9 shows the time series of total billed hours for the 42-week period investigated.

The control group data could be analyzed in several ways. The easiest and most straightforward was to perform an ARIMA analysis using the control variable (total weekly billed hours) as the dependent variable and the dummy-coded intervention variable as the independent variable. Following the procedure outlined above. The initial plots of the ACF and PACF indicated a stationary series making differencing unnecessary. No clear pattern was otherwise indicated. Therefore, both ARIMA (1,0,0) and ARIMA (0,0,1) models were fitted. The ARIMA (1,0,0) model was slightly superior based on residual variance and AIC and SBC statistics. The residuals analysis also indicated a good fit. This analysis indicates that the case management software was unrelated to total billed hours. It is a plausible inference that hours logged to case management changed because of the system not because of a change in workload in the firm during the period investigated. Figure 10 shows the output of the ARIMA (1,0,0)

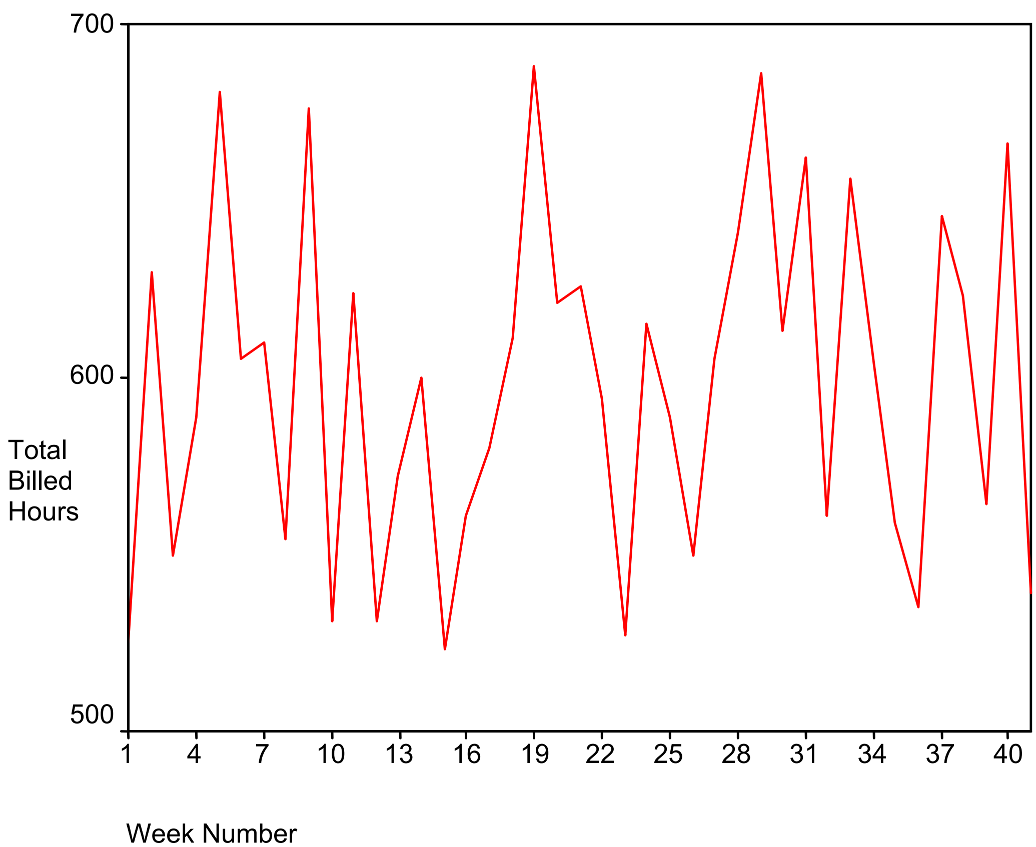


Figure 9. Total Billed Hours

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	38	90702.625	2384.9952

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
AR1	-.17997	.163759	-1.098990	.27868688
DUMMY	13.85079	13.259172	1.044619	.30279890
CONSTANT	592.88344	8.498477	69.763490	.00000000

Figure 10. ARIMA (1,0,0) Output

In Figure 10 the AR1 variable is statistically insignificant. It is clear that the total hours variable is a random walk model, ARIMA (0,0,0). Since at least one parameter must be non-zero for the ARIMA analysis to be performed, the autoregressive component is included since it produced the best fit.

A second way to analyze the control group variable is to include it as an additional independent variable in the analysis of clerical hours. This process would be similar to stepwise regression analysis using clerical hours as the dependent variable and total hours and the dummy-coded intervention as independent variables. The results of the ARIMA (0,1,1) model with total hours and the intervention variable as independent variables is shown in Figure 11. Two things are clear from this analysis. First, the impact of the case management software included in the model as the dummy variable is statistically reliable even with the effects of total hours included in the equation. Secondly, total hours are unrelated to clerical hours. Consequently, this model makes a compelling case for the efficacy of the case management software in reducing clerical hours independent of changes in total billed hours.

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	36	67.782835	1.8286539

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
MA1	.8301455	.10309088	8.0525606	.00000000
TOTHRS	-.0028940	.00412604	-.7013895	.48756958
DUMMY	-1.9262672	.91593851	-2.1030530	.04251560
CONSTANT	-.0261258	.05064614	-.5158495	.60911339

Figure 11. Results of ARIMA (0,1,1) Model

CASE 3

The final case involves a medium-sized medical services organization with nearly 1000 employees and an annual budget in excess of $50 million. The IT implemented here was an inventory control system for medical/surgical supplies. The new system was imposed externally on the organization because it had been part of the package when they changed medical/surgical supply vendors. Two primary outcome variables were selected for this analysis:

- weekly cost of medical/surgical supplies and
- satisfaction with the new inventory control system as reported on weekly satisfaction surveys conducted by patient care supervisors.

The first outcome variable was one that the Board of Directors and executive team was very interested in, and the second outcome variable was of great concern to the VP for Patient Services and the Director of Human Resources. Important control variables included patient census, overall full time equivalent (FTE) staffing levels in patient care, and general job satisfaction levels on weekly surveys. Data were collected for 15 weeks prior to the implementation of the new system and 18 weeks afterwards.

The system implementation variable was coded as before as a dichotomous dummy variable using “0” for the first 15 weeks and “1” for the 18 weeks following implementation. For the two outcome variables, medical/surgical supply costs were determined compiling the cost of supplies used weekly, and staff satisfaction with the system was measured using a 5-point response scale taken from weekly staff surveys. Operationally defining the two control variables, weekly census figures were computed as weekly averages of daily patient counts, and general job satisfaction was taken from a question addressing this issue from the weekly staff surveys.

This case added some important dimensions to outcome-based system evaluation. First, it involved two primary outcome variables. In complex organizations, IT implementations may have two or more primary outcomes. Outcome-based application of intervention analysis can easily respond to these situations. In addition, this case provided several different kinds of variables:

- financial variables from the cost figures,
- aggregate output variables from census records,
- staffing levels, and
- survey data reflecting general job satisfaction and specific satisfaction with the new vendor system.

The survey data was gathered specifically to evaluate that component of the system evaluation representing a true social experiment as proposed in the program evaluation literature. These variables were combined into an outcome-based evaluation to understand the impact of a specific IT on an organization.

The analysis began by examining the two primary outcome variables. First the medical/surgical supply costs outcome variable was used as the time series in the ARIMA analysis with hospital census, staffing level, and the dummy-coded intervention variable as independent variables. This analysis was performed as outlined in Cases 1 and 2 and resulted in the best fit of the supply cost time series. The results are shown in Figure 12.

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	26	107154101.4	4113873.5

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
AR1	-.23337	.19661	-1.1869442	.24598178
CENSUS	24.01836	26.78525	.8967009	.37810659
STAFFING	-31.83184	24.80998	-1.2830256	.21080265
DUMMY	-2060.49835	613.26768	-3.3598678	.00241709
CONSTANT	32449.34490	10436.54815	3.1092028	.00450761

Figure 12. Analysis of the Two Outcome Variables

The results displayed in Figure 12 indicate that the dummy-coded inventory control system impacted total supply costs significantly even allowing for any impact of census and staffing levels. It is understood that it is impossible to separate the cost benefit derived from the new information system from those derived from the new vendor because they were conjoined. However, the impact of the inventory system distinct from the vendor was assessed using the second outcome variable.

For the second analysis, the staff satisfaction level with the new inventory control system was used as the time-series variable with overall job satisfaction used as a control variable. An ARIMA (1,0,0) model was indicated following the procedure discussed in cases 1 and 2. As in the earlier analyses the intervention variable was entered last to measure the impact of the system over and above the impact of the control variable. Figure 13 shows the results of this analysis. The the system impacted staff satisfaction positively independent of overall satisfaction levels.

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	27	2.3966055	.08867517

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
AR1	-.1740155	.19574432	-.8889939	.38185858
JOBSATIS	.1042598	.20334134	.5127327	.61230785
DUMMY	.3071714	.09343884	3.2874055	.00280871
CONSTANT	3.4149819	.74306437	4.5958091	.00009024

Figure 13. Analysis of Staff Satisfaction with System

III. DISSCUSSION

The use of interrupted time-series design as part of an outcome-based evaluation of information systems proved highly beneficial. In this method of evaluation, systems were seen as interventions designed and implemented to address a problem or opportunity noted in an existing data series normally collected by the organization. The evaluation method seeks to identify a discontinuity in an important time series concomitant with the implementation of the system. Simply put, if a system is implemented to address some problem or opportunity as shown in a data series, effectiveness or success of the system must be assessed in terms of the impact on the very data series serving as the initial rationale for the implementation of that system. Although it is acknowledged that an information system may have many unanticipated benefits, to evaluate a system based on a criterion other than its stated goals is not evaluation; rather it is post-hoc justification.

The three cases presented here, include several research findings:.

1. Outcome-based evaluation using intervention analysis is effective on a wide range of outcome variables. Any variable that can be measured, quantified, and reliably collected can serve as a possible outcome variable for this type of system evaluation. While it is evident that if the outcome variable is not identified and adequately measurable, this methodology would not work. However, this would be a fatal flaw for all other evaluation methods!

2. The adaptability of this outcome-based approach to multiple systems outcomes was established. Of course, some practical limit exists for the number of outcomes examined given the possible interactions among outcome variables. Nonetheless, the requirement that an evaluation method must accommodate multiple outcomes was met successfully in these cases.

3. The impact of intervening or moderating variables can be assessed by including them in the ARIMA model as additional independent variables. This moves outcome-based evaluation to a level of making causal inferences about the impact of IT as Cook and Campbell [8] have clearly demonstrated. Business enterprises, then, would be able to assess systems flaws or benefits from among several complex results

A wide array of tools can potentially make outcome-based systems evaluation even more sophisticated. If one understands that IT systems are interventions in the same way that social programs, medical treatments, and managerial decisions are interventions, many tools used to assess the impacts of those interventions become available. For example, any textbook on time-series analysis demonstrates how the coefficients resulting from the ARIMA analysis can be used to quantify the level and rate of change in an outcome variable, given a specific level of application of an intervention variable. If the outcome variable is monetary in nature, the exact financial impacts can be assessed and used in computing payback periods, break-even analysis, and financial ratios resulting from IT implementation.

The whole idea of outcome-based evaluation and intervention analysis is rooted in program evaluation and evaluation research in which interventions of all types are viewed as social experiments. Similarly, the application of IT to address specific problems or opportunities must be viewed as organizational experiments in field settings if effective systems evaluation is to occur. In this way, the IS body of knowledge is enhanced through theoretical models of assessment with subsequent “real-world’ applied tests. Outcome-based evaluation is a powerful tool in this endeavor.

Editor’s Note: This article was received on June 29, 2001. It was with the author for 10 months for one revision. It was published on June15, 2003.

REFERENCES

Amoake-Gyampah, K. and White, K.B. (1993) "User Involvement and User Satisfaction: An Exploratory Contingency Model," *Information & Management*, (25) 1, pp 1-10.

Bailey, J.E. and Pearson, S.W. (1983) "Development of a Tool for Measuring and Analyzing Computer User Satisfaction," *Management Science*, (29) 5, pp 530-544.

Baroudi, J.J., Olson, M.H., and Ives, B. (1986) "An Empirical Study of the Impact of User Involvement on System Usage and Information Satisfaction," *Communications of the ACM* (29) 3, pp 232-237.

Boehm, B. and Bell, T. (1977) "Issues in Computer Performance Evaluation-Some Consensus, Some Divergence," *Performance Evaluation Review* (4) 2, pp 4-40.

Borovits, I. and Giladi, R. (1993) "Evaluating Cost/Utilization of Organizations' Information Systems and Users," *Information and Management* (25) 3, pp 273-280.

Chandler, J. (1982) "A Multiple Criteria Approach for Evaluating Information Systems," *MIS Quarterly* (6) 1, pp 61-74.

Cook, T.D. and Campbell, D.T. (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Boston MA: Houghton Mifflin, pp 207-292.

DeLone, W.H. and McLean E.R. (1992) "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research*, (3) 1, pp 60-95.

Dennis, A.R., George, J.F., Jessup, L.M., Nunamaker Jr., J.F. and Vogel, D.R. (1988) "Information Technology to Support Electronic Meetings," *MIS Quarterly* (12) 4, pp 591-624.

DeSanctis, G. and Gallupe, R.B. (1987) "A Foundation for the Study of Group Decision Support Systems," *Management Science* (33) 5, pp 589-609.

Evans, G.E. (1999) "Information Systems as Interventions, The Case for Outcomes Based Evaluation," *Journal of Computer Information Systems* (39) 2, pp 69-74.

Gatian, A.W. (1994) "Is User Satisfaction a Valid Measure of System Effectiveness?" *Information & Management* (26) 1, pp 119-131.

George, J.F., Nunamaker, Jr., J.F., and Vogel, D.R. (1994) "Group Decision Support Systems and Their Implications for Designers and Managers: The Arizona Experience," *Transactions of the Eighth International Conference on Decision Support Systems*, Boston, MA, pp 13-25.

Iivari, J., and Ervasti, I. (1983) "User Information Satisfaction: IS Implementability and Effectiveness," *Information & Management*, (27) 2, pp 205-220.

Ives, B., Olson, M.H., & Baroudi, J.J. (1983) "The Measurement of User Information Satisfaction," *Communications of the ACM* (26) 10, pp 785-793.

Keen, P.G.W. and Scott-Morton, M (1977) *Decision Support Systems: An Organizational Perspective*, Reading, Mass: Addison-Wesley.

McDowall, D., McCleary, R., Meidinger, E.E., and Hay, R.A. Jr. (1980) *Interrupted Time Series Analysis*, Beverly Hills: Sage Publications.

McKeen, J.D., Guimaraes, T., & Wetherbe, J.C. (1994) "The Relationship Between User Participation and User Satisfaction: An Investigation of Four Contingency Factors," *MIS Quarterly*, (18) 4, pp 427-449.

Poavac, E.J. and Carey, R.G. (1980) *Program Evaluation Methods and Case Studies*, Englewood Cliffs, NJ: Prentice Hall.

Senge, P.M. (1990) *The Fifth Discipline: The Art and Practice of The Learning Organization*, New York: Currency Doubleday.

Shannon, C.E. and Weaver, W. (1949) *The Mathematical Theory of Communication*, Urbana, IL: University of Illinois Press.

Sprague, R.H. and Carlson, E.D. (1982) *Building Effective Decision Support Systems*, Englewood Cliffs, NJ: Prentice Hall.

SPSS (1993) *SPSS for Windows, Trends, Release 6.1*, Chicago: SPSS Inc., pp 137-151.

Wichern, D.W. and Jones, R.H. (1977) "Assessing the Impact of Market Disturbances Using Intervention Analysis," Management Science (24) 3, pp 329-337.

ABOUT THE AUTHORS

Gerald E. Evans earned a Ph.D. in 1985 from The Claremont Graduate University School of Behavioral and Organizational Studies and completed post-doctoral study in Information Sciences at Claremont in 1986. His undergraduate degree is from The University of Great Falls and he earned two Masters Degrees from Montana State University. He currently teaches in the areas of Information Systems, Leadership, and The Contemporary Organization. He won several teaching awards including his department undergraduate teaching award in 2002-2003 and 1991-1992, and the Outstanding MBA Faculty Award in 1998-1999, 1996-1997, and 1995-1996. His research areas include Systems Analysis and Design and Information Systems Evaluation. Dr. Evans is active in consulting in the business and professional community and serves on the Board of Directors of the Montana Chamber of Commerce Foundation and of Partners in Homecare.

Bruce A. Costa received a B.S. in Industrial Engineering/Operations Research from The University of Massachusetts/Amherst and a Ph.D. in Business Administration with an emphasis in Finance from Florida State University. His academic research focuses on mutual fund performance. He joined the faculty at the University of Montana in August of 2000 and teaches courses at both the graduate and undergraduate level in finance. He was presented the Outstanding Faculty Award for the Department of Accounting and Finance for the academic year 2001 - 2002. He spent 13 years in the Navy on active duty in various positions as a helicopter pilot, test pilot, and assistant navigator aboard the USS Saratoga during the Gulf War. He piloted over 25 different aircraft and logged over 2800 hours of accident-free flying. In 1990 he was selected for membership in the *Society of Experimental Test Pilots.*

Communications of the Association for Information Systems

ISSN: 1529-3181