Accidents

Working Paper Series

Economics and Econometrics

Analysis of the Car Accident Indexes in Spain: A Multiple Time Series Approach

Antonio Garcia-Ferrer
and
Juan del Hoyo*
Universidad Autónoma de Madrid
Cantoblanco, 28034 Madrid
Spain

March, 1985

Working Paper 85-16

University of Chicago

Graduate School of Business
Department of Economics
H.G.B. Alexander Foundation

Address inquiries and orders to:
H.G.B. Alexander Foundation
Graduate School of Business, University of Chicago
1101 East 58th Street, Chicago, IL 60637
(312) 962-7145

Analysis of the Car Accident Indexes in Spain: A Multiple Time Series Approach

Antonio Garcia-Ferrer
and
Juan del Hoyo*
Universidad Autónoma de Madrid
Cantoblanco, 28034 Madrid
Spain

March, 1985

Working Paper 85-16

Abstract

The main objective of this paper is to specify and estimate a model for the car accident rates in Spain in order to improve input for the decision-making process for insurance companies and provide useful information for traffic authorities. The prediction performance of the model is also analyzed trying to verify the improvement in prediction that takes place when we go to more elaborate models.

---

*Antonio Garcia-Ferrer and Juan del Hoyo are professors of Econometrics at the Universidad Autónoma de Madrid, Cantoblanco, 28034 Madrid, Spain. The research for this paper was supported by a research grant from RUMASA S.A. We are grateful to Arthur Treadway and Arnold Zellner for useful comments on previous drafts of this paper. An earlier version of it was presented at the 1984 European Meeting of the Econometric Society in Madrid.

4. Main Body of Article

1. INTRODUCTION

The main objective of this paper is to specify and estimate a model for automobile accident rates in Spain in order to improve input for the decision-making process for insurance companies and provide useful information for traffic authorities.

An increasing number of automobile accidents not only means a considerable loss of human lives but also important economic costs to society. In Spain, the number of vehicles increased from around 200 thousand in 1950 to close to 11 million in 1982. The number of accidents has grown during these years as a consequence of the greater number and use of cars.

The insurance sector in Spain represented 2.1% of GNP in 1981 and the share of automobile insurance was around 35% of the whole sector. Recent economic results for the insurance companies have been negatively influenced by rapid increases both in medical care and mechanical repair costs, leading to difficult financial conditions for firms in the insurance sector. As a consequence of the increasing divergence between revenues and expenses in recent years, many companies have started to cancel auto insurance policies and concentrate activities in other areas. The share of automobile insurance has decreased from 40.8% in 1970 to 34.1% of insurance companies' sales in 1981.

On examination of insurance company reports, we noticed that increasing mechanical and medical costs are not the only causes of companies' financial difficulties. As a matter of fact, both internal organizational problems and inability to forecast well, are also responsible for the present situation.

The paper is organized as follows. In section 2, we propose a general framework relating the variables in the model. Section 3 discusses the sources, definitions and characteristics of the data. In section 4, we present the empirical results. Starting with the analysis of the data, we proceed first to find univariate and intervention models as starting points to construct transfer functions and multivariate time series models. In section 5, we analyze the predictive performance of the models, and in section 6 we present our conclusions and policy recommendations.

2. GENERAL FRAMEWORK

In this section, we try to analyze systematically the economic relationships that explain the behavior of the automobile insurance sector. Due to the lack of published data on individual companies, we will concentrate our analysis on the aggregate sector, although the analytical framework can also be used for individual firms.

The initial specification of the accident's model is shown in Figure 1. The first assumption is that the number of accidents depends on the use of cars and some exogenous variables such as road and weather conditions, police activity, age and physical condition of automobiles and the legal regulations that control urban and nonurban traffic (e.g., speed limits, degree of enforcement, driver education, etc,). With good forecasting of the number of accidents, insurance companies can obtain reasonable predictions of the costs of accidents and together with revenues from insurance policies determine net sector revenue.

Another basic assumption of our model is that the level of economic activity affects variation in the stock of cars, as well as degree of utilization. Additionally, the use of cars will depend upon certain exogenous variables such as the length and time of vacation periods, the number of holidays during the year, etc. While data on the stock of cars are readily available, we have not been able to find good data for the degree of car utilization. Consequently, we have decided to use the consumption of gasoline as a proxy for degree of utilization.

* (Figure 1)

At this point, our initial theoretical model, very much in the spirit of previous works [e.g., Lave and Weber (1970) and Peltzman (1975)] can be summarized in the following three structural equations:

$$\text{Accidents} \equiv ACC_t = f_1 \text{ (gasoline consumption, economic activity, } Z_1)$$

$$\text{Stock of cars} \equiv VEH_t = f_2 \text{ (}VEH_{t-1}\text{, new purchases, number of scrapped cars)}$$

$$\text{Gasoline consumption} \equiv GAS_t = f_3 \text{ (gasoline prices, income, } Z_3)$$

where Z_1 and Z_3 represents exogenous variables not used in the empirical implementation of the model (because of the lack of monthly data) but worth noting. For instance, Z_1 represents factors such as driving speeds, driver ages and alchohol consumption which have long been mentioned as leading contributors to highway fatalities [e.g., Peltzman (1975) and DGT(1981)]. Some additional factors are also worth noting. These included the average age of cars, the ratio of new cars to all cars (because it has been suggested that while drivers familiarize themselves with their new cars, accident risk may increase), traffic density, expenditures on roads, the ratio of imports to total cars (because there is evidence that small cars are more lethal than large cars if an accident occurs), education of the population, and the availability of hospital care (which might reduce deaths if injury occurs). As have been argued in other energy studies [Holmes (1975) and Tsurumi (1980) among others], Z_3 represents a score of factors ranging from engine-tuning to driving habits that can be considered to be stable in the short run since they are determined by habits. On the other hand, one may argue that the threat of the gasoline shortage during the oil crisis and the large increase in the gasoline price has made drivers conscious of maintaining their cars' fuel efficiency more than before through frequent engine tuning and tire changing.

Initially, we considered a classification of accidents by broad groups such as urban and non-urban, with and without victims, according to the driver's age, etc. These classifications of accidents are very important if we want to verify whether or not the explanation of each type of accident has certain variables in common. Unfortunately, the length and disaggregation of some of the time series needed to perform these tests were unavailable, and the present empirical analysis is carried using just aggregate data.

Modelling the variables of the automobile insurance sector is not an easy task. Although companies are familiar with certain facts relating to automobile accident rates, e.g., accident rates are greater in the north of the country, the effects of drivers' ages on rates, etc., there are many other factors that considerably complicate the analysis. Some of these are:

i) Strong seasonality is present in accident rates which must be taken into account in formulating a dynamic model and in forecasting. Seasonality is further complicated by the presence of "moving" holidays that change from year to year.

ii) Seasonality is sometimes "obscured" by the lack of coincidence of the time the accident is reported to the company, and the date of a car's repair. In many situations, the dates differ by up to six or seven months.

iii) There is a potential influence of a large number of factors. For instance, according to the companies' experts, 1981 was an exceptionally good year (few accidents) because of low rainfall. Thus rainfall, other weather conditions, driver's experience and ability levels, etc., may affect accident rates.

3. DESCRIPTION OF THE DATA

The need for good data is the first requirement if we want to study some of the effects mentioned earlier. At the same time, it is convenient to work with a large number of observations in order to help assure that the statistical inferences will be reasonably accurate. Strictly speaking, some of the hypotheses we want to test would require data on a weekly and daily basis. Unfortunately, we have only been able to collect data on a monthly basis at an aggregate level. For some important variables such as automobile and road conditions, the lack of time series data has forced us to omit them from the analysis. While in this case the omission of these variables will not likely produce changes in the direction of "causality" [e.g., Lütkepohl (1982)] and the resulting models are still useful for forecasting purposes, we are aware that superior forecasts may be obtained from a higher dimensional system.

The data for the variables used in the empirical implementation of the model have been included in four groups:

i) Variables related to accident rates [Monthly data from January 1962, to December, 1981, (240 observations)]

Number of accidents with injured passengers: ACC

Number of accidents with fatalities: ACCD

Number of injured persons in accidents: VIC

Number of fatalities in accidents: DTHS

ii) Variables representing the stock of vehicles and its variations [Monthly Data from January, 1962, to December, 1981, (240 observations)].

Stock (number) of vehicles: VEH

New registration of vehicles: NUVE

iii) Variables representing the evolution of economic activity:

Index of Industrial Production (April, 1972 = 100): IIP [January, 1965 - December, 1981, (204 observations)].

Gasoline Consumption in millions of litres: GAS [January, 1969 - December, 1981, (156 observations)].

Rate of change of nominal gasoline prices: PRIGAS

iv) Variables representing exogenous factors:

Number of litres of rain per squared meter: RAIN [January, 1965 - June, 1982, (186 observations)].

Plots of the main variables in the model (ACC, GAS, VEH, NUVE, and IIP) are shown in figures 2 to 6. With the exception of the stock of vehicles, the remaining variables in the model are flow variables showing nonstationarity as well as strong seasonality. In most cases, stationarity is achieved by appropriate seasonal and nonseasonal differencing by use of the $(1-B)(1-B^{12})$ operator.

* (Figures 2 to 6)

4. EMPIRICAL RESULTS

4.1 Univariate and Intervention Models

The main object of this section is to obtain dynamic representations of individual variables in the model using only their own past history, through the information provided by the autocorrelation (acf) and partial autocorrelation (pacf) functions, using the well known approach described in Box and Jenkins (1976). We shall also use univariate models to provide a valuable tool for screening data during the early stages of our analysis and for taking appropriate action if abnormal events produce large residuals whose causes can be identified.

The identification and estimation results for the univariate models are shown in Table 1 (absolute values of t ratios are in parenthesis). In spite of apparent differences in the resulting models of the variables related to accident rates, these models are very similar. As a matter of fact, a more parsimonious MA representation of models (1) and (3) can be obtained by noting that the non-seasonal AR(3) could be approximated by expanding an MA(1) [Garcia-Ferrer and del Hoyo (1983)]. When the estimation period for those models is restricted from 1/69 to 12/81 (to be coincident with the data base of the GAS variable), the MA(1) $MA_{12}(1)$ becomes an obvious parametrization.

With the exception of the VEH and RAIN variables, the plots of residuals of the estimated univariate models showed the presence of large outliers as the consequence of "moving" Easter holidays between March and April, as well as certain abnormal effects that can be easily identified. By defining a dummy variable EAST in the following way:

$$EAST_t = \begin{cases} 1 & \text{the month Easter holiday takes place} \\ 0 & \text{otherwise} \end{cases}$$

we reestimated models (1) to (8) using <u>intervention analysis</u> as suggested by Box and Tiao (1975). The results are shown in Table 2. For the DTHS variable, the estimated residuals of model (4) show the well known Easter effect and another high value corresponding to the month of July in 1978, when the Alfaques camping accident took place. Consequently, in that equation, a new dummy variable is defined such that:

$$ALF_t = \begin{cases} 1 & \text{in July, 1978} \\ 0 & \text{otherwise} \end{cases}$$

From the results in Tables 1 and 2 it is seen that the algebraic signs of the coefficient estimates are reasonable. The coefficients of the dummy variables are significantly different from zero at conventional significance levels. Also there is a considerable reduction in the residual variance in all models in Table 2 as compared with Table 1. For some variables such as VIC or DTHS, the reduction is remarkable.

* (Tables 1 and 2)

4.2 Transfer Function Models

The explanatory ability and predictive performance of the univariate models may be improved by using additional information from our theoretical model presented in Section 2. If we are able to identify and estimate the dynamic relationships among the variables in the model, it is likely that our modelling results can be improved. Therefore, the main objective of this section is to obtain dynamic transfer function models for the accident rate variable and some of the variables included in the initial specification of the model. The cross-correlation functions (ccf) between the univariate models' residual series after prewhitening will be used to identify transfer function models.

4.2.1 Dynamic Relationship between ACC and ACCD

Having in mind that the definition of "fatality" in our data is deaths occuring 24 hours after the accident, it is not surprising that the transfer function identification using the ccf indicated an instantaneous relationship between ACC and ACCD, with ACC the explanatory variable and ACCD the dependent variable.

The transfer function estimated using the backasting method is:

$$\nabla\nabla_{12} \ln ACCD_t = \underset{[10.4]}{0.951}\, \nabla\nabla_{12} \ln ACC_t + (1 - \underset{[26.3]}{0.875}B)(1 - \underset{[45.1]}{0.896}B^{12})\, a_t \quad (10)$$

$T = 240,\ \hat{\sigma}^2_a = 4.96 \times 10^{-3},\ Q(12) = 11,\ Q(24) = 28,\ Q(36) = 31$

where Q(.) represent the Ljung and Box statistics. First, note that the EAST variable disappears from equation (10), as a result of the fact that the Easter effect is present in both variables. Second, the reduction of the residual variance with respect to equations (2) and (2b) is 42% and 32% respectively, implying a considerable improvement in the explanatory power of the model. Third, equation (10) provides an estimate of the elasticity between ACC and ACCD. It is positive and less than one indicating a "less" than proportional increase in ACCD as ACC increases. Note, however that we cannot reject the hypothesis unitary elasticity at 5% level.

Information about elasticities and the effects of regulatory actions is a very important topic for insurance companies and traffic authorities wanting to know whether Security Road Plans are really effective in reducing ACC and ACCD (e.g., effects of mandatory seat belts on fatalities or speed limits on accidents). Technological studies [Huelke and Gikas(1968), Levine and Campbell (1971) and Joksch and Wuerderman (1972), among others] imply that annual highway deaths would be 20% greater without legally mandated installation of various safety devices on automobiles. However, as pointed out by Peltzman (1975), this literature ignores offsetting effects of nonregulatory demand for safety and driver response to the devices. We tested this hypothesis for the Spanish case obtaining inconclusive results as a consequence of the short period of time since the Traffic Security Plan started in 1980 [Garcia-Ferrer and del Hoyo (1983, pp. 130-134)].

4.2.2 Dynamic Relationship between GAS and ACC

One of the maintained hypotheses of our theoretical model is that economic variables affect accident rates. Therefore, we have chosen gasoline consumption (GAS) as a proxy for the use of the stock of cars. With more cars "competing" for a fixed reduced space (road) the probability of accident increases.

The identification of the transfer function through the ccf suggested a unidirectional causality relationship from GAS to ACC. The estimated model was:

$$\nabla\nabla_{12} \ln ACC_t = \underset{[5.35]}{(0.383} + \underset{[2.62]}{0.187B)} \nabla\nabla_{12} \ln GAS_t + (1 - \underset{[13.1]}{0.729B})(1 - \underset{[36.3]}{0.910B^{12}}) a_t \qquad (11)$$

$T = 156,\ \hat{\sigma}^2_a = 1.30 \times 10^{-3},\ Q(12) = 9,\ Q(24) = 12,\ Q(36) = 19$

where the expected positive sign and statistical significance of the coefficients of the input variable are confirmed. Although, model (11) represents a considerable reduction on $\hat{\sigma}_a^2$ with respect to either (1) or (1b), residual variances are not strictly comparable due to the large differences in samples sizes.

4.2.3 Dynamic Relationship for Gasoline Consumption

When forecasting economic times series, situations often occur in which it is important to know whether a step change in a model input variable, say income or price, affects an output variable, such as consumption, in (i) a permanent way, that is, it has a long term effect or (ii) a transient way, having an effect which lasts only for a short period. However, in many cases consumption - price elasticities are usually estimated from transfer function models of the form

$$\ln C_t = w(B) \ln P_t + N_t$$

implying that a permanent change in price produces a permanent change in consumption. With models of this form, a short term effect lasting for say 1 to 3 periods could give the impression that a weak long term or permanent price effect is in operation. [e.g., Abbie et al (1983)]. Hence, when there is doubt from a priori considerations as to whether an input variable has a short or long term effect, this hypothesis should be tested when analyzing the data.

Because of the fact that the evolution of the gasoline price level in Spain is determined by government authorities, its nature resembles an intervention variable in which each "step" represents the corresponding price increase. Therefore, our transfer function model is, in this case, like an intervention model with two variables: gasoline price (PRIGAS) and EAST. Strictly speaking, however, we should expect a nonlinear response between GAS and PRIGAS. As an approximation for this relationship, we tried to disentangle effects for large and small gasoline price increases. The reason for this is obvious, since we expect that the price elasticity associated with large increases will be higher than the corresponding one for small changes in the price. Consequently we define two new variables: PRIGAS 1 and PRIGAS 2 measuring the rate of growth of gasoline prices corresponding to a price increase above 10% and below 10% respectively. The 10% line was chosen because it roughly represents the mean of the price increases.

Due to the lack of monthly data on income, we first only considered price effects. The estimated model was:

$$\nabla\nabla_{12} \ln GAS_t = \frac{\underset{[6.66]}{-0.0035}}{1 + \underset{[7.88]}{0.672}\, B} \nabla_{12} PRIGAS\ 1_t - \underset{[2.54]}{0.0038}\, \nabla_{12} PRIGAS\ 2_t +$$

$$+ \underset{[5.12]}{0.0549}\, \nabla\nabla_{12} EAST_t + \frac{1}{(1 + \underset{[8.54]}{0.688}\, B + \underset{[3.82]}{0.307}\, B^2)(1 + \underset{[7.58]}{0.548}\, B^{12} + \underset{[2.80]}{0.215}\, B^{24})} a_t \qquad (12)$$

$$T = 156, \quad \hat{\sigma}_a^2 = 1.14 \times 10^{-3}, \quad Q(12) = 19, \quad Q(24) = 28, \quad Q(36) = 37$$

Note first, that the expected negative signs are confirmed for both price variables, although the statistical significance associated with PRIGAS 1 seems much higher than that corresponding to PRIGAS 2. On the other hand the dynamic gain responses for each variable (standard error in parenthesis) are:

$$g(PRIGAS\ 1) = \frac{-0.0035}{1 + 0.672} = \underset{(0.00016)}{-0.0021}$$

$$g(\text{PRIGAS } 2) = -\underset{(0.0014)}{0.0038}$$

According to these results, it seems that the effect of PRIGAS 2 is stronger than the effect of PRIGAS 1. However, if we consider the frequency content of the dynamic response of both variables, the PRIGAS 1 response is considerably stronger that the PRIGAS 2 response [e.g., del Hoyo and Terceiro (1983)]. We can see this by considering the square of the spectral gains for PRIGAS 1 and PRIGAS 2. For PRIGAS 1

$$G_1^2(w) = \frac{w_0^2}{1 + \delta_1^2 - 2\delta_1 \cos w} = \frac{12.25 \times 10^{-6}}{1.452 + 1.344 \cos w} \; ; \; 0 < w < \pi$$

which reaches its maximum at $w = \pi$, therefore

$$G_1^2(0) = 4.38 \times 10^{-6}$$
$$G_1^2(\pi/2) = 8.44 \times 10^{-6}$$
$$G_1^2(\pi) = 113.4 \times 10^{-6}$$

For PRIGAS 2, the spectral gain is:

$$G_2^2(w) = (0.0038)^2 = 1.44 \times 10^{-5}$$

* (Figures 7 and 8)

With monthly data, π represents 2 months' effects and $\pi/2$ and $\pi/6$ 4 months and one year respectively, we can see that while PRIGAS 1 has a short run effect on gasoline consumption (considerably stronger than its long run effect), PRIGAS 2 has its influence on gasoline consumption more uniformly distributed. Additionally, the effect of PRIGAS 1 (measured through its contribution to the residual variance reduction) is considerably greater than the effect of PRIGAS 2. A verification of this fact is confirmed when we carried out the estimation of the transfer function using PRIGAS 1 and PRIGAS 2 separately. In the first case (using PRIGAS 1), $\hat{\sigma}_a^2 = 1.35 \times 10^{-3}$, while in the second (using PRIGAS 2), $\hat{\sigma}_a^2 = 1.70 \times 10^{-3}$. If we compare these with the value of $\hat{\sigma}_a^2 = 1.28 \times 10^{-3}$ obtained from a variant of model (12) (without the EAST variable), we can see how the reduction in the residual variance is basically due to the inclusion of the PRIGAS 1 variable.

An improvement of equation (12) would involve the addition of a variable measuring disposable income. Due to the lack of monthly income data, we have used a proxy variable which captures the growth of the stock of cars (income effect) causing greater gasoline consumption. Such a variable is NUVE, and the estimation is shown in equation (13).

$$\nabla\nabla_{12} \ln GAS_t = \frac{\underset{[5.70]}{-\,0.0030}}{1 + \underset{[8.74]}{0.717}\, B} \nabla_{12} PRIGAS\,1_t - \underset{[2.97]}{0.0043}\, \nabla_{12} PRIGAS\,2_t +$$

$$\underset{[5.60]}{0.0592}\, \nabla\nabla_{12} EAST_t + \underset{[4.11]}{0.0935}\, \nabla\nabla_{12} \ln NUVE_t +$$

$$\frac{1}{(1 + \underset{[8.0]}{0.663}\, B + \underset{[3.5]}{0.298}\, B^2)(1 + \underset{[9.31]}{0.718}\, B^{12} + \underset{[4.74]}{0.394}\, B^{24})} a_t \qquad (13)$$

$T = 156$, $\hat{\sigma}_a^2 = 1.02 \times 10^{-3}$, $Q(12) = 21$, $Q(24) = 27$, $Q(36) = 36$

Note in this case, that due to the orthogonality characteristics of the input variables, the usual problems of orthogonality among inputs [e.g., Liu and

Hanssens (1982)] are not important. It is useful, however, to note the statistical significance of all coefficients in the model, as well as the confirmation of the expected signs. Note also that the inclusion of the income variable has changed very little the estimated coefficients of the price variables. The reduction on the residual variance estimation with respect to equation (12) is around 12%.

4.2.4 Dynamic Relationship between ACC and RAIN

It is a well known opinion among the insurance companies that certain atmospheric phenomena (especially rain) tend to increase car accidents. Trying to model such relationship, we expected an instantaneous effect between ACC and the RAIN index. However, when we proceeded to the analysis of the ccf of the residuals of the univariate models, we found no evidence at all of any relationship between both variables, as is shown in Figure 9.

* (Figure 9)

The conclusion that experts are wrong is not necessarily true. This lack of "causal relationship" between RAIN and ACC is, primarily, due to contemporaneous dynamic and spatial aggregation problems [e.g., del Hoyo and Antoñanzas (1983)]. A correct strategy for dealing with these problems would imply better statistical information on rain fallings to avoid the aggregation bias as a consequence of the "disguise effect" of dry and rainy days within each month. Obtaining such a disaggregated information would be of primary concern to insurance firms in order to design an optimal strategy of prices according to the structural differences of each region.

4.3 Simultaneous Equation Models

The last step of our modelling exercise is centered on a simultaneous dynamic equations approach. Several methods for specifying lag structures have been proposed in recent years. Restricting the discussion to those techniques that can handle multiple-equation models with feedback effects, two basic approaches have emerged (with several variants): pairwise cross-correlation methods on prewhitened data [e.g., Haugh and Box (1976), Granger and Newbold (1977), McLeod (1989)] and vector ARMA methods [e.g., Jenkins and Alavi (1981), Tiao and Box (1981), Tiao and Tsay (1983)]. The first technique is very useful when the direction of causality between two variables is a priori unclear; however, it is difficult to generalize for larger models. Additionally, there are some methodological problems [e.g., Maravall (1981), Newbold (1981)] that seriously question its use. On the other hand, initial experience with vector ARMA methods is encouraging, in the sense that it facilitates better understanding of the interrelationships among variables. We will follow this second approach.

4.3.1 Bivariate Model between ACC and ACCD

Extending the basic ideas in Box and Jenkins (1976) for the univariate series, Tiao and Box (1981) proposed an iterative procedure consisting of three phases of tentative specification, estimation and diagnostic checking using the sample cross-correlation (SCM) and partial autoregression (PAM) matrices as the basic tools for initial identification.

Using the computer program developed by Tiao et al. (1983), the estimated bivariate model between ACC and ACCD is reported in equation (14) on Table 5. The basic features of this model are the following:

i) The triangular structure of the model implies a unidirectional "causality" from ACC to ACCD, confirming the results obtained with the transfer function model of equation (10).

ii) By comparing the empirical results obtained in (10) and (14) and using some algebra, it can be shown that both structures are extremely similar.

iii) The differences in the residual variance from univariate to bivariate model is shown in Table 3. From Table 3, it becomes clear that while for the ACC variable little change has taken place when we go from the univariate to the bivariate model, for the ACCD there is a considerable reduction when "more information" is used.

iv) No evidence of autocorrelation between a_{1t} and a_{2t} was found.

*(Table 3)

4.3.2 Bivariate Model between ACC and GAS

As pointed out earlier, the main interest in this model rests on the hypothesis that certain economic variables (by measuring car utilization indexes) may influence accident rates. In this case, the final model (15) appears in Table 5 and deserves some comments:

i) Both the univariate model for GAS and the transfer function model for ACC are very similar to those obtained in (8) and (11) respectively.

ii) As in the previous section, the reduction in the residual variance is very little for the "exogenous" variable and very spectular for the "endogenous" variable, as can be seen in Table 4.

iii) No evidence of autocorrelation was found between a_{1t} and a_{2t}.

* (Tables 4 and 5)

4.3.3 General Simultaneous Dynamic Model

The purpose of this section is to build a multiple time series model connecting accident rates variables (ACC) with those representing the evolution of the stock of automobiles (VEH and NUVE), the car utilization index (GAS), and the IIP index.

The model (16) was estimated using the conditional maximum likelihood algorithm, and it is shown in Table 6. From its structure, it is easily seen that GAS, NUVE, and IIP are exogenous while ACC and VEH depend on GAS and NUVE. While these results confirm our initial assumptions, it is hard to accept the lack of "causation" from IIP to GAS and NUVE, especially when there is evidence of positive correlation between the residuals of IIP and NUVE and the residuals of IIP and GAS. When we estimated these possibilities, we found the expected signs but no statistical significance on the coefficients.One of the reasons for the poor performance of the IIP variable in the general model is related with problems created by the change which occurred in the construction of the index after 1976. See interalia, Espasa (1983) and Garcia-Ferrer and del Hoyo (1984).

*(Table 6)

5. PREDICTIVE PERFORMANCE OF THE MODEL

The main interest of this section is to study the prediction performance of some of the models that we have analyzed earlier. In particular, we examine:

i) the predictive accuracy of the univariate models

ii) the improvement in prediction that takes place when we go to more elaborate models

For this purpose, forecasts have been evaluated according to two different forecasting measures: first, the percentage prediction error defined as:

$$\% \text{ PE} = \frac{F\ (t + s) - A\ (t + s)}{A\ (t + s)}$$

where A(t) is the actual value, F(t) denotes the forecast value and s = 1,2, . . . 12 is the total number of forecasts. Given this definition, a positive error indicates an overforecast and a negative error an underforecast. As a measure of aggregate forecasting performance, we also have computed the percentage root mean square error (%RMSE) defined as:

$$\%\text{RMSE} = \left\{ 100 \times \sum_{s=1}^{12} \frac{[F(t + s) - A(t + s)]^2}{[12\ A(t + s)^2]} \right\}^{\frac{1}{2}}$$

The forecasting results for ACC, ACCD, GAS, and NUVE (we have not included IIP because of the problems mentioned earlier, neither VEH because it can be easily obtained from NUVE) are summarized in Table 7. The following remarks are relevant:

*(Table 7)

i) For the ACC variable, the predictive performance of all models is reasonable (see Figure 10). For all models, the greatest % PE value takes place in April as a consequence of the fact that Easter holidays fell between the last week of March and the first week of April. Univariate forecasting is good for short period horizons, and only for longer horizons is there a slight improvement in %PE with respect to more elaborated models. As happens with GAS, the bivariate models forecasts are very similar to those obtained by the TF models. (see Sections 4.3.1 and 4.3.2)

ii) Similar comments concern the predictive results for the ACCD variable, albeit the forecasting performance is slightly worse than in the previous case. For this variable, however, the forecasting improvement for longer horizons obtained with more elaborated models is more evident than for the ACC variable. With more data available on Security Road Plans, it would have been very interesting to test whether this systematic "overprediction" for the ACCD variable is a successful outcome of the plan.

* (Figures 10 and 11)

iii) For the GAS variable, the univariate forecasting is so good that it could hardly be beaten by any other model in the paper. Note however, that GAS appears as exogenous in the multivariate model (16), and therefore it should not have many differences with its univariate forecasting. On the other hand, the transfer function model captures two basic features of the GAS variable (see Figure 11):

- The reduction of %PE in April and March as a consequence of the intervention analysis of the Easter effect
- The reduction of %PE in December (%PE = 0.004) as a consequence of a large price increase of gasoline (21%) which took place on that month and that considerably affected gasoline consumption in the short run.

iv) For the NUVE variable, the univariate and intervention models show similar results, although there is a significant reduction in the %PE in March and April as a consequence of the intervention analysis of the Easter effect. Also, note that the multivariate forecasting is not strictly comparable with the previous ones, because of the large difference in the data base, 240 and 156 observations respectively.

6. CONCLUSIONS

The main objective of this work has been the specification and estimation of a dynamic simultaneous equation model in order to explain the evolution of the automobile accident indexes in Spain from our initial theoretical model developed in section 2. It should be clear that the final outcome is not the result of a purely empirical approximation, but rather the interaction of the different stages of specification, estimation and verification of the dynamic relationships.

From the univariate analyses of the different series in the model, it is clear that seasonal components are present as shown through the ∇_{12} operator as well as the seasonal autoregressive and moving average components. This finding - in sharp contrast with some experts' opinions - is the result of the different definitions used by insurance companies and the sources of data that we have used. [2]

The outliers around Easter have been taken into account by means of a properly defined intervention analysis. The estimated relationships have

---

[2] For most insurance companies the accident rate is defined as the product of two variables: frequency and cost. Since Lower frequency of accidents tends to compensate with larger costs during seasonal periods, most companies do not pay attention to the seasonal aspects of their data.

coefficient estimates with expected algebraic signs, and the intervention coefficients were in all cases statistically significant. The same type of analysis was applied to test the success of the Traffic Security Plan started in 1980, obtaining inconclusive statistical results as a consequence of the small number of observations available.

The results for single equation dynamic models were in accordance with a priori expectations. The only exception was the equation relating accidents with the rainfall index. As explained in 4.2.5, this is the consequence of both contemporaneous dynamic and spatial aggregation problems. Among the remaining transfer functions, we find particularly interesting the results obtained for gasoline consumption in terms of its own price and new registration of vehicles. The effect of price is nonlinear, and is stronger - as shown by means of spectral gain analysis - for high increases (PRIGAS 1) than for small increases (PRIGAS 2).

The simultaneous equation models, especially the bivariate ones, confirmed the structure of the single equation models. In the general model, however, the expected relationships are not completely confirmed due to the problems associated with changes which occurred in the elaboration of the industrial production index, after 1976.

Finally, the predictive performance of the univariate models is very good. The improvement in prediction with more elaborated models is more evident for longer time horizons. In any case, there is a considerable informative gain by using more elaborate models which allow the estimation of elasticities and dynamic multipliers.

5. References

Abbie, E.G., Everest, J.T., Jenkins, G.M., and Paulley, N.J. (1983). "The Analysis of Competition Between Rail and Air on Four London to Scotland Routes," in Jenkins and McLeod (eds.) Case Studies in Time Series Analysis. Volume 1, Lancaster: GJP Publication. pp. 38-72.

Box, G.E.P., and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control, San Francisco, Holden-Day.

Box, G.E.P., and Tiao, G.C. (1975). "Intervention Analysis with Applications to Economic and Environmental Problems", Journal of the American Statistical Association, 70, 70-79.

Direccion General de Tráfico. (1981). "Informe Anual sobre Accidentes en Carretera". Ministerio de Transportes. Madrid, 1981.

Espasa, A. (1983). "Deterministic and Stochastic Seasonality: A Univariate Study of the Spanish Industrial Production Index"; Paper presented at the European Meeting of the Econometric Society in Pisa.

Garcia-Ferrer, A. and del Hoyo, J. (1983). Análisis Econométrico de los Indices de Siniestralidad del Sector del Automovil. Becas de Investigación RUMASA, Madrid.

Garcia-Ferrer, A. and del Hoyo, J. (1984). "Dynamic Behaviour of the Industrial Production Index in Spain from 1964 to 1983". Unpublished working paper. Departamento de Econometría. Universidad Autónoma, Madrid.

Granger, C.W.J., and Newbold, P. (1977). Forecasting Economic Time Series. New York: Academic Press.

Haugh, L.D., and Box, G.E.P. (1976). "Identification of Dynamic Regressions (Distributed Lags) Models Connecting Two Times Series", Journal of the American Statistical Association, 72, 121-130.

Holmes, J. (1975). "The Relative Burden of Higher Gasoline Prices", in: Duncan and Morgan (eds). Analysis of the First Six Years of the Panel Study of Income Dynamics, vol. II. University of Michigan, 183-199.

del Hoyo, J. and Antoñanzas, F. (1983). "Identification of Dynamic Models between Contemporaneous Aggregate Variables". Unpublished working paper, Departamento de Econometría. Universidad Autónoma de Madrid.

del Hoyo, J. and Terceiro, J. (1983). "On the Identification of Long-Run and Short-Run Effects on Transfer Function Models". Paper presented at the II Meeting on Economics. Universidad Menendez y Pelayo, Sitges.

Huelke, D., and Gikas, P. "Causes of Deaths in Automobile Accidents", Journal of the American Medical Association, vol. 25, 1100-1107.

Jenkins, G.M., and Alavi, A.S. (1981). "Some Aspects of Modelling and Forecasting Multivariate Time Series", Journal of Time Series Analysis, 2, 1-48.

Joksch, H., and Wuederman, H. (1971). "Estimating the Effects of Crash Phase Injury Countermeasures", Accident Analysis and Prevention, 4, 89-108.

Lave, L., and Weber, W. (1970). "A Benefit-Cost Analysis fo Auto Safety Measures", Applied Economics, 2, #4, 215-275.

Levine, D., and Compbell, B. (1971). "Effectiveness of Lap Seat Belts and the Energy Absorbing Steering System in the Reduction of Injuries", Highway Safety Research Center, University of North Carolina.

Liu, L.M., and Hanssens, D.M. (1982). "Identification fo Multiple-Input Transfer Function Models", Communications in Statistics-Theory and Methods, 11, 297-314.

Lütkepohl, H. (1983), "Non-Causality Due to Omitted Variables", Journal of Econometrics, 19, pp. 367-378.

Maravall, A. (1981). "A Note on Identification of Multivariate Time-Series Models", Journal of Econometrics, 16, 237-247.

McLeod, A.O. (1979). "Distribution of Residual Cross-Correlation in Univariate ARMA Time Series Models". Journal of the American Statistical Association, 74, 849-855.

Newbold, P. (1981). "Model Checking in Time Series Analysis" ASA-CENSUS-NBER Conference on Applied Time Series of Economic Data. Washington D.C.

Peltzman, S. (1975). "The Effects of Automobile Safety Regulation", Journal of Political Economy, 83, #4, 677-725.

Tiao, G.C., and Box, G.E.P. (1981). "Modelling Multiple Time Series with Applications", Journal of the American Statistical Association, 76, 802-816.

Tiao, G.C., Liu, L.M., Hudak, G.B., Muller, M.E., and Box, G.E.P. (1983), The SCA System for Univariate-Multivariate Time Series and General Statistical Analysis, Scientific Computing Associates, P.O.Box 625, DeKalb, IL 60115.

Tiao, G.C., and Tsay, R.S. (1983). "Multiple Time Series Modelling and Extended Sample Cross-Correlations", Journal of Business and Economic Statistics, 1, 43-56.

Tsurumi, H. (1980). "A Bayesian Estimation of Structural Shifts by Gradual Switching Regressions with an Application to the U.S. Gasoline Market", in A. Zellner (ed). Bayesian Analysis in Econometrics and Statistics, Amsterdam, North Holland, 213-240.

6. Tables and Figure Titles:

6.a. Tables

Table 1. Univariate Models for Monthly Data

Table 2. Intervention Models for Monthly Data

Table 3. Residual Variances for Several Models

Table 4. Residual Variances for Several Models

Table 5. Bivariate Models for Monthly Data

Table 6. General Simultaneous Dynamic Model

Table 7. Descriptive Measures of Forecast Errors of Levels of ACC, ACCD, GAS, and NUVE

6.b. Figures

Figure 1. Initial Specification of the Model

Figure 2. Number of Accidents

Figure 3. Gasoline Consumption

Figure 4. Stock of Vehicles

Figure 5. Industrial Production Index

Figure 6. New Registration of Vehicles

Figure 7. Spectral Gain for PRIGAS 1

Figure 8. Spectral Gain for PRIGAS 2

Figure 9. ccf between ACC and RAIN

Figure 10. Actual and Predicted Values for the ACC variable

Figure 11. Actual and Predicted Values for the GAS variable

Table 1: Univariate Models for Monthly Data

Variable	Estimation Period (T)	Model	$\hat{\sigma}^2_a$	Correlation Coefficients	Ljung-Box Statistics
ACC	62/1 81/12 (240)	$(1 + 0.546B + 0.189B^2 + 0.143B^3)\nabla\nabla_{12} \ln ACC_t = (1 - 0.742B^{12}) a_t$ (1) [8.5] [2.6] [2.3] [18.6]	$1.72.10^{-3}$	$\rho(\hat{\phi}_1\hat{\phi}_2) = 0.46$ $\rho(\hat{\phi}_2\hat{\phi}_3) = 0.47$	Q(12) = 13 Q(24) = 29 Q(36) = 41
ACCD	62/1 81/12 (240)	$\nabla\nabla_{12} \ln ACCD_t = (1 - 0.806B)(1 - 0.801B^{12}) a_t$ (2) [21.1] [22.1]	$7.04.10^{-3}$		Q(12) = 7.8 Q(24) = 25 Q(36) = 32
VIC	62/1 81/12 (240)	$(1 + 0.706B + 0.273B^2 + 0.141B^3)\nabla\nabla_{12} \ln VIC_t = (1 - 0.767B^{12}) a_t$ (3) [11.0] [3.6] [2.2] [20.1]	$2.46.10^{-3}$	$\rho(\hat{\phi}_1\hat{\phi}_2) = 0.56$ $\rho(\hat{\phi}_2\hat{\phi}_3) = 0.57$	Q(12) = 18 Q(24) = 29 Q(36) = 47
DTHS	62/1 81/12 (240)	$(1 + 0.181B)\nabla\nabla_{12} \ln DTHS_t = (1 - 0.763B)(1 - 0.801B^{12}) a_t$ (4) [2.34] [15.2] [21.2]	$9.65.10^{-3}$	$\rho(\hat{\phi}_1\hat{\theta}_1) = 0.55$	Q(12) = 3.2 Q(24) = 13 Q(36) = 17
VEH	62/1 81/12 (240)	$\nabla^2\nabla_{12} \ln VEH_t = (1 - 0.702B)(1 - 0.906B^{12}) a_t$ (5) [15.0] [45.9]	$1.37.10^{-6}$		Q(12) = 17 Q(24) = 29 Q(36) = 38
NUVE	62/1 81/12 (240)	$(1 - 0.186B^{12} + 0.167B^{24})\nabla\nabla_{12} \ln NUVE_t = (1 - 0.626B)(1 - 0.881B^{12}) a_t$ (6) [2.7] [2.4] [11.9] [33.0]	$9.65.10^{-3}$	$\rho(\hat{\Phi}_{12}\hat{\Theta}_{12}) = 0.24$	Q(12) = 8 Q(24) = 16 Q(36) = 25
IIP	65/1 81/12 (204)	$\nabla\nabla_{12} \ln IIP_t = (1 - 0.612B)(1 - 0.448B^{12}) a_t$ (7) [10.7] [6.61]	$1.54.10^{-3}$		Q(12) = 16 Q(24) = 35 Q(36) = 47
GAS	69/1 81/12 (156)	$(1 + 0.746B + 0.230B^2)(1 + 0.760B^{12} + 0.385B^{24})\nabla\nabla_{12} \ln GAS_t = a_t$ (8) [8.83] [2.73] [9.54] [4.74]	$1.76.10^{-3}$	$\rho(\hat{\phi}_1\hat{\phi}_2) = 0.60$ $(\Phi_{12}\Phi_{24}) = 0.57$	Q(12) = 13 Q(24) = 23 Q(36) = 32
RAIN	65/1 82/6 (186)	$(1 - 0.307B)(1 - 0.226B^{12}) RAIN_t = a_t$ (9) [4.38] [3.14]	$9.44.10^{4}$		Q(12) = 15 Q(24) = 24 Q(36) = 36

Table 2: Intervention Models for Monthly Data

Variable	Estimation Period (T)	Model	$\hat{\sigma}_a^2$	Ljung-Box Statistics
ACC	62/1 81/12 (240)	$\nabla\nabla_{12}\ln ACC_t = \underset{[6.33]}{0.060}\,\nabla\nabla_{12} EAST_t + \dfrac{(1 - \overset{[17.6]}{0.726}B^{12})}{(1 + \underset{[6.9]}{0.448}B + \underset{[3.52]}{0.243}B^2 + \underset{[1.88]}{0.122}B^3)}\, a_t$ (1.b)	$1.52.10^{-3}$	Q(12) = 8.1 Q(24) = 27 Q(36) = 33
ACCD	62/1 81/12 (240)	$\nabla\nabla_{12}\ln ACCD_t = \underset{[4.3]}{0.103}\,\nabla\nabla_{12} EAST_t + (1 - \underset{[20.1]}{0.793}B)(1 - \underset{[20.6]}{0.784}B^{12})\, a_t$ (2.b)	$6.60.10^{-3}$	Q(12) = 6.1 Q(24) = 24 Q(36) = 33
VIC	62/1 81/12 (240)	$\nabla\nabla_{12}\ln VIC_t = \underset{[8.5]}{0.099}\,\nabla\nabla_{12} EAST_t + \dfrac{(1 - \overset{[17.6]}{0.735}B^{12})}{(1 + \underset{[8.5]}{0.541}B + \underset{[3.9]}{0.243}B^2)}\, a_t$ (3.b)	$2.07.10^{-3}$	Q(12) = 11 Q(24) = 28 Q(36) = 38
DTHS	62/1 81/12 (240)	$\nabla\nabla_{12}\ln DTHS_t = \underset{[4.5]}{0.125}\,\nabla\nabla_{12} EAST_t + \underset{[4.2]}{0.356}\,\nabla\nabla_{12} ALF_t + (1 - \underset{[21.0]}{0.804}B)\ (1 - \underset{[19.0]}{0.782}B^{12})\, a_t$ (4.b)	$8.57.10^{-3}$	Q(12) = 6.8 Q(24) = 13 Q(36) = 20
NUVE	62/1 81/12 (240)	$\nabla\nabla_{12}\ln NUVE_t = \underset{[2.59]}{-0.068}\,\nabla\nabla_{12} EAST_t + \dfrac{(1 - \overset{[11.4]}{0.609}B)(1 - \overset{[32.6]}{0.880}B^{12})}{(1 - \underset{[2.83]}{0.199}B^{12} + \underset{[2.19]}{0.880}B^{24})}\, a_t$ (6.b)	$9.43.10^{-3}$	Q(12) = 6.7 Q(24) = 17 Q(36) = 24
GAS	69/1 81/12 (240)	$\nabla\nabla_{12}\ln GAS_t = \underset{[4.23]}{0.056}\,\nabla\nabla_{12} EAST_t + \dfrac{1}{(1 + \underset{[8.33]}{0.716}B + \underset{[2.78]}{0.239}B^2)(1 + \underset{[8.69]}{0.721}B^{12} + \underset{[4.09]}{0.369}B^{24})}\, a_t$ (8.b)	$1.61.10^{-3}$	Q(12) = 18 Q(24) = 26 Q(36) = 35

	Univariate Model	Transfer Function Model	Bivariate Model
ACC	1.77×10^{-3}		1.83×10^{-3}
ACCD	7.04×10^{-3}	4.96×10^{-3}	4.92×10^{-3}

Table 3: Residual Variances for several models.[1]

[1] Numerical results are not strictly comparable because of the use of different estimation algorithms for the transfer function and the model/bivariate. Specifically, the TF model has been estimated with the BMDP package which is only a "backasting method" while the bivariate model has been estimated by exact MLE.

	Univariate Model	Transfer Function Model	Bivariate Model
GAS	$1.76\ 10^{-3}$		$1.70\ 10^{-3}$
ACC	$1.77\ 10^{-3}$	$1.30\ 10^{-3}$	$1.12\ 10^{-3}$

Table 4: Residual Variances for Several Models

Table 5: Bivariate Models

Variables	Estim. Period (T)	Models	$\hat{\Sigma}_a$
ACC ACCD	1/62 12/81 (240)	$\nabla\nabla_{12} \ln \begin{bmatrix} ACC_t \\ ACCD_t \end{bmatrix} = \begin{bmatrix} \underset{[10.9]}{(1 - 0.575\,B)} & \\ \underset{[2.77]}{0.213\,B} & \underset{[26.1]}{(1 - 0.862\,B)} \end{bmatrix} \begin{bmatrix} \underset{[14.7]}{(1 - 0.692\,B^{12})} & \\ \underset{[3.8]}{0.345\,B^{12}} & \underset{[23.8]}{(1 - 0.904\,B^{12})} \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$ (14)	$10^{-3} \begin{bmatrix} 1.83 & \\ 1.68 & 6.46 \end{bmatrix}$
GAS ACC	1/69 12/81 (156)	$\begin{bmatrix} (1 + \underset{[8.67]}{0.780\,B} + \underset{[2.44]}{0.220\,B^{2}}) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} (1 + \underset{[9.4]}{0.799\,B^{12}} + \underset{[4.5]}{0.396\,B^{24}}) & 0 \\ 0 & 1 \end{bmatrix} \nabla\nabla_{12} \ln \begin{bmatrix} GAS_t \\ ACC_t \end{bmatrix} =$ $= \begin{bmatrix} 1 & 0 \\ \underset{[1.70]}{0.077\,B} & \underset{[8.14]}{(1 - 0.586\,B)} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \underset{[14.9]}{(1 - 0.820\,B^{12})} \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$ (15)	$10^{-3} \begin{bmatrix} 1.70 & \\ 0.50 & 1.42 \end{bmatrix}$

Table 6: General Simultaneous Dynamic Model

$$
\begin{bmatrix}
(1 + \underset{[9.40]}{0.748}\,B + \underset{[2.10]}{0.164}\,B^2) & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\nabla\nabla_{12}
\begin{bmatrix}
\ln GAS_t \\
\ln ACC_t \\
\nabla \ln VEH \\
\ln NUVE \\
\ln IIP
\end{bmatrix}
=
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
\underset{[2.25]}{0.170B} & (1 - \underset{[11.1]}{0.676B}) & 0 & 0 & 0 \\
0 & 0 & (1 - \underset{[32.7]}{0.932B}) & \underset{[4.31]}{0.028B} & 0 \\
0 & 0 & 0 & (1 - \underset{[9.66]}{0.592B}) & 0 \\
0 & 0 & 0 & 0 & (1 - \underset{[11.5]}{0.0658B})
\end{bmatrix}
$$

$$
\begin{bmatrix}
(1 - 0.683B^{12}) & 0 & 0 & 0 & 0 \\
0 & (1 - \underset{[12.3]}{0.704B^{12}}) & 0 & 0 & 0 \\
0 & 0 & (1 - \underset{[7.3]}{0.433B^{12}}) & 0 & 0 \\
0 & 0 & 0 & (1 - \underset{[11.3]}{0.619B^{12}}) & 0 \\
0 & 0 & 0 & 0 & (1 - \underset{[5.1]}{0.370B^{12}})
\end{bmatrix}
\begin{bmatrix}
a_{1t} \\
a_{2t} \\
a_{3t} \\
a_{4t} \\
a_{5t}
\end{bmatrix}
\qquad (16) \qquad
\begin{matrix} 1/69 - 12/81 \\ T = 156 \text{ obs.} \end{matrix}
$$

$$
\hat{\Sigma}_a =
\begin{bmatrix}
1.73\ 10^{-3} & & & & \\
4.98\ 10^{-4} & 1.86\ 10^{-3} & & & \\
7.90\ 10^{-5} & -6.63\ 10^{-6} & 1.98\ 10^{-6} & & \\
1.37\ 10^{-3} & -3.47\ 10^{-4} & 1.22\ 10^{-4} & 1.37\ 10^{-2} & \\
7.13\ 10^{-5} & -1.13\ 10^{-4} & 2.24\ 10^{-5} & 1.94\ 10^{-3} & 1.77\ 10^{-3}
\end{bmatrix}
$$

Table 7. Descriptive Measures of Forecast Errors of Levels of ACC, ACCD, GAS, and NUVE for 1982

Variable / Model	ACC			ACCD			GAS			NUVE		
	UN[a] (1)	TF[b] (11)	MT[c] (16)	UN (2)	TF (10)	MT (14)	UN (8)	TF (12)	MT (16)	UN (6)	TF (66)	MT (16)
% PE												
January	-0.40	3.30	0.95	-2.61	-1.96	-4.25	1.60	3.52	2.57	-4.72	-5.31	-5.46
February	-0.92	-1.19	-0.31	1.14	1.89	1.14	0.45	0.58	0.77	-4.51	-4.67	-6.28
March	3.52	5.83	5.13	13.7	9.89	9.89	-2.92	-0.82	-2.25	-14.1	-12.2	-15.9
April	7.95	9.69	8.51	17.5	16.0	13.7	-1.93	-0.89	-2.40	3.92	1.66	-0.55
May	-1.87	2.05	-1.31	0.66	-0.66	-5.57	3.32	4.32	3.37	3.57	3.58	5.20
June	6.23	6.02	6.89	12.8	12.8	12.1	-1.40	0.91	0.57	3.15	2.78	-5.05
July	3.02	3.76	3.66	10.8	8.13	6.50	-1.39	4.81	3.23	-3.68	-3.83	-8.19
August	1.86	1.80	2.23	7.28	4.85	2.67	-2.24	1.37	0.58	7.10	6.73	-3.06
September	2.76	2.06	3.06	11.8	8.70	8.39	-2.28	3.51	1.02	6.74	6.64	4.64
October	0.96	0.02	1.29	0.84	-0.28	-1.12	0.16	4.71	1.89	-4.73	-4.61	-3.75
November	2.29	2.57	2.73	15.4	14.4	14.7	-3.01	1.08	-1.96	-14.7	-14.9	-16.8
December	3.53	2.41	4.24	10.0	8.24	9.41	1.57	-0.004	3.39	-15.1	14.9	21.8
%RMSE	0.361	0.418	0.408	1.049	0.893	0.904	0.187	0.195	0.224	0.7623	0.7332	0.8809

a UN denotes univariate model

b TF denotes transfer function model

c MT denotes multivariate model

8. Figures

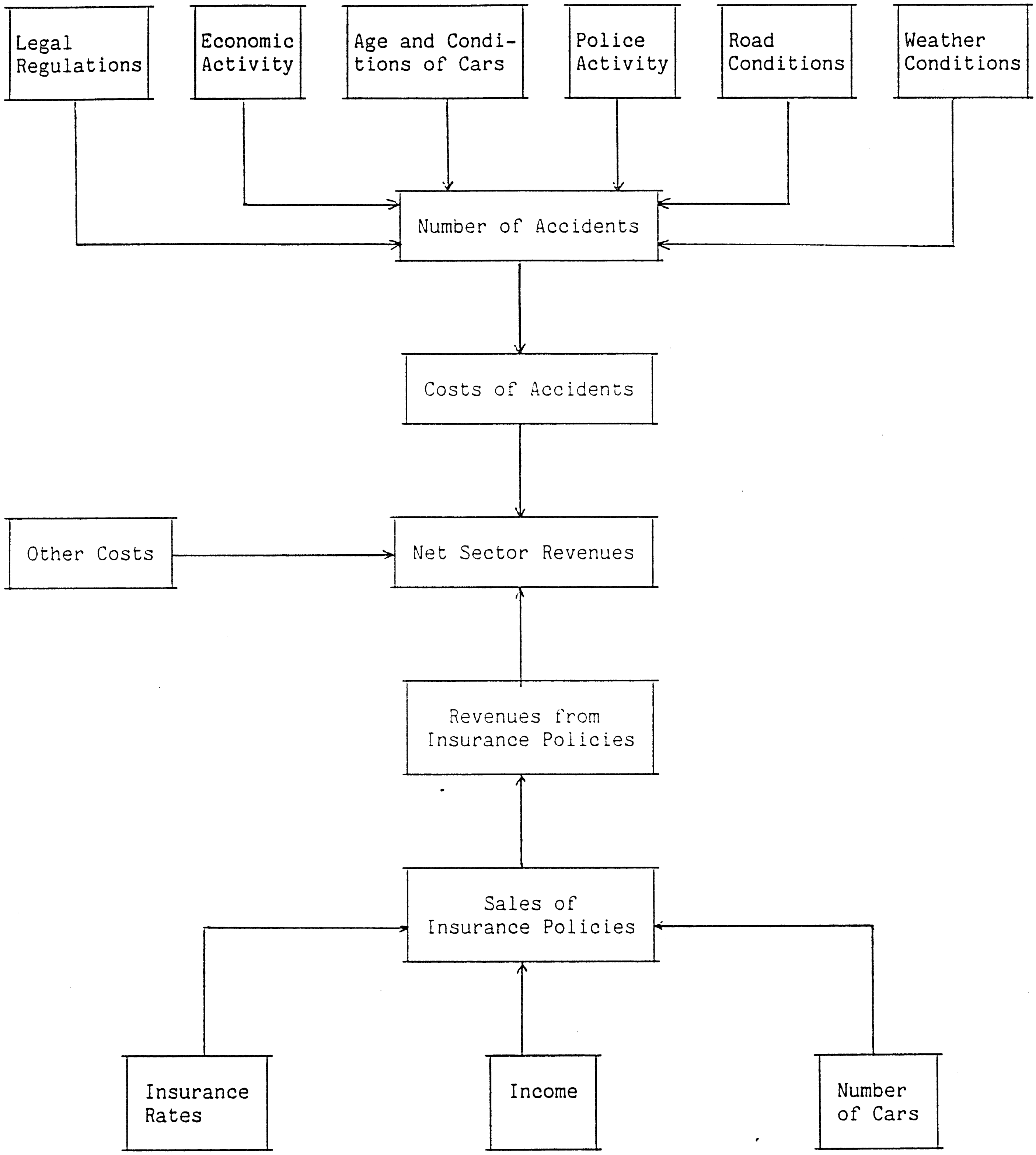

Figure 1: Initial Specification of the Accidents Model.

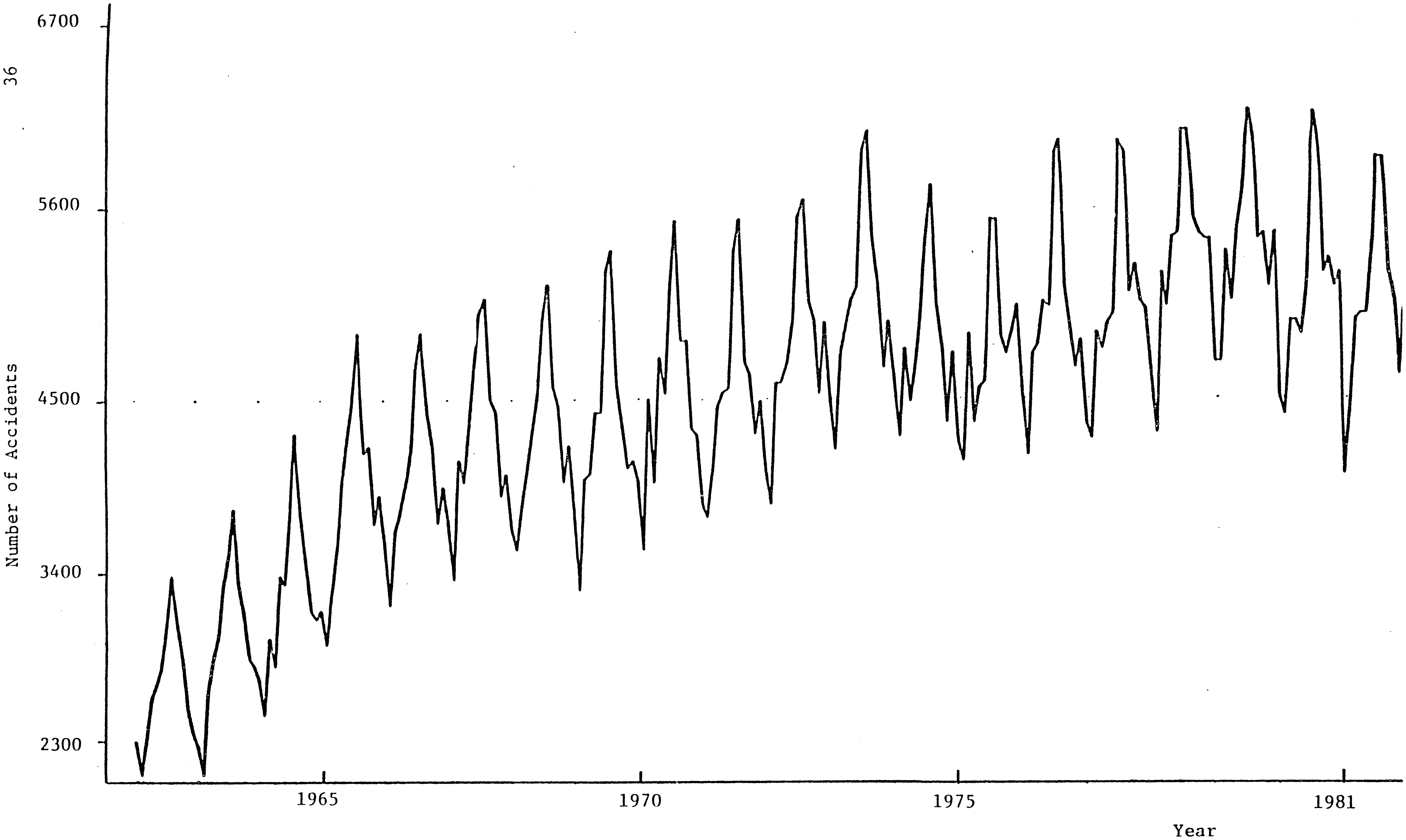

Figure 2. Number of Accidents (ACC)

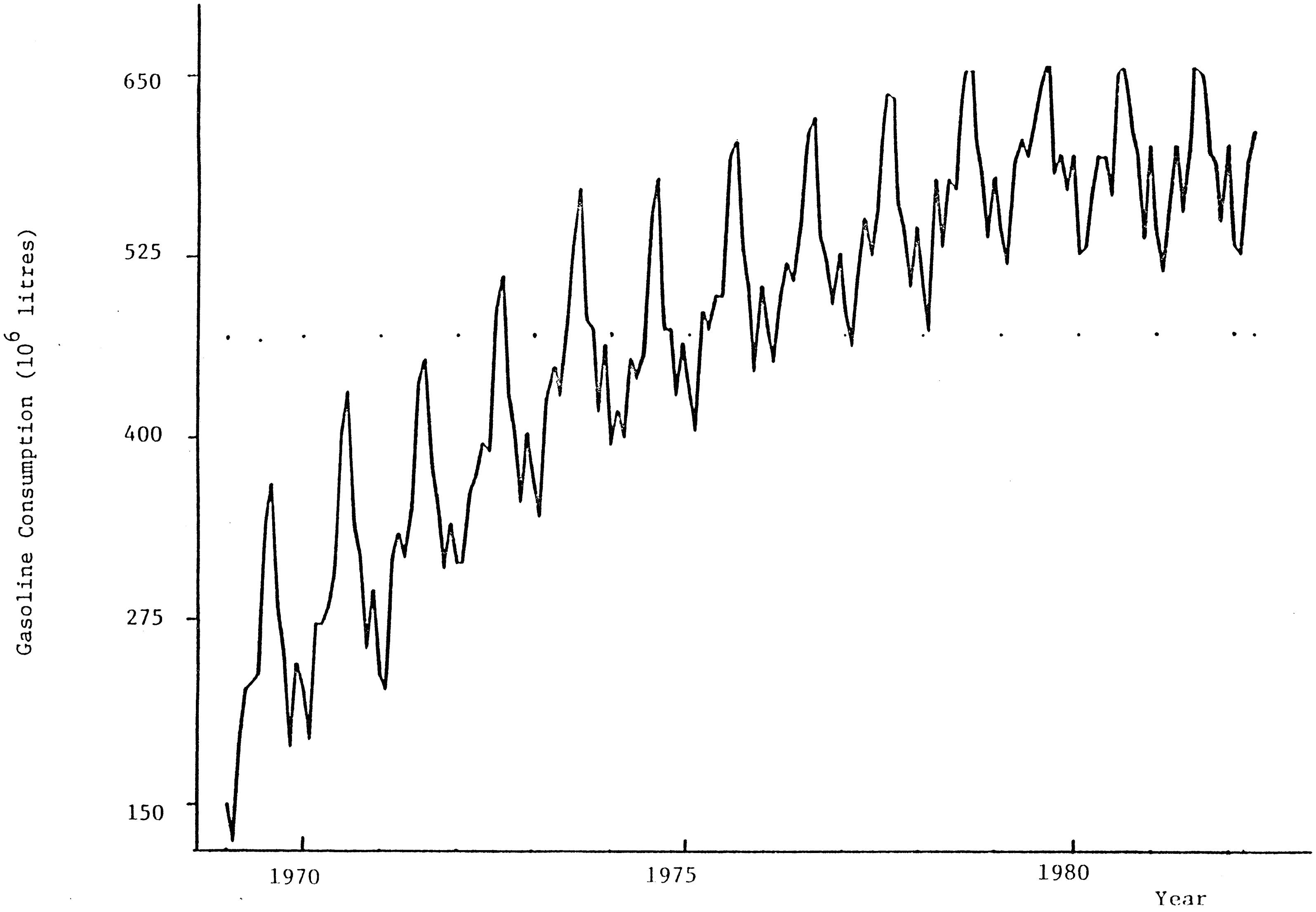

Figure 3. Gasoline Consumption (GAS)

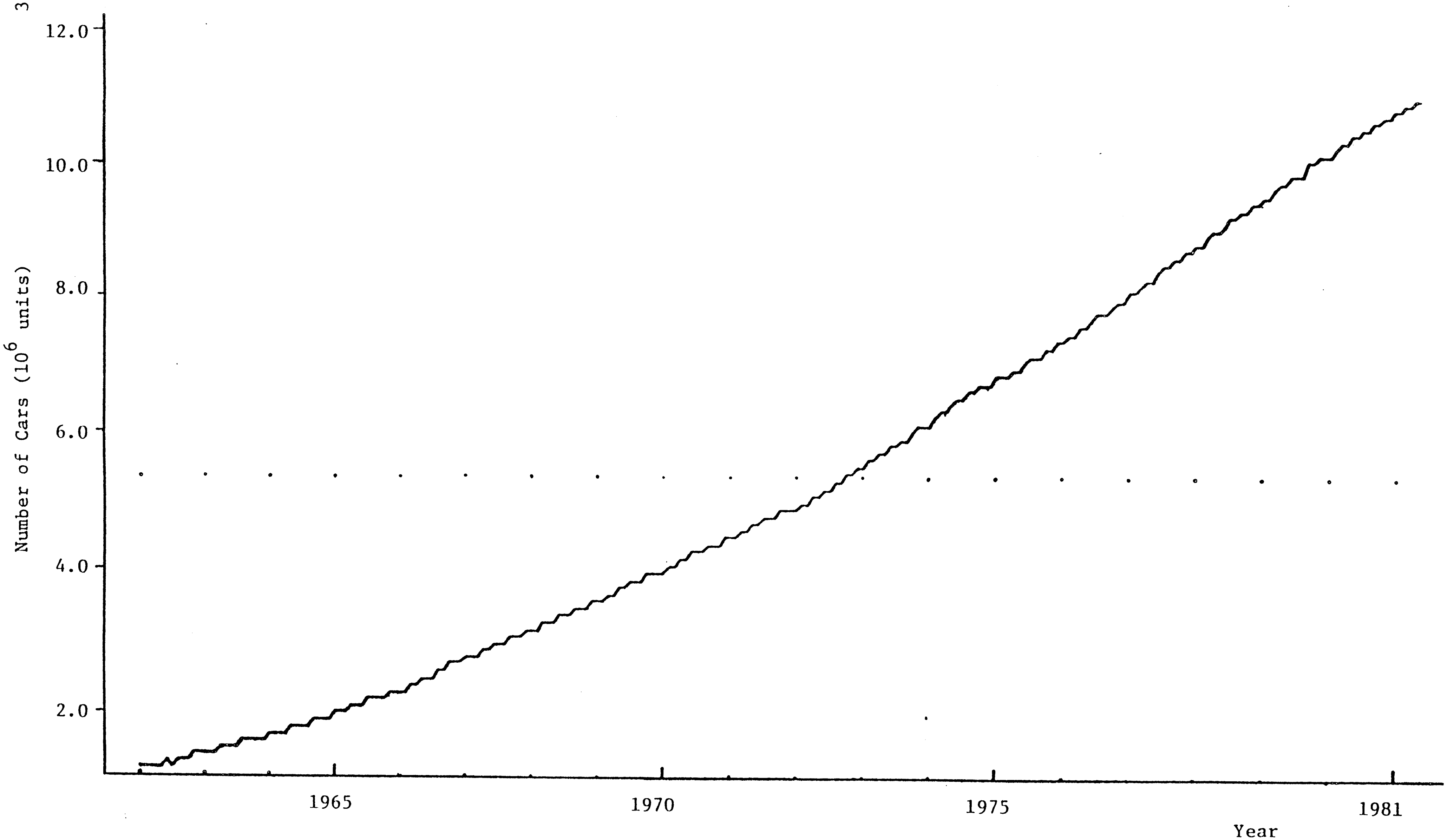

Figure 4. Stock of Vehicles (VEH)

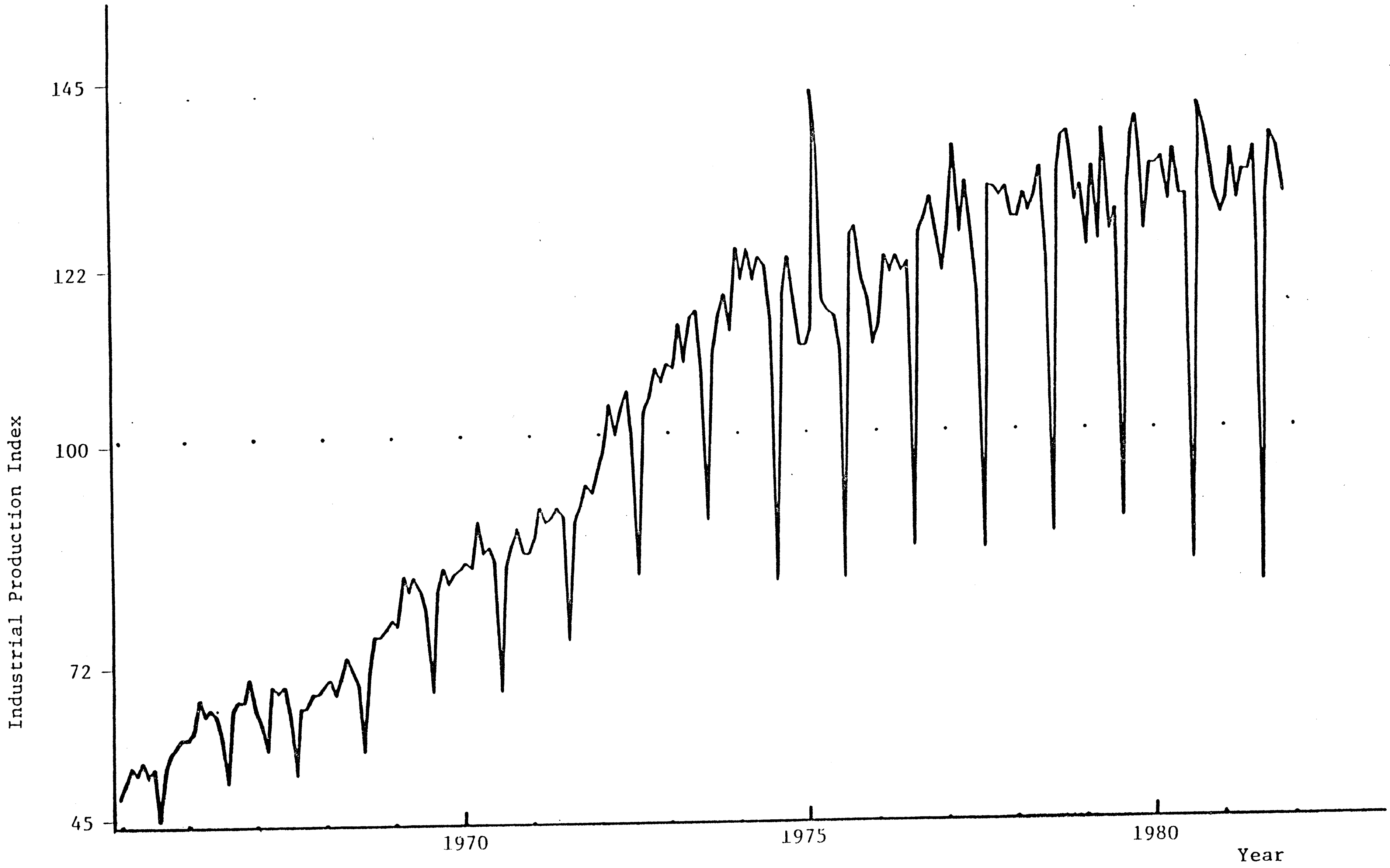

Figure 5. Industrial Production Index (IIP), 1972 = 100

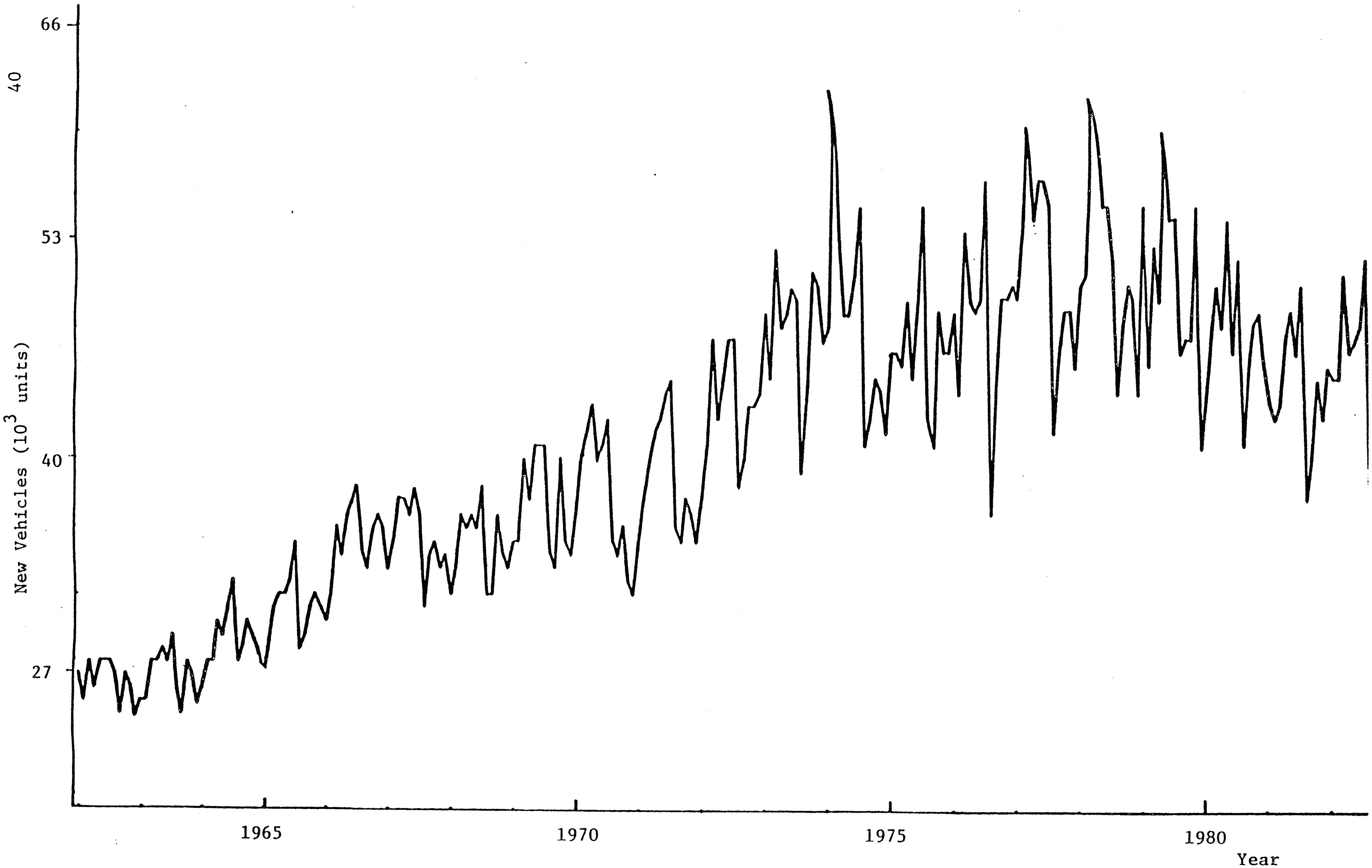

Figure 6. New Registration of Vehicles (NUVE)

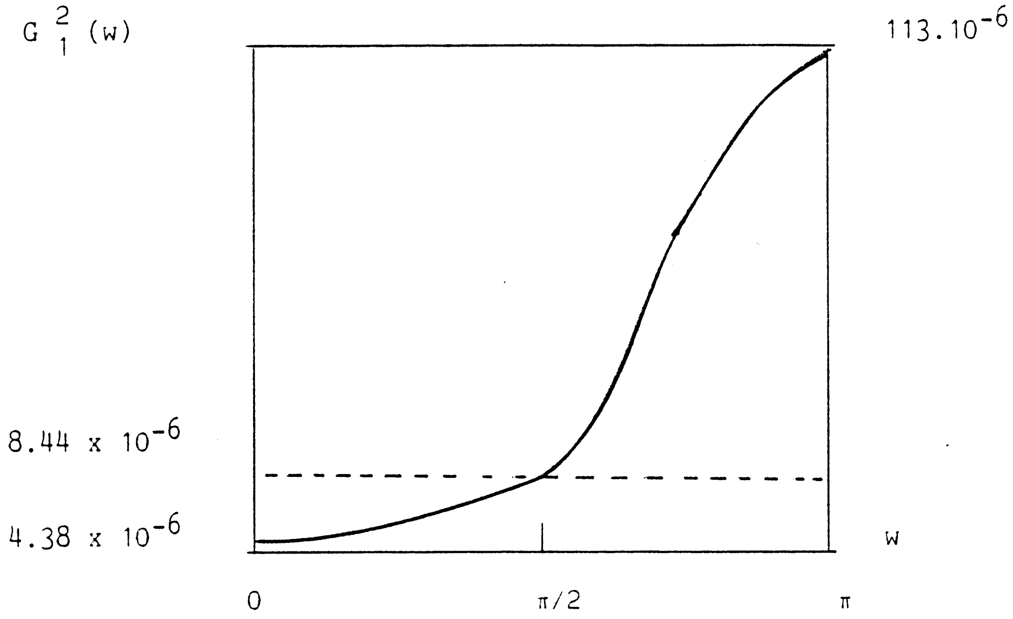

Figure 7: Spectral Gain for PRIGAS 1

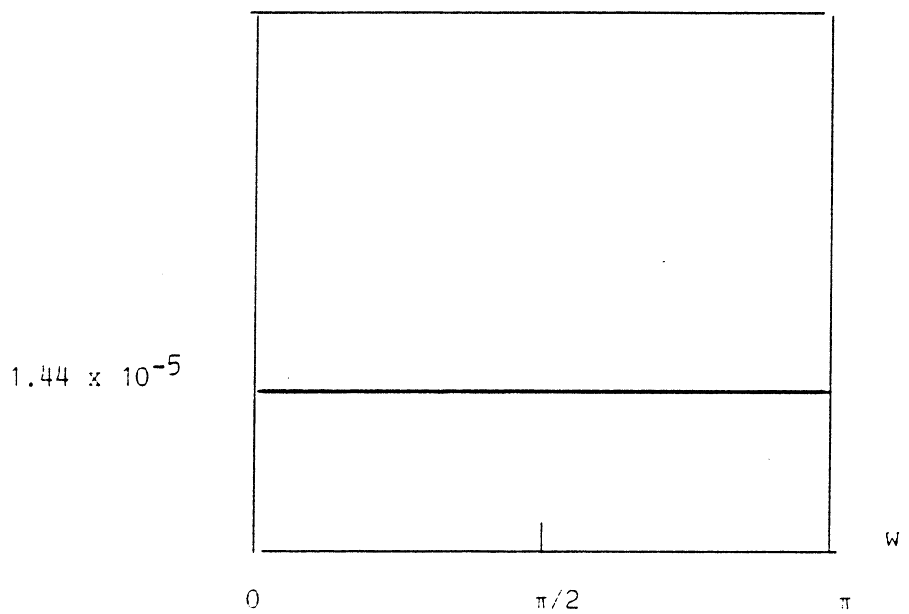

Figure 8: Spectral Gain for PRIGAS 2

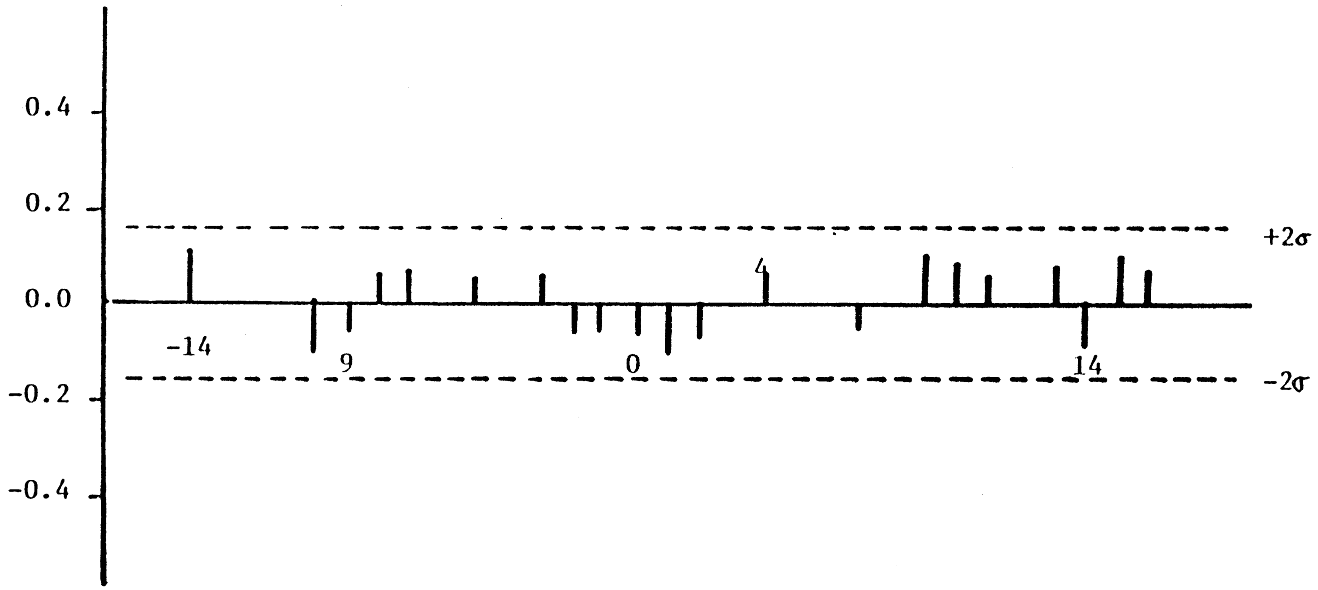

Figure 9. CCF between ACC and RAIN.

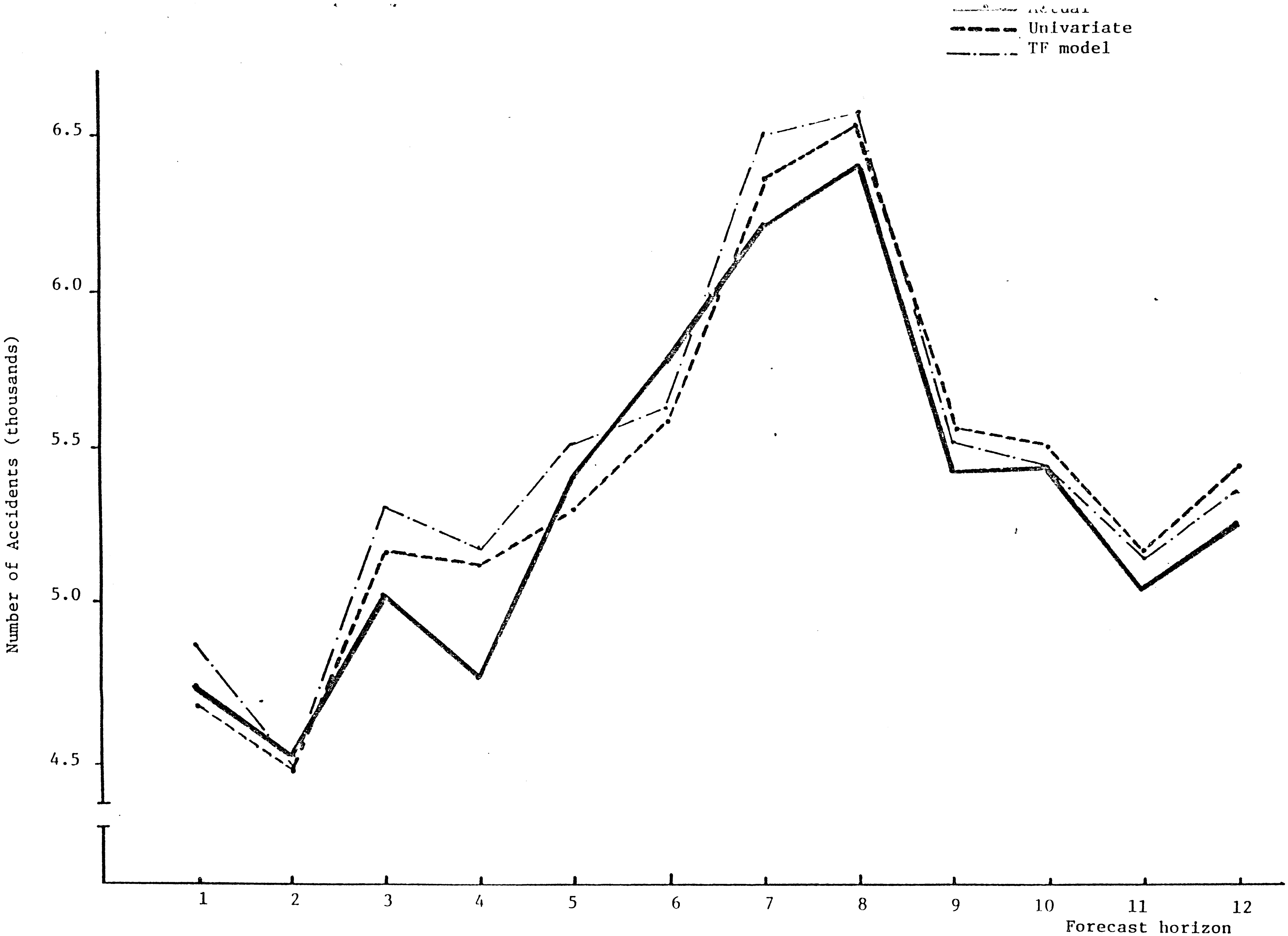

Figure 10. Actual and predicted values for the ACC variable

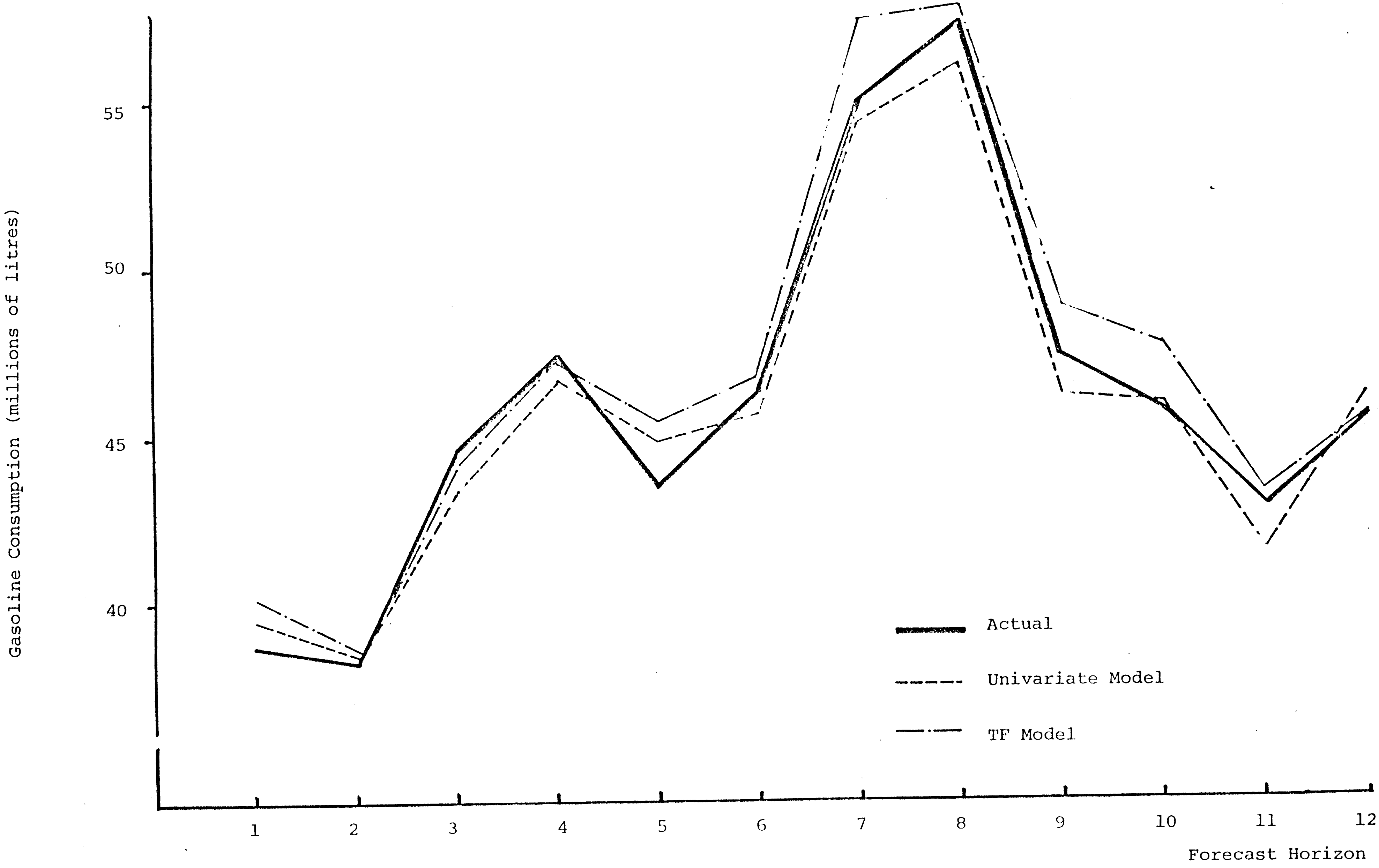

Figure 11. Actual and predicted values for the GAS variable