THE HECKMAN BINDERY, INC.
North Manchester, Indiana

KRI

H or V	JUST	FONT	SLOT	TITLE
H	CC	1W	22	BEBR
			21	FACULTY
			20	WORKING
			19	PAPER
H	CC	1W	8	1989
			7	NO.1555-1572
H	CC	1W		330
				B385<"CV">
				no.1555-1572
				cop.2
				<IMPRINT>
H	CC	7W		U. of ILL.
				LIBRARY
				URBANA

BINDING COPY

PERIODICAL ☐ CUSTOM ☐ STANDARD ☐ ECONOMY ☐ THESIS
BOOK. ☐ CUSTOM ☐ MUSIC ☐ ECONOMY AUTH. 1ST ☐
NO. VOLS. THIS TITLE — LEAD ATTACH

ACCOUNT: 66672 — LIBRARY: 001 — NEW — RUB OR SAMPLE — TITLE I.D.: 6632 — COLOR: FOIL: WHI — MATERIAL: 488

ACCOUNT NAME: UNIV OF ILLINOIS — ISSN.

ACCOUNT INTERNAL I.D.: B01912400 — NOTES — BINDING FREQUENCY — WHEEL: 1 — SYS. I.D: 3 — 39256

I.D. #2: STX3

COLLATING: 35

ADDITIONAL INSTRUCTIONS
Dept=STX3 Lot=#20 Item=143 HNM=1E14
1=R2ST3CP MARK BY # B4 91

SEP. SHEETS | PTS. BD. PAPER | TAPE STUBS | CLOTH EXT | GUM | FILLER | STUB | LEAF ATTACH.

POCKETS: PAPER | BUCK | CLOTH — SPECIAL PREP

INSERT MAT. — ACCOUNT LOT NO.: #20 — JOB NO.: HV363

PRODUCT TYPE: 11 — ACCOUNT PIECE NO.: 143 — PIECE NO: 85

HEIGHT: 11.2 — GROUP CARD: 1 — VOL THIS TITLE

COVER SIZE: X

001247549

BEBR
FACULTY WORKING
PAPER NO. 89-1570

Forecasting Daily Cash Receipts and Disbursements: A General Statistical Approach

Monzurul Hoque
James A. Gentry
Paul Newbold

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 89-1570

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

May 1989

Forecasting Daily Cash Receipts and Disbursements: A General Statistical Approach

Monzural Hoque
Loyola University of Chicago

James A. Gentry, Professor
Department of Finance

Paul Newbold, Professor
Department of Economics

FORECASTING DAILY CASH RECEIPTS AND DISBURSEMENTS:
A GENERAL STATISTICAL APPROACH

Abstract

This study provides a theory and application of a general statistical model for forecasting daily cash receipts and disbursements. The contributions of the cash forecasting literature are reviewed and structured into a three dimension problem space. A general model is presented that generates the daily cash flow forecast using ARIMA and/or regression methods. Whenever a time series structure is present, the forecast is obtained by combining both methods, otherwise the regression specification is considered adequate. Following the explanation of the general model, daily cash receipt and disbursement data from a large public pension fund and a Fortune 500 company were used to demonstrate the application of the proposed model. When daily cash receipt data for a public pension system were used, a time series structure was identified, and the general model was applied to this data to generate a forecast. As expected, the forecast based on the general model performed better than the pure regression model. However, daily cash receipt and disbursement data for a Fortune 500 company did not exhibit a time series structure, thus the regression technique was considered adequate for forecasting. In summary, to ensure an efficient and reliable daily cash forecast, this study recommends the use of a general approach.

FORECASTING DAILY CASH RECEIPTS AND DISBURSEMENTS: A GENERAL STATISTICAL APPROACH

Daily cash forecasting is an extremely important practical problem that has historically lacked a rigorous methodology. The leading research on daily cash forecasting has been completed by Stone [15], Miller and Stone [9], Stone and Miller [16, 17, 18, 19] and Stone and Wood [20]. This study provides a theory and an application of a model for forecasting daily cash receipts and disbursements. The theory is based on a methodology that combines two statistical techniques, regression and a time series analysis, ARIMA. The application of the general model is based on 58 months of daily cash receipt data from a Public Pension System and 12 months of daily cash flow data from a Fortune 500 company.

In theory, shareholder wealth is created by the efficient management of the flow of cash receipts and disbursements as well as the amount of cash balances maintained. Corporate cash management is broadly divided into two interrelated value creating activities, namely, cash balance management and cash flow management. Cash balances are maintained within a specified range in order to meet transaction, precautionary, speculative and compensating balance requirements. Experience indicates that the success of cash management is closely related to reliable forecasts of future cash flows and balances. If the timing and amount of future cash inflows match the cash outflows, it is not necessary to maintain a cash buffer. Likewise, under perfect financial market conditions a company does not need cash balances. Problems posed by the lack of the synchronization of cash inflows and outflows in a perfect market can be settled immediately by lending or borrowing at the market rate of interest. Beranek [1], Cohn and Pringle [3], and Morris [10] have indicated that if a perfect market prevails, then the cash flow problem generated by conditions of uncertainty can be solved along the lines mentioned above. Hence, the optimal strategy would be to do without cash balances and rely on the perfect market to satisfy cash needs.

Cash balance management becomes relevant when financial markets are imperfect. That is, the importance of cash rises when there is a cash flow short-fall and short-term borrowing is not readily available. In contrast, there is an opportunity cost to the firm if it chooses to carry excessive idle cash balances. The challenge, therefore, is to keep cash balances at a minimum. Under conditions of uncertain cash flows, cash balance management is dependent on an accurate forecast of receipts and disbursements.

Fundamentally, value creation is closely related to cash flow management, which, in turn, is dependent on the accuracy of the forecasts of daily cash receipts and disbursements. The flow of cash into and out of a firm's cash pool is determined by management decisions associated with credit terms to customers and suppliers as well as cash gathering, mobilization, and concentration activities, Srinivasan and Kim [14]. These activities play a pivotal role in forecasting the level and speed of cash flow, as well as affecting cash flow stability and patterns.

A primary objective of this study is to provide a generalized model that will enrich and improve the cash forecasting process. In the next section we provide a brief review of the cash forecasting literature in a unique three dimensional format. A general cash forecasting model, which combines regression and ARIMA techniques, is presented in Section II. Section III provides the results of an empirical analysis, and conclusions are developed in Section IV.

I. REVIEW OF THE CASH FORECASTING LITERATURE

Cash forecasting is generally divided into quantitative and qualitative components, Kallberg and Parkinson [7]. However, the academic literature has focused primarily on the quantitative approach to cash forecasting. The structure of the cash forecasting literature encompasses three distinct conceptual frameworks based on accounting, optimization and statistical information and techniques. Exhibit 1 uses a three dimensional framework to structure the forecasting concepts and models. This approach not only will enable us to review the existing literature, but also will unfold possible unchartered territory for future research. Corner 1 shows a qualitative approach to forecasting that does not use an optimizing or statistical framework, but rather relies on accounting information. In general, corner 1 forecasts are based on the qualitative judgment of experts, e.g., the Delphi method. In corner 2 the cash forecast is based on an accounting based approach that utilizes pro forma income statements, balance sheets, and cash budgets. The accounting approach to cash forecasting assumes that a company can estimate with reasonable accuracy its inflows and outflows for future periods, and these financial statements generate the firm's cash balance position. The accounting approach is dependent on the reliability of the forecast of cash receipts and disbursements and knowledge of the simultaneous interrelationship that exists among the numerous financial components.

Optimizing models are located at corner 4 and are founded on well established theoretical relationships and make a significant analytical contribution to the cash forecasting process. Models by Baumol [2] and Miller and Orr [8] recognize the linkage between cash balances, on the one hand, and cash flow on the other, but for reasons of tractability each makes unique assumptions concerning the cash flow generating process. Baumol [2] was the first to recognize that the flow of cash resembled the flow associated with inventories and hence, if cash is idle, there is a carrying cost. He built an inventory theoretic model under the assumption that a firm's disbursement rate remains constant. The model determines the level of cash that balances the cost of holding extra dollars against the cost of investing extra dollars in marketable securities. Miller and Orr [8], on the other hand, assumed that cash patterns were purely random and proceeded to determine the optimal cash level for the firm. Naturally, the realism of the assumptions in an inventory theoretic optimizing model determines the quality of the cash forecasts that are generated.

Morris [10] used the capital asset pricing model to design an aggregate cash management system in a market risk adjusted valuation framework. Morris's contribution represents a purely theoretical optimizing model. In contrast the cash balance models of Baumol [2] and Miller and Orr [8] focus only on the cost side, which ignore the effect of risk on the value of the firm. Therefore, these two models are suboptimal from the viewpoint of financial theory which considers the trade-off between risk and return. Applying the CAPM, Morris considers both risk and return and thus could derive optimal cash management policies.

Optimizing cash forecasting models assume a specific cash generating process, while a statistical based approach theorizes there is a well defined relationship that exists between cash flow and various independent variables. Corner 3 represents a statistical framework that was pioneered by Stone and Wood [20]. In a regression based forecasting approach, the variables involved in generating the cash forecasts are assumed to be stable. Stone and Wood (SW) do not explicitly forecast cash flows, but rather use a dummy variable regression technique to distribute the estimated monthly cash receipts and disbursements into a daily cash forecast. The SW approach is not a pure forecast of daily cash flows because it is dependent on a company providing estimates of total cash receipts and disbursements from its monthly budgeting process. The SW methdology generates estimated daily values based on a day-of-the-week (DOW) and day-of-the-month (DOM) structure. The SW approach can be divided into three steps. First, major cash receipts and disbursements are separated from the data because they are assumed to be highly predictable. The remaining non-major receipts and disbursements are regressed to measure day-of-the-week and day-of-the-month coefficients. Finally, the regression coefficients are used to distribute the forecasted monthly totals over the work days in a specified month. Because the SW statistical approach to daily cash forecasting was incomplete, Miller and Stone

[9] and Stone and Miller [15] subsequently introduced an array of specifications that improved the SW forecast performance.

Another corner 3 alternative to regression modelling is a time series forecast of cash receipts and disbursements. Instead of explicitly expounding the relationship between dependent and independent variables, time series analysis assumes the underlying data generating process is stable and establishes a solid relationship between current and past performance of the dependent variable. Stone and Wood [20], and later, Stone and Miller [16] observed that a straightforward application of the time series analysis, ARIMA, to cash forecasting is inappropriate because autoregressive models rely on the periodicity of data. They showed cash flow data contain large flows which are non-periodic and generally cannot be estimated from past history. Stone and Miller [16] present income tax payments as an illustration of a major cash flow. Tax payments occur four to six times per year without being periodic, and a particular tax payment is not strongly related to past cash flows or tax payments. Hence, application of ARIMA to cash flow data containing tax payments or any other major flow is inappropriate, therefore, they rejected the ARIMA technique in favor of regression.

Corner 7 combines an accounting and an optimizing framework based on a distribution approach. For example, Robichek, et al. [12] developed a linear programming model to forecast cash flows and used information from the financial statements to formulate various constraints. Later, Orgler [11] constructed a comprehensive linear programming model with several constraints, which he was able to forecast using uneven time periods. Each model optimizes operating decisions subject to financial constraints including the opportunity cost of long term funds.

The Stone and Wood [20] distribution approach to cash forecasting is located at corner 5 and is based on combining accounting information into a statistical framework. Corner 8 uses all three approaches. To date no system has integrated all three frameworks simultaneously, however, Stone's [15] cash management model is closely related to corner 8. He introduces accounting based cash forecasts into the optimizing framework of Miller and Orr [8] and the cash forecast is based on Stone and Miller's [16] distribution approach. Thus, Stone's [15] cash management model indirectly uses all three frameworks.

In this study we concentrate on statistical modelling located at corner 3. The traditional approach to cash forecasting has generally used either a regression or an ARIMA model. However, it is our judgment that such a separation is unnecessary. Combining both approaches eliminates individual limitations and utilizies the best features of each approach. The primary task of this paper is to develop a general approach to statistical cash forecasting.

II. A GENERAL DAILY CASH FORECASTING MODEL

Regression modelling is useful in cash forecasting as long as the theoretical relationships specified reflect what actually happens. The specification of the relationship between dependent and independent variables is crucial to achieving accurate cash forecasting. Hence, there is a need to identify the true relationship among the components used in daily cash forecasting. The process of identifying true relationships in the process of daily cash forecasts can be carried out in one of the following ways:

1. An ad hoc depiction of relationships that ignores interdependencies, but tests these relationships with actual data. We shall refer to this approach as an ad hoc single equation regression approach.

2. The theoretical depiction of relationships that ignores interdependencies and the testing of these relationships with actual data. This approach is called a theory based single equation regression approach.

3. An ad hoc simultaneous equation regression approach[1] is similar to (1) except that it takes into account interdependencies of relationships.

4. A theory based simultaneous equation approach is similar to (2) except that it takes into account interdependencies of relationships.

Stone and Wood [20] basically follow the single equation approach. Experience indicates that cash management components are interrelated and hence, a theory based simultaneous equation approach is most desirable. This requires a rich mathematical understanding and development of these relationships. Currently, these relationships are neither well understood nor developed, therefore, we can only attempt to specify true relationships by focusing on the unspecified segment in the equation. Step 1 in a general daily cash forecasting model requires application of the Stone and Wood [20] approach.

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \varepsilon_t \qquad (1)$$

where

y_t = cash receipts or disbursements;

x_{1t} = dummy variables for days of the week, days of the month, and holiday effects;

ε_t = error term.

Stone and Miller observed that the specification in equation (1) can be improved by using a multiplicative form. The benefit of a multiplicative specification must be weighed against the costs of modelling effort and loss of data.

The next step is to analyze the error term in equation (1) by using ARIMA procedures. Because the error term in a regression is considered as unexplained information, an effort is made to reduce the unexplained portion by increasing the explained portion, e.g., as shown in Stone and Miller. In the generalized approach, the objective is to reduce the unexplained variance by expounding solid theoretical relationships and identifying a specification of the model through analysis of residuals in a time series framework. If the error term exhibits a time series relationship, the inference is that the modelling in the first step was incomplete. Otherwise, the first step is considered adequate.

The time series pattern can be characterized in three ways, viz. (a) autoregressive (AR), (b) moving average (MA), and (c) autoregressive moving average (ARMA). The autoregressive model (AR) assumes the current residual observation ε_t is a linear combination of the past p observations plus a random term. The following equation represents an AR structure:

$$\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \ldots + \phi_p\varepsilon_{t-p} + e_t \tag{2a}$$

where

$\phi_{i's}$ = AR coefficients to be estimated,

$\varepsilon_{t-i's}$ = p observations of the time series for the residuals,

e_t = a random disturbance.

The MA model represents the current observation as a linear combination of the past random disturbances plus a random term.

$$\varepsilon_t = -\Theta_1 e_{t-1} - \Theta_2 e_{t-2} - \ldots - \Theta_q e_{t-q} + e_t \tag{2b}$$

where

$\Theta_{i's}$ = MA coefficients to be estimated,

$e_{t's}$ = random disturbances.

A generalization of the AR(p) and MA(p) models that includes both AR(p) and MA(q) models as special cases is the mixed ARMA(p,q) model:

$$\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \ldots - \phi_p\varepsilon_{t-p} - \Theta_1 e_{t-1}$$

$$- \Theta_2 e_{t-2} - \ldots - \Theta_q e_{t-q} + e_t \tag{2c}$$

where each term has been defined earlier. For the ARMA (4,4) case:

$$\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \phi_3\varepsilon_{t-3} + \phi_4\varepsilon_{t-4} - \Theta_1 e_{t-1}$$

$$- \Theta_2 e_{t-2} - \Theta_3 e_{t-3} - \Theta_4 e_{t-4} + e_t \tag{2d}$$

Focusing on a pure AR process in equation (2a) allows us to illustrate the general model.

If equation (2a) can be identified, the inference is that equation (1) is an incomplete specification and, thus, in determining the final forecast, the estimates of equation (2a) are substituted into equation (1). On the other hand, if no such time series relationship can be identified, the inference is that the specification in equation (1) is efficient enough to generate an adequate forecast. Thus, Stone and Wood's [20] distribution approach to cash forecasting is a special case of the general approach as shown in Exhibit 2.

Assuming that equation (2a) is identified, the next step is to estimate it[2] and obtain equation (3).

$$\hat{\varepsilon}_i = \hat{\phi}_0 + \hat{\phi}_1 \varepsilon_{t-1} + \hat{\phi}_2 \varepsilon_{t-2} + \ldots + \hat{\phi}_p \varepsilon_{t-p} + \hat{e}_t \qquad (3)$$

where " represents estimates.

Finally, equation (3) is substituted into equation (1) to obtain

$$y_t = \hat{\alpha} + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t} + \ldots + \hat{\beta}_n x_{nt} + \hat{\varepsilon}_t \qquad (4)$$

Our forecast then would be $\hat{y}_{t+h}$ corresponding to $\hat{\varepsilon}_{t+h}$ where h is the number of periods to be forecast.

In summary, the general approach to forecasting cash receipts or disbursements is:

Step 1. Specify and regress the cash flow component on various explanatory variables including dummy variables;[3]

Step 2. Identify time series component of the residual term;

Step 3. Estimate, if necessary, time series relationships;

Step 4. If Step 3 is necessary and an appropriate ARIMA structure can be selected, use the ARIMA information with the Step 1 result to determine the final forecast. If Step 3 is unnecessary or no time series structure is accepted by model selection criterion, then use the Step 1 result to determine a final forecast.

III. EMPIRICAL TEST

In this section two applications of the general model are presented: one with respect to a Public Pension System (PPS) data and another with respect to a Fortune 500 company data. The initial tests are based on available daily cash receipts of the PPS data.

Major flows were removed from the data which covers the period September 1, 1983 through June 30, 1988. Thus we had 58 months of daily nonmajor cash receipts (CR). The cash disbursement occurred once each week, therefore, they were not of any value in searching for a daily or monthly effects.

Initially, cash receipts (CR) were regressed on holiday dummy variables, D1-D20 on the printout. The dummy variables are: DL = Labor Day, DLA = day after Labor Day, DC = Columbus Day, DCA = day after Columbus Day, DV = Veteran's Day, DVA = day after Veteran's Day, DT = Thanksgiving Day, DTA = day after Thanksgiving Day, DX = Christmas Day, DXA = day after Christmas Day, DN = New Year's Day, DNA = day after New Year's Day, DO = Martin Luther King Day, DOA = day after Martin Luther King Day, DW = Washington's birthday, DWA = day after Washington's birthday, DM = Memorial Day, DMA = day after Memorial Day. The plots of the residuals showed three outliers which were removed by introducing three more dummy variables (D21-D23).

An analysis of the plots showed significant autocorrelations at lags of 5, 10, 15 and 20 days. This observation suggested the introduction of day of the week (DOW) dummy variables. When regression on holidays and days after holidays (D1-D20), outliers (D21-D23) and days of the week (D24-D27) dummy variables was run, the autocorrelation of residuals at a lag of 5 became insignificant. Even though it is recognized in the literature, e.g., [16], [20], that days of the month (DOM) have a significant influence on cash receipts or disbursements, the presence of DOM effect was tested first through graphical analysis. The plot of residuals against DOM indicated there was no day of the month effect present in the cash receipts of the PPS. Nevertheless, an ANOVA test was run to determine analytically the presence of a DOM effect on CR. The results are given in Exhibit 3.

The hypothesis that DOM has no impact on the residuals cannot be rejected at a generally accepted level since the probability value is .20, i.e., the hypothesis could only be rejected at 20 percent or above which is an unacceptable type I error. Therefore, it is inferred analytically that DOM has no significant effect on CR. Thus the analysis made certain that DOM dummy variables were not essential.

The next step determined if there was a Month effect in the CR of the public pension system. Following the above procedure, it was discovered through graphical and ANOVA analysis, that the actual Month had a significant effect on CR. In Exhibit 4 the plot shows a systematic variation of residuals with respect to different months. If month has no impact on residuals, a random plot would be expected, indicating one month is as good as the other. However, Exhibit 4 shows a plot of a systematic variation of residuals with respect to different months, which indicates that the month does matter.

Hence, dummy variables for Months (D28-D38) were introduced into the regression. This completed the first stage of our analysis as stated in Section II.

Exhibit 5 provides the estimated coefficients for the regression of CR on holidays, days after holidays, outliers, days of the week and month dummy variables. Even though some of the dummy variables are insignificant, they cannot be dropped on the basis of t-statistics until it is certain that residuals are white noise. If the errors are otherwise, it is well known that t-statistics are unreliable. Therefore, in the second stage, the residuals from the above regression were checked for autocorrelations. Exhibit 6 provides the plots of the residual autocorrelations functions (ACF), inverse ACF, partial ACF and an analysis for white noise. The probability value of the chi-square test, that is given under "Autocorrelations Check for Residuals for White Noise" in Exhibit 6, rejects the hypothesis that residuals are white noise at conventional levels. Thus it suggests that simple regression on dummy variables is not adequate. The autocorrelation plots indicate that the series was stationary since the ACF declined exponentially. The stationarity was double checked by differencing the data. The differencing created a slowly decaying inverse ACF, indicating that the series is overdifferenced. Therefore, it was concluded that the analysis should use the original series.

Looking at the residual ACFs, a spike at a lag of 20 days can be observed. This suggests AR structure at lag of 20 days. On the other hand the spikes at lags of 1, 2, 4 and 5 on the inverse ACF plot suggest an MA structure at those lags. It is thus inferred that MA structure at lags of 1, 2, 4 and 5 and AR structure at lag of 20 might be present.

In the third stage, we estimated the time series structure: first with pure MA structure, then with pure AR structure and finally with a general ARMA structure. Because the Schwarz Bayesian Criterion (SBC) presented in Exhibit 7 are lowest for AR, it supports the pure AR time series structure.[4] On the other hand, the Akaike Information Criterion (AIC) supports a general ARMA time structure because the AIC results in Exhibit 7 are lowest for ARMA.[5] It was decided to extend the analysis not with a single chosen model but with all three possibilities. By now, however, two things are apparent: (1) The theory behind the general model is supported because a time series structure in the residuals can be identified, and (2) the estimates of regression in Exhibit 5 that ignored time series structure in the residuals are inefficient.

The fourth stage used a regression that recognized the specific time series structure in the residuals. Both SBC and AIC supported the ARMA model as these yielded lowest values for ARMA. The Maximum Likelihood Estimates of the regression incorporating ARMA structure in the residuals are presented in Exhibit 8. The t-statistics given in Exhibit 8 to determine the significance of the variables can now be relied on because estimates had been obtained after explicitly incorporating the information on the particular time structure present in the residuals. Finally, another regression with time series structure was run after dropping the insignificant variables. The estimates are provided in Exhibit 9. Comparing the results of Exhibit 9 with Exhibit 8, it is apparent that both SBC, AIC and standard error have decreased. This suggests that dropping the insignificant variables improved the fit of the model. An analysis of the plots of residual ACF, IACF and PACF confirmed that the data is pure white noise. Hence, it can be concluded the modeling was adequate.

Forecasts Using Pension Data

The focus now turns to forecasting. The sample forecasts generated were based on the estimates of two different models. First, for the regression model, the following equation is used to produce forecasts.

$$\hat{y}_t = a + b_1X_{1t} + b_2X_{2t} + \ldots + \varepsilon_t \qquad (5)$$

where a and b_i's are least square estimates and ε_t is assumed to be white noise. In sample forecats are generated by incorporating the values for X_1 through X_n, i.e., for D1-D38, for the respective observations. Prediction errors $(y_t - \hat{y}_t)$ are calculated for the entire sample.

Second, the general model makes use of information contained in the residual (ε_t) and incorporates that information into producing a final prediction for y. It uses equation 4 that was presented earlier:

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}_1X_{1t} + \hat{\beta}_2X_{2t} + \ldots + \hat{\varepsilon}_t \qquad (4)$$

where $\hat{\alpha}$ and $\hat{\beta}_i$'s are estimates generated after adjusting for ARMA structure in the residual which is represented by $\hat{\varepsilon}_t$. In sample forecasts are generated again by incorporating the values for independent variables for the respective observations. The prediction errors $(y_t - \hat{y}_t)$ are calculated for the entire sample. Root mean square error (RMSE) or standard error of the estimate were used to assess the relative predictive performance of the two models. The RMSE is the square root of the prediction error of all observations averaged by the degrees of freedom and it takes account of the size of the forecasting performance for the entire sample. In Exhibit 10 RMSE values are provided for both models. It is apparent that the RMSE of the general model is lower than that of the regression model. Therefore, the results show that on average the forecasts using the general model perform better than those that only use the regression model. The result is not surprising because regression model wastes information contained in the error structure.

Thus the general model is supported by the public pension system data in two ways: first, the point that residuals might indicate the presence of time series structure in a cash forecasting model proved to be true and second, cash forecasts based on the general model are on the average more precise than forecasts generated from a pure regression model.

Forecasting Using Fortune 500 Company Data

The analysis of the Fortune 500 company data set focuses on all three components of cash flows, viz., total disbursement (TD), cash receipts and net cash flow (NCF). As with the previous data, major flows were removed from the data which covers the period August 1, 1986 through July 31, 1987. The previous analysis is duplicated exactly for the Fortune 500 company data. After clearing holiday effects and outliers, the effects of DOW and DOM were checked graphically. In contrast to earlier data, a systematic DOM pattern was observed. Also the presence of DOW effect was found. The next step was the ANOVA analysis and the DOM effect was found to be significant. Therefore, a regression was run with holidays, days after holidays, outliers, DOW and DOM dummy variables. An analysis of the residuals from this regression showed the series to be stationary. Further, it was inferred that there might be an AR structure present in the residuals. However, SBC rejected any time series structure in the data. Therefore, it was concluded that unlike the pension fund data, the regression approach to cash forecasting was adequate for the data of the Fortune 500 company.

IV. CONCLUSION

The study provides a theory and an application of a model that combines regression and time series techniques. This methodology enables users to determine analytically the adequacy of the regression approach. If the regression analysis is adequate, as was the case for the data of the Fortune 500 company, the cash forecast that use regression are efficient. If, on the other hand there is a time series structure in the residual data, as was found in the Public Pension System data, the regression forecasts alone are inadequate.

That is the regression approach wasted information contained in the error structure. Therefore a cash flow forecast should utilize both techniques. Thus we see that the generalized methodology is a useful tool to ensure an efficient cash forecast.

NOTES

[1] The "Statement of Cash Flows" based on the direct method recognizes that cash inflows and outflows are generated through the interaction of all major balance sheet and income statement components. The information on the cash flow statement generates the data for modelling the cash flow process in a simultaneous equation framework. In the future, a simultaneous equation approach can be used to determine the complex relationships that exist among various cash flow components, which would be a substantive contribution to the cash forecasting literature.

[2] To estimate equation (2), the literature suggests that maximum likelihood estimators perform better (Judge et al. [5]).

[3] Simultaneous equation is superior and will yield better estimates due to specifications that capture interdependencies.

[4] Schwarz proposes that the particular values (p,q) for which

$$\log \sigma^2_{p,q} + (p+q)\log n/n$$

is smallest, be chosen.

[5] Akaike proposes that the particular value K, the order of the approximating autoregression, for which

$$\log \sigma^2_K + 2K/n$$

REFERENCES

[1] W. Beranek, "Towards a Positive Theory of Working Capital," Working Paper, September 26, 1986.

[2] W. J. Baumol, "The Transaction Demand for Cash: An Inventory Theoretic Approach," Quarterly Journal of Economics, Vol. 66 (November 1952), pp. 545-556.

[3] R. A. Cohn and J. J. Pringle, "Steps Toward an Integration of Corporate Financial Theory," Keith V. Smith, ed., Readings on the Management of Working Capital, St. Paul, West Publishing Company, 1980, pp. 35-42.

[4] C. W. J. Granger and P. Newbold, Forecasting Economic Time Series, Second Edition, Orlando, FL: Academic Press, Inc., 1986.

[5] V. Hodgson, "Cash Forecasting by Box-Jenkins," presented at TIMS/ORSA meetings, May 1979.

[6] G. G. Judge, R. C. Hill, W. Griffiths, H. Lütkepohl, and T. C. Lee, Introduction to the Theory and Practice of Econometrics, New York: John Wiley & Sons, 1982.

[7] J. E. Kallberg and K. Parkinson, Current Asset Management, New York: John Wiley and Sons, 1984.

[8] M. A. Miller and D. Orr, "A Model of the Demand for Money by Firms," Quarterly Journal of Economics, Vol. 80 (August 1966), pp. 413-435.

[9] T. W. Miller and B. K. Stone, "Daily Cash Forecasting: Alternative Models and Techniques," Journal of Financial and Quantitative Analysis (September 1985), pp. 335-351.

[10] J. R. Morris, "The Role of Cash Balances in Firm Valuation," Journal of Financial and Quantitative Analysis (December 1983), pp. 533-545.

[11] Y. Orgler, Cash Management: Methods and Models, Belmont, CA: Wadsworth Publishing Company, Inc., 1970.

[12] A. A. Robichek, D. Teichroew and J. M. Jones, "Optimal Short-Term Financing Decision," Management Sciences (September 1965), pp. 1-36.

[13] G. Schwarz, "Estimating the Dimension of a Model," Ann. Statist., Vol. 6, pp. 461-464.

[14] V. Srinivasan and Y. H. Kim, "Deterministic Cash Flow Management: State of the Art and Research Direction," Omega, Vol. 14 (2) (1986), pp. 145-166.

[15] B. K. Stone, "The Use of Forecasts and Smoothing in Control Limit Models for Cash Management," Financial Management, Vol. 1 (Spring 1972), pp. 72-84.

[16] B. K. Stone and T. W. Miller, "Daily Cash Forecasting with Multiplicative Models of Cash Flow Patterns," Financial Management (Winter 1987), pp. 45-54.

[17] ___________, "Daily Cash Forecasting," Chapter 7, pp. 120-41 in Corporate Cash Management: Techniques and Analysis, F. J. Fabozzi and L. N. Mansonson, editors, Homewood, IL: Dow Jones-Irwin, 1985.

[18] ___________, "Forecasting Disbursement Funding Requirements: The Clearing Pattern Approach," Journal of Cash Management (October/November 1983), pp. 67-78.

[19] ___________, "Daily Cash Forecasting: A Structural Framework," Journal of Cash Management (October 1981), pp. 35-50.

[20] B. K. Stone and R. A. Wood, "Daily Cash Forecasting: A Simple Method for Implementing the Distribution Approach," Financial Management (Fall 1977), pp. 40-50.

EXHIBIT 1
CASH FORECASTING PROBLEM SPACE
Concepts and Models

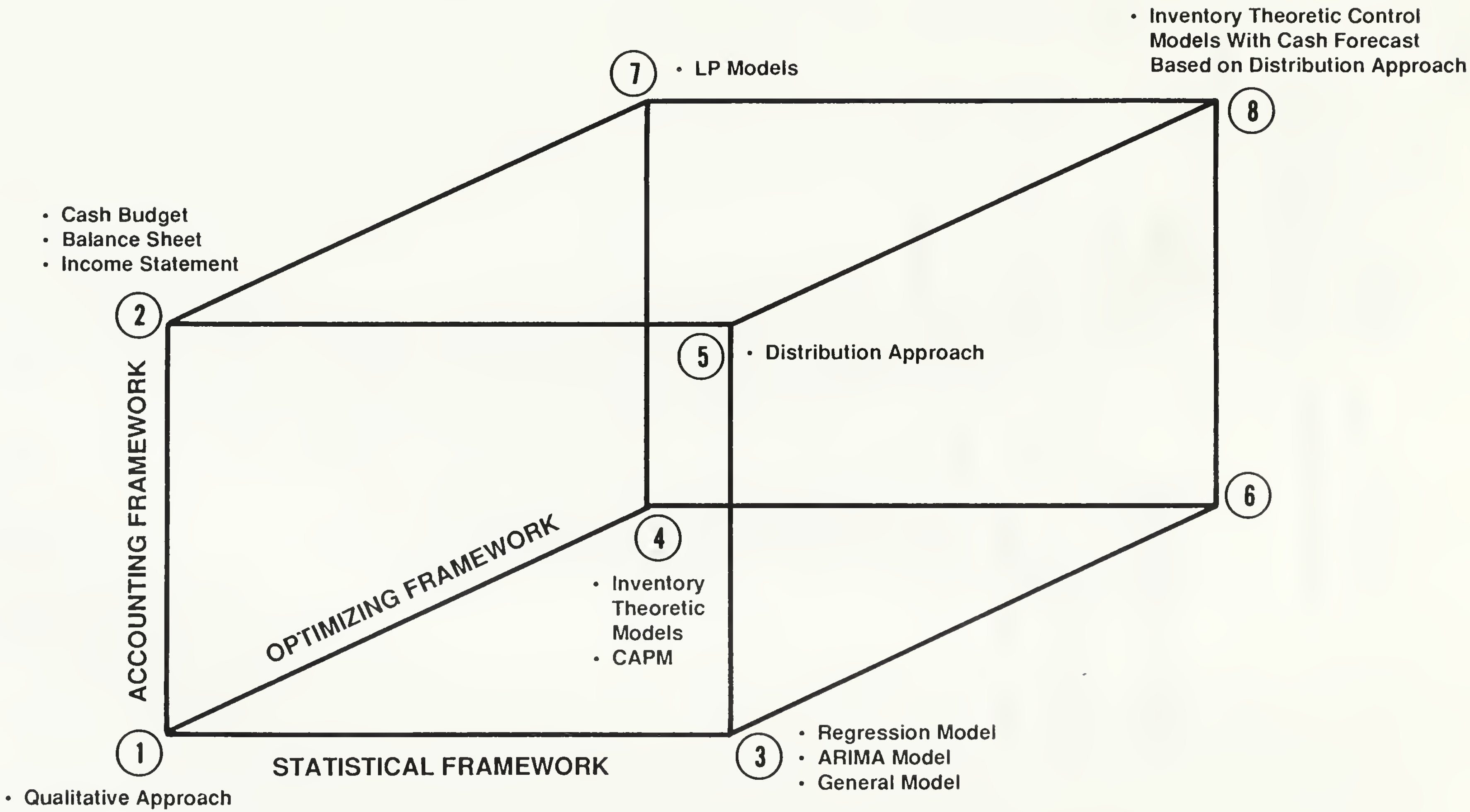


Exhibit 2

Statistical Cash Forecasting

	Equation 1	Equation 2
Regresion Modelling (Stone and Wood [20])	yes	no
ARIMA Modelling (Hodgson [5])	no	yes
General Modelling	yes	yes

Exhibit 3

ANOVA Test for DOM Effect
in Public Pension System Data

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: RSDULS RESIDUALS

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE
MODEL	30	23657166587874.40000	788572219595.81500
ERROR	1230	801364209257900.00000	651515617282.84500
CORRECTED TOTAL	1260	825021375845774.00000	

MODEL F =	1.21	PR > F = 0.2019

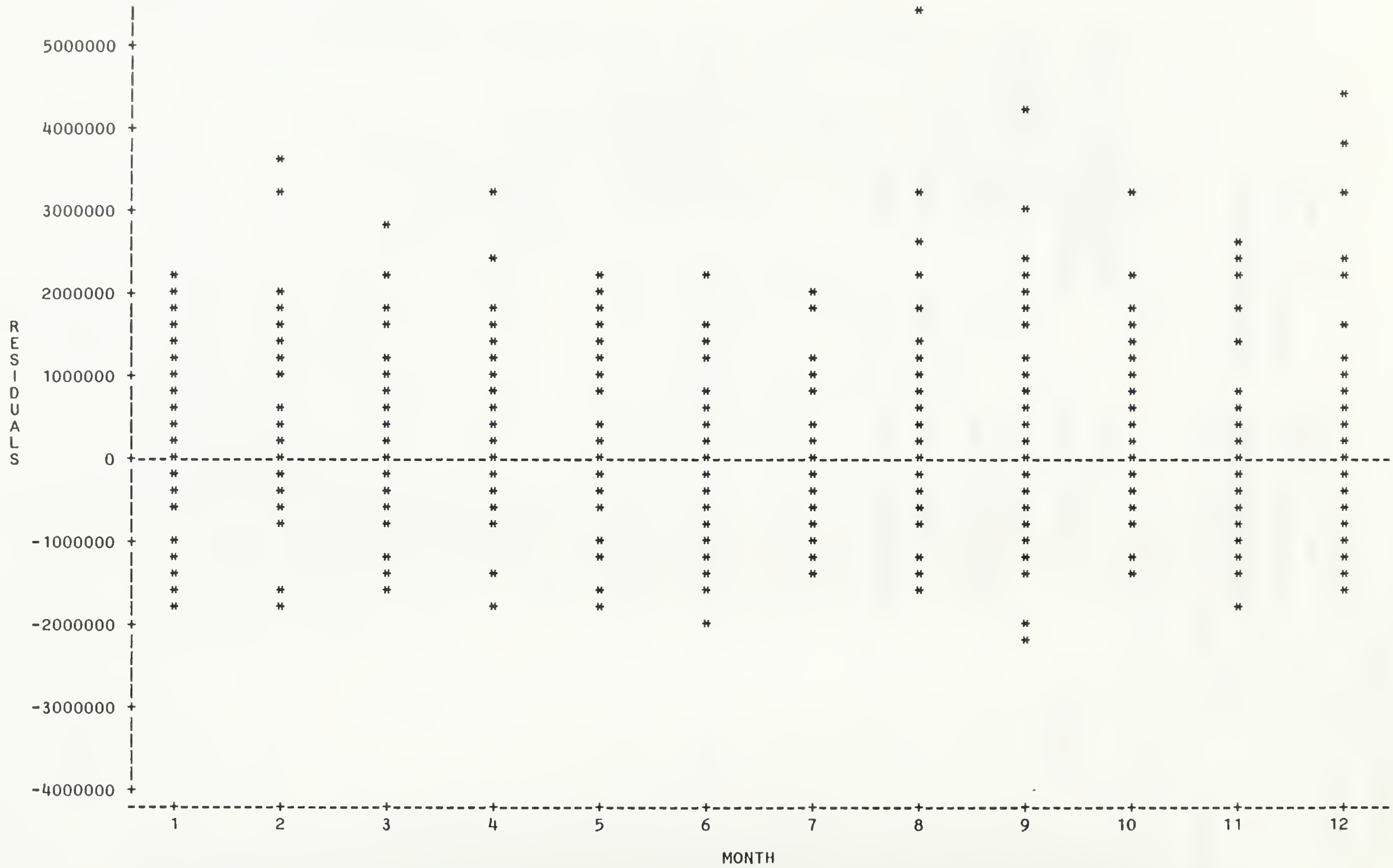
R-SQUARE	C.V.	ROOT MSE	RSDULS MEAN
0.028675	50680.6317	807165.17348	1592.65018140

SOURCE	DF	TYPE I SS	F VALUE	PR > F
DOM	30	23657166587874.40000	1.21	0.2019

SOURCE	DF	TYPE III SS	F VALUE	PR > F
DOM	30	23657166587874.40000	1.21	0.2019

Graphical Test for Month Effect in
Public Pension System Data

PLOT OF RSDULS*MONTH   SYMBOL USED IS *



NOTE: 1 OBS HAD MISSING VALUES   1036 OBS HIDDEN

Maximum Likelihood Estimates for Regression Model

DEP VARIABLE: CR
ANALYSIS OF VARIANCE

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	38	5.15788E+14	1.35734E+13	20.023	0.0001
ERROR	1223	8.29058E+14	677888774613		
C TOTAL	1261	1.34485E+15			

ROOT MSE	823340	R-SQUARE	0.3835
DEP MEAN	733588.4	ADJ R-SQ	0.3644
C.V.	112.2346		

PARAMETER ESTIMATES

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > |T|
INTERCEP	1	2008331.88	100646.47	19.954	0.0001
D1	1	-712726.86	377583.21	-1.888	0.0593
D2	1	-87988.08709	420222.86	-0.209	0.8342
D3	1	27140.53597	588964.18	0.046	0.9633
D4	1	-525904.59	483105.92	-1.089	0.2765
D5	1	-1873320.67	381331.76	-4.913	0.0001
D6	1	819300.32	381442.96	2.148	0.0319
D7	1	-1824713.23	380592.33	-4.794	0.0001
D8	1	1025353.16	380545.16	2.694	0.0071
D9	1	-374062.33	421742.79	-0.887	0.3753
D10	1	-253135.86	421729.23	-0.600	0.5485
D11	1	-1803045.54	380739.18	-4.736	0.0001
D12	1	2419607.45	380925.09	6.352	0.0001
D13	1	-1749907.77	380439.75	-4.600	0.0001
D14	1	1296675.82	380468.20	3.408	0.0007
D15	1	-966590.88	378962.94	-2.551	0.0109
D16	1	582462.15	421610.76	1.382	0.1674
D17	1	-28645.68653	381778.74	-0.075	0.9402
D18	1	-374480.23	381785.02	-0.981	0.3269
D19	1	-764683.89	377127.50	-2.028	0.0428
D20	1	-19151.86838	419864.87	-0.046	0.9636
D21	1	-1396240.99	76415.91716	-18.272	0.0001
D22	1	-1273147.12	74957.97875	-16.985	0.0001
D23	1	-1266789.66	75324.83677	-16.818	0.0001
D24	1	-1279257.72	75461.96494	-16.952	0.0001
D25	1	2365078.38	586093.10	4.035	0.0001
D26	1	6744606.12	829468.80	8.131	0.0001
D27	1	7093316.12	828966.71	8.557	0.0001
D28	1	-205286.33	121882.30	-1.684	0.0924
D29	1	-258424.11	120185.30	-2.150	0.0317
D30	1	-166221.33	124896.54	-1.331	0.1835
D31	1	-200560.89	120181.61	-1.669	0.0954
D32	1	-252517.92	121576.82	-2.077	0.0380
D33	1	-135011.21	123150.48	-1.096	0.2732
D34	1	-220517.67	118383.20	-1.863	0.0627
D35	1	-250194.25	118626.07	-2.109	0.0351
D36	1	-183618.65	120462.08	-1.524	0.1277
D37	1	-408442.17	119247.26	-3.425	0.0006
D38	1	-477152.67	126564.74	-3.770	0.0002

Exhibit 6

Autocorrelations, Inverse Autocorrelations,
Partial Autocorrelations and White Noise Test for Residuals

```
ARIMA PROCEDURE

NAME OF VARIABLE       =    RSDULS
MEAN OF WORKING SERIES= 3.491E-09
STANDARD DEVIATION     =    810518
NUMBER OF OBSERVATIONS=       1262
AUTOCORRELATIONS

LAG COVARIANCE CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  0  6.569E+11     1.00000 |                    |********************|
  1 -3.156E+10    -0.04804 |                   *|.                   |
  2 -3.511E+10    -0.05345 |                   *|.                   |
  3 -1.715E+10    -0.02610 |                   *|.                   |
  4 -2.866E+10    -0.04363 |                   *|.                   |
  5 -3.545E+10    -0.05395 |                   *|.                   |
  6  1.824E+10     0.02777 |                   .|*                   |
  7 -1.058E+10    -0.01610 |                   .|.                   |
  8 -2.248E+10    -0.03422 |                   *|.                   |
  9 -3.967E+10    -0.06039 |                   *|.                   |
 10  1.256E+10     0.01912 |                   .|.                   |
 11  2.171E+10     0.03304 |                   .|*                   |
 12 -2.609E+10    -0.03971 |                   *|.                   |
 13 7439278010     0.01132 |                   .|.                   |
 14 6838484769     0.01041 |                   .|.                   |
 15  1.957E+10     0.02979 |                   .|*                   |
 16 -3.902E+09    -0.00594 |                   .|.                   |
 17  1.071E+10     0.01630 |                   .|.                   |
 18 5506433453     0.00838 |                   .|.                   |
 19  2.441E+10     0.03715 |                   .|*                   |
 20  5.931E+10     0.09028 |                   .|**                  |
 21  2.344E+10     0.03568 |                   .|*                   |
 22  1.262E+10     0.01922 |                   .|.                   |
 23 -1.233E+10    -0.01877 |                   .|.                   |
 24 -6.768E+09    -0.01030 |                   .|.                   |
                           '.' MARKS TWO STANDARD ERRORS
```

Autocorrelations, Inverse Autocorrelations,
Partial Autocorrelations and White Noise Test for Residuals

INVERSE AUTOCORRELATIONS

```
 LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
   1     0.10435  |                    .|**                 |
   2     0.09759  |                    .|**                 |
   3     0.07031  |                    .|*                  |
   4     0.08834  |                    .|**                 |
   5     0.08800  |                    .|**                 |
   6    -0.00265  |                    .|.                  |
   7     0.03955  |                    .|*                  |
   8     0.04106  |                    .|*                  |
   9     0.06727  |                    .|*                  |
  10    -0.01730  |                    .|.                  |
  11    -0.04133  |                    *|.                  |
  12     0.01806  |                    .|.                  |
  13    -0.02637  |                    *|.                  |
  14    -0.02481  |                    .|.                  |
  15    -0.06305  |                    *|.                  |
  16    -0.02466  |                    .|.                  |
  17    -0.04227  |                    *|.                  |
  18    -0.04489  |                    *|.                  |
  19    -0.07221  |                    *|.                  |
  20    -0.11417  |                   **|.                  |
  21    -0.05587  |                    *|.                  |
  22    -0.04910  |                    *|.                  |
  23    -0.01186  |                    .|.                  |
  24    -0.01502  |                    .|.                  |
```

PARTIAL AUTOCORRELATIONS

```
 LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
   1    -0.04804  |                    *|.                  |
   2    -0.05589  |                    *|.                  |
   3    -0.03167  |                    *|.                  |
   4    -0.04992  |                    *|.                  |
   5    -0.06268  |                    *|.                  |
   6     0.01530  |                    .|.                  |
   7    -0.02375  |                    .|.                  |
   8    -0.04038  |                    *|.                  |
   9    -0.07214  |                    *|.                  |
  10     0.00443  |                    .|.                  |
  11     0.02472  |                    .|.                  |
  12    -0.04710  |                    *|.                  |
  13     0.00020  |                    .|.                  |
  14     0.00244  |                    .|.                  |
  15     0.03444  |                    .|*                  |
  16    -0.00709  |                    .|.                  |
  17     0.00997  |                    .|.                  |
  18     0.01320  |                    .|.                  |
  19     0.04684  |                    .|*                  |
  20     0.10391  |                    .|**                 |
  21     0.04919  |                    .|*                  |
  22     0.04816  |                    .|*                  |
  23     0.01029  |                    .|.                  |
  24     0.01592  |                    .|.                  |
```

AUTOCORRELATION CHECK FOR WHITE NOISE

TO LAG	CHI SQUARE	DF	PROB	AUTOCORRELATIONS					
6	14.49	6	0.025	-0.048	-0.053	-0.026	-0.044	-0.054	0.028
12	24.82	12	0.016	-0.016	-0.034	-0.060	0.019	0.033	-0.040
18	26.73	18	0.084	0.011	0.010	0.030	-0.006	0.016	0.008
24	41.67	24	0.014	0.037	0.090	0.036	0.019	-0.019	-0.010

Exhibit 7

Selection of Time Series Model

	SBC	AIC
White Noise	37928.6	37923.5
MA only	37941.1	37915.4
AR only	37925.3	37915.1
ARMA	37934.4	37903.5

Exhibit 8

Maximum Likelihood Estimates
with Time Series Structure in Residuals

ARIMA: MAXIMUM LIKELIHOOD ESTIMATION

PARAMETER	ESTIMATE	APPROX. STD ERROR	T RATIO	LAG	VARIABLE	SHFT
MU	2013683	84857.2	23.73	0	CR	0
MA1,1	0.0758169	0.028757	2.64	1	CR	0
MA1,2	0.0755152	0.0289272	2.61	2	CR	0
MA1,3	0.064055	0.0287658	2.23	4	CR	0
MA1,4	0.0735852	0.0290856	2.53	5	CR	0
AR1,1	0.119027	0.0293337	4.06	20	CR	0
NUM1	-697867	368642	-1.89	0	D1	0
NUM2	-218165	410847	-0.53	0	D2	0
NUM3	-17110.2	578447	-0.03	0	D3	0
NUM4	-678439	477943	-1.42	0	D4	0
NUM5	-1660386	382098	-4.35	0	D5	0
NUM6	680020	381179	1.78	0	D6	0
NUM7	-1651782	374995	-4.40	0	D7	0
NUM8	1100737	371573	2.96	0	D8	0
NUM9	-255039	414222	-0.62	0	D9	0
NUM10	-179668	415570	-0.43	0	D10	0
NUM11	-1997114	373372	-5.35	0	D11	0
NUM12	2303113	373692	6.16	0	D12	0
NUM13	-1478717	378477	-3.91	0	D13	0
NUM14	1244144	377852	3.29	0	D14	0
NUM15	-903394	380383	-2.37	0	D15	0
NUM16	634377	418891	1.51	0	D16	0
NUM17	-87672.9	376687	-0.23	0	D17	0
NUM18	-487019	374597	-1.30	0	D18	0
NUM19	-810571	374211	-2.17	0	D19	0
NUM20	-221148	414357	-0.53	0	D20	0
NUM21	-1375277	84455.2	-16.28	0	D21	0
NUM22	-1253056	80598.7	-15.55	0	D22	0
NUM23	-1248376	80876.8	-15.44	0	D23	0
NUM24	-1257556	83438.4	-15.07	0	D24	0
NUM25	2214858	576075	3.84	0	D25	0
NUM26	6981502	812053	8.60	0	D26	0
NUM27	7164165	808815	8.86	0	D27	0
NUM28	-220634	88775.1	-2.49	0	D28	0
NUM29	-274338	91714.1	-2.99	0	D29	0
NUM30	-210893	96858.7	-2.18	0	D30	0
NUM31	-195422	91267.2	-2.14	0	D31	0
NUM32	-262991	92120.6	-2.85	0	D32	0
NUM33	-164407	94505.3	-1.74	0	D33	0
NUM34	-250435	88850	-2.82	0	D34	0
NUM35	-273587	88932.4	-3.08	0	D35	0
NUM36	-213193	91184	-2.34	0	D36	0
NUM37	-428214	89185.5	-4.80	0	D37	0
NUM38	-524329	93665.9	-5.60	0	D38	0

CONSTANT ESTIMATE = 1774001

VARIANCE ESTIMATE = 6.629E+11
STD ERROR ESTIMATE = 814202
AIC = 37976.5
SBC = 38202.7
NUMBER OF RESIDUALS= 1262

Exhibit 9

Maximum Likelihood Estimates with Time Series Structure in Residuals and Insignificant Variables Are Eliminated

ARIMA: MAXIMUM LIKELIHOOD ESTIMATION

PARAMETER	ESTIMATE	APPROX. STD ERROR	T RATIO	LAG	VARIABLE	SHIFT
MU	2016350	86375.[illegible]	23.34	0	CR	0
MA1,1	0.0750317	0.0286319	2.62	1	CR	0
MA1,2	0.0753007	0.0287976	2.61	2	CR	0
MA1,3	0.0644122	0.0286532	2.25	4	CR	0
MA1,4	0.0728862	0.0290122	2.51	5	CR	0
AR1,1	0.116567	0.0291934	3.99	26	CR	0
NUM1	-715552	366945	-1.95	0	D1	0
NUM2	-686302	476456	-1.44	0	D4	0
NUM3	-1666133	381345	-4.37	0	D5	0
NUM4	680705	380293	1.79	0	D6	0
NUM5	-1655706	374380	-4.42	0	D7	0
NUM6	1100195	370988	2.97	0	D8	0
NUM7	-1993350	372774	-5.35	0	D11	0
NUM8	2301243	373069	6.17	0	D12	0
NUM9	-1494023	377301	-3.93	0	D13	0
NUM10	1242006	377192	3.29	0	D14	0
NUM11	-1277689	378435	-3.37	0	D15	0
NUM12	637565	417170	1.53	0	D16	0
NUM13	-473029	373765	-1.27	0	D18	0
NUM14	-811941	373359	-2.17	0	D19	0
NUM15	-1376033	85933.5	-16.0[illegible]	0	D21	0
NUM16	-1254267	80114.4	-15.65	0	D22	0
NUM17	-1252136	80072.5	-15.64	0	D23	0
NUM18	-1259790	82959.5	-15.19	0	D24	0
NUM19	[illegible]	[illegible]74.53	3.30	0	D25	0
NUM20	696692	210573	3.57	0	D26	0
NUM21	716874	80747.[illegible]	[illegible]	0	D27	0
NUM22	-221573	88724.1	-2.50	0	D28	0
NUM23	-274[illegible]	91356.[illegible]	-3.00	0	D29	0
NUM24	-213843	94683.9	-2.26	0	D30	0
NUM25	-205565	89983.7	-2.28	0	D31	0
NUM26	-271217	90323.5	-3.00	0	D32	0
NUM27	-165233	94441.1	-1.75	0	D33	0
NUM28	-251523	83895.2	-2.34	0	D34	0
NUM29	-274959	88779.5	-3.09	0	D35	0
NUM30	-214114	91036.1	-2.35	0	D36	0
NUM31	-430667	88999.0	-4.8[illegible]	0	D37	0
NUM32	-545993	89559	-6.05	0	D38	0

CONSTANT ESTIMATE = 1751270

VARIANCE ESTIMATE = 6.60E+11
STD ERROR ESTIMATE = 812619
AIC = 37955.3
SBC = 38151.1
NUMBER OF RESIDUALS= 1242

Exhibit 10

Forecast Performance

	RMSE
Regression Only	823340
General Model	812619