FINANCIAL FORECASTING THROUGH SAS MACRO PROCESSES

Heydar Pourian, Western Carolina University

Introduction

I. Forecasting Needs of a Financial Manager

II. Modeling Cyclical External Factors

III. SAS MACRO Programming

IV. The Data

V. Application and Error Evaluation

Notes

References

One of the most serious problems faced by forecasters in recent years is that external factors have been large, unprecedented errors associated with forecasting financial variables [see, e.g., Penner, 1984]. A possible explanation for unsatisfactory results may be due to forecasters avoiding methods which rely on complex computation procedures which, historically speaking, have been costly to use.

This study contains a forecasting method based on a computationally-lengthy, yet analytically-simple, procedure. By use of SAS Macros, the difficulty of the computations is greatly reduced. The included method is referred to as the Cyclical Indexing Analogy Method (or ClAM) and is based on the assumption that the financial-forecast variables exhibit a large degree of continuous cyclical variation. The objective of this study is to provide a workable guide to the formulation of a SAS Macro-based forecasting technique. In addition, it will demonstrate that SAS Macros offer potentially broad processes in the sense that a variety of alternative judgemental and systematic rules can be introduced within its framework to fit and forecast any variable with cyclical characteristics.

The plan of the study is to discuss briefly in the first section the superiority of the ClAM procedure as compared to many other methods used to forecast financial variables. Such superiority is due to ClAM's inherent simplicity and theoretical logic. The following section describes a step-by-step ClAM modeling formulation. SAS Macro programming of this method is presented in Section III. An application is then provided by generating "half-cycle ahead" yield forecasts of U.S. prime commercial paper. For comparative purposes, Box-Jenkins forecasts are presented using the SAS ARIMA procedure. An evaluation of forecast errors reveals the superiority of ClAM, SAS Macro-based forecasts.

I. FORECASTING NEEDS OF A FINANCIAL MANAGER

Forecasting needs of a financial manager, a security analyst, or a portfolio manager are spread over a wide spectrum. They range from forecasting sales per se to a requirement for effective financial planning to forecasting interest rates for hedging purposes and portfolio diversification strategies [see, e.g., Brealey and Myers, 1984 pp. 596-597 and (Chicago Mercantile Exchange, 1985, p. 81). Normally, forecasting needs can be approached by either "external" or "internal" forecasting methods [see, e.g., Gitman, 1985, p. 158-159]. Each general approach to forecasting, in turn, generates errors particular to that technique. But errors in internal forecasts are not normally as serious as errors associated with external forecasts. This is because many factors in the internal operation of a business firm are endogenous and under the control of the enterprise itself. Factors influencing the external forecasts, however, are exogenous and by definition beyond the control of a firm. Many external factors in a business environment, on the other hand, have a cyclical nature. They include sales, interest rates, cash receipts, among others. External forecast errors observed in recent years suggest the exploration of alternative forecasting schemes. The method proposed herein is an external method of forecasting which can satisfy the cyclical forecasting needs of a business enterprise in regard to financial planning, capital budgeting and financial hedging.

Cyclical Indexing Analogy Method is an index-based technique which partitions time into variable-length cycles and aggregates the variation of the forecast variable into a "forecast estimator array." The first question is, however, why should one use an index-based forecasting procedure?

It is established in economics and finance that cyclical patterns do occur. In fact, the analyses of these patterns have been central to economic and business studies: from growth theory and macroeconomics to inventory control. In application, however, various "smoothing" methods (which are also available in the SAS System) are utilized to study, fit, simulate, and forecast these patterns. The main problem with such methods is that the resultant smoothed projection into the future does not contain any of the inherent cyclical properties! In other words, the smoothed projection of any given variable removes and decomposes its cyclical properties!

It is therefore suggested that averaging and smoothing procedures should be combined with "cyclical indexing" to preserve the cyclical properties of the forecast series. The proposed method will preserve the theoretical properties of the time series by permitting the forecast series to fluctuate also. On the other hand, its utilization through SAS Macro greatly reduces the complexity and the costs of the computations. Moreover, the underlying advantage lies in the fact that ClAM will let "a data set speak for itself." In this sense, it is "parsimonious" and thus conforms to a foundation of modern forecasting and econometrics.
Technically speaking, CIAM makes an attempt to forecast a financial variable on an "x-cycle ahead" basis, rather than on the basis of a point in time. As in the Box-Jenkins method, CIAM entertains lead time in forecasting. More specifically, this method first produces a forecast estimator array. Then by use of a given actual value at the beginning of the array, CIAM generates an x-cycle ahead forecast of the variable in question, where "x" represents a selected time duration expressed as a fraction of a full cycle. The next section contains the construction of such an array.

II. MODELING CYCLICAL EXTERNAL FACTORS

The construction of a SAS Macro-based Cyclical Indexing Analogy "forecast estimator array" which would focus on factors external to a business enterprise can be accomplished through the following steps.

Step 1: Cyclic Specification

Consider a financial time series variable \( V_{ij} \) during the recurring business cycles \( i = 1 ... n \), where \( n \) is the number of cycles in the included history, and \( j \) specifies the timing of this variable within the \( i \)th cycle. Let \( \delta \) be the number of cycles in a full cycle. Let variation of \( j \) commence from an initial trough \( (q_0) \) of a cycle, continue to the peak \( (p) \) of the same cycle, and end at the terminal trough \( (q) \). Within each cycle, therefore, \( j \) takes the following sequence, forming an array:

\[ q_0, q_0+1, ... , p, p+1, ... , q \]

It will be useful to rename the time series variable according to the two major phases (or halves) of the business cycle. Thus, let:

\[ E_{ij} = V_{ij} \text{ for } i = 1 ... n \text{ while } q_0 \leq j \leq p-1 \]

\[ C_{ij} = V_{ij} \text{ when } i = 1 ... n \text{ and } p \leq j \leq q-1 \]

where \( E \) and \( C \) reflect, respectively, the time series variable during the expansion and contraction phases of each business cycle.

Step 2: Indexing

In an index-based array forecasting method, the movement, rather than the level, is the point of focus. Accordingly, let indexing take the following format. In expansions,

\[ E_{ij} = 1 \text{ for } i = 1 ... n \text{ while } j = q_0 \]

\[ = E_{1j}/E_{q_0} \text{ when } i = 1 ... n \text{ and } q_0+1 \leq j \leq p-1 \]

and in contractions,

\[ C_{ij} = 1 \text{ for } i = 1 ... n \text{ while } j = p \]

\[ = C_{1j}/C_{p} \text{ when } i = 1 ... n \text{ and } p+1 \leq j \leq q-1 \]

where \( E_{1j} \) and \( C_{1j} \) are the Indexes of the financial variable, respectively, in expansions and contractions.

Step 3: Smoothing

Based on this \( n \)-cycle history, now we develop dynamic half-cycle indexes (DIE's) by averaging or smoothing. One could use SAS ARIMA, AUTOREG, or some other procedure. Another method, which we shall utilize here, may rely on simple averaging in the following sense:

\[ \text{DIE}_j = \frac{1}{n} \sum_{i} E_{ij} \text{ for } q_0 \leq j \leq p-1 \]

\[ \text{DIC}_j = \frac{1}{n} \sum_{i} C_{ij} \text{ for } p \leq j \leq q-1 \]

Notice that \( \text{DIE}_j = 1 \) for \( j = q_0 \), and \( \text{DIC}_j = 1 \) when \( j = p \), always.

Step 4: Averaging and Aggregating

Since the length of \( \delta \) (that is, the number of periods within each cycle) is variable, "average durations" for expansions and contractions should be computed. As one alternative, simple averaging can be used. Thus,

\[ p = \frac{1}{n} \sum_{i} P_i \]

\[ q = \frac{1}{n} \sum_{i} Q_i \]

Given these average durations of expansions and of contractions, the above indexing and averaging produce two dynamic arrays, which then, in order to produce a full-cycle array, are merged in the following manner.

\[ I_j = \text{DIE}_j \times \text{DIC}_j \text{ for } j = q_0 \ldots \bar{p} \ldots q \]

where now

\[ \text{DIE}_j = \text{DIC}_j \text{ for } j = \bar{p} \text{ only} \]

\[ = \text{DIC}_j / \text{DIE}_p \text{ while } \bar{p}+1 \leq j \leq q-1 \]

Step 5: Forecasting

Based on the constructed dynamic index, a forecast array \( F_j \) can now be generated at any given period \( t \) within any business cycle simply by using the most recent actual value \( A_j \) and by letting

\[ F_j = A_0 \times I_j \]

where now

\[ I_j = 1 \text{ for } j = t \]

\[ = I_j / I_t \text{ while } t+1 \leq j \leq q-1 \]

If, for instance, \( t = q_0 \) is chosen and the entire full-cycle array \( I_0 \) is used, then the forecast can be labeled as "one-cycle ahead" forecast. On the other hand, the forecast can be referred to as "half-cycle," if only the array \( \text{DIE}_j \) is used.
We are now ready to program the CrAM formulation according to the model of the last section. The following SAS program with several MACROS would accomplish the task of subsetting and processing the indexed data. Let us begin by describing the data set.

Our multi-variable data set, which is stored on a magnetic tape, has the following characteristics. It contains 20 financial variables, the first of which will be referred to as V16 and the last as V108. Each financial variable has three types of observation: annual (with time code = 1), quarterly (with time code = 2) and monthly (with time code = 3). Suppose we are interested in generating monthly forecasting arrays. Moreover, since the annual code is not ready-to-use for reporting purposes, we must first reformat it. Immediately below, the monthly data is subset and the variableR are defined. In addition, a new variable named "TIME" is created by concatenation. This variable will be used subsequently in Macro specifications.

```
DATA MONTHLY;
  LENGTH T 7 TIME 5;
  INFILE MAGTAPE;
  INPUT @9 YEAR $ 3. @ 12 MONTH $ 2. @14 TIMECODE 1. @105 (VI6-VI8) (9.3) .. ;
/* ...more variables are specified and/or defined ... */
  VI08M24 = V105 - V24;
  IF TIMECODE = 3; IF YEAR = 9999 THEN YEAR = 89906; TIME = INPUT(YMD, YYMMDD8.); FORMAT T MONYY7.;
  LENGTH TIME $ 5; TIME = YEARMONTH;
```

A SAS Array is now defined in order to specify the missing data. Subsequently, the unnecessary variables are dropped for the rest of the process.

```
ARRAY A V16--VI08M24;
  DO OVER A;
  IF A = -9999 THEN A = .;
  END;
DROP VI6--VI08M24;
```

The following SAS Calls would now subset and process the data for the U.S. business cycles. The first %CYC Call, for instance, specifies that the 21st expansion (E21) in the U.S. history starts on June, 1897 (89706), and ends on June, 1899 (89906). The computations continue until SAS Macros process the most recent officially dated cycle.

```
%MACRO CYC(CN,F,L);
  DATA &CN;
  SET M;
  IF &F=<TIME=<&L;
  DATA &ACN;
  SET &ACN; IF _N_=1;
  DATA &CN;
  IF _N_=1 THEN SET &B&C; SET &ACN; N = _N_;
  ARRAY A V16--V10824;
  ARRAY B B1-B20;
  DO OVER A;
  B = 100/A;
  END;
  DATA &CN;
  IF N =1 THEN SET &B&C; SET A&CN; N = _N_; ARRAY A V16--V10824;
  ARRAY B B1-B20;
  ARRAY C C16&CN .. ;
  /* ... Note the suffix; and more variables such as C17&CN are defined ... */
  DO OVER C;
  C = A*B;
  END;
  DROP VI6--VI0824 B1-B20;
%MEND eyc;
```

The resultant data sets with prefix 'E,' which are created in MACRO CYC by DATA &CN and contain the index values for the expansion phase of each cycle must now be merged in DATA EAGG. Subsequently, the arithmetic mean of each index variable is computed by use of the MEAN function. It is easier to use another Macro at this stage (MACRO E). The Macro parameters now have the following definitions: VN = Variable Number, CNS = Cycle number at beginning date, and CNE = Cycle Number at ending date of the process.

```
%MACRO E(VN,CNS,CNE);
  DATA EAGG;
  MERGE E21-ES3;
  KEEP N G&VN;
  /* ... more calls if desired ... */
```

Thus, for expansions,

```
%MACRO E(VN,CNS,CNE);
  DATA EAGG;
  MERGE E21-E53;
  BY N;
  GAVN = MEAN(OP GAVNCRS-GAVNCRF);
  KEEP N GAVN;
  ZMEND E;
```
The same procedure can be used for contractions. Now, the following Calls would produce the desired results. Processing all the raw yields, transforming them to indexes such that the beginning of each cycle is set to 100 account as to for their movement within the cycle, and computing their aggregate Index-mean for all the included history. For instance, if variable V16 is the selected forecast variable and we desired to include expansions number 23 to 53 of the U.S. business cycles, then SAS Call would take the following format:

```
%E(16,E21,E53)
```

IV. THE DATA

The yield data analyzed in this paper are a part of a large data set compiled by the National Bureau of Economic Research (NBER) for the United States, Britain, France and Germany. The original sources primarily include F.R. Macaulay, Standard and Poor's, and The Economist. In the series used in this study, the Commercial-Paper yield is the rate observed on prime commercial paper in New York City from June, 1897 to December, 1969. This series is not seasonally adjusted. The data set is then updated by use of the Federal Reserve Bulletin, the Federal Reserve's H.15 Release, and the U.S. Financial Data, which is published by the Federal Reserve Bank of St. Louis. The most recent data included in the analysis of this paper is data corresponding to November, 1982. The cyclical behaviors of these rates were analyzed during a total of 20 U.S. business cycles over approximately 85 years. The business cycle dates used are the same as "officially" designated by the NBER and used by the Federal Government. The original sources primarily include F.R. Macaulay, Standard and Poor's, and The Economist. In the series used in this study, the Commercial-Paper yield is the rate observed on prime commercial paper in New York City from June, 1897 to December, 1969. This series is not seasonally adjusted. The data set is then updated by use of the Federal Reserve Bulletin, the Federal Reserve's H.15 Release, and the U.S. Financial Data, which is published by the Federal Reserve Bank of St. Louis. The most recent data included in the analysis of this paper is data corresponding to November, 1982. The cyclical behaviors of these rates were analyzed during a total of 20 U.S. business cycles over approximately 85 years. The business cycle dates used are the same as "officially" designated by the NBER and used by the Federal Government.

V. APPLICATION AND ERROR EVALUATION

As far as financial planning and hedging strategies are concerned, the yield on the U.S. prime commercial paper provides an excellent application of the proposed method due to the long-term availability of its historical data. The following yield forecasts use the actual yield on December 1982 (198212) as the given rate. ClAM forecasts are then generated by multiplying this given rate to the financial-forecast estimator array generated through running the SAS program of the previous section. The ClAM methodology presented in this study is largely based on a paper by the author presented at the Fifth International Symposium on Forecasting held in Montreal, Canada, [Pouliot, 1985]. The substance of the present method and the included SAS procedures have been previously communicated on a consultation basis to the Socioeconomic Division of the Department of Environment, Canadian Government, Ottawa. Part of the present study was undertaken when the author served as Assistant Professor of Economics at the University of Missouri, St. Louis. The reader should be reminded at this stage that on an ex ante basis, the "better identified" ARIMA model will not necessarily perform better. In addition, the ex ante ARIMA's have either very low or negative correlation with the actual series! In investment and hedging strategies such a negative correlation would be of very limited use. Indeed, it could lead to disastrous results in financial planning. Thus, ex ante ARIMA forecasts applied in this paper have very little cyclical variation—a characteristic which ClAM clearly takes into consideration. Accordingly, since the minimum forecast error is achieved and cyclical property is preserved, in addition to its theoretical soundness, ClAM is shown to be superior in application. Once ClAM's analytical simplicity is mixed with the utilization of SAS Macro processes in order to ease its complex and lengthy computations, ClAM can provide a useful and a more precise forecasting tool for security analysts as well as financial and portfolio managers.

NOTES

1. The ClAM methodology presented in this study is largely based on a paper by the author presented at the Fifth International Symposium on Forecasting held in Montreal, Canada, [Pouliot, 1985]. The substance of the present method and the included SAS procedures have been previously communicated on a consultation basis to the Socioeconomic Division of the Department of Environment, Canadian Government, Ottawa. Part of the present study was undertaken when the author served as Assistant Professor of Economics at the University of Missouri, St. Louis. I acknowledge valuable comments by Professors Carroll D. AbY, Jr., Patrick A. Hays and Austin Spencer on this paper. Responsibility of remaining errors is mine.

2. Naturally, the "estimation memory" itself is subject to selection by the forecasters. For instance, ClAM could be based on all twentieth-century data, or on post-war data, and so on.
### Table 1

**MONTHLY YIELDS ON U.S. PRIME COMMERCIAL PAPER FORECASTS**

<table>
<thead>
<tr>
<th>OBS</th>
<th>DATE</th>
<th>ACTUAL RATE</th>
<th>BEST</th>
<th>BEST</th>
<th>BEST</th>
<th>BEST</th>
<th>BASED ON</th>
<th>BASED ON</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIAM</td>
<td>ARIMA</td>
<td>ARIMA</td>
<td>CIAM</td>
<td>EX ANTE</td>
<td>EX POST</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>198212</td>
<td>8.52</td>
<td>8.59</td>
<td>8.71</td>
<td>8.82</td>
<td>8.69</td>
<td>8.92</td>
<td>8.81</td>
</tr>
<tr>
<td>2</td>
<td>198301</td>
<td>8.18</td>
<td>8.44</td>
<td>8.69</td>
<td>8.97</td>
<td>8.45</td>
<td>9.23</td>
<td>8.94</td>
</tr>
<tr>
<td>3</td>
<td>198302</td>
<td>8.38</td>
<td>8.30</td>
<td>8.67</td>
<td>9.11</td>
<td>8.34</td>
<td>9.48</td>
<td>9.07</td>
</tr>
<tr>
<td>4</td>
<td>198303</td>
<td>8.51</td>
<td>8.62</td>
<td>8.65</td>
<td>9.24</td>
<td>9.06</td>
<td>9.65</td>
<td>9.18</td>
</tr>
<tr>
<td>5</td>
<td>198304</td>
<td>8.52</td>
<td>8.87</td>
<td>8.64</td>
<td>9.36</td>
<td>9.99</td>
<td>9.75</td>
<td>9.29</td>
</tr>
<tr>
<td>6</td>
<td>198305</td>
<td>8.32</td>
<td>8.94</td>
<td>8.62</td>
<td>9.46</td>
<td>10.15</td>
<td>9.87</td>
<td>9.38</td>
</tr>
<tr>
<td>7</td>
<td>198306</td>
<td>9.02</td>
<td>8.91</td>
<td>8.60</td>
<td>9.56</td>
<td>10.00</td>
<td>9.85</td>
<td>9.47</td>
</tr>
<tr>
<td>8</td>
<td>198307</td>
<td>9.32</td>
<td>9.06</td>
<td>8.58</td>
<td>9.65</td>
<td>10.00</td>
<td>9.87</td>
<td>9.55</td>
</tr>
<tr>
<td>10</td>
<td>198309</td>
<td>9.27</td>
<td>9.73</td>
<td>8.54</td>
<td>9.81</td>
<td>10.67</td>
<td>9.89</td>
<td>9.69</td>
</tr>
<tr>
<td>12</td>
<td>198311</td>
<td>9.10</td>
<td>8.85</td>
<td>8.50</td>
<td>9.94</td>
<td>10.26</td>
<td>9.89</td>
<td>9.82</td>
</tr>
<tr>
<td>15</td>
<td>198402</td>
<td>9.32</td>
<td>9.73</td>
<td>8.44</td>
<td>10.09</td>
<td>11.10</td>
<td>9.89</td>
<td>9.96</td>
</tr>
<tr>
<td>16</td>
<td>198403</td>
<td>9.83</td>
<td>9.93</td>
<td>8.43</td>
<td>10.13</td>
<td>11.38</td>
<td>9.89</td>
<td>10.00</td>
</tr>
<tr>
<td>17</td>
<td>198404</td>
<td>10.18</td>
<td>10.16</td>
<td>8.41</td>
<td>10.17</td>
<td>12.09</td>
<td>9.89</td>
<td>10.04</td>
</tr>
<tr>
<td>18</td>
<td>198405</td>
<td>10.65</td>
<td>10.25</td>
<td>8.39</td>
<td>10.21</td>
<td>12.27</td>
<td>9.89</td>
<td>10.08</td>
</tr>
</tbody>
</table>

### Table 2

**AN EVALUATION OF FORECAST ERRORS**

<table>
<thead>
<tr>
<th></th>
<th>BEST</th>
<th>BEST</th>
<th>BEST</th>
<th>BEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CIAM</td>
<td>ARIMA</td>
<td>ARIMA</td>
<td>CIAM</td>
</tr>
<tr>
<td>MEAN ERROR</td>
<td>-.0704</td>
<td>.9720</td>
<td>-.3988</td>
<td>-1.3532</td>
</tr>
<tr>
<td>MEAN ABSOLUTE ERROR</td>
<td>.4312</td>
<td>1.0960</td>
<td>.6956</td>
<td>1.3564</td>
</tr>
<tr>
<td>MEAN SQUARED ERROR</td>
<td>.3318</td>
<td>2.0111</td>
<td>.6445</td>
<td>2.3603</td>
</tr>
<tr>
<td>RMS ERROR</td>
<td>.5760</td>
<td>1.4181</td>
<td>.8028</td>
<td>1.5363</td>
</tr>
<tr>
<td>CORR COEFF</td>
<td>.7930</td>
<td>-.6670</td>
<td>.7020</td>
<td>.8390</td>
</tr>
</tbody>
</table>

585
3. SAS INDEXING was initially designed and published by the author; see (Pourian, 1984).

4. Once again, it is possible to use more sophisticated averaging methods such as the moving average, etc. As it is well known in the forecasting literature, however, more complex methods need not necessarily minimize forecast errors.

5. The first financial variable is designated as V16 due to the fact that it is the sixteenth variable in the data set regardless of being financial or non-financial (e.g., time) variable. Also notice that this variable starts at column 105 as the subsequent program’s input line reveals. On the other hand, the variable V108 is the twentieth variable chosen in our study—it is the 108th variable in the data set.

6. Best ex post and best ex ante ARIMA forecasts have the following meanings. An ex ante ARIMA is “identified” during the estimation period by use of guidelines suggested in the Box-Jenkins literature. Ex post ARIMA model, on the other hand, is a model which has produced better forecasts without “proper” ex ante identification.

7. For the problems encountered in a priori and posterior identification of ARIMA models see (Pourian, 1985b).

REFERENCES


CHICAGO MERCANTILE EXCHANGE [1985]. Inside S&P Stock Index Futures.


TRADEMARK NOTE

SAS is the registered trademark of SAS Institute Inc., Cary, NC, USA.

AUTHOR’S ADDRESS

The author is Assistant Professor of Finance in the Department of Economics and Finance, School of Business, Western Carolina University, Cullowhee, N.C. 28723, U.S.A.; Telephone Numbers: (704) 227-7403 and (704) 293-3914. Suggestions and questions are welcome.