Working Paper

Research Department
Federal Reserve Bank
of Minneapolis

Federal Reserve Bank of Minneapolis
Research Department

A BVAR FORECASTING MODEL FOR THE CHILEAN ECONOMY

Richard M. Todd*

Working Paper 407

August 1988

NOT FOR DISTRIBUTION
WITHOUT AUTHOR APPROVAL

ABSTRACT

Doan, Litterman, and Sims have described a method for estimating Bayesian vector autoregressive (BVAR) forecasting models. The method has been successfully applied to the U.S. macroeconomic dataset, which is relatively long and stable. Despite the brevity and volatility of the post-1976 Chilean macroeconomic dataset, this paper shows that a straightforward application of the DLS method to this dataset, with simple modifications to allow for delays in the release of data, also appears to satisfy at least one criterion of relative forecasting accuracy suggested by Doan, Litterman, and Sims. However, the forecast errors of the Chilean BVARs are still large in absolute terms. Also, the model's coefficients change sharply in periods marked by policy shifts, such as the floating of the peso in 1982.

*Federal Reserve Bank of Minneapolis. Some of the research for this paper was performed while the author was a consultant to the Economics Department of the University of Santiago, whose support is gratefully acknowledged. Thanks also to Felipe G. Morande for conceiving this project and supplying many of the ideas. Comments and encouragement from Felipe Montt and Valentin Carril were also quite helpful. The author, naturally, takes responsibility for the paper's errors and shortcomings.

Doan, Litterman, and Sims (1984, hereafter DLS) have described a method for estimating Bayesian vector autoregressive (BVAR) forecasting models. The method has been successfully applied to the U.S. macroeconomic dataset, which is relatively long and stable. Despite the brevity and volatility of the post-1976 Chilean macroeconomic dataset, a straightforward application of the DLS method to this dataset also appears to satisfy at least one criterion of relative forecasting accuracy suggested in DLS. However, the Chilean BVAR's forecast errors are still large in absolute terms. An improved dataset or modifications to the basic DLS method may be needed to significantly improve the model's forecasting performance.

The DLS Method Has Produced Successful Forecasting Models for U.S. Data

The DLS method mainly consists of procedures for choosing a Bayesian prior distribution for the coefficients of a vector autoregression. A vector autoregression is a multivariate time series model where an $n \times 1$ vector of time-indexed elements is regressed on its own lagged values. Typically, the value of the vector at time t is regressed on its values at $t - 1$ through $t - k$. This means that the model contains n equations, each with a constant, a disturbance term, and k lags of each variable on the right side. DLS describe how to choose a prior probability distribution for the disturbance term variance and the $nk + 1$ coefficients of each equation.

To simplify the task of choosing the $nk + 1$ means and $(nk+1) \times (nk+1)$ covariances of each equation's coefficients, DLS first propose that the prior distributions be chosen from a particular family of distributions. Many aspects of the prior distribution are common to all members of the family. For example, in each member of the DLS family, the prior means of the coefficients are set to values associated with a random walk. Also, the prior variance of the coefficient on the k^{th} lag of a variable declines as k increases, indicating increasing confidence that the coefficient should be close to its prior mean.

The members of the DLS family do differ in a few dimensions. For example, members could differ by the degree of confidence they express in the random walk prior means of the coefficients, in the zero prior mean of the constant term, or in the importance of time variation in the coefficients.

Each dimension by which the members differ is indexed by a so-called hyperparameter. By specifying a value for each hyperparameter, the model builder would select a particular prior from the DLS family. The number of hyperparameters is typically small (in DLS, 8), and each has an economic or statistical interpretation, and usually a numerical scale as well, that does not vary with the model to which it is applied. This means that forecasters probably can develop beliefs about the best values of these hyperparameters more readily than they can develop beliefs directly about the numerous means and covariances of the prior distribution of the coefficients of a particular model.

In a fully Bayesian implementation of the DLS method, the forecaster would also specify a prior distribution over the hyperparameters. In principle, at least, it would then be possible to construct the prior distribution of the model's coefficients as a mixture of the distributions associated with each individual hyperparameter setting, where the weights used to form the mixture are taken from the prior over the hyperparameters. Conventional Bayesian procedures--integration and scaling of the likelihood times the prior--would yield a posterior distribution for the coefficients. In general, this involves intractable integrals and thus cannot be done. An obvious exception occurs when the forecaster's prior over the hyperparameters is degenerate, putting unit mass on a single member of the DLS family. In this case the Kalman filter will easily compute the posterior distribution of the coefficients.

DLS do not expect, however, that forecasters should be able to compute the difficult integrals of the general case or identify a single member of the DLS family of distributions as their own prior beliefs about their model's coefficients. DLS propose a tractable alternative that, under certain assumptions, should approximate the fully Bayesian procedure just described. For each hyperparameter setting, the forecaster computes how well a model with those hyperparameters would have forecasted in the past. The criterion DLS use to evaluate each model's simulated forecasting performance can be interpreted as a likelihood function relating the data and the hyperparameters. DLS recommend using the hyperparameter setting that maximizes this likelihood function.

Since the data are used to pick the prior, this is clearly not a strict Bayesian procedure. However, DLS note that if (a) the forecaster's priors over the hyperparameters are nearly flat, (b) the DLS likelihood statistic is high within a region R and low elsewhere, and (c) the important features of the estimated models are not too sensitive to variations of the hyperparameters within R, then picking the hyperparameters that maximize the DLS likelihood gives a model whose important features are approximately the same as the model implied by the fully Bayesian procedure. DLS then argue that condition (a) is plausible and that conditions (b) and (c) seem to hold, at least for the U.S. macroeconomic dataset they examine. This rationale justifies the use of a non-Bayesian procedure to estimate a "Bayesian" vector autoregression.

Whatever its rationale, the DLS method for estimating BVARs has produced models that forecast U.S. economic data reasonably accurately. In the simulated out-of-sample forecasts, DLS (p. 22) observed "an average of about 2 percent improvement in the one-step-ahead forecast errors in going from (a system of univariate autoregressions for each variable) to the final (BVAR)." They claim (p. 24) that,

> "Despite the small absolute gain in forecast accuracy, it is significant that we have documented a consistent gain from the use of a formally explicit multivariate method in a system of this size. This has not been done before, to our knowledge. The difference in accuracy that we find between multivariate and univariate methods is substantial relative to differences in forecast accuracy ordinarily turned up in comparisons across methods, even

> though it is not large relative to total forecast error. Moreover, if we think of a decomposition of movements in the data into signal and noise, with noise being the dominant component, then a 2 percent increase in forecast accuracy must represent a much larger percentage increase in the amount of signal that is being captured."

Litterman (1986) and McNees (1986) present evidence that the actual forecasts generated in the early 1980s by a small BVAR of the U.S. macroeconomy were also at least as accurate, for real variables like real GNP and unemployment, as the forecasts of the major U.S. economic consulting firms.

Chilean Macroeconomic Data Pose Severe Difficulties to Any Forecasting Method

Compared to the U.S. macroeconomic data series that DLS used, the Chilean macroeconomic data series are short and volatile. Current practice among analysts of the Chilean economy, I am told, is to regard all data available for periods before 1976 as unreliable, incompatible with current data, or both. The validity of this practice needs to be examined, but I have adopted it here. As a result, I have about 144 monthly observations, enough to encompass only about 3 or 4 normal business cycles.

During this short period, however, business cycles were not, at least by U.S. standards, normal. The period began at the tail end of a rapid disinflation, and growth rates of the seasonally adjusted M1 money stock (M1NPS) and the wholesale price level (WPI) continued to drift down from the 8-12 percent range in 1976 to slightly negative rates by 1981 (Figures 2c and 3c). Domestic interest rates (DIR) began the period at levels far above the highest values in the DLS dataset, fell precipitously for a year, but remained high (for example, relative to percentage changes in the WPI) throughout the period (Figure 6a; nominal rates are shown). Policy regarding the foreign exchange value of the Chilean peso (XCH) shifted twice, from floating to fixed in 1979 and back to floating in 1982 (Figure 4a). The first exchange rate shift was roughly contemporaneous with a peak in the price of a major export good, copper (PCOB, see Figure 8a), and a liberalization of capital controls. It was followed by surges in Chilean capital inflows (KINF; Figure 5a) and international nominal rates of interest (LIBOR; Figure 7a). The second exchange rate shift was preceded by sharp declines in the (copper) terms of trade (Figure 8a), capital inflows (Figure 5a), and seasonally adjusted industrial production (IPINSS; Figure 1a). It was followed by a burst of inflation (Figure 3c), a spike in domestic interest rates (Figure 6a), and partial rebounds in the (copper) terms of trade (Figure 8a) and industrial production (Figure 1a). Variables like these also vary in the United States, but generally to a much milder degree and in a dataset whose greater length allows more precise measurement of any associated changes in the relationships among variables.

The brevity of the Chilean dataset, the volatility of the Chilean data series, and the possibility that policy changes significantly affected the relationships among Chilean variables all pose difficulties for any forecasting methodology. Successful forecasting models may require longer datasets, the imposition of many coefficient restrictions derived from structural econometric models of the recent Chilean experience, or more sophisticated modeling of time variation in both the coefficients and the disturbance term distributions. In this paper, I will not pursue those possible avenues of improvement. Instead I will show that DLS's BVAR technique, applied to the existing Chilean dataset with no significant modifications to take account of the data's volatility or the effects of policy shifts, can still at least match the forecasting performance of univariate time series models while capturing some relationships among variables.

The DLS BVAR Method at Least Matches a System of Univariate Equations

I have applied a slightly modified version of the DLS BVAR methodology to the January 1976 through December 1987 monthly values of the eight data series discussed in the previous section. The method yields a quasi-univariate system of equations. That is, under apparently optimal hyperparameter settings, the estimated forecasting equations forecast about as accurately as univariate equations and allow only moderately more interaction among the variables. Experimentation with other hyperparameter settings suggests there may be a tradeoff between optimizing the model to predict industrial production and optimizing it to predict other variables, such as inflation.

The Chilean BVAR was originally specified with six endogenous and two exogenous variables. The six original endogenous variables were IPINSS (seasonally adjusted industrial production), MINPS (seasonally adjusted M1 money), WPI (wholesale price index), XCH (peso-dollar exchange rate), KINF (capital inflows), and DIR (domestic interest rates). The two original exogenous variables were LIBOR (international interest rates, represented by the London interbank offer rate) and PCOB (the price of copper). The original model had six equations for the endogenous variables, each with a constant term and six lags of each of the endogenous and exogenous variables on the right side. The two purely exogenous variables were represented by unrestricted univariate autoregressions, with 2 lags for LIBOR and 4 lags for PCOB.

I also experimented with models in which capital inflows were treated as exogenous. The model reported here is somewhat intermediate. Technically it treats capital inflows as endogenous and has six multivariate equations (plus univariate equations for LIBOR and PCOB). However, the prior distribution of the coefficients in the equation for capital inflows causes most of the coefficients on other variables in that equation to be nearly zero. Except for contemporaneous correlations between its disturbance term and other disturbance terms, capital inflows are nearly exogenous.

The model presented here differs from the initial model by allowing for varying delays in the release of data for its variables. Data on the money supply, the wholesale price index, the exchange rate, domestic and foreign interest rates, and the price of copper in month t are assumed to all first be available in month $t + k$, $k \geq 1$. Data on industrial production and capital inflows in month t are assumed to be released two months later (t+k+2). Accordingly, the equations for IPINSS and KINF are augmented by terms for the contemporaneous and lead one values of the other variables.

The prior means of the coefficients of the six endogenous equations are set according to the continuous-time random walk prior. That is, the prior means are set to conform to a model in which each variable evolves as a continuous function of time such that at each instant its expected value at any future date equals its current value. The variables are observed only as discrete monthly averages, however. Discrete time averages of continuous random walks are generated by autoregressive processes with an infinite number of lags, where the coefficient on the k^{th} lag is given by

$$(1-\alpha)\alpha^{(k-1)},$$

where $\alpha = \sqrt{3} - 2$ [see Working (1960) or Christiano and Eichenbaum (1987)]. In each equation of the Chilean BVAR, the coefficients on the six lags of the dependent variable are given values according to this formula, with k = 1, 2, 3, 4, 5, 6. All other coefficients have a prior mean of zero.

The prior variances of the coefficients in the Chilean BVAR are governed by ten hyperparameters--the eight discussed in DLS and two more subsequently introduced by Sims. Five of these hyperparameters affect the variances of each coefficient individually. Three--including the two new ones--affect the variances of linear combinations of coefficients. Two control the nature of the time variation in the coefficients.

Before allowing for restrictions on linear combinations of coefficients, the prior variances of the model's coefficients have the following form:

1. For the variance of the k^{th} lag of the i^{th} variable in the i^{th} equation,

$$\mathrm{var}(a^i_{i,k}) = \frac{\mathrm{TITE} \times \mathrm{OWN}}{k \times \exp\left(\mathrm{WT}\times\mathrm{WEIGHT}(i,i)\right)}.$$

2. For the variance of the k^{th} lag of the j^{th} variable in the i^{th} equation $(i\neq j, k\neq 0)$,

$$\mathrm{var}(a^i_{j,k}) = \frac{\mathrm{TITE} \times \mathrm{CROSS} \times \sigma_i^2}{|k| \times \exp\left(\mathrm{WT}\times\mathrm{WEIGHT}(i,j)\right) \times \sigma_j^2}.$$

3. For the variance of the coefficient on the contemporaneous value of the j^{th} variable in the i^{th} equation ($i\neq j$; i=IPINSS or i=KINF),

$$\mathrm{var}(a^i_{j,k}) = \frac{2 \times \mathrm{TITE} \times \mathrm{CROSS} \times \sigma_i^2}{\exp\left(\mathrm{WT}\times\mathrm{WEIGHT}(i,j)\right) \times \sigma_j^2}.$$

4. For the variance of the constant term in the i^{th} equation,

$$\mathrm{var}(c^i) = \mathrm{TITE} \times \mathrm{CON} \times \sigma_i^2.$$

In these expressions, σ_i is the variance of the disturbance term in equation i. The σ_i are treated as though they are known but are in fact estimated as 0.9 times the standard error of the residual in a regression of variable i on six lags of itself. The terms WEIGHT(i,j) come from the 6×8 matrix

$$\begin{bmatrix} 0.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\ 1.0 & 0.0 & 0.5 & 0.5 & 0.5 & 1.0 & 1.0 & 1.0 \\ 1.0 & 0.5 & 0.0 & 0.5 & 1.0 & 1.0 & 1.0 & 1.0 \\ 2.0 & 2.0 & 2.0 & 1.0 & 2.0 & 2.0 & 2.0 & 2.0 \\ 1.0\times KW & 1.0\times KW & 1.0\times KW & 1.0\times KW & 0.0\times KW & 1.0\times KW & 1.0\times KW & 1.0\times KW \\ 2.0 & 2.0 & 2.0 & 2.0 & 2.0 & 1.0 & 2.0 & 2.0 \end{bmatrix}.$$

Along with the hyperparameter WT, this matrix allows selective alteration of the variances of a given variable's coefficients in a given equation.

I exploited this possibility by experimenting (in the initial model only) with various values of KW, which controls the influence other variables have on capital inflows. Treating KW as, in effect, an eleventh hyperparameter, I selected a value of KW = 8, which makes capital inflows nearly a univariate process.

Except for the KW factor, the rest of the WEIGHT matrix is patterned after the one used by DLS. Rows 1 and 5, for IPINSS and KINF, have the basic DLS pattern of zeroes on own lags and ones on other variables. Rows 4 and 6, for XCH and DIR, have ones on own lags and twos on other variables, the pattern DLS suggest for variables especially likely to follow random walks. Rows 2 and 3, for MINPS and WPI, have a modified form of the 0-1 pattern. Because they are likely to be sensitive to each other, to the exchange rate and, in the case of MINPS, to KINF, these other variables are given the intermediate downweighting factor 0.5 in rows 2 and 3.

The terms OWN, CROSS, and CON are hyperparameters governing the size of the variances of the coefficients of, respectively, own lags (lags of the dependent variable), cross lags (lags of variables other than the dependent variable), and the constant terms. Finally, the hyperparameter TITE is used to scale all prior variances up or down simultaneously.

Three hyperparameters govern the tightness with which three linear restrictions are imposed on the coefficients. SUM is used to control the tightness of a restriction that the sums of the coefficients on own lags should be one and the sums of the coefficients on cross lags should be zero. BEGWT controls the tightness of another restriction that even if the SUM restriction is violated, the coefficients on all variables should collectively imply that the best forecast of a variable is given by the random walk prior. NOMWT controls the tightness of a long-run superneutrality constraint. This constraint allows the sums of individual nominal variables to deviate from one in their own equations and from zero in other equations, but requires that the sum of all nominal variable coefficients be approximately one in nominal variable equations and approximately zero in other equations. (Nominal here means variables measured in units of domestic currency and hence likely to inflate at about the same rate in the long run. MINPS and WPI were treated as nominal here. XCH was not, since it partly depends on inflation outside Chile and because experiments suggested little gain from treating it as nominal.)

The eight hyperparameters discussed so far determine the prior mean, β_0, and prior variance-covariance matrix, $\sum_0$. Two more hyperparameters govern how the posterior mean and variance evolve as the model is estimated by applying the Kalman filter to the observations one by one. The hypothesized law of motion of the coefficients, which must be supplied to the Kalman filter, is

$$(\beta_t - \beta_0) = \text{DECAY} \times (\beta_{t-1} - \beta_0) + \mu_t,$$

where μ_t is taken to be normally distributed with mean zero and variance-covariance matrix TVAR $\times \sum_0$. DECAY and TVAR are the two hyperparameters governing, respectively, the rate at which the coefficients decay toward their prior mean and the extent to which they vary around their expected path.

I selected values for the ten hyperparameters by attempting to maximize the likelihood statistic developed by DLS. For the i^{th} endogenous variable, the likelihood statistic for any given hyperparameter setting is computed as a weighted average of variable i's one-step-ahead forecast errors when the model is estimated with those hyperparameters. The one-step-ahead forecast errors are computed recursively, with each forecast based on coefficients estimated only through the data that would have been available when the forecast was made. Forecasts of exogenous variables, which are needed to forecast the endogenous variables, are computed in the same recursive fashion, using their univariate equations. The weights on the individual forecast errors in the average are given by the forecast's conditional variance (conditional on the data and estimated probability distribution of the coefficients at the time the forecast was made) divided into the geometric mean of all conditional forecast variances for variable i. The overall likelihood statistic for the model is ordinarily the sum of the likelihoods for each endogenous variable, but I also experimented with maximizing the likelihoods of individual equations (see below).

In attempting to find the hyperparameter setting with the highest likelihood, I searched over hundreds of possibilities. Computing the likelihood for a given hyperparameter setting takes about 10 seconds on an Amdahl dual 580 mainframe computer system, and a large-scale search on a personal computer would take days. Even on the mainframe, it is not practical to thoroughly search all interesting hyperparameter settings. I chose a method called axial search, which searches over one hyperparameter at a time while keeping the others fixed at their best (up till then) values. With about ten values for each hyperparameter and ten hyperparameters, each axial search iteration covered about 100 settings.

The success of axial search depends on the shape of the likelihood function and the order in which the hyperparameters are searched. If the likelihood is symmetric (around lines through its peak and parallel to the axes) as illustrated in Figure 20a, then axial search will probably find a nearly maximizing setting for the hyperparameters no matter in what order they are searched. However, if the likelihood has the asymmetric shape of Figure 20b, then results of axial search may depend on the order in which the hyperparameters are searched, and some orders may not find settings that are close to optimal. Even repeating the axial search from the best point of a previous axial search may not get around this problem if, for example, the likelihood is asymmetric and has multiple local peaks. To lessen the possibility of such a result, I sometimes varied the starting values and search order of the hyperparameters.

For the model I have described, with KW = 8 implying a very nearly univariate equation for endogenous KINF and exogenous LIBOR and PCOB having strictly univariate equations, the highest likelihood value I found in my initial hyperparameter search is associated with estimated equations for the other five endogenous variables that are also not far different from univariate. The chosen hyperparameters, shown in Table 2, are not much different from those typically found in applications of the DLS method to U.S. data, although the TITE*CROSS product of 0.0001 implies a relatively high degree of confidence that the coefficients of variable j in equation i ($i \neq j$) are zero. However, as also shown in Table 2, the BVAR model's root mean squared forecasting errors 1, 6, and 12 months ahead during 1981-87 are generally similar to, and for IPINSS worse than, those of the system of univariate equations shown in Table 1. The similarity of the univariate and BVAR models is also evident in the histories of their forecast errors, shown in the upper panels of Figures 9-14.

In addition, a decomposition of the sources of forecast error indicates that the BVAR model attributes a very high percentage of the variance of each variable's forecast error to the variable itself (that is, to its own disturbance term). For forecasts of MINPS, WPI, and KINF 1-6 months ahead, nearly all of the variance of the forecast errors is attributed to the variable's own disturbance term. This is somewhat less true of IPINSS, XCH, and DIR. The slightly lower degree of autonomy displayed by these variables apparently reflects contemporaneous correlation between their disturbance terms and the disturbance terms of the other variables, for the univariate system shows roughly the same degree of autonomous variation in these variables. After one year, some of the variance decompositions show stronger cross-variable effects, but these figures are subject to wide confidence bands and thus may not be significant. [See Runkle (1987) for a discussion of this point. See Doan and Litterman (1986, p. 19-4) for a procedure for computing confidence bands for BVARs.] Though not shown in the tables, estimated coefficients on variables other than the dependent variable or constant term are also small. Similarly, the response of variable j to a surprise movement in variable i (the impulse response of j to i) is generally small.

Despite the similarity of the BVAR and univariate models, the likelihood statistic favors the BVAR. The discrepancy between the likelihood statistic, which favors the BVAR, and the root mean squared errors, which show mixed results, may be due to several factors. One obvious reason is that the system likelihood, as the sum of the individual equation likelihoods, balances the BVAR's inferior performance in forecasting IPINSS against its superior performance in forecasting the other five endogenous variables. Another possible reason is that the equation likelihood, unlike the root mean squared error, does not necessarily give equal weight to two errors of the same magnitude. An error in a period for which the conditional variance of the forecast error was high will depress the likelihood less than an error of the same size occurring when the conditional forecast error variance was low.

Some of the forecast error histories in the upper panels of Figures 9-14 also indicate some possible advantages for the BVAR. In particular, the univariate model's forecast errors have on average been more biased than the BVAR's errors in recent years, as shown by the greater tendency of running totals of univariate model errors to drift up or down since about 1984.

Changing the likelihood criterion by omitting the likelihood of one or more endogenous variables can lead to somewhat different results. I experimented with maximizing just the likelihood of WPI and just the likelihood of IPINSS. In the former case, shown in Table 3, the overall system likelihood is actually higher than in Table 2, where the hyperparameters were chosen in an attempt to maximize the system likelihood. The axial search procedure for Table 2 obviously failed to maximize the system likelihood. This suggests that a technique like Sims's Bayesian interpolation of the likelihood surface may be useful (Sims 1986). The optimal values in Table 3 are quite extreme, especially for SUM and NOMWT. (SUM was always the last hyperparameter whose values were searched, and its value was set to zero during the initial searches over values of the other hyperparameters. NOMWT, by contrast, was generally among the first three or four hyperparameters searched over.) Together they keep the sums of coefficients fixed at their prior means. The rapid rate at which parameters decay toward their prior means (DECAY) is also unusual.

There are other anomalies in Tables 2 and 3. Despite the tight priors and rapid decay toward the prior means of its coefficients, the model of Table 3 shows slightly more cross-variable interaction in its variance decompositions than the first BVAR. Also, despite a stronger system likelihood and generally stronger equation likelihoods, the Table 3 model's root mean squared errors are generally higher than for the BVAR of Table 2.

Maximizing solely the likelihood of the IPINSS equation appears to imply modest changes in forecast performance but more substantial changes in the coefficients of the models. As shown in Table 4, the system likelihood for the IPINSS optimized model is lower than in Tables 2 or 3. Nonetheless, the equation likelihood for IPINSS and some root mean squared errors are superior in Table 4. The lower panels of Figures 9-14 also suggest that the model of Table 4 predicts IPINSS somewhat better, and other variables somewhat worse, than the other BVARs.

The running totals of one-step-ahead forecast errors in Figures 9-14 give a somewhat different perspective on the model of Table 4. For all variables except DIR, the Table 4 model has less of a tendency to consistently under or over predict during the 1981-83 period, as shown by the gaps that open up at that time in panel I of Figures 9-14. Thereafter, the lines in the I panels are roughly parallel, suggesting nearly equal tendencies to under or over predict. The superiority of the Table 4 model in 1981-83 may be just a fluke, attributable to the small sample size. Or it may be evidence that the Table 4 model captures useful information about turning points that the Table 2 model misses. Perhaps time will tell.

The difference in the Table 4 model's forecast performance appears to be small, however, compared to the change in the coefficients of the model. Figures 15-19, for example, generally show that the coefficients of the Table 4 model have evolved very differently from those of the Table 2 model (and the coefficients of the Table 3 model don't evolve at all). One difference is that the coefficients optimal for forecasting IPINSS allow much more interaction among variables. Table 4 reveals much lower degrees of autonomy in its variance decompositions of all endogenous variables, at least after one year. This may be due to the zero values of the hyperparameters SUM and NOMWT chosen in the maximization of the likelihood of IPINSS. The relatively high degree of time variation in the model's coefficients could also play a role. Also note that in both models changes in coefficients were especially rapid in about 1982 and, to a lesser degree, about 1985.

The generally moderate changes in the BVAR models' forecasting performance as the hyperparameters are varied around the optimal values is encouraging in one sense. As discussed above, this is one of the conditions necessary for interpreting the likelihood maximization performed here as an approximately Bayesian procedure. The forecasts of all of the models above have been fairly similar historically, and any one of them thus approximates reasonably well the mode of a Bayesian posterior distribution over future events. This convenient result may not extend to questions about the structure of the Chilean economy, given the wide variation the models show in the relationships among variables and the evolution of coefficients.

Conclusion

In some ways, this initial attempt to estimate a Chilean BVAR has been successful. With no significant modifications, the DLS method produces a multivariate model that captures at least a small degree of interaction among key macro variables while achieving much higher DLS likelihoods and perhaps slightly lower root mean squared forecast errors than a system of univariate equations. DLS suggest that this is not a trivial accomplishment.

At the same time, the Chilean BVAR of Table 2 is not a lot better than or even very different from a system of univariate equations. Further research on Chilean BVARs should probably look for improvements in three directions. One path toward possible improvements would be to tailor the DLS method to the Chilean situation. This could be done, for example, by modifying the time variation of the coefficients to make them more stable within but less stable across policy regimes. (The tendency for the models' coefficients to change rapidly during periods of well-known policy shocks, such as during 1982, recommends this path. See Figures 15-19.) It could also be attempted through the specification of restrictions on the variance-covariance matrix of the disturbance terms, perhaps with an eye toward achieving the kind of structural identification discussed by Sims (1987). An alternative way to improve Chilean BVARs would be to reconstruct a longer macroeconomic dataset. Finally, Sims has suggested further modifications (beyond the BEGWT and NOMWT priors used here) to the DLS method to allow for nonnormality and conditional heteroscedasticity in the distributions of the equation error terms.

Table 1

Performance Statistics for a BVAR Model With
An Approximately Univariate Hyperparameter Setting (2 lags)

Hyperparameters

OWN	= 100,000,000.00	SUM	= 0.0
CROSS	= 0.000,000,000,1	BEGWT	= 0.0
CON	= 100,000,000.00	NOMWT	= 0.0
WT	= 0.0	TVAR	= 0.0
TITE	= 1.0	DECAY	= 1.0

Performance Statistics

DLS Likelihood

SYSTEM	IPINSS	M1NPS	WPI	XCH	KINF	DIR
+389.72	+200.61	+216.28	+237.63	+201.84	-382.33	-84.32

Typical Forecast Errors* (percent)

IPINSS	M1NPS	WPI	XCH	KINF	DIR	
1-month-ahead	3.21	2.94	2.50	3.31	136.56	19.99
6-months-ahead	6.48	11.65	13.06	18.53	163.88	38.94
12-months-ahead	9.58	21.11	23.83	35.28	193.57	47.23

Autonomous Portion of Forecast Error Variance** (percent)

LIBOR	PCOB	M1NPS	WPI	XCH	KINF	DIR	IPINSS	
1-month-ahead	100.0	98.3	99.5	92.2	90.6	98.5	87.5	85.8
6-months-ahead	100.0	98.3	99.5	92.2	90.6	98.5	87.5	85.8
12-months-ahead	100.0	98.3	99.5	92.2	90.6	98.5	87.5	85.8

---

*Root-mean-squared errors in simulated out-of-sample forecasts from June 1981 to November 1987.

**Portion of forecast error variance attributed to own innovations. Computed from coefficients that were estimated over the full 1976-87 period and variance-covariance matrix of disturbances that was estimated over the October 1978 to December 1987 period. Choleski decomposition of variance-covariance matrix performed with variables ordered as here (from LIBOR to IPINSS).

Table 2

Performance Statistics for a BVAR Model With
Hyperparameters Set to Maximize the DLS Likelihood

Hyperparameters

OWN	= 0.01	SUM	= 25.0
CROSS	= 0.00001	BEGWT	= 0.0
CON	= 1000.0	NOMWT	= 2.0
WT	= 1.0	TVAR	= 0.0
TITE	= 10.0	DECAY	= 1.0

Performance Statistics

DLS Likelihood

SYSTEM	IPINSS	M1NPS	WPI	XCH	KINF	DIR
+472.10	+197.36	+225.73	+296.36	+206.86	-372.22	-81.98

Typical Forecast Errors* (percent)

IPINSS	M1NPS	WPI	XVH	KINF	DIR	
1-month-ahead	3.31	2.74	1.73	3.20	128.60	19.66
6-months-ahead	6.59	8.47	10.18	17.86	151.98	35.10
12-months-ahead	10.32	12.39	16.27	33.34	181.52	39.88

Autonomous Portion of Forecast Error Variance** (percent)

LIBOR	PCOB	M1NPS	WPI	XCH	KINF	DIR	IPINSS	
1-month-ahead	100.0	98.2	99.6	99.6	84.2	98.0	85.8	91.7
6-months-ahead	100.0	98.2	99.3	99.4	84.3	98.0	85.8	91.5
12-months-ahead	100.0	98.2	96.1	98.7	84.3	98.0	85.6	90.5

*Root-mean-squared errors in simulated out-of-sample forecasts from June 1981 to November 1987.

**Portion of forecast error variance attributed to own innovations. Computed from coefficients that were estimated over the full 1976-87 period and variance-covariance matrix of disturbances that was estimated over the October 1978 to December 1987 period. Choleski decomposition of variance-covariance matrix performed with variables ordered as here (from LIBOR to IPINSS).

Table 3

Performance Statistics for a BVAR Model With
Hyperparameters Set to Maximize the DLS Likelihood of WPI Only

Hyperparameters

OWN	= 0.08	SUM	= 100,000.0
CROSS	= 0.01	BEGWT	= 3.0
CON	= 1.0	NOMWT	= 10,000,000.0
WT	= 3.0	TVAR	= 0.000,000,001
TITE	= 2.0	DECAY	= 0.999

Performance Statistics

DLS Likelihood

SYSTEM	IPINSS	M1NPS	WPI	XCH	KINF	DIR
+486.37	+202.15	+225.04	+301.85	+213.94	-371.57	-85.05

Typical Forecast Errors* (percent)

IPINSS	M1NPS	WPI	XCH	KINF	DIR	
1-month-ahead	3.29	2.95	1.75	3.04	127.69	20.14
6-months-ahead	7.01	13.14	11.85	16.24	151.16	39.42
12-months-ahead	12.53	30.42	22.67	29.09	182.93	50.04

Autonomous Portion of Forecast Error Variance** (percent)

LIBOR	PCOB	M1NPS	WPI	XCH	KINF	DIR	IPINSS	
1-month-ahead	100.0	98.2	99.7	99.5	83.5	98.6	86.8	91.7
6-months-ahead	100.0	98.2	97.1	93.8	83.4	98.6	86.7	87.9
12-months-ahead	100.0	98.2	91.7	86.4	83.3	98.6	86.5	80.9

*Root-mean-squared errors in simulated out-of-sample forecasts from June 1981 to November 1987.

**Portion of forecast error variance attributed to own innovations. Computed from coefficients that were estimated over the full 1976-87 period and variance-covariance matrix of disturbances that was estimated over the October 1978 to December 1987 period. Choleski decomposition of variance-covariance matrix performed with variables ordered as here (from LIBOR to IPINSS).

Table 4

Performance Statistics for a BVAR Model With
Hyperparameters Set to Maximize the DLS Likelihood of IPINSS Only

Hyperparameters

OWN	=	0.7	SUM	=	0.0
CROSS	=	0.001	BEGWT	=	0.0
CON	=	2000.0	NOMWT	=	0.0
WT	=	1.0	TVAR	=	0.00001
TITE	=	10.0	DECAY	=	1.0

Performance Statistics

DLS Likelihood

SYSTEM	IPINSS	M1NPS	WPI	XCH	KINF	DIR
+451.61	+211.63	+211.29	+289.14	+200.43	-372.31	-88.56

Typical Forecast Errors* (percent)

IPINSS	M1NPS	WPI	XCH	KINF	DIR	
1-month-ahead	2.98	2.91	1.77	3.26	126.36	20.23
6-months-ahead	5.59	10.95	11.01	19.62	150.86	36.53
12-months-ahead	9.06	20.73	20.45	38.04	173.00	42.38

Autonomous Portion of Forecast Error Variance** (percent)

LIBOR	PCOB	M1NPS	WPI	XCH	KINF	DIR	IPINSS	
1-month-ahead	100.0	98.2	99.4	99.9	83.8	96.5	84.9	93.5
6-months-ahead	100.0	98.2	65.0	53.9	79.5	96.5	82.6	83.3
12-months-ahead	100.0	98.2	35.4	27.4	59.6	96.5	73.7	68.5

*Root-mean-squared errors in simulated out-of-sample forecasts from June 1981 to November 1987.

**Portion of forecast error variance attributed to own innovations. Computed from coefficients that were estimated over the full 1976-87 period and variance-covariance matrix of disturbances that was estimated over the October 1978 to December 1987 period. Choleski decomposition of variance-covariance matrix performed with variables ordered as here (from LIBOR to IPINSS).

References

Christiano, Lawrence J., and Eichenbaum, Martin. 1987. "Temporal Aggregation and Structural Inference in Macroeconomics." Carnegie-Rochester Conference Series on Public Policy 26, 63-130.

Doan, Thomas A., and Litterman, Robert B. 1986. User's Manual-RATs-Version 2.00. VAR Econometrics.

Doan, Thomas A., Litterman, Robert B., and Sims, Christopher A. 1984. "Forecasting and Conditional Projection Using Realistic Prior Distributions." Economic Reviews 3, 1-100.

Litterman, Robert B. 1986. "Forecasting with Bayesian Vector Autoregressions--Five Years of Experience." Journal of Business and Economic Statistics 4, 25-38.

McNees, Stephen K. 1986. "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts." Journal of Business and Economic Statistics 4, 5-15.

Runkle, David E. 1987. "Vector Autoregressions and Reality." Journal of Business and Economic Statistics 5, 437-42.

Sims, Christopher A. 1986. "Bayesmth: A Program for Multivariate Bayesian Interpolation." University of Minnesota, Center for Economic Research, Discussion Paper 234 (September).

Sims, Christopher A. 1987. "A Rational Expectations Framework for Short-Run Policy Analysis." William A. Barnett and Kenneth J. Singleton, eds. (1987), New Approaches to Monetary Economics. Cambridge University Press.

Working, Holbrook. 1960. "Note on the Correlation of First Difference of Averages in a Random Chain." Econometrica 28, 916-18.

FIGURE 1: IPINSS

(Industrial production, seasonally adjusted)

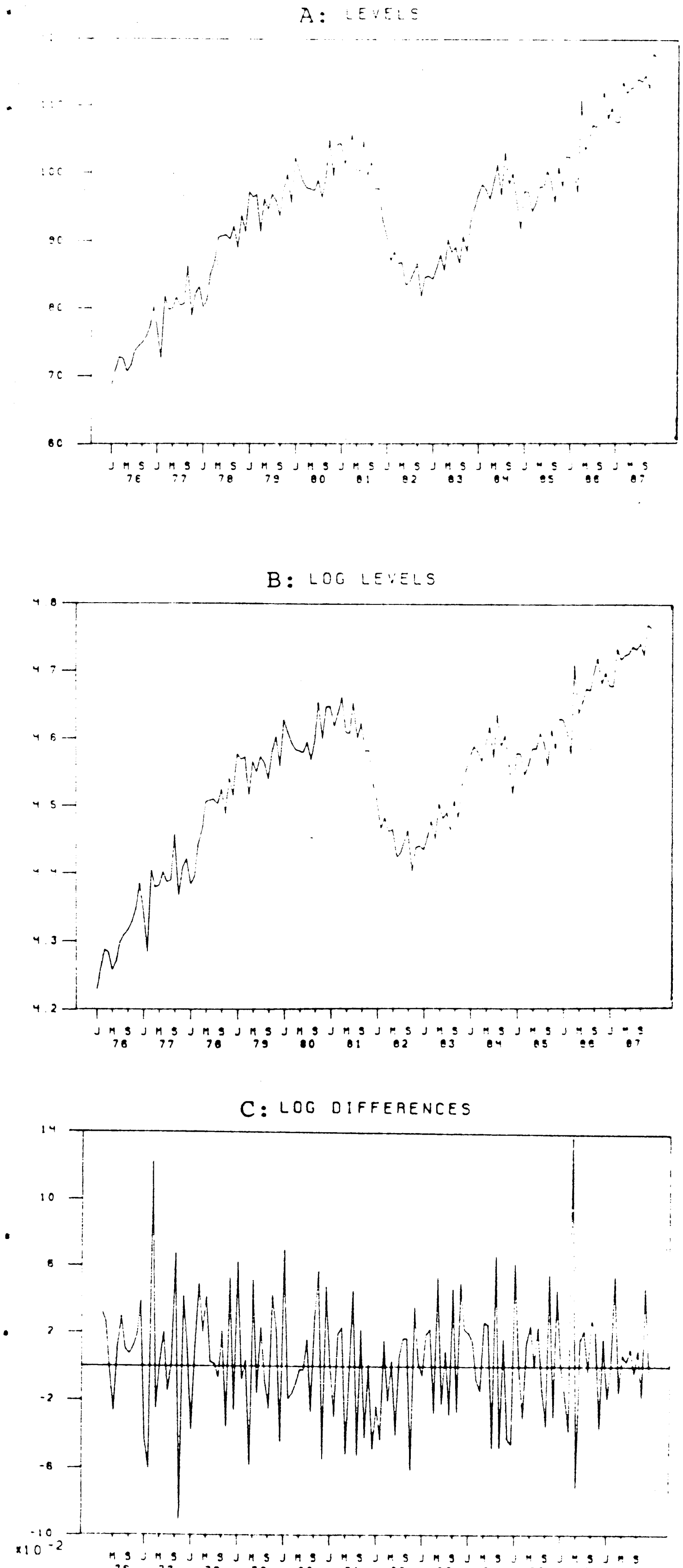


FIGURE 2: MINPS

(M1 money supply, seasonally adjusted)

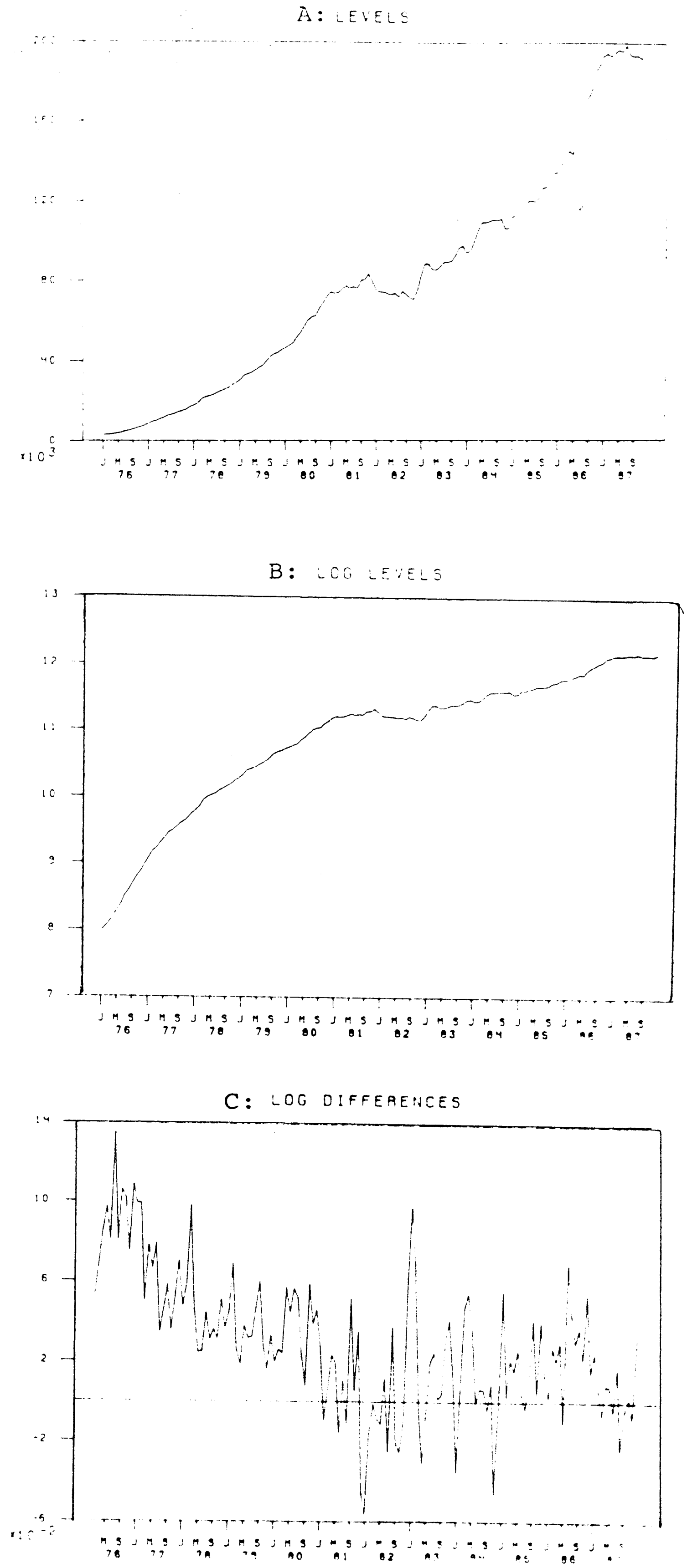


FIGURE 3: WPI
(Wholesale price index)

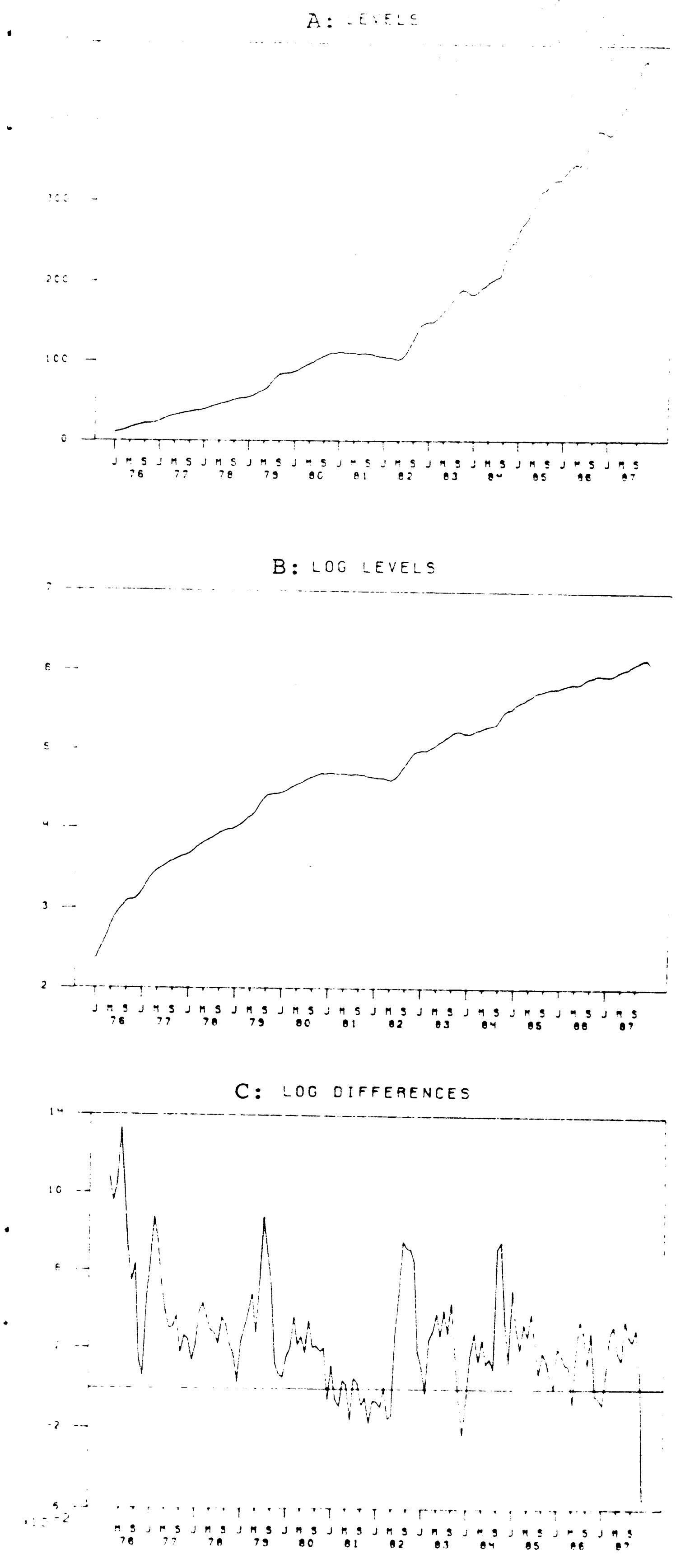


FIGURE 4: XCH
(Exchange rate, pesos per dollar)

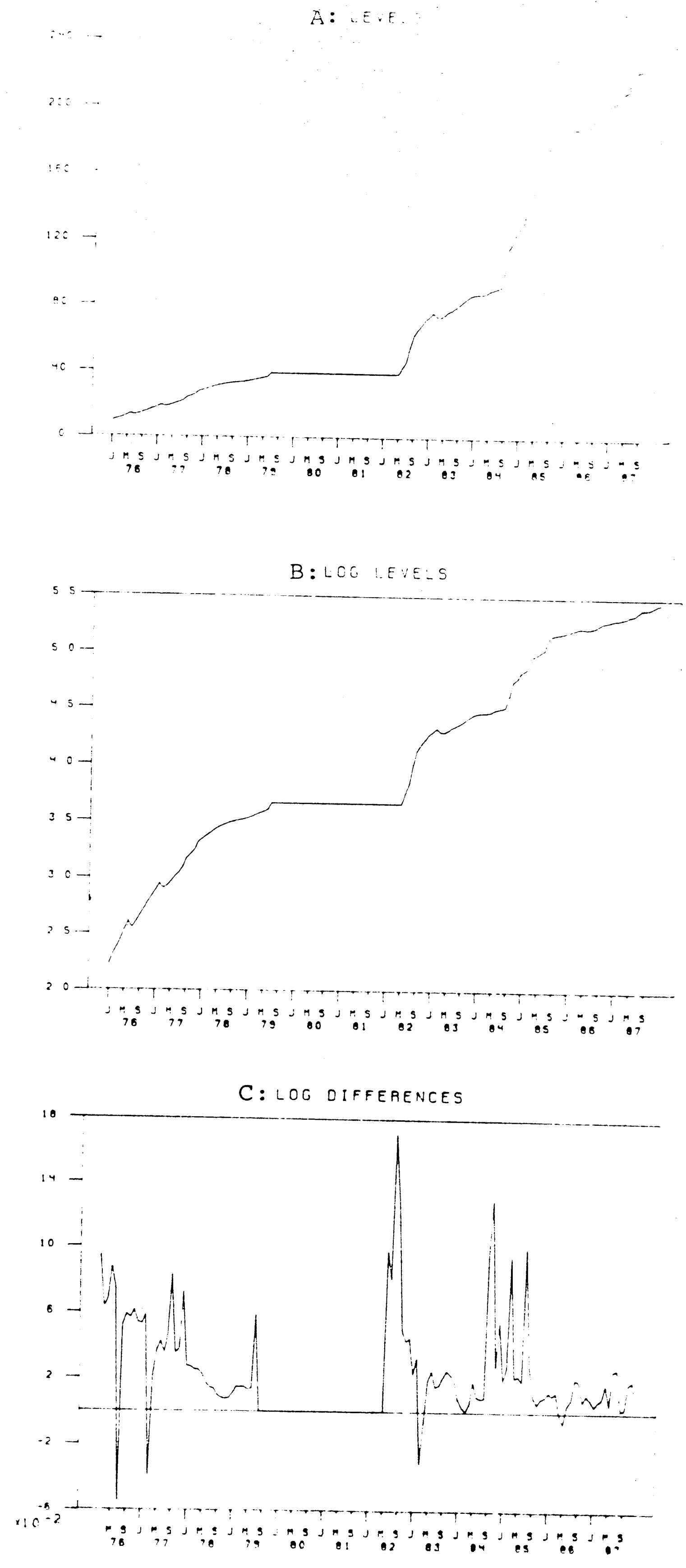


FIGURE 5: KINF
(CAPITAL INFLOWS)

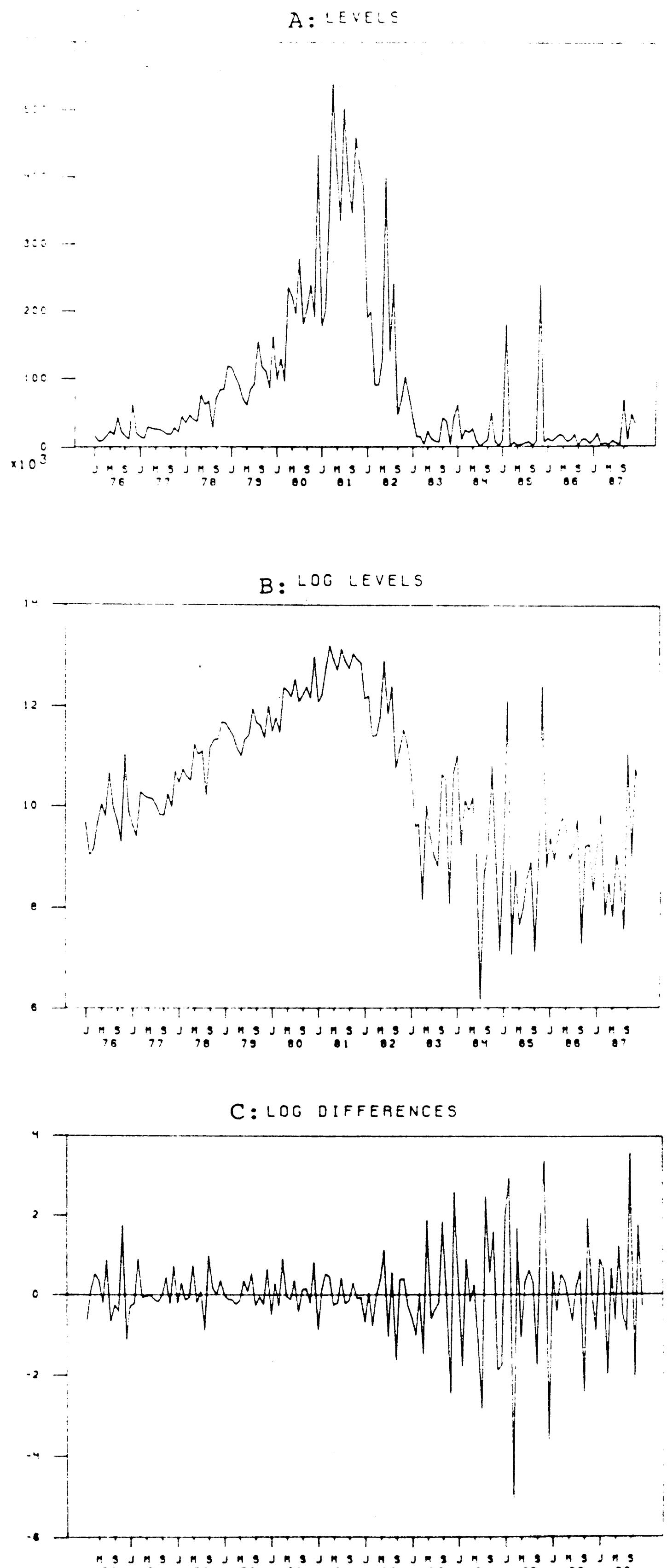


FIGURE 6: DIR
(DOMESTIC INTEREST RATES)

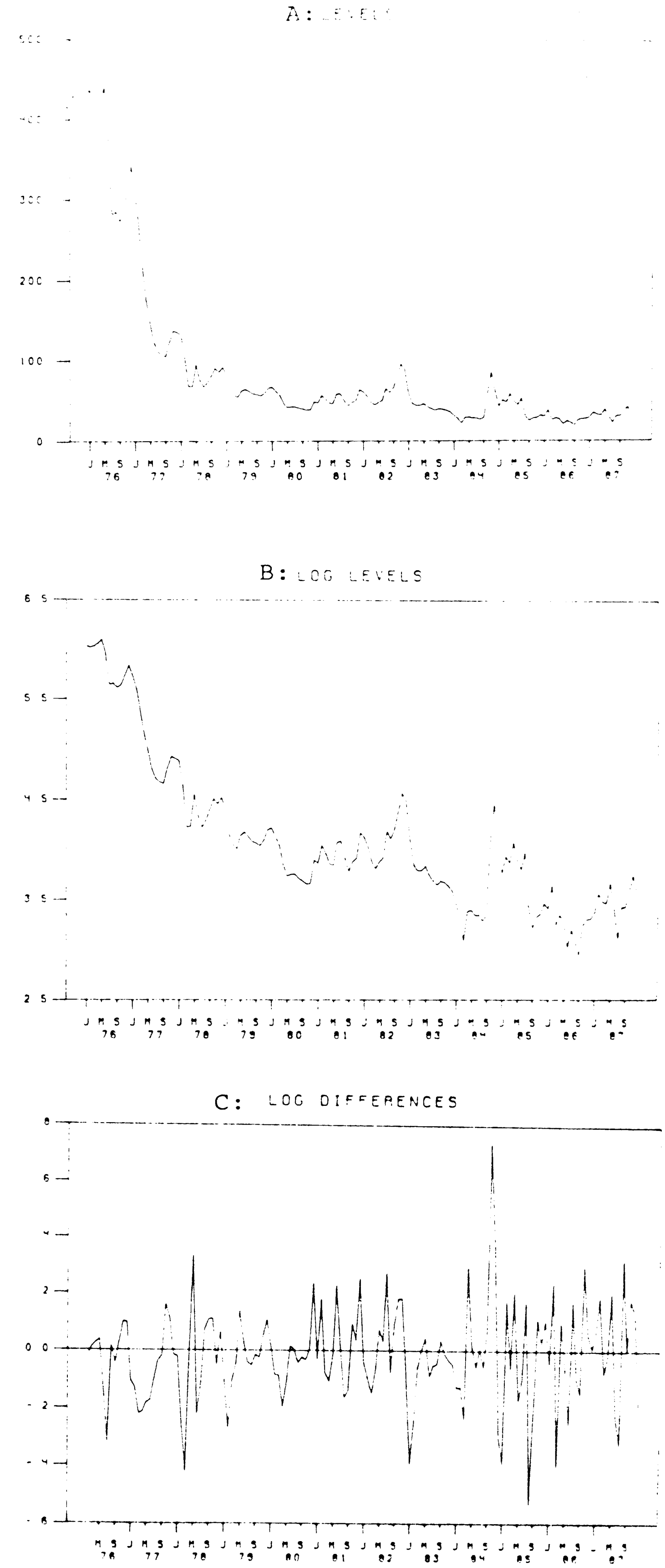


FIGURE 7: LIBOR

(London interbank rate of interest)

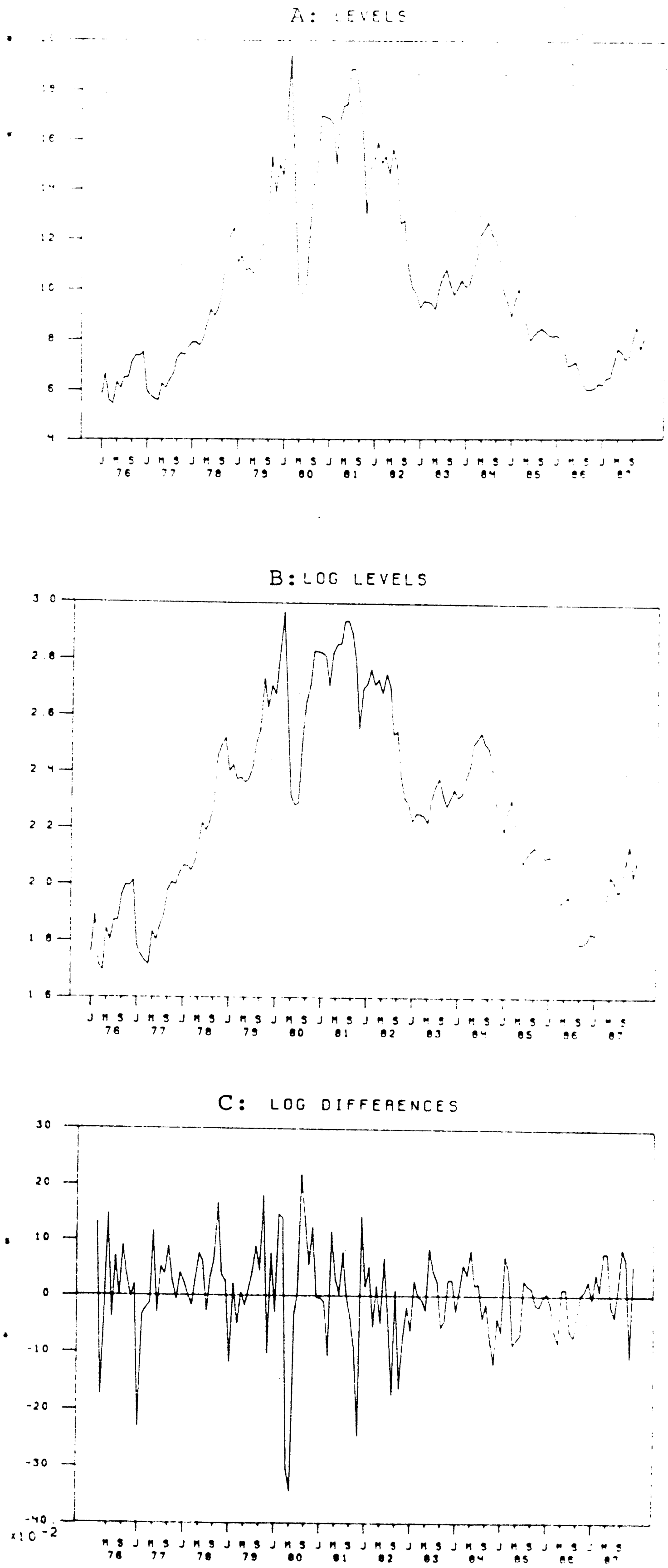


FIGURE 8: PCOB

(Price of copper)

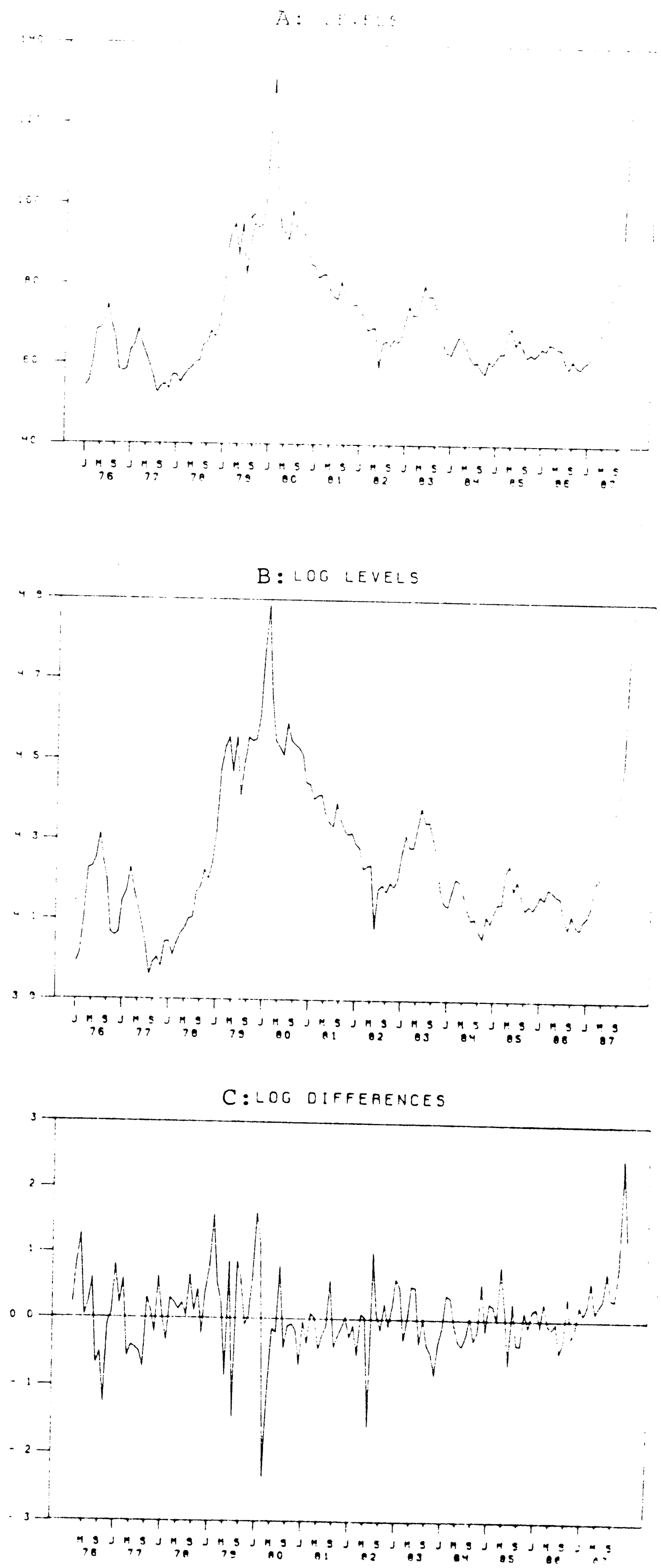


FIGURE 9:

HISTORY OF ERRORS IN FORECASTING IPINSS

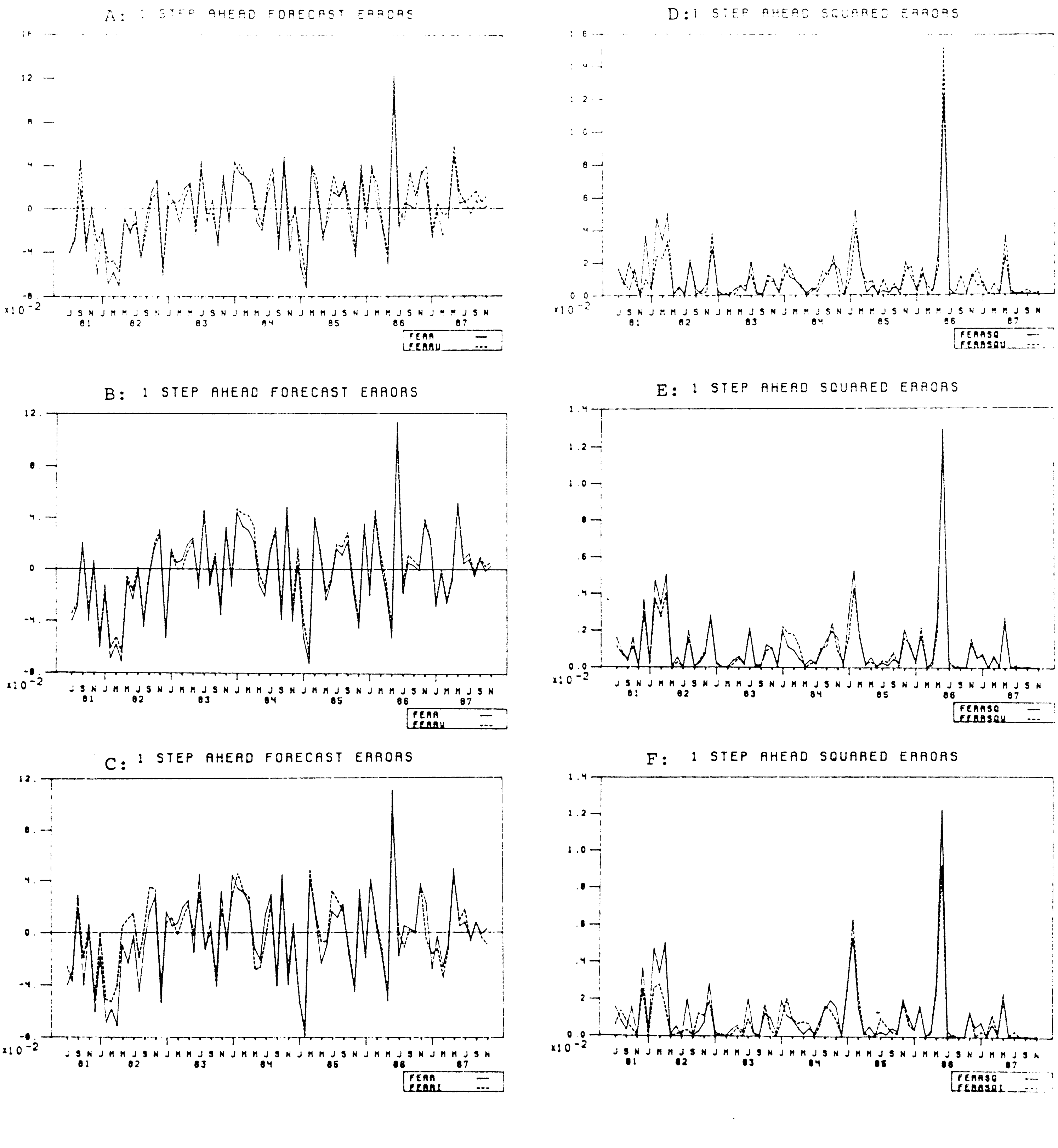


FIGURE 9 (CONTINUED)

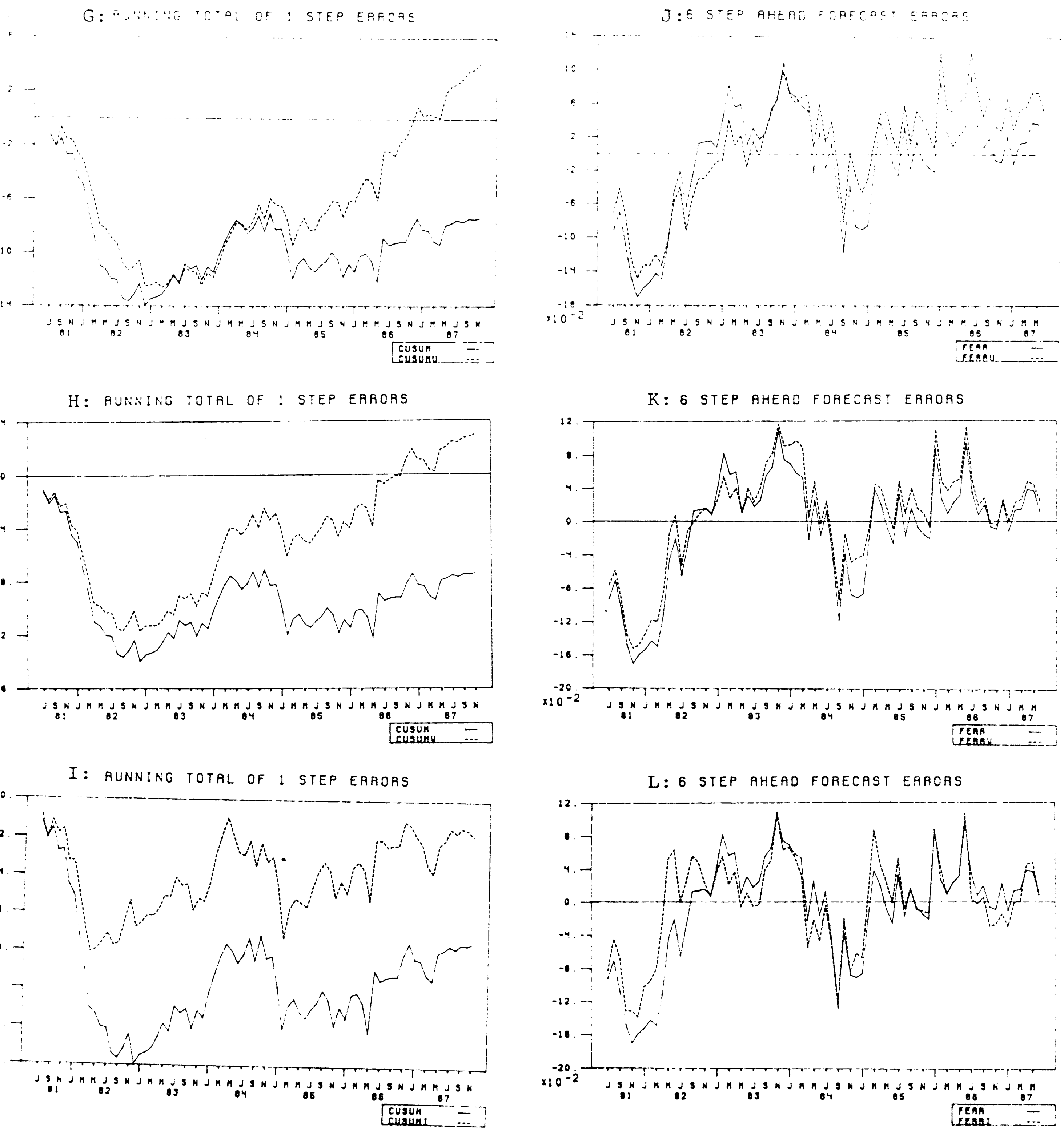


FIGURE 10:

HISTORY OF ERRORS IN FORECASTING M1NPS

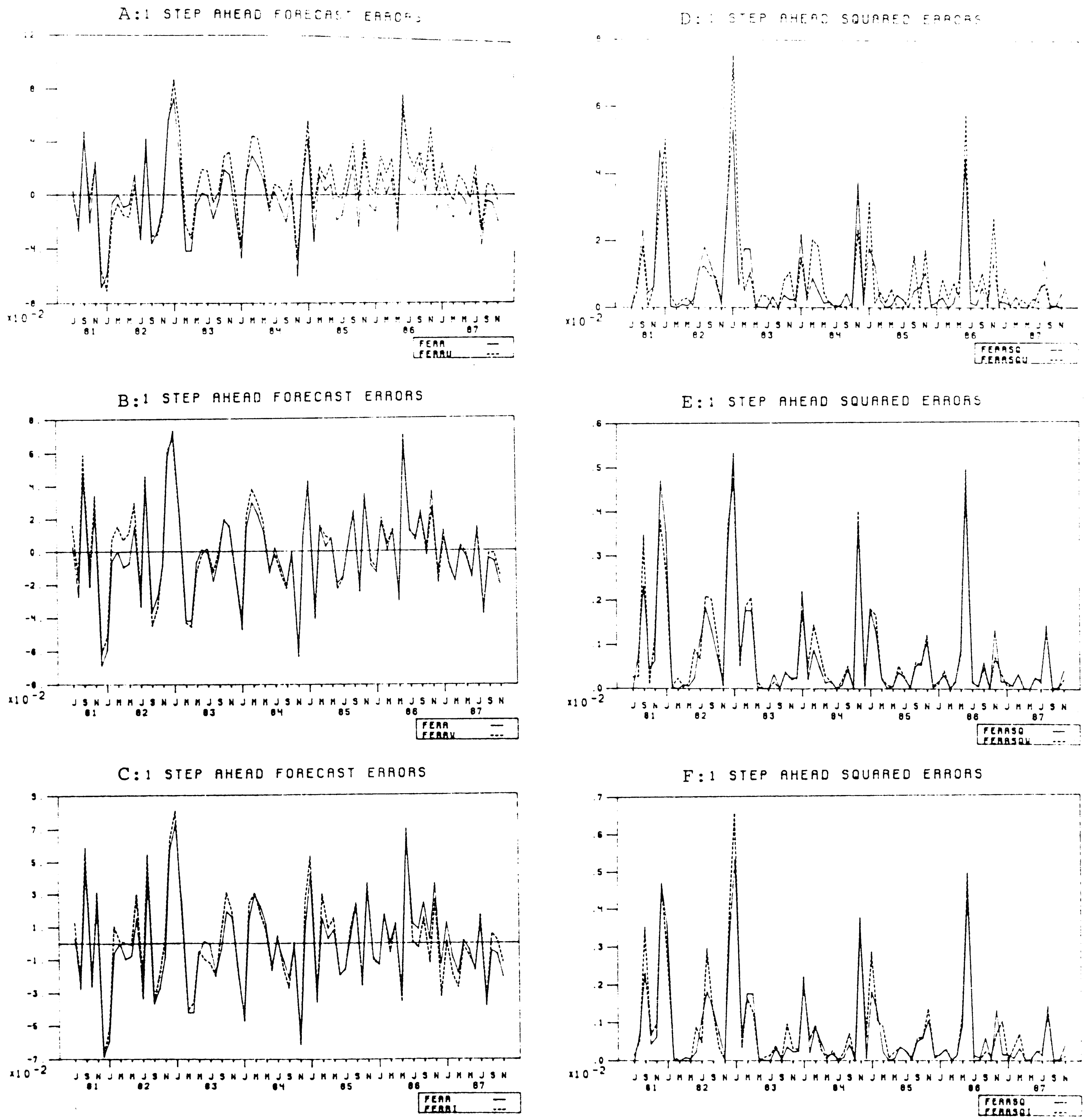


FIGURE 10 (CONTINUED)

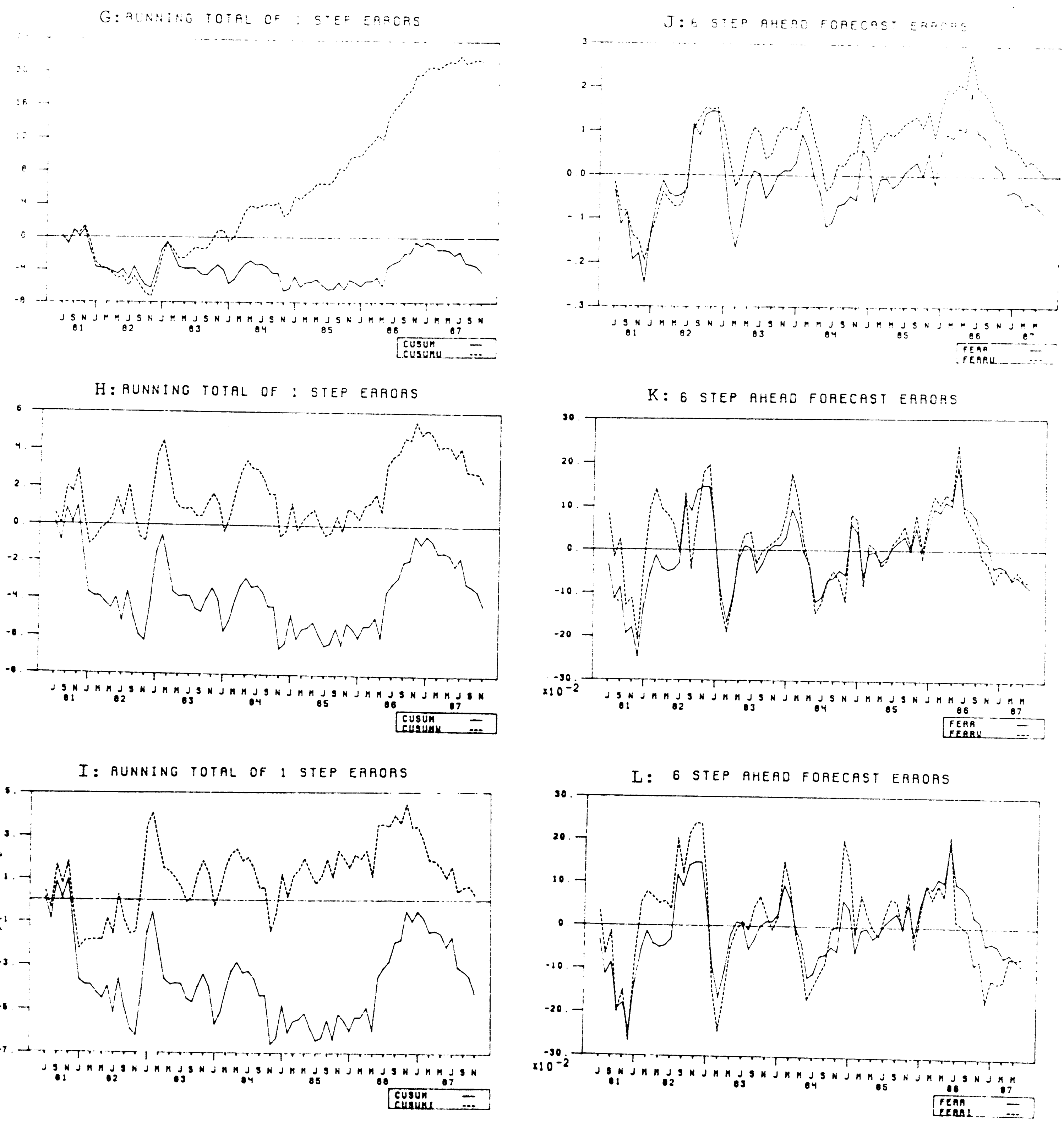


FIGURE 11:

HISTORY OF ERRORS IN FORECASTING WPI

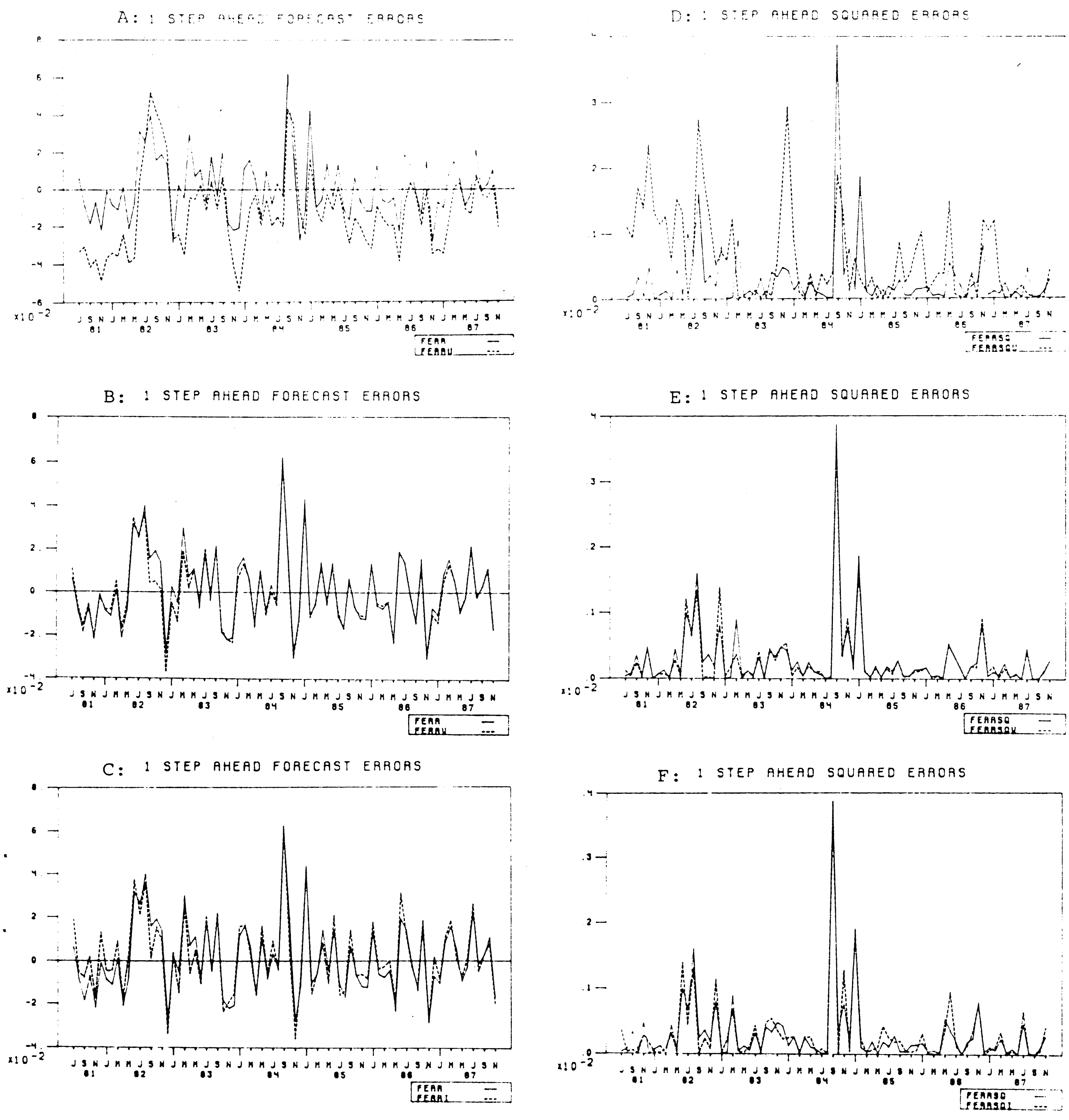


FIGURE 11 (CONTINUED)

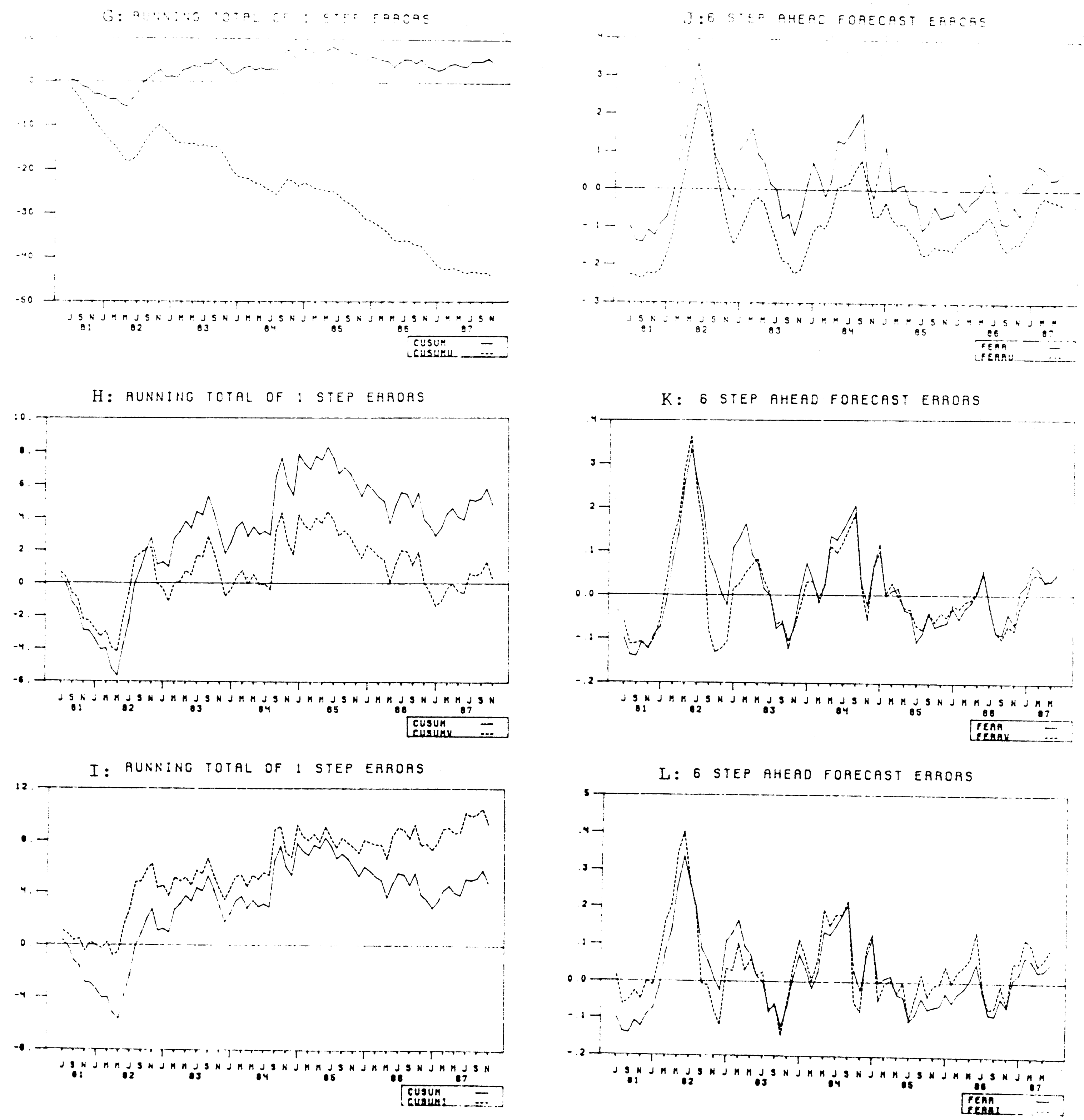


FIGURE 12:

HISTORY OF ERRORS IN FORECASTING XCH

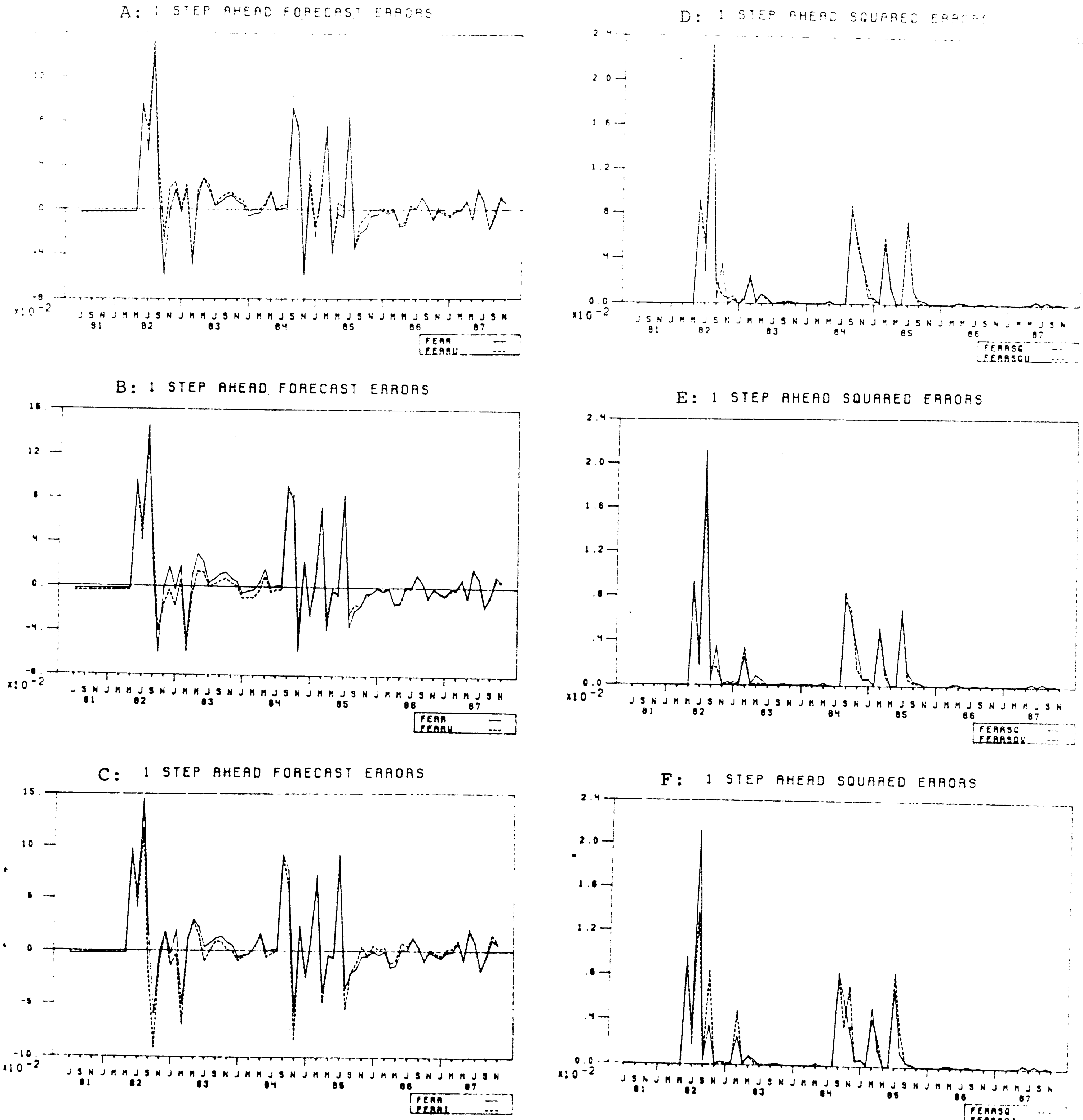


FIGURE 12 (CONTINUED)

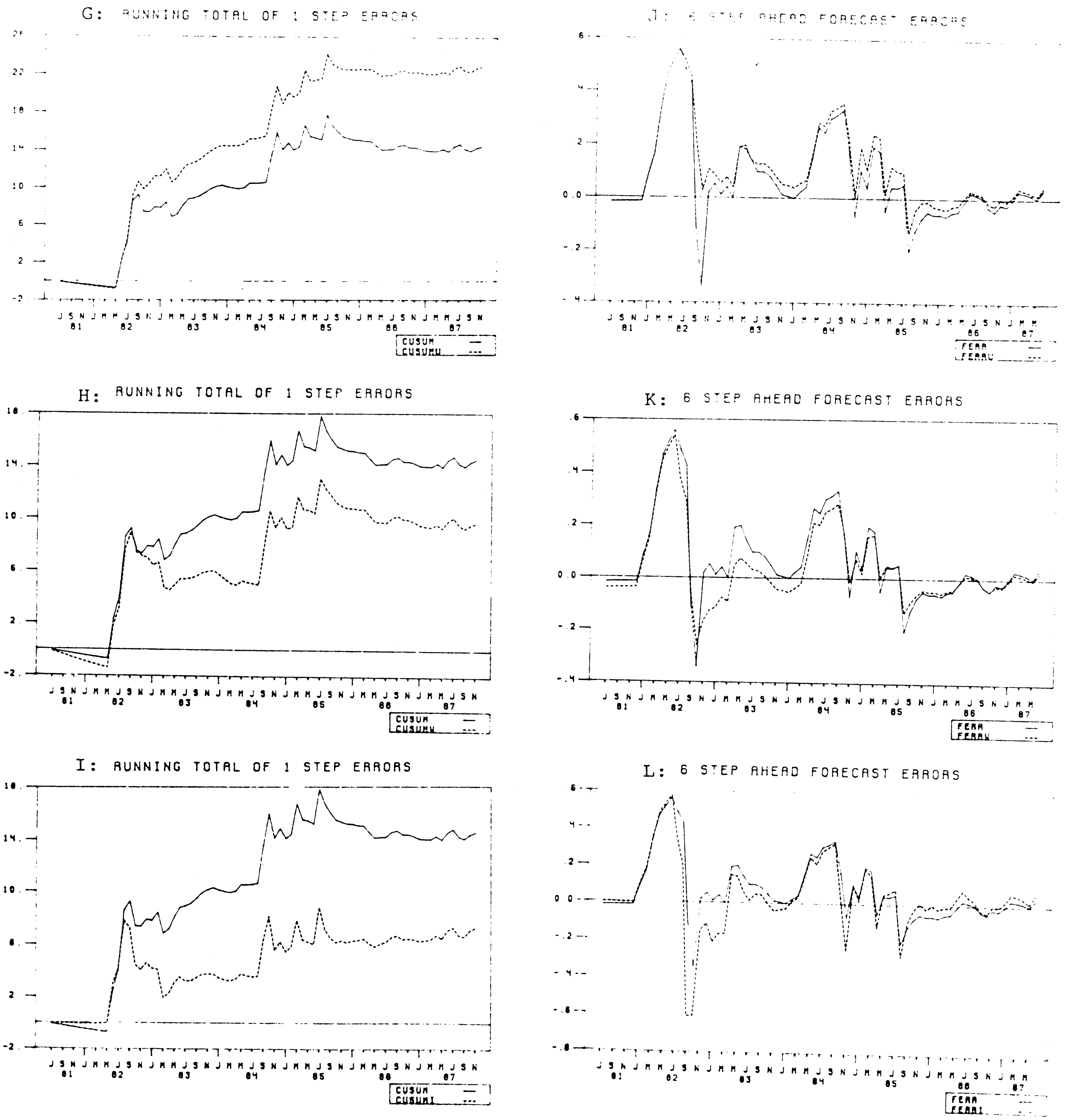


FIGURE 13:

HISTORY OF ERRORS IN FORECASTING KINF

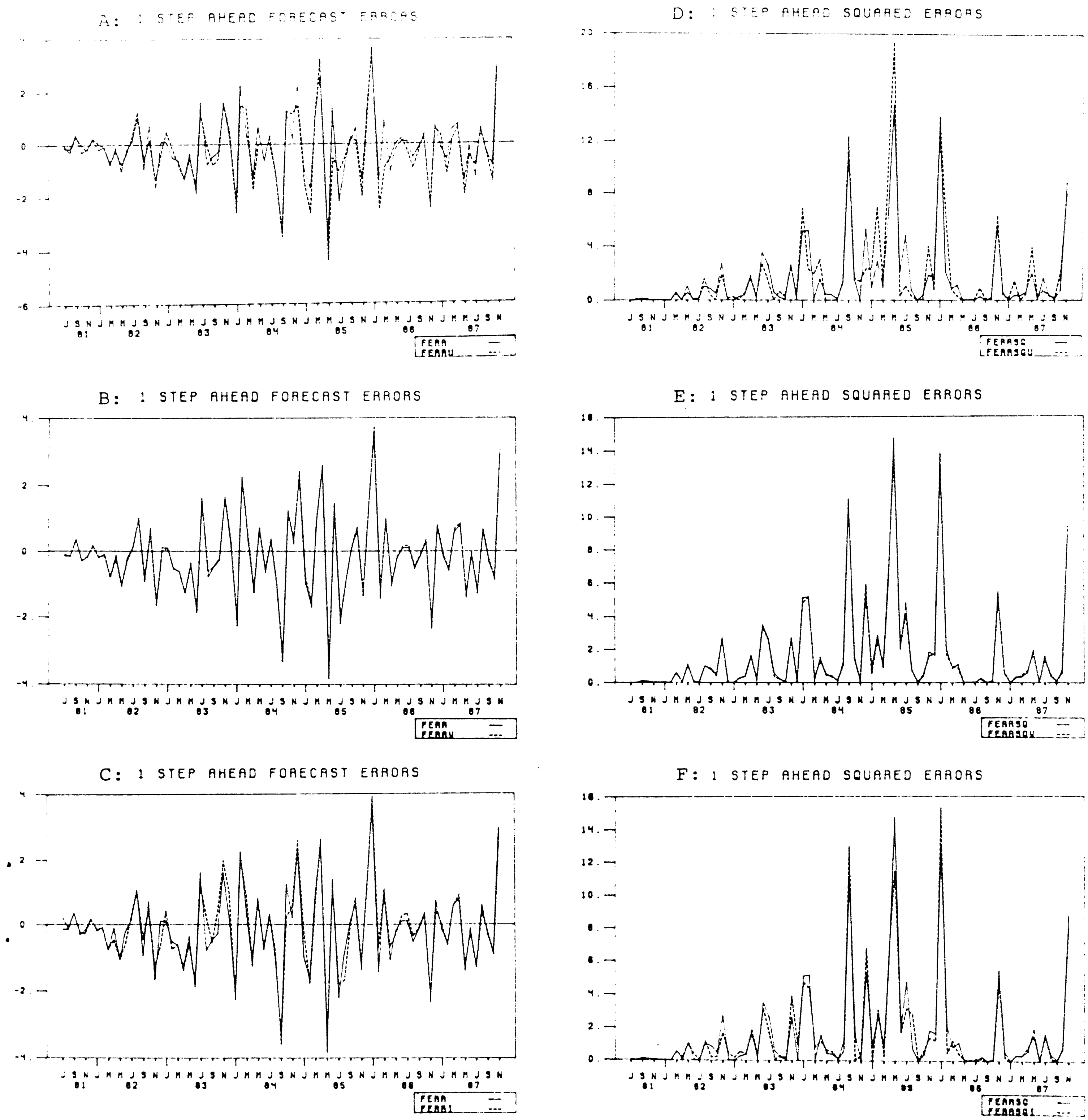


FIGURE 13 (CONTINUED)

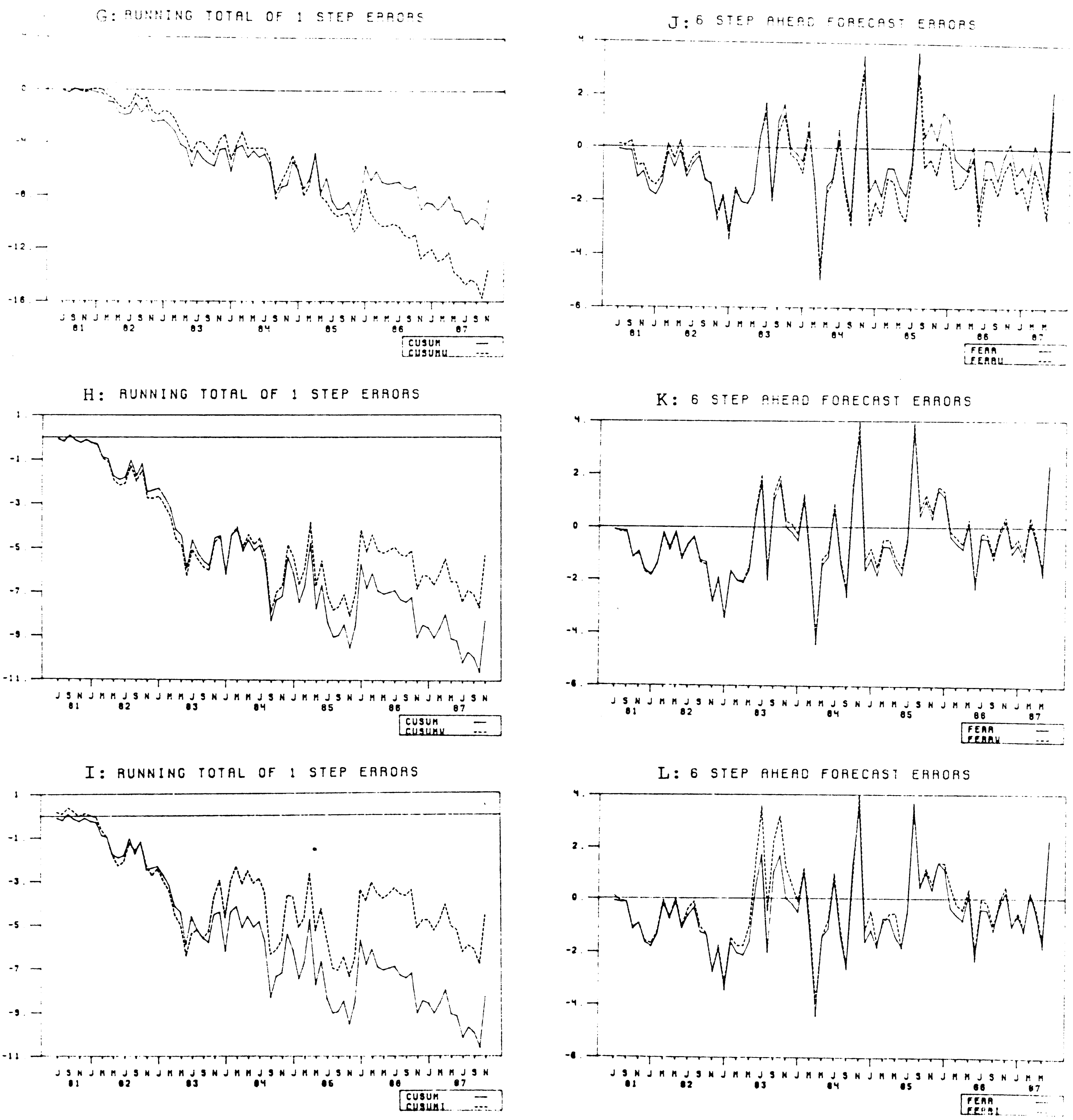


FIGURE 14:

HISTORY OF ERRORS IN FORECASTING DIR

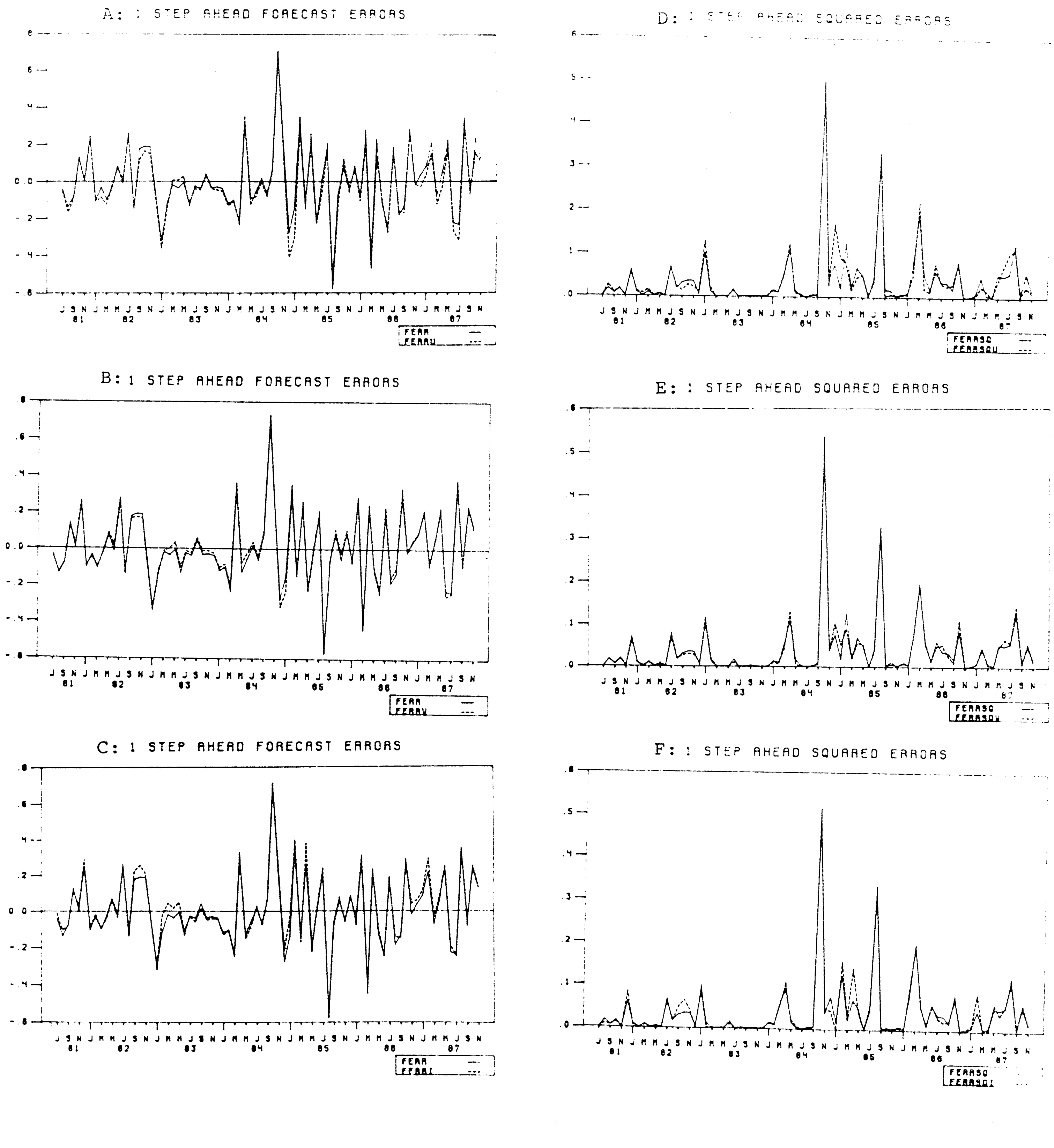


FIGURE 14 (CONTINUED)

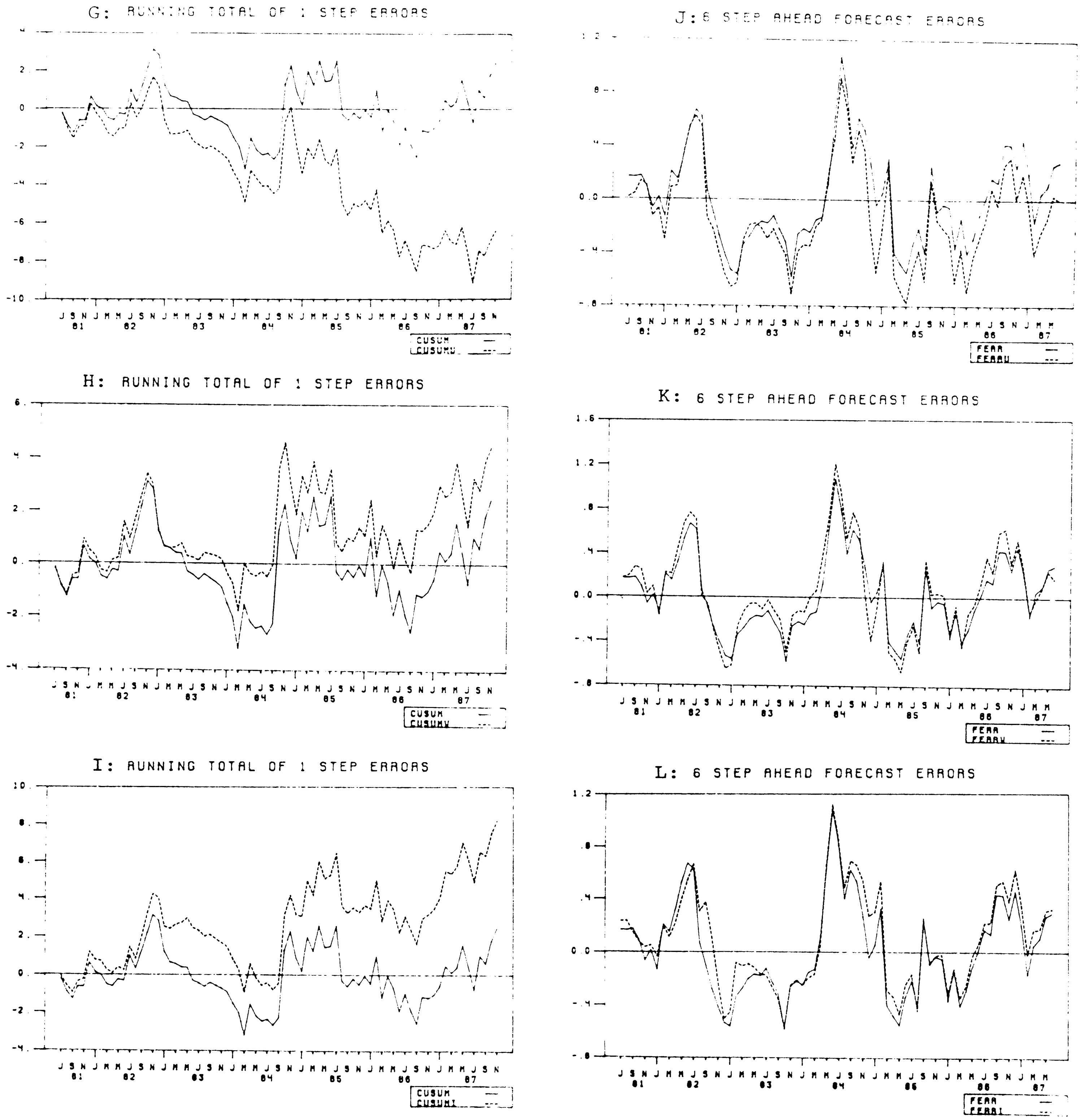


Figure 15: Evolution of Coefficients in the IPINSS Equation

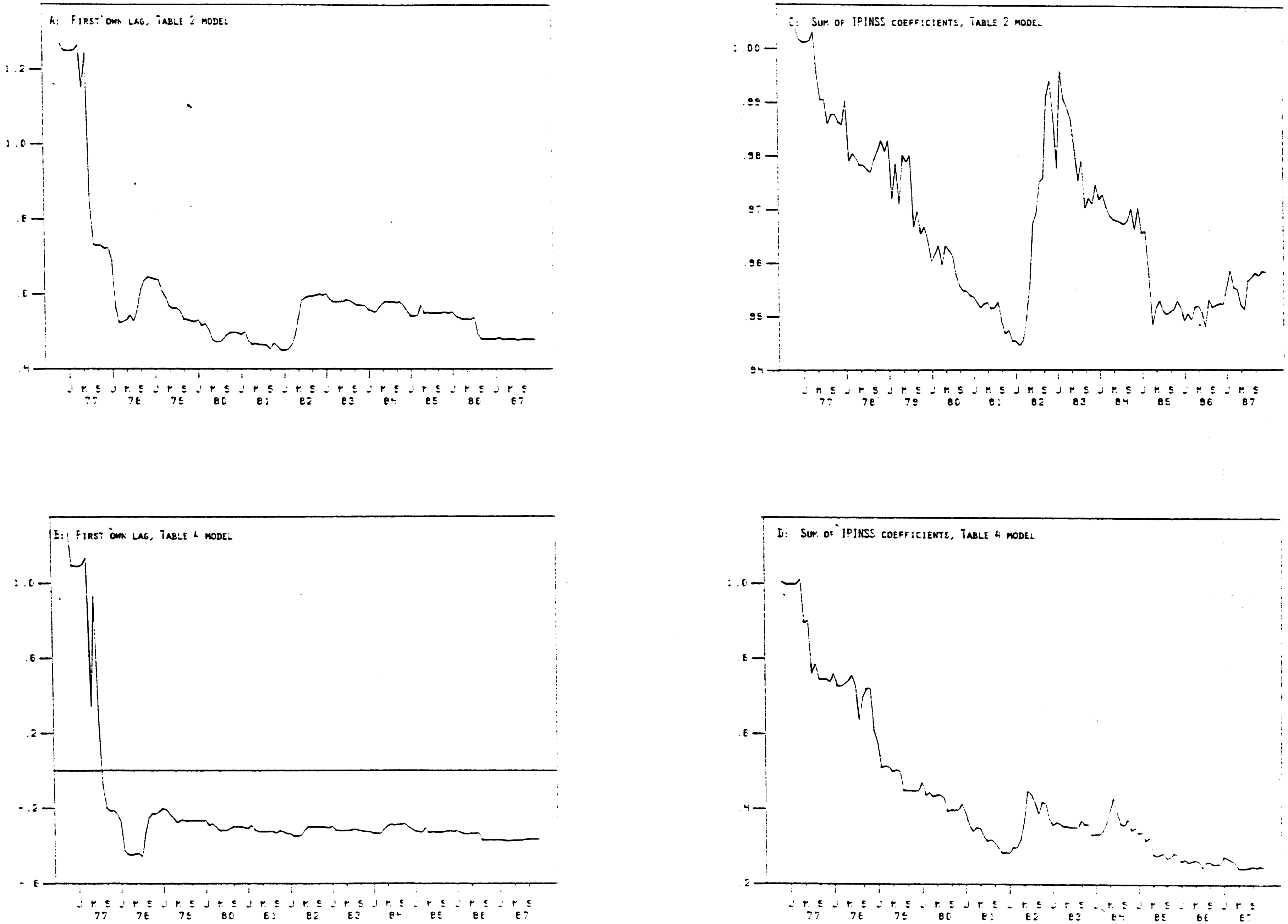


FIGURE 15 (IPINSS EQUATION, CONTINUED)

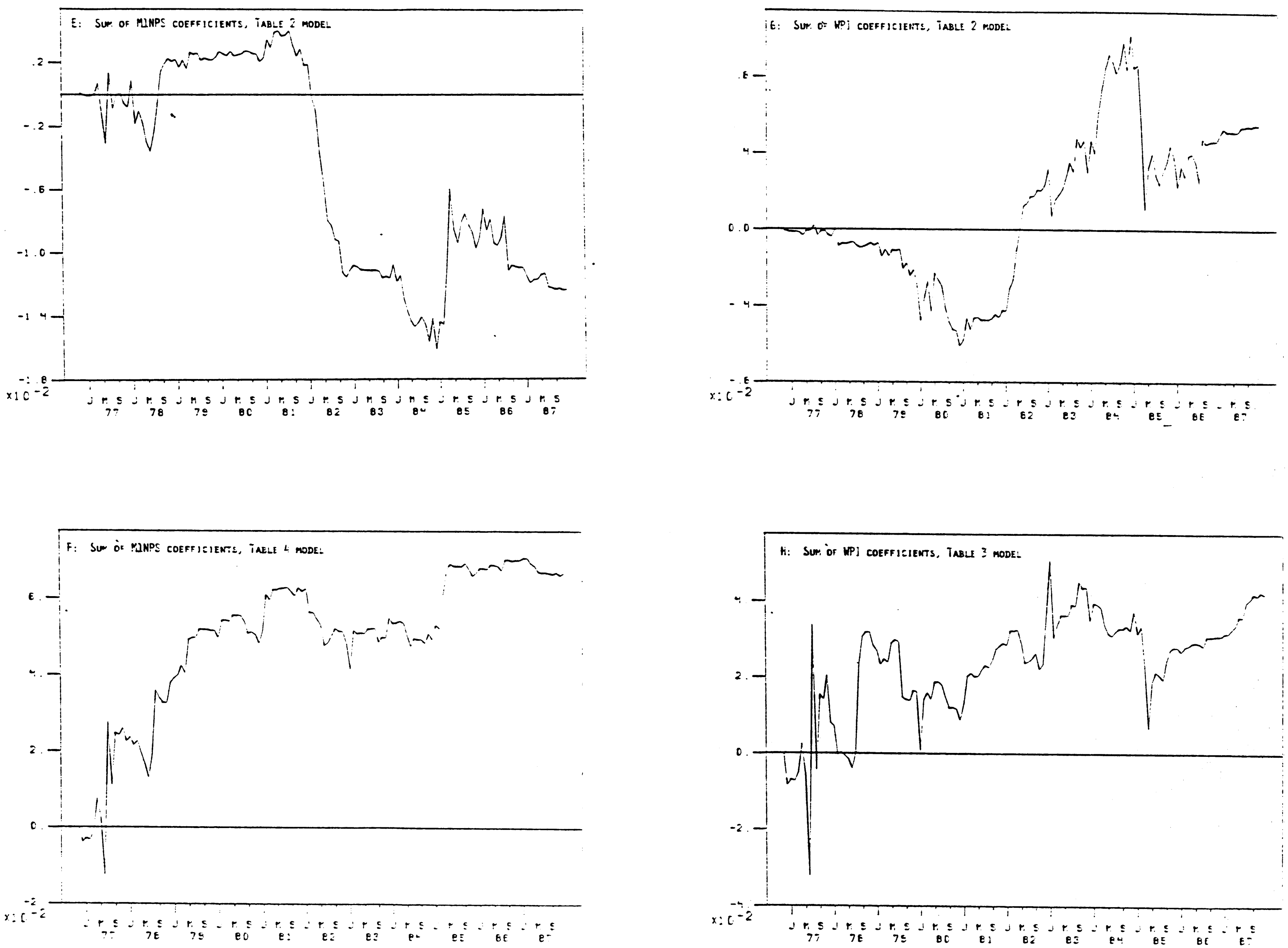


FIGURE 15 (IPINSS EQUATION, CONTINUED)

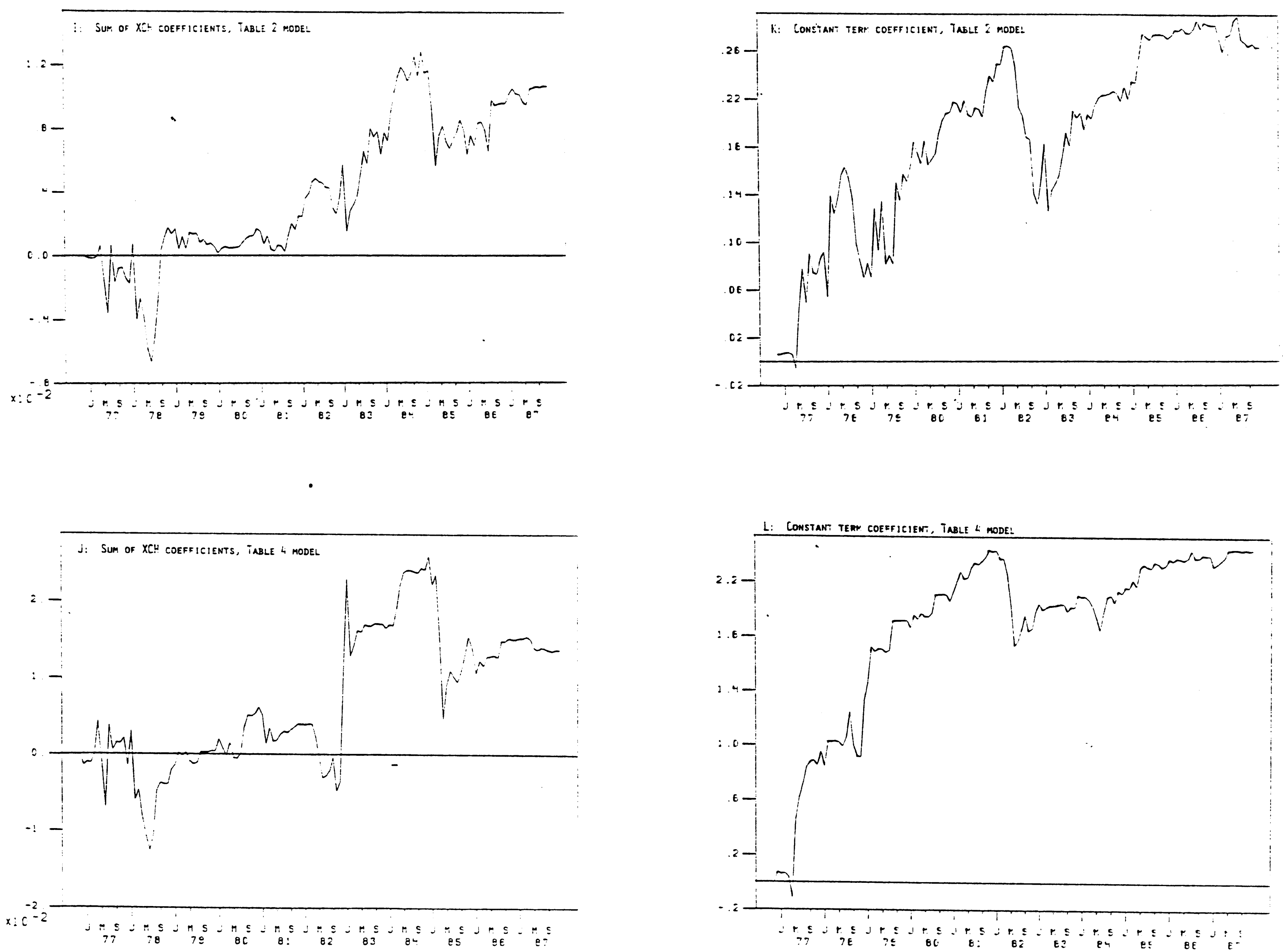


Figure 16: Evolution of Coefficients in the Equation for M1NPS

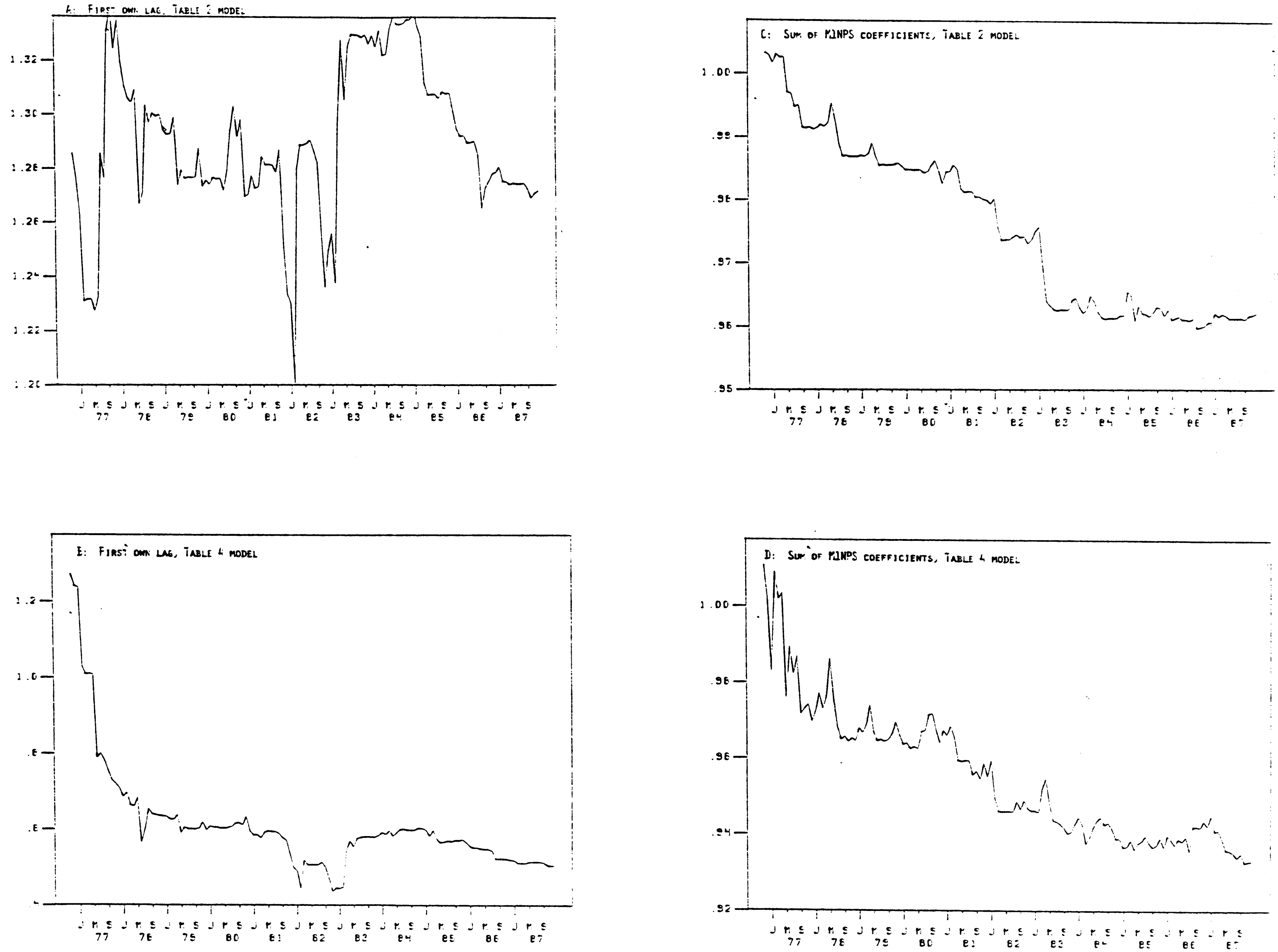
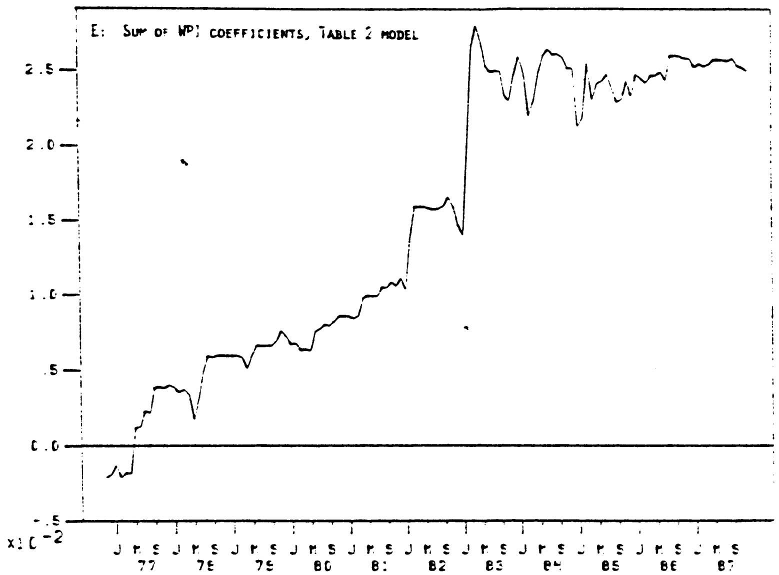
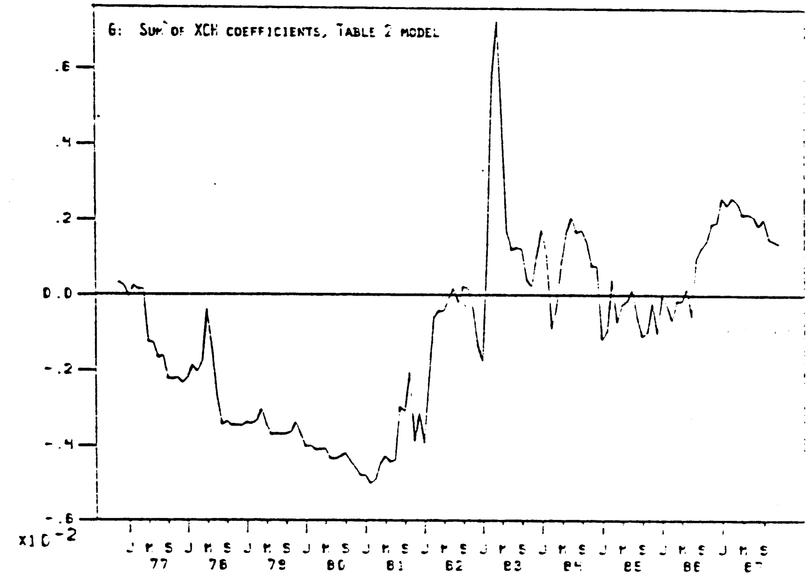


FIGURE 16 (M1NPS EQUATION, CONTINUED)





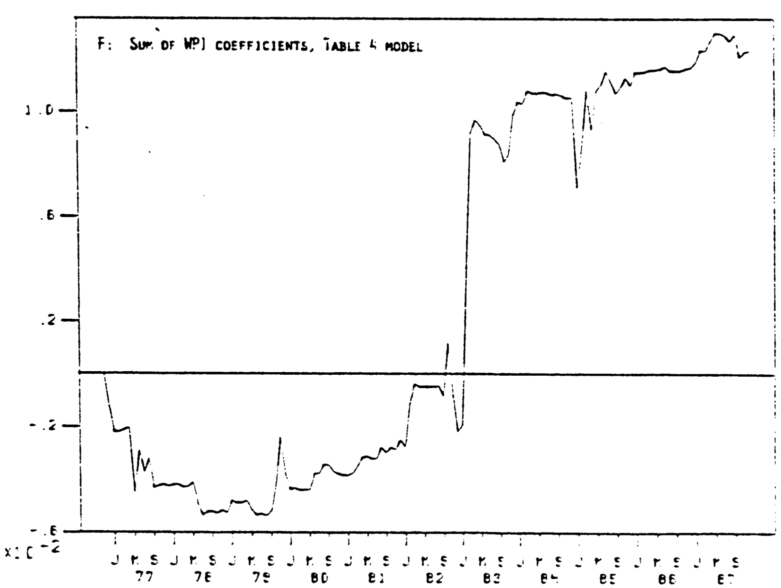


FIGURE 16 (M1NPS EQUATION, CONTINUED)

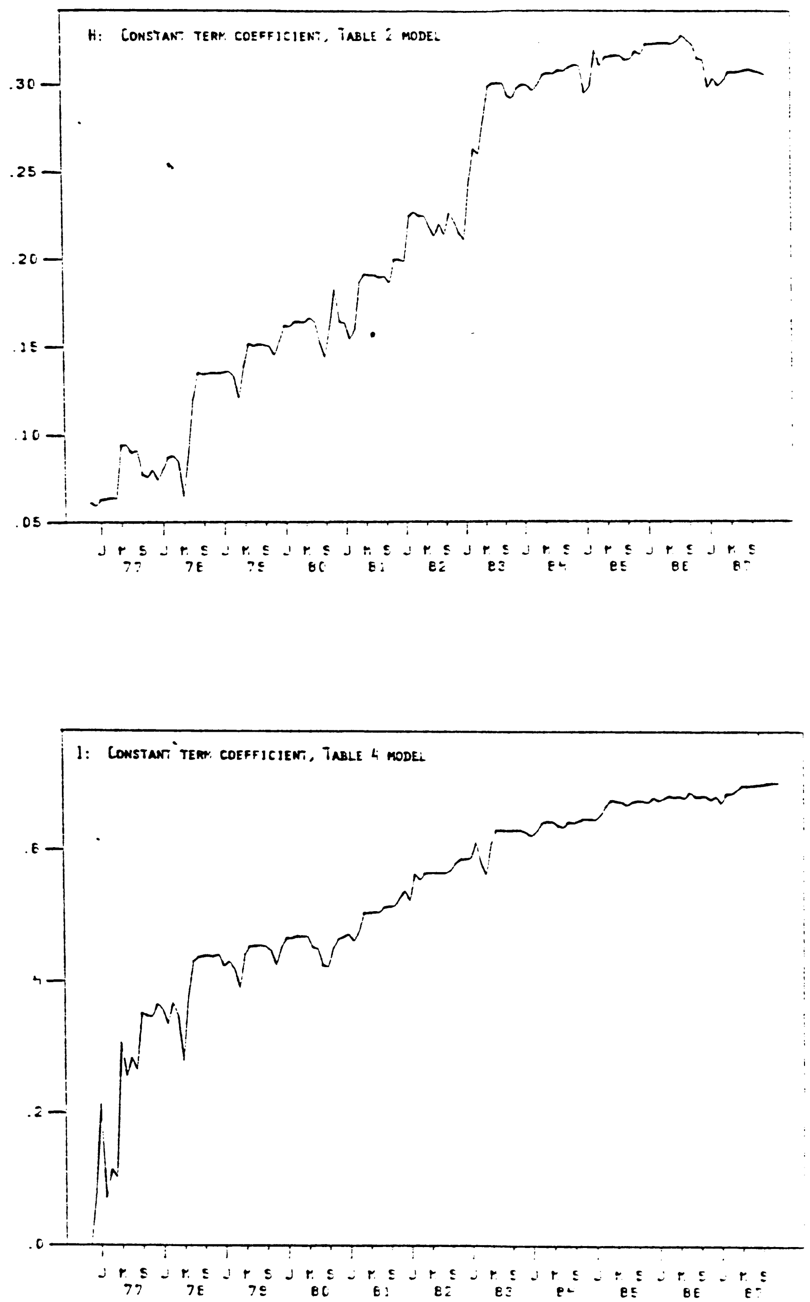


FIGURE 17: EVOLUTION OF COEFFICIENTS IN THE EQUATION FOR WPI

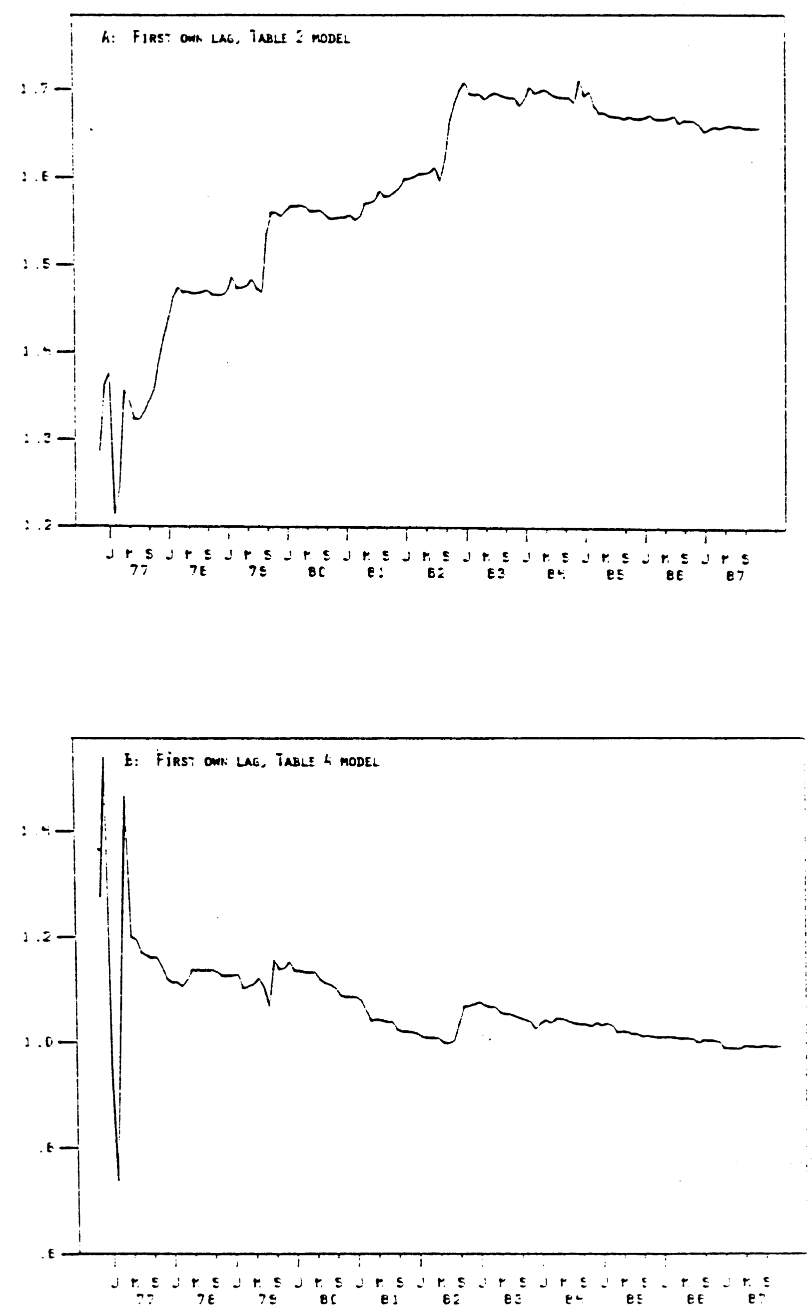
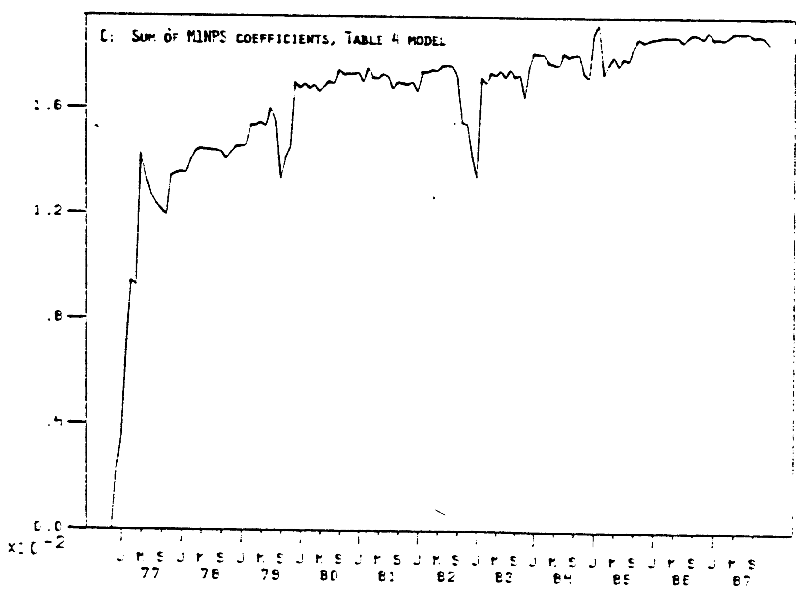


FIGURE 17 (WPI EQUATION, CONTINUED)



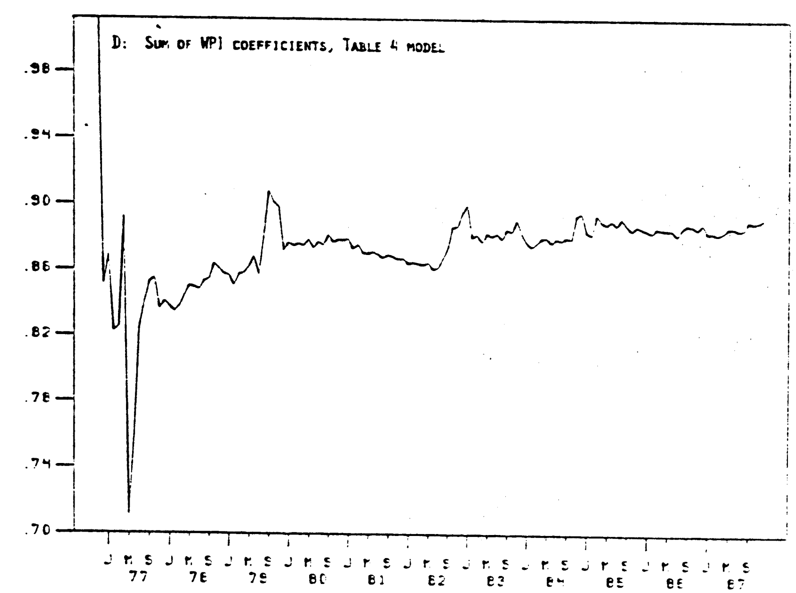


Figure 17 (WPI equation, continued)

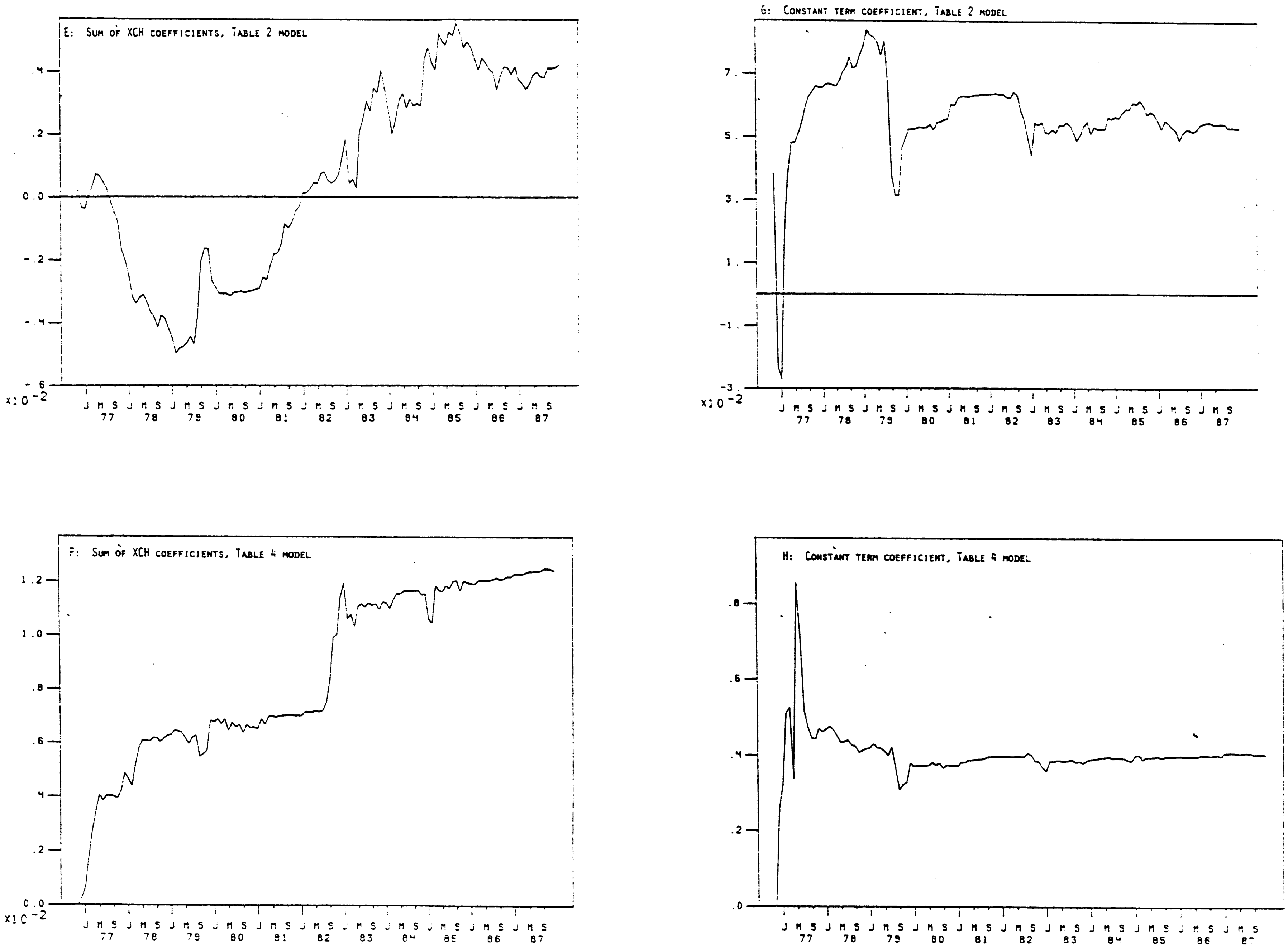


Figure 18: Evolution of the Coefficients in the Equation for XCH

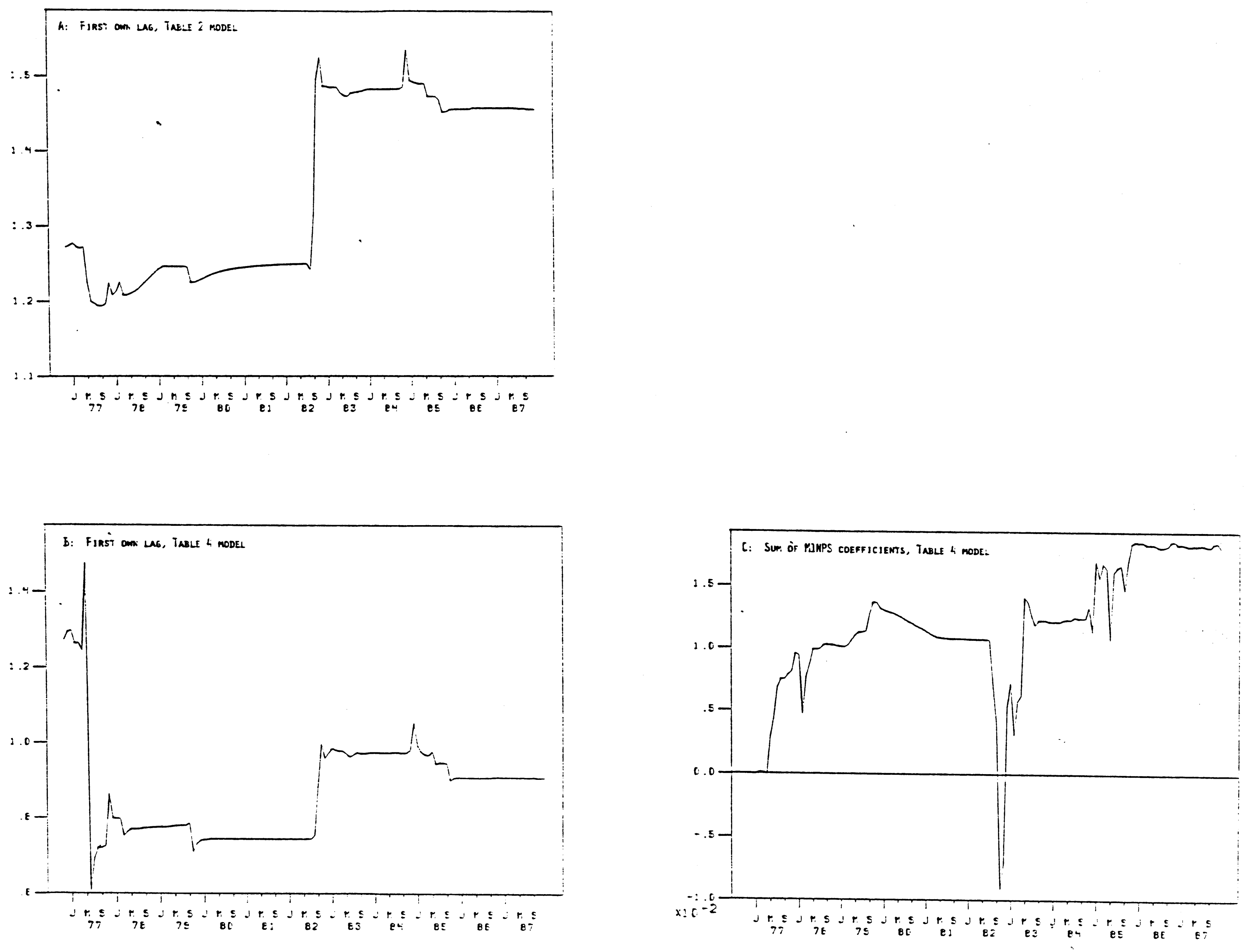
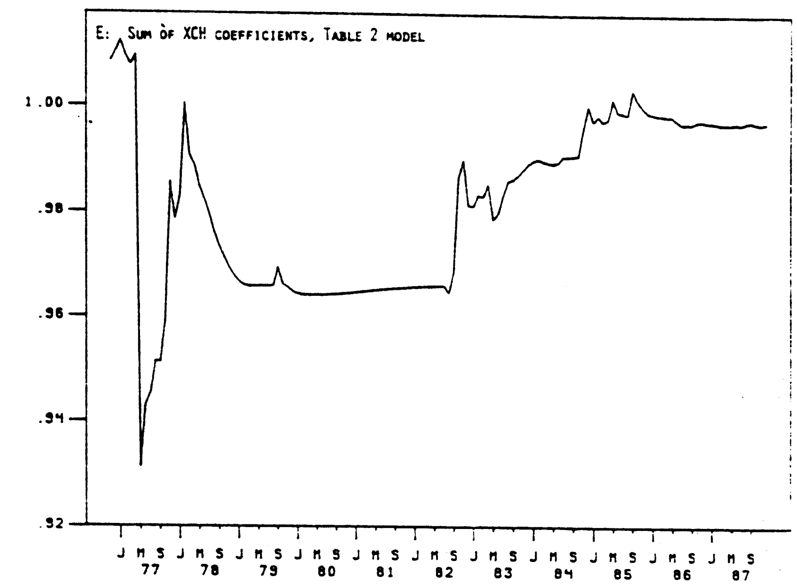
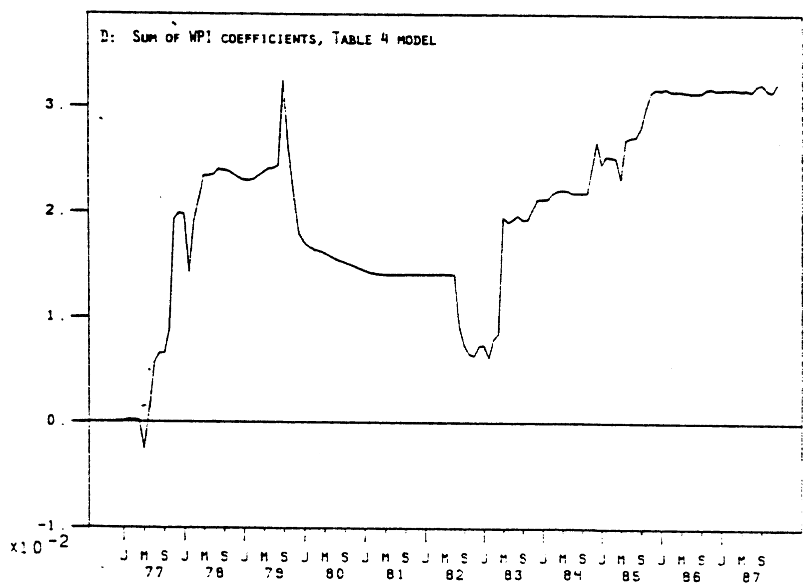


FIGURE 18 (XCH EQUATION, CONTINUED)





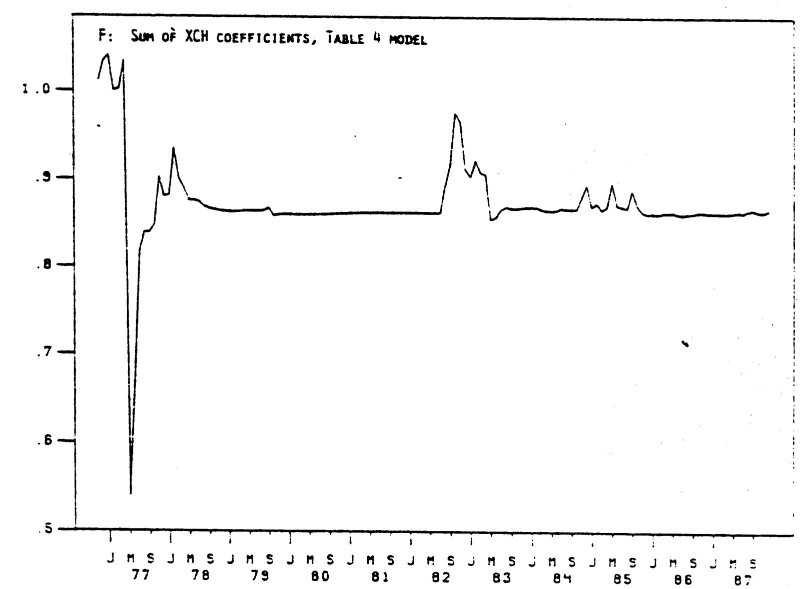
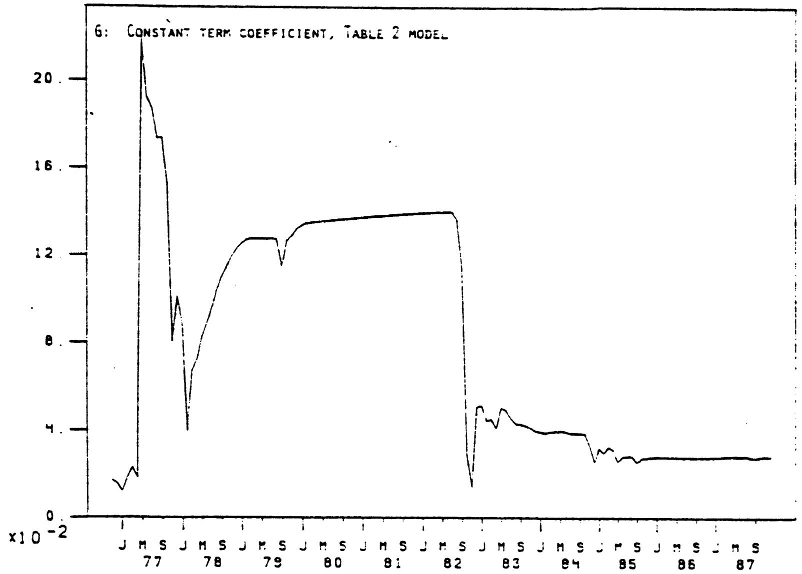


Figure 18 (XCH equation, continued)



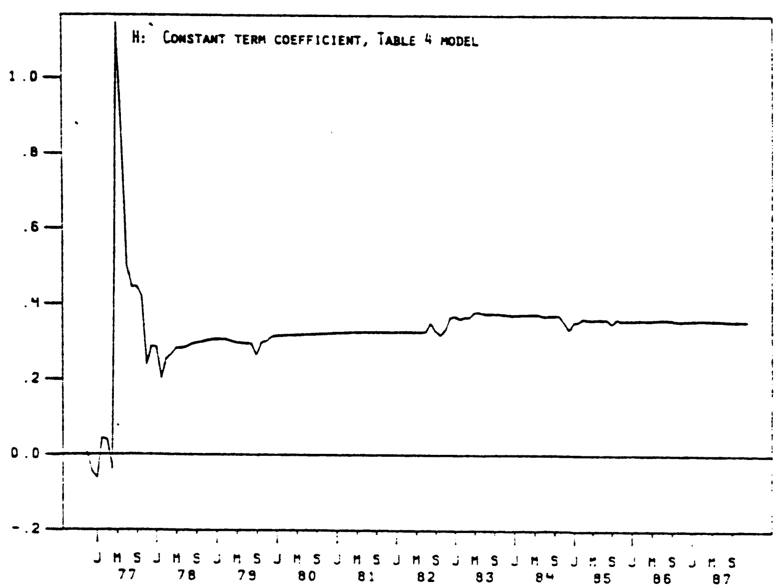
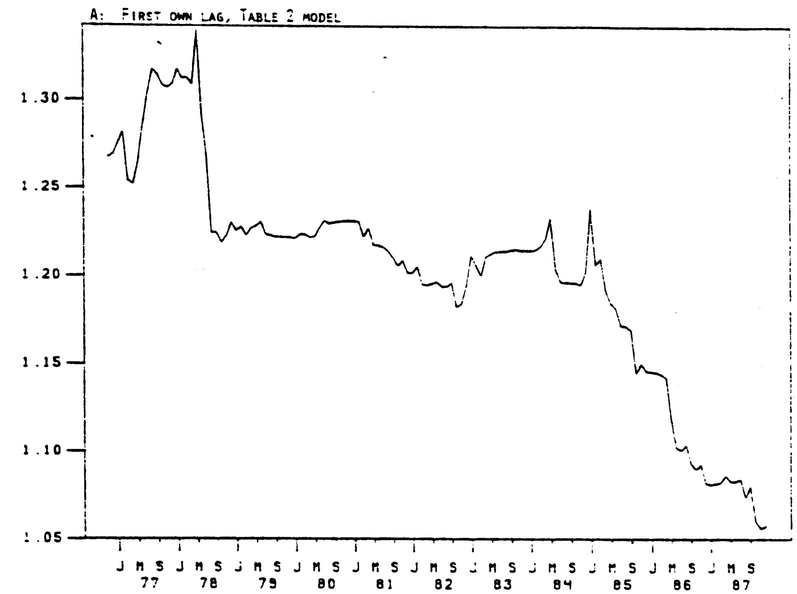


Figure 19: Evolution of Coefficients in the Equation for DIR



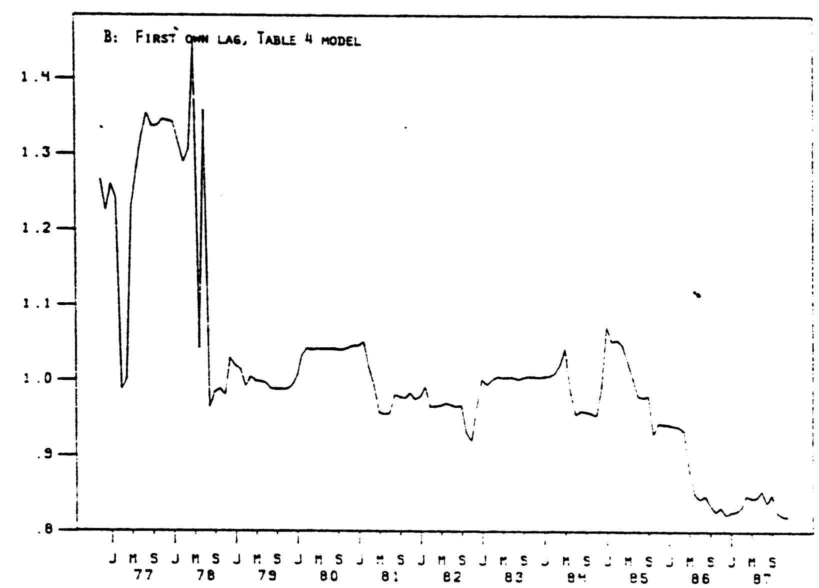


Figure 19 (DIR equation, continued)

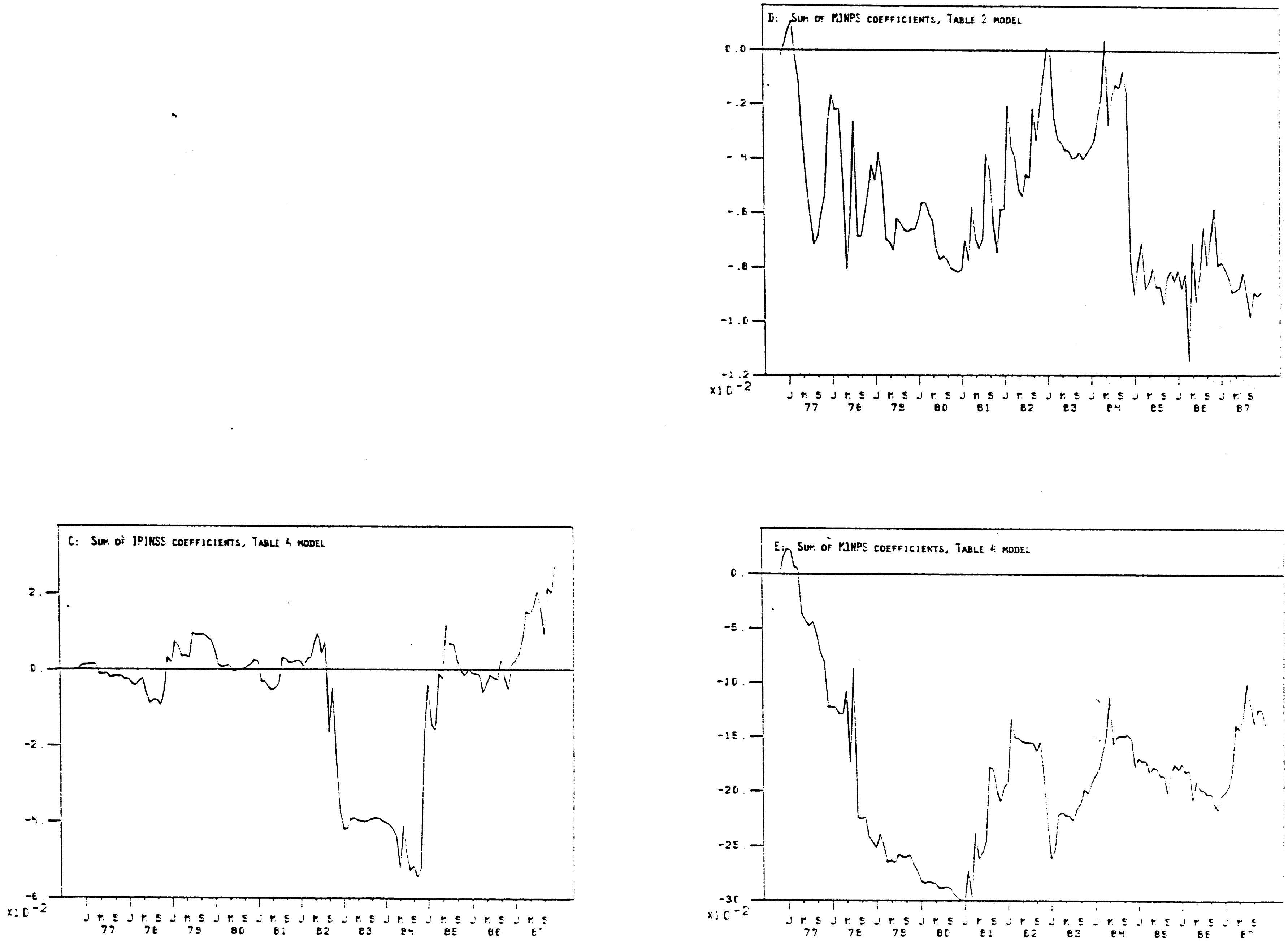
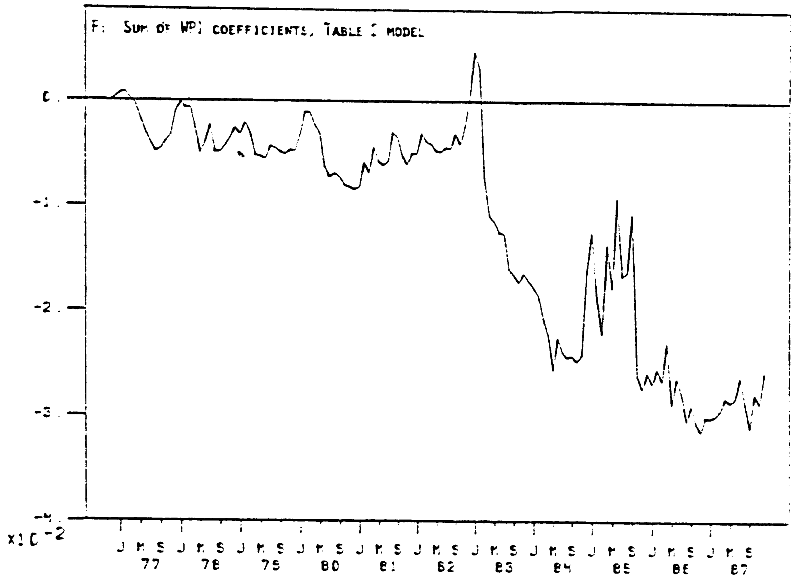
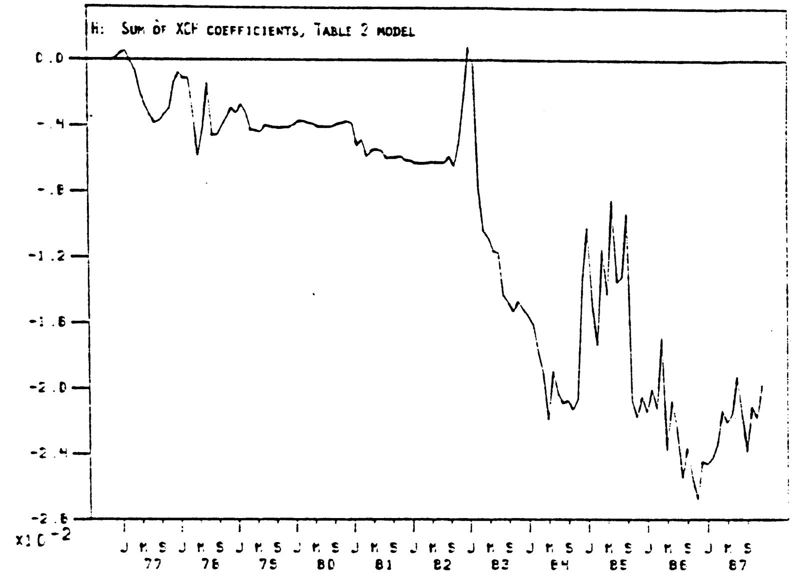
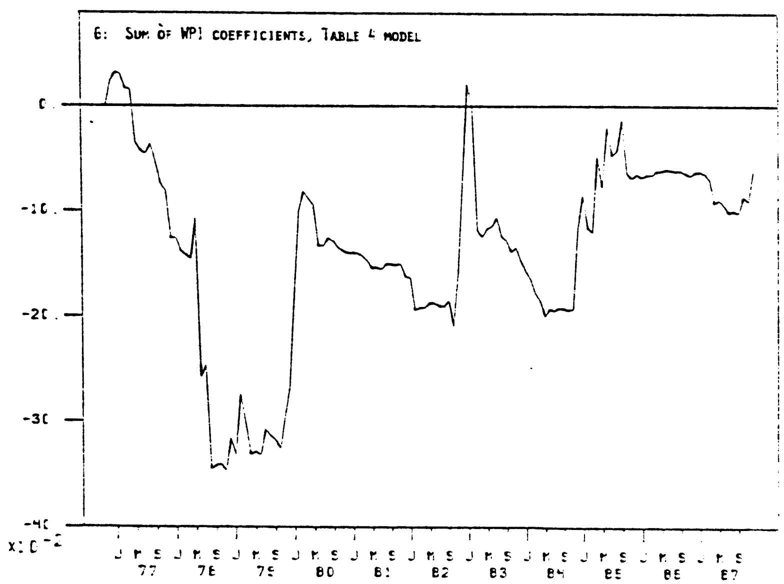


FIGURE 19 (DIR EUQATION, CONTINUED)







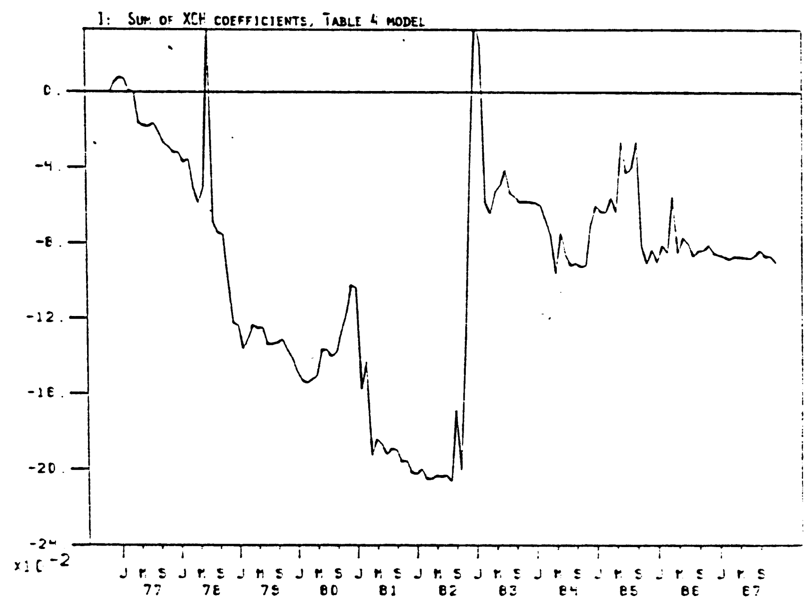


FIGURE 19 (DIR EQUATION, CONTINUED)

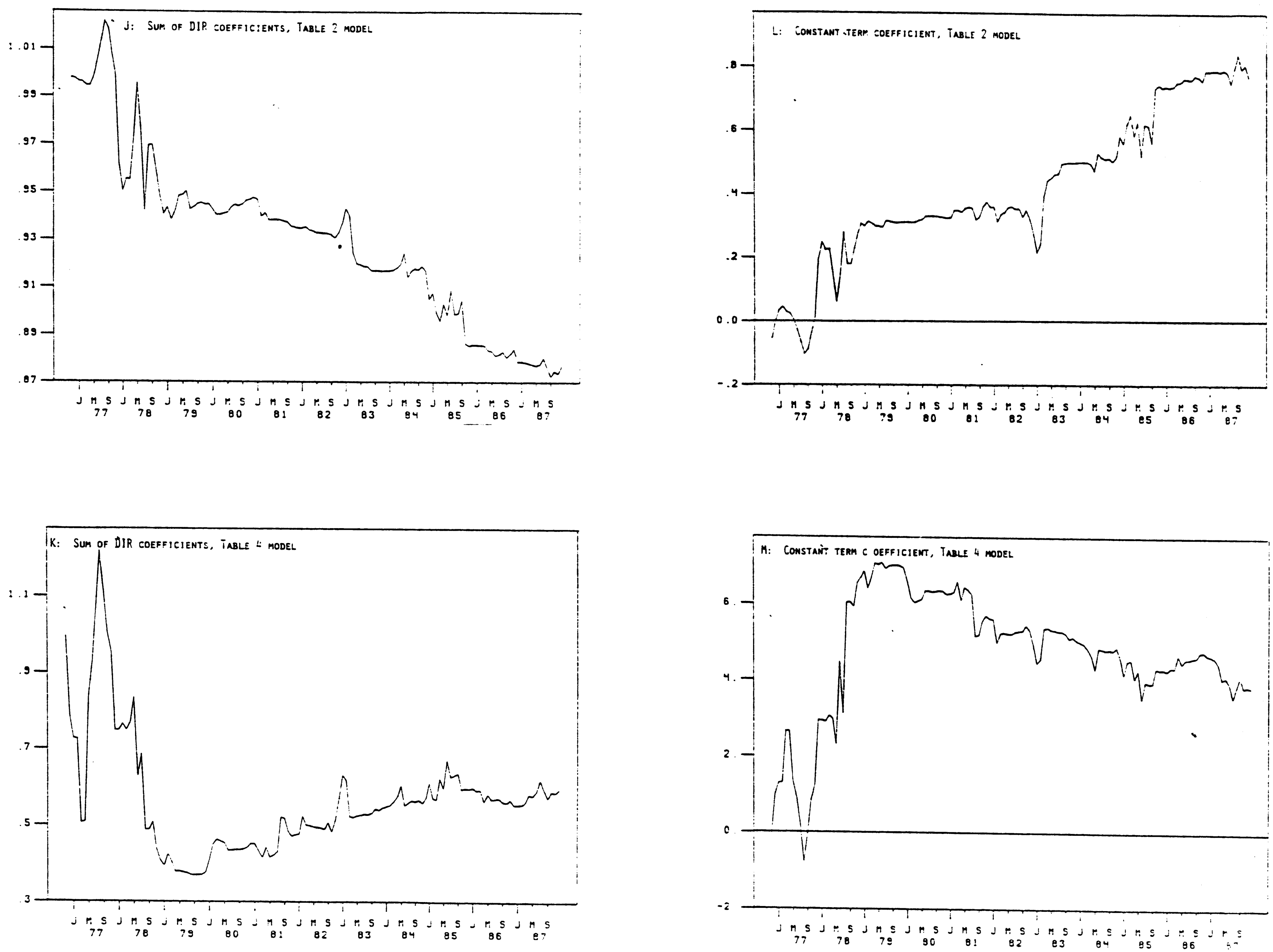


FIGURE 20A: SUCCESSFUL AXIAL SEARCH

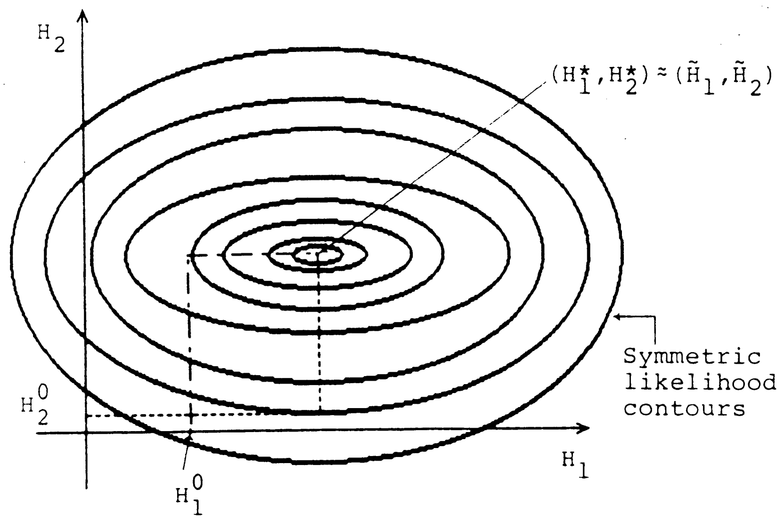
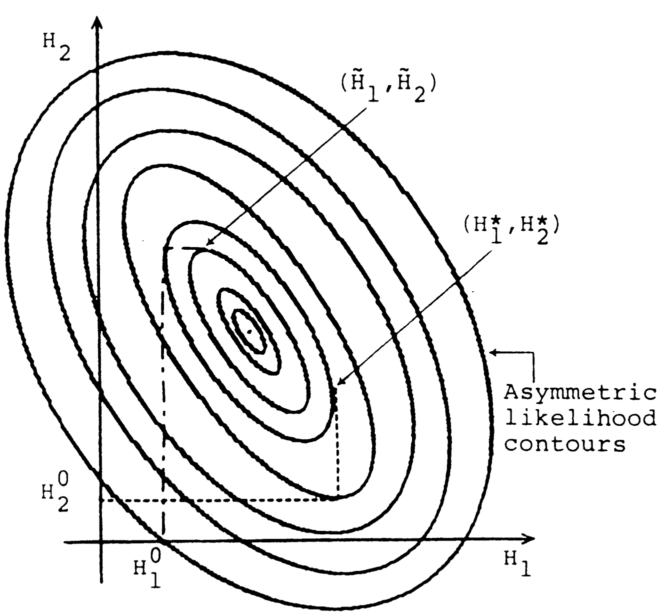


FIGURE 20B: UNSUCCESSFUL AXIAL SEARCH



Explanation: Each graph shows two searches. Each search begins at the initial guess (H_1^0, H_2^0). The search denoted by -------------- and * first optimizes H_1 while fixing H_2 at H_2^0. It then optimizes H_2 while fixing H_1 and H_1^*. The search denoted by — · — · — · — or ˜ first optimizes H_2 (with $H_1 = H_1^0$) and then optimizes H_1 (with $H_2 = \tilde{H}_2$).