DISAGGREGATION AND FORECASTING: A BAYESIAN ANALYSIS

Enrique de Alba
ITAM

Yolanda Mendoza
SEMIP

DOCUMENTO DE TRABAJO

Núm. IX - 1985

DISAGGREGATION AND FORECASTING

A BAYESIAN ANALYSIS

Enrique de Alba

ITAM

and

Yolanda Mendoza

SEMIP

1985

RESUMEN

Se analiza el problema de la desagregación temporal de series mediante Métodos Bayesianos. Los valores desagregados se obtienen por medio de una distribución posterior que se deriva a partir de una distribución previa no-informativa sobre los parametros. Se demuestra que los resultados son equivalentes a los que se obtienen con métodos clásicos. Se obtienen intervalos de pronósticos Bayesianos. Los pronósticos de los valores desagregados futuros se obtienen suponiendo una distribución previa conjugada para el valor agregado futuro.

ABSTRACT

The problem of temporal disaggregation of time series is analysed by means of Bayesian methods. The disaggregated values are obtained through a posterior distribution derived using a diffuse prior on the parameters. The results are shown to be equivalent to some sampling theory results. Bayesian prediction intervals are obtained. Forecasts for future disaggregated values are derived assuming a conjugate prior for the future aggregated value.

AUTHOR FOOTNOTE*

This work was carried out while the author was visiting at the Graduate School of Business at the University of Chicago. Funds were provided by Consejo Nacional de Ciencia y Tecnología (CONACyT) and Secretaría de Hacienda y Crédito Público, both in Mexico City.

The author thanks Arnold Zellner for very enlightening comments and suggestions made during the development of the present work.

DISAGGREGATION AND FORECASTING

A BAYESIAN ANALYSIS

Enrique de Alba*
ITAM

Yolanda Mendoza
SEMIP

I. INTRODUCTION

When preparing data to carry out short term economic analysis economists are sometimes faced with the problem of having time series data of different periodicities, i.e. one series may be made up of monthly observations while another may only be available in quarterly form. In the more developed countries this will generally not be a frequent ocurrence. However in countries with incomplete data bases, with large gaps in the time series, or non existent quarterly or monthly data, it will be more likely to happen. The development of econometric models, or any form of econometric analysis involving the inconsistent series, cannot be carried out with the usual procedures. Under these circunstances a decision has to be made as to whether to continue with the analysis but using only the more time-aggregated data, say yearly, or to find procedures which allow optimal use of all the available information. If only aggregates are used some of the data will be left out. This will reduce the possibility of short-run analysis and prediction. Temporal disaggregation of time series is among the procedures which may allow the use of all the data. In some countries this may be the only way to generate statistics for subperiods.

The problem of temporal disaggregation of time series is not new. It has appeared in the literature under different names. The same basic problem has been referred to as that of "interpolation of time series" by

Friedman (1962), "adjustment of monthly or quarterly series to annual totals" by Denton (1971), "derivation of quarterly figures consistent with annual data" by Chow and Lin (1971), etc. All are concerned with constructing estimates for subperiods (months, quarters) of a time series for which only period values (years) are known. The period value is the "aggregated" value and the estimates for the subperiods are required to be consistent with it. The main approaches to the problem may be classified according to the restrictions that the subperiods must satisty in relation to the period data. The three alternatives are: interpolation, distribution and extrapolation. Interpolation consists in finding values for the component subperiods, i.e., if data exists for the end (or beginning) of a period, say a quarter, then it is required to find the values of the series at the end of each of the intermediate subperiods: the months. In Economics this would be referred to as interpolating stock data. No constraint is imposed on the estimates. Distribution is concerned with finding the subperiod values when the total or the average is known for the whole period. In this case there is a restriction on the estimates with regard to the period value: their total, or their average, must equal the aggregate. This is the case of flow data in Economics. Extrapolation is used in the usual sense of estimating values beyond the sample period. A further distinction may be made between the different procedures depending on whether or not they make use of "related" series, known for all the subperiods, to solve the problem.

The different approaches have been explored by de Alba (1979) and by Sanz (1982), both of which present a unified view of them. This is particularly important to the practitioner who finds himself facing the double problem of inadequate statistical data on one hand, and on the other hand a host of methods to solve it. There has not usually been sufficient information as to which is really best suited for particular needs. Vandaele (1978) discusses, from a Bayesian point of view, the special case of a multiple regression model where all the data is quarterly, except for one or more of the independent variables, which are only observed annually. Hsiao (1979) has reported similar results using a maximum likelihood approach. Stram and Wei (1982) present an ARIMA-model-based disaggregation procedure. In this paper we use a Bayesian approach to the problem. We also consider the question of forecasting disaggregated values. These must satisfy certain constraints similar to those imposed on the estimated values over the sample period.

2. THE PROBLEM

We shall state the problem for a particular situation, i.e. we shall specifically refer to the period and disaggregated values of the variables as the annual and quarterly data, respectively. However, the results are applicable to alternative formulations with minor changes.

Let $\underline{y}'_a = (y_{a1}, y_{a2}, \ldots, y_{an})$ a vector of annual data of the variable to be disaggregated where n is the number of years. Let X_a be a $n \times k$ matrix of annual data for each of k related, or auxiliary, variables, i.e. $X_a = \{x_{aij}\}$, $i=1,\ldots,n$, $j=1,\ldots k$. Both $\underline{y}_a$ and X_a are known.

Let $\underline{y}'_t = (y_1, y_2, \ldots, y_{4n})$ be a vector of quarterly values corresponding to the $\underline{y}_a$. These are the unknown values we want to estimate and/or forecast. Similarly X is a $4n \times k$ matrix of quarterly values for each one of the k related variables, i.e. $X = \{x_{ij}\}$, $i = 1,\ldots,4n, j=1,\ldots,k$.

We assume $\underline{y}$ and X are related by the linear model.

$$\underline{y} = X\underline{\beta} + \underline{u}, \quad \underline{u} \sim N(\underline{0}, \sigma^2 I_n), \qquad (1)$$

where $\underline{u}$ is a vector of errors, I_n is a n-dimension identity matrix, σ^2 is the unknown variance and $\underline{\beta}$ is a k-dimensional vector of unknown parame-ters. In addition it is required that the y_t, $t=1,\ldots,4n$ satisfy restrictions of the form

$$\sum_{t=4(s-1)+1}^{4s} c_t y_t = y_{as}, \qquad s = 1,\ldots,n,$$

or in matrix notation,

$$C_a \underline{y} = \underline{y}_a, \qquad (2)$$

where C_a is an $n \times 4n$ matrix of constants. In the specific case of interest $C_a = I_n \otimes \underline{\ell}'$, with $\underline{\ell}$ a 4×1 vector of constants and $\otimes$ denoting Kronecker product.

The specific form of $\underline{\ell}$ depends on whether we are doing interpolation, or distribution: In the first case $\underline{\ell}' = (1\ 0\ 0\ 0)$ and in the latter $\underline{\ell}' = (1111)$, if we want the total of the quarters to equal the annual, or $\underline{\ell}' = \frac{1}{4}(1111)$ if the average of the quarters must equal the annual. Weighted average schemes of the quarters can be implemented by choosing $\underline{\ell}$ accordingly.

The problem is then to estimate and forecast $\underline{y}$ given $\underline{y}_a$, X and X_a; subject to the constraint (2). Chow and Lin (1971) obtain a solution to the problem deriving best linear unbiased predictions subject to certain constraints. Their approach falls within the classical framework. A survey of other ad hoc procedures can be found in de Alba (1979) and in Sanz (1982).

The Bayesian solution which we present here goes beyond previous results in the sense that it not only provides estimates for $\underline{y}$. In addition we derive their distribution, as well as forecasts for future values incorporating a priori information relative to future annual values. Although the main focus is on $\underline{y}$, highest posterior density (HPD) intervals could be derived for the parameters $\underline{\beta}$ and σ^2.

3. BAYESIAN SOLUTION

Let C_a be defined as before, i.e.

$$C_a = I_n \otimes \underline{\ell}' \quad \text{with } \underline{\ell}' = (1111), \tag{3}$$

and let

$$C_o = C_2 (I_{4n} - C_a' C_a / 4) \tag{4}$$

with

$$C_2 = I_n \otimes U, \tag{5}$$

where

$$U= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} . \qquad (6)$$

If we define the $4n\times 1$ vector $\underline{y}^*$ as follows

$$\underline{y}^* = \begin{pmatrix} \underline{y}_a \\ \underline{y}_o \end{pmatrix} = C\underline{y}, \qquad (7)$$

where C is an $4n\times 4n$ matrix of the form

$$C= \begin{pmatrix} C_a \\ C_o \end{pmatrix} . \qquad (8)$$

Using $C_aC_a' = 4I_n$ it can be verified that $C_{ov}(\underline{y}_a, \underline{y}_o) = 0$ so that, given the assumption of normality, $\underline{y}_o$ and $\underline{y}_a$ are independent. We shall continue the use of normality although it is not essential for the results to hold. It is possible to work instead with a multivariate Student-t distribution, and the derivations essentially hold, Zellner (1976).

Pre multiplying (1) by C we get

$$\underline{y}^* = CX\underline{\beta} + C\underline{u} = X^*\underline{\beta} + \underline{u}^*, \qquad (9)$$

where

$$X^* = CX = \begin{pmatrix} C_a \\ C_o \end{pmatrix} X = \begin{pmatrix} X_a \\ X_o \end{pmatrix} , \qquad (10)$$

i.e., we assume that the related variables satisfy the same type of constraint as $\underline{y}$. The error term in (9) will now be $N(\underline{0},\sigma^2 CC')$ and the distribution of $\underline{y}^*$ is

$$f\ (\underline{y}^*|X^*,\underline{\beta},\sigma) = |CC'|^{-1/2}\ (^1/_2\ \pi\sigma^2)^{4n/2} \times$$

$$\exp\ \{-(\underline{y}^*-X^*\underline{\beta})'(CC')^{-1}(\underline{y}^*-X^*\underline{\beta})/2\sigma^2\}. \qquad (11)$$

But from (8) we have

$$CC' = \begin{pmatrix} C_aC_a' & 0 \\ 0 & C_oC_o' \end{pmatrix},$$

so that

$$f\ (\underline{y}_a|X^*,\underline{\beta},\sigma) = |C_aC_a'|^{-1/2}\ (^1/_2\ \pi\sigma^2)^{n/2} \times$$

$$\exp\ \{-(\underline{y}_a-X_a\underline{\beta})'(C_aC_a')^{-1}(\underline{y}_a-X_a\underline{\beta})/2\sigma^2\}.$$

Given the available data, $\underline{y}_a$ and X^*, this becomes our likelihood function. Using the fact that $C_aC_a' = 4I_n$, if we combine it with non-informativa pri ors for $\underline{\beta}$ and σ we get the posterior

$$f\ (\underline{\beta},\sigma|\underline{y}_a,X^*) \propto (1/\sigma^{n+1})\exp\{-(\underline{y}_a-X_a\underline{\beta})'(\underline{y}_a-X_a\underline{\beta})/2\times 4\sigma^2\}. \qquad (12)$$

If $\underline{y}_o$ is considered as a vector of unknown parameters we have, by definition, that

$$f\ (y_o,\beta,\sigma|y_a,X^*) = f(y_o|\beta,\sigma,Y_a)\ f(\beta,\sigma|y_a,X^*)$$

which is equivalent to

$$f\ (\underline{y}_o,\underline{\beta},\sigma|\underline{y}_a,X^*)\propto(1/\sigma^{4n+1})\ \exp\ \{-(\underline{y}_o-X_o\underline{\beta})'$$

$$(C_oC_o')^{-1}(\underline{y}_o-X_o\underline{\beta})+(\underline{y}_a-X_a\underline{\beta})'(\underline{y}_a-X_a\underline{\beta})/4\}/2\sigma^2. \tag{13}$$

On integrating out σ and $\underline{\beta}$, in that order, we have

$$f\ (\underline{y}_o|\underline{y}_a,X^*)\propto\{\underline{y}_o'(C_oC_o')^{-1}\underline{y}_o+(\underline{y}_a'\underline{y}_a/4)\ -$$

$$[\underline{y}_o'(C_oC_o')^{-1}X_o+(\underline{y}_a'X_a/4)]M^{-1}[X_o(C_oC_o')^{-1}\underline{y}_o+$$

$$(X_a'\underline{y}_a/4)]\}^{-\frac{4n-k}{n}} \tag{14}$$

with $\quad M= X_o'(C_oC_o')^{-1}X_o+(X_a'X_a/4).$ (15)

This is the posterior distribution of $\underline{y}_o$. After lengthy algebraic manipulations and simplification it can be reexpressed in the following form:

$$f\ (\underline{y}_o|\underline{y}_a,X^*)\propto\{\underline{y}_a'[W_a-X_aM^{-1}X_o'(C_oC_o')^{-1}W_o{}^{-1}$$

$$(C_oC_o')^{-1}X_oM^{-1}X_a'/16]\ \underline{y}_a+[\underline{y}_o-X_o(X_a'X_a)^{-1}X_a'\underline{y}_a]'$$

$$W_o[\underline{y}_o-X_o(X'_aX_a)^{-1}X_a'\underline{y}_a]\}^{-\frac{3n+v}{2}} \tag{16}$$

where

$$W_o= (C_oC_o')^{-1}-(C_oC_o')^{-1}X_oM^{-1}X_o'(C_oC_o')^{-1}. \tag{17}$$

Now (16) has the form of a 3n-variate Student-t with v=n-k degree of freedom from which we get the Bayesian estimator for $\underline{y}_o$,

$$E\,(\underline{y}_o|\underline{y}_a,X^*)= X_a(X_a'X_a)^{-1}X_a'\underline{y}_a= X_o\hat{\underline{\beta}}_a \tag{18}$$

where $\hat{\underline{\beta}}_a$ is the L.S. estimate of $\underline{\beta}$ based on the annual data only, and

$$\mathrm{Var}\;(\underline{y}_o|\underline{y}_a,X^*)=(vs^2/(v-2))W_o^{-1}, \tag{19}$$

with

$$vs^2= (\underline{y}_a-X_a\hat{\underline{\beta}}_a)'(\underline{y}_a-X_a\hat{\underline{\beta}}_a)/4. \tag{20}$$

Notice that we can find HPD intervals for individual elements of $\underline{y}_o$, say y_{oi}, by using their marginal posterior p.d.f.'s which can be seen to satisfy.

$$(y_{oi}-X_{o(i)}\hat{\underline{\beta}}_a)/\sqrt{h^{ii}} \sim t_v, \tag{21}$$

where $X_{o(i)}$ is the i-th row of X_o and h^{ii} is the (i,i)-th element of W_o^{-1}.

We usually want results in terms of the original observations, $\underline{y}$, rather than the deviations from the annual values, $\underline{y}_o$. These can be obtained as a Bayesian estimator, using equation (7), as

$$\hat{\underline{y}}= DE(\underline{y}^*|\underline{y}_a,X^*) \tag{22}$$

where $D= (C'C)^{-1}C'$

and

$$E\,(\underline{y}^*|\underline{y}_a, X^*) = \begin{pmatrix} \underline{y}_a \\ \\ E(\underline{y}_o|\underline{y}_a, X^*) \end{pmatrix}. \tag{23}$$

The derivation of the posterior variance of $\hat{\underline{y}}$ is straightforward keeping in mind that we are dealing with a singular distribution. It turns out to be

$$\mathrm{Var}(\hat{\underline{y}}) = D\ \mathrm{Var}(\underline{y}^*|\underline{y}_a, X^*)\ D' = \frac{vs^2}{v-2} D\, W_o^{-1} D' \tag{24}$$

with

$$W_o^{-1} = (C_oC_o') + 4X_o(X'_aX_a)^{-1}X_o'.$$

Substituting (18) into (23) and recalling the definition of $\underline{y}_a$, we have

$$\hat{\underline{y}} = D\begin{pmatrix} C_a\underline{y} \\ \\ X_o\hat{\underline{\beta}}_a \end{pmatrix} = D\begin{pmatrix} C_a\underline{y} \\ \\ C_oX\hat{\underline{\beta}}_a \end{pmatrix} = D\begin{pmatrix} C_a 0 \\ \\ 0\ C_o \end{pmatrix}\begin{pmatrix} \underline{y} \\ \\ X\hat{\underline{\beta}}_a \end{pmatrix}$$

which can be shown to be

$$\hat{\underline{y}} = C_a'\underline{y}_a/4 + X\hat{\underline{\beta}}_a - C_a'X_a\hat{\underline{\beta}}_a/4\,. \tag{25}$$

This expression is equivalent to the result given by Chow and Lin (1971), when the covariance matrix is $\sigma^2 I$. In Section 5 we use (24) and (25) with specific data to derive HPD intervals for each of the disaggregated values.

4. FORECASTING

It frequently happens that we are not only interested in disaggregation of the series alone, but would like to forecast the quarterly values for the next years, perhaps having some information about the yearly data.

To solve this problem we need to derive the predictive distribution for values outside the observed range. Let $\underline{y}_a$, C_a, and $\underline{y}$ be as defined above, y_{n+1} is the (n+1)-st annual data (unknown), and $\underline{y}_f$ is the vector of future values for the quarters in that year. Assuming the same relation between $\underline{y}$ and X as before, equation (1), and augmenting the future values, we have

$$\begin{pmatrix} y_{n+1} \\ \underline{y}_a \end{pmatrix} = C_A \begin{pmatrix} \underline{y}_f \\ \underline{y} \end{pmatrix} = C_A \begin{pmatrix} X_f \\ X \end{pmatrix} \underline{\beta} + C_A \begin{pmatrix} \underline{u}_f \\ \underline{u} \end{pmatrix} = \begin{pmatrix} \underline{x}'_{n+1} \\ X_a \end{pmatrix} \underline{\beta} + \begin{pmatrix} \underline{e}_f \\ \underline{e} \end{pmatrix} \tag{26}$$

where $C_A = \begin{pmatrix} 0 & \underline{\ell}' \\ C_a & 0 \end{pmatrix}$,

X_f is the matrix of future values of the independent variables, assumed known, $\underline{u}_f$ is the vector of future errors, and $\underline{x}'_{n+1} = \underline{\ell}' X_f$ is the vector of future annual values of the auxiliary variables. It is readily seen that $\underline{e}$ and $\underline{e}_f$ are independent.

Considering y_{n+1} as an unknown parameter and assuming normality of $\underline{e}$ and $\underline{e}_f$ in (26) we can write

$$f(\underline{y}_a|y_{n+1},\underline{\beta},\sigma^2)\propto(1/\sigma^{n+1})\exp\{-[(\underline{y}_a-X_a\underline{\beta})'(\underline{y}_a-X_a\underline{\beta})+(y_{n+1}-\underline{x}'_{n+1}\underline{\beta})^2/8\sigma^2\} \quad (27)$$

This is the likelihood function of the parameters given the annual data. Now assume we have some information on y_{n+1} which takes the form of a natural conjugate prior, independent of $\underline{\beta}$ and σ, i.e.

$$f(y_{n+1},\underline{\beta},\sigma)= f(\underline{\beta},\sigma)f(y_{n+1})\propto(1/\sigma)(1/\sigma_o\sqrt{2\pi})\exp\{-(y_{n+1}-\mu)^2/2\sigma^2\} \quad (28)$$

where μ is the known prior mean.

Combining (27) and (28) we get

$$f(\underline{\beta},\sigma,y_{n+1}|\underline{y}_a)\propto(1/\sigma^{n+2}\sigma_o)$$

$$\exp\{-[(\underline{y}_a-X_a\underline{\beta})'(\underline{y}_a-X_a\underline{\beta})/4+(y_{n+1}-\underline{x}'_{n+1}\underline{\beta})^2/4+(y_{n+1}-\mu)^2/\lambda]/2\sigma^2\} \quad (29)$$

which is the posterior of the parameters given the annual data and $\lambda=\sigma_o{}^2/\sigma^2$.

Let $\underline{y}_{fo}= C_f\underline{y}_f$ with $C_f= U(I_4-\underline{\ell}\underline{\ell}'/4)$ and

$$U=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\end{pmatrix}.$$

Notice that while $\underline{y}_f$ is a 4×1 vector of future quarterly values, $\underline{y}_{fo}$ is the 3×1 vector of deviations of the first three elements in $\underline{y}_f$ from their mean, we lose one degree of freedom. We can apply a similar transformation to the future values of the independent variables. Hence using $X_{fo} = C_f X_f$, we have the predictive distribution for $\underline{y}_{fo}$

$$f(\underline{y}_{fo}|y_{n+1},\sigma^2,\underline{\beta},X_{fo}) \propto (1/\sigma^3)\exp\{-(\underline{y}_{fo}-X_{fo}\underline{\beta})'(C_f C_f')^{-1}(\underline{y}_{fo}-X_{fo}\underline{\beta})/2\sigma^2\} \quad (30)$$

Combining formulas (29) and (30) we get, after integrating out y_{n+1}, σ and β,

$$f(\underline{y}_{fo}|\mu,X_{fo},\underline{y}_a) \propto \{\underline{\overline{Y}}(\underline{\overline{V}}^{-1}-\underline{\overline{X}}(\underline{\overline{X}}'\underline{\overline{V}}^{-1}\underline{\overline{X}})^{-1}\underline{\overline{X}}')\underline{\overline{Y}}+(\underline{y}_{fo}-X_{fo}\underline{\tilde{\beta}}_A)'W_o(\underline{y}_{fo}-X_{fo}\underline{\tilde{\beta}}_A)\}^{-\frac{v+3}{2}}$$

This is a multivariate Student-t distribution with $v=n-k+1$ degrees of freedom. In these expressions we have

$$\underline{\tilde{\beta}}_A = (\underline{\overline{X}}'\underline{\overline{V}}^{-1}\underline{\overline{X}})^{-1}\underline{\overline{X}}'\underline{\overline{V}}^{-1}\underline{\overline{Y}},$$

$$\underline{\overline{Y}}' = (\mu\ \underline{y}_a'),$$

$$\underline{\overline{X}}' = (\underline{x}_{n+1}'X_a),$$

$$\underline{\overline{V}}^{-1} = \text{diag}\ \{1/(\lambda+4)\cdot I_n \quad (1/4)\ I_n\},$$

$$W_o^{-1} = V_{fo}+X_{fo}(\underline{\overline{X}}'\underline{\overline{V}}^{-1}\underline{\overline{X}})^{-1}X_{fo}' \text{ and } V_{fo} = I_3-(1/4)\ G_3, \text{ with}$$

$$G_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Notice that in this case

$$vs^2 = \underline{\overline{Y}}'[\underline{\overline{V}}^{-1} - \underline{\overline{X}}(\underline{\overline{X}}'\underline{\overline{V}}^{-1}\underline{\overline{X}})^{-1}\underline{\overline{X}}']\underline{\overline{Y}}, \tag{31}$$

and $\underline{\tilde{\beta}}_A$ can be shown to be a GLS estimator for $\underline{\beta}$ obtained by using μ instead if y_{n+1} and weighting the observed annual values by $(1/4)$ and the mean of the prior distribution of y_{n+1} by $1/(\lambda+4)$. The predictive variance of $\underline{y}_{fo}$ is $[v/(v-2)]s^2 W_o^{-1}$. Under the conjugate prior assumption, further analysis shows that the mean of the predictive distribution for the original quarterly values $\underline{y}_f$ is

$$E(\underline{y}_f|\underline{\overline{X}}^*) = X_f\underline{\tilde{\beta}}_A$$

where $\underline{\overline{X}}^*$ is the set of all the sample information available, i.e. $\underline{\overline{X}}^* = \{X_f, X, X_a, y_a\}$. The variance of the predicted quarterly values is

$$\mathrm{Cov}(\underline{y}_f|\underline{\overline{X}}^*) = \{[I_4(\lambda+4)-G_4][I_4 + \frac{1}{\lambda}G_4 + 4X_f(\underline{\overline{X}}'\underline{\overline{V}}^{-1}\underline{\overline{X}})^{-1} X_f'][I_4(\lambda+4)-G_4]$$

$$/(\lambda+4)\}(v/(v-2))s^2 .$$

Furthermore the distribution of $\underline{y}_f$ is a multivariate Student-t distribution with v degrees of freedom, v=n-k+1. We can use this to obtain marginal predictive intervals for the individual y_f's.

5. AN EXAMPLE

In this section we apply the results derived so far to a particular situation. The matrix computations for the disaggregation were carried out using GEM while the forecasts were obtained by means of TSP version 3.5, both on an IBM-370 computer. The series we want to disaggregate is GNP of Mexico and the related series is an Index of Industrial Production, which is a component of GNP. Since the average of the quarterly values of the index must equal the yearly figure, and the above results are given for the case where the SUM equals the annual, we simply consider a yearly index which is the sum of the quarters. This does not affect the results. Table 1 shows the data.

Using formula (25) to disaggregate GNP we obtain the results of Table 2 below. Application of (24) yields a variance of approximately 7.270 for each one of the values. This almost constant variance can be explained as follows. The variance formula is equivalent to the following expression

$$\mathrm{Var}(\underline{y}) = [vs^2/(v-2)][I_n \otimes (I_4 - G_4/4)][I_{4n} + X(X_a'X_a)^{-1}X'/4][I_n \otimes (I_4 - G_4/4)] \cdot$$

With the particular set of data we are using it turns out that the term $X(X_a'X_a)^{-1}X'$ is rather small so that the remaining terms, which are constant, dominate. Hence the 95% Bayesian intervals have a length of 5.28. This is very reasonable considering the order of magnitude of the quantities. The fact that the Bayesian intervals do not get wider as we move away from the mean, as in regression, may also be due to the fact that in this method the X_a and y_a are given. The intervals are plotted along with the estimates in Figures 2 to 5.

We now ilustrate the results on forecasting. We could, of course, use some standard procedures and derive future values without any regard to available prior information nor restrictions. This was actually done using the Holt-Winters method to obtain forecasts for the next eight periods. They are shown in Figure 1. As is usual in these cases, they reflect a slight seasonality and the strong trend underlying the historical observations.

Assume now that we do have some prior information with respect to GNP in the following year. This is not a very restrictive assumption, since in most countries government economists usually have a fairly good idea of what the growth of GNP will be in the coming year. In the specific case of Mexico we assume that in 1981 the prevailing belief was that in 1982 GNP would decrease by 2%. We use this figure to set the prior mean, μ, in formula (28), i.e. $\mu=889.4$. An additional piece of information we need to specify is the value of σ_0^2, the prior variance of the future yearly figure. A reasonable value for λ would be 4. This is based on the assumption that $Var(y_{n+1})$ equals the variance of the historical annual data or, equivalently, four times the quarterly variance. However we consider alternative values in order to get a better feeling of the solution. We also need forecasts for the related variables, X_f, which we take as given, see Table 3. This table shows the predicted values for the disaggregated series under different assumptions. The upper portion of the table gives results assuming a decrease of 2% in Mexican GNP for 1982. The lower portion considers the alternative of zero growth in GNP. On the other hand, in each portion we present results for alternative values of λ, the variance ratio, along with those obtained under standard procedures.

Notice that with a large value of λ, say 100, the adding up restriction for the future disaggregated values is ignored. This is the result of having a kind of "very low" information prior, since we have a very large prior variance on y_{n+1}. In fact, if we look at the yearly value obtained by adding the quarterly figures it is even larger than that of the unrestricted case. On the other hand, when λ is small, say .01, the constraint is satisfied. Here we have strong prior beliefs about our future yearly value. Figures 2 to 5 show the different effects graphically. It is clear thar strong priors on our future annual values induce a change in the trend of the predicted values with respect to the historical data. This is precisely what we would expect. So what we have here can be seen as a means of predicting a change in trend by means of prior information.

REFERENCES

1. Chow, G.C. and S. Lin (1971), "Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series", Review of Economics and Statistics, Vol. 53, 372-75.

2. de Alba, E. (1979), "Temporal Disaggregation of Time Series, a Unified Approach", Proc. of the Bus. and Eco. Statistics Section, ASA-Washington, D.C.

3. Denton, F. (1971), "Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization", Journal of the American Statistical Association, Vol. 66, 99-102.

4. Friedman, M. (1962), "The Interpolation of Time Series by Related Series", Journal of the American Statistical Association, Vol. 57, 729-57.

5. Hsiao, C. (1979), "Linear Regression Using both Temporally Aggregated and Temporally Disaggregated Data", Journal of Econometrics, Vol. 10, 243-252.

6. Sanz, R. (1981), "Métodos de Desagregación Temporal de Series Económicas", Servicio de Estudios, Banco de España.

7. Vandaele, W. (1978), "Stock and Flow Unobservables. A Bayesian Approach", Grad. School of Bus. Adm., Harvard University Working Paper HBS78-56.

8. Zellner, A. (1976), "Bayesian and non-Bayesian Analysis of the Regression Model with Multivariate Student-t Error Terms", J.A.S.A. 71, 400-405.

TABLE 1

		Industrial Production Index (1970=100)		GNP Year Total (1970 constant pesos)
		Quarter	Year	
1970	1	96.65		
	2	103.52		
	3	100.00		
	4	99.83	400.00	444.271
1971	1	103.54		
	2	101.17		
	3	100.87		
	4	102.66	408.24	462.804
1972	1	108.10		
	2	114.66		
	3	113.33		
	4	113.55	449.68	502.086
1973	1	118.75		
	2	122.61		
	3	125.42		
	4	128.16	494.96	544.307
1974	1	132.10		
	2	133.5		
	3	131.38		
	4	134.25	531.24	577.568
1975	1	131.49		
	2	143.79		
	3	140.32		
	4	140.67	556.28	609.976
1976	1	144.27		
	2	145.67		
	3	143.92		
	4	137.61	571.48	635.831
1977	1	139.35		
	2	149.90		
	3	151.42		
	4	150.50	591.16	657.721
1978	1	151.20		
	2	166.55		
	3	167.38		
	4	165.10	650.24	711.211
1979	1	172.10		
	2	178.50		
	3	182.18		
	4	184.71	717.48	776.642
1980	1	187.20		
	2	194.07		
	3	197.27		
	4	198.87	777.40	841.103
1981	1	199.67		
	2	214.17		
	3	215.40		
	4	207.63	836.88	907.550

TABLE 2

QUARTERLY GNP

(BILLION 1970 PESOS)

1970	1	107.6	1976	1	160.4
	2	114.7		2	161.9
	3	111.1		3	160.1
	4	110.9		4	153.5
1971	1	117.2	1977	1	155.6
	2	114.8		2	166.6
	3	114.4		3	168.2
	4	116.3		4	167.3
1972	1	121.0	1978	1	165.9
	2	127.9		2	182.0
	3	126.5		3	182.8
	4	126.7		4	180.5
1973	1	130.9	1979	1	186.6
	2	134.9		2	193.2
	3	137.8		3	197.1
	4	140.7		4	199.7
1974	1	143.7	1980	1	202.8
	2	145.1		2	210.0
	3	142.9		3	213.3
	4	145.9		4	215.0
1975	1	144.6	1981	1	216.9
	2	157.4		2	232.1
	3	153.8		3	233.4
	4	154.2		4	225.2

TABLE 3

ΔGNP= -2%

		λ=.01	.5	1	100	WINTERS	INDEX
1982	I	214.5	220.9	226.0	254.4	233.6	217.7
	II	226.0	232.5	237.5	266.6	246.0	221.0
	III	226.9	233.4	238.4	267.5	247.0	224.3
	IV	222.6	229.1	234.2	263.0	244.0	227.6
TOTAL		890.0	915.9	936.1	1 051.5	970.6	890.6

ΔGNP= 0%

		λ=.01	.5	1	100	WINTERS	INDEX
1982	I	218.9	224.4	228.6	254.3	233.6	217.7
	II	230.6	236.0	240.3	266.5	246.0	221.0
	III	231.4	236.9	241.2	267.4	247.0	224.3
	IV	227.2	232.7	236.9	262.9	244.0	227.6
TOTAL		908.1	930.0	947.0	1 051.1	970.6	890.6

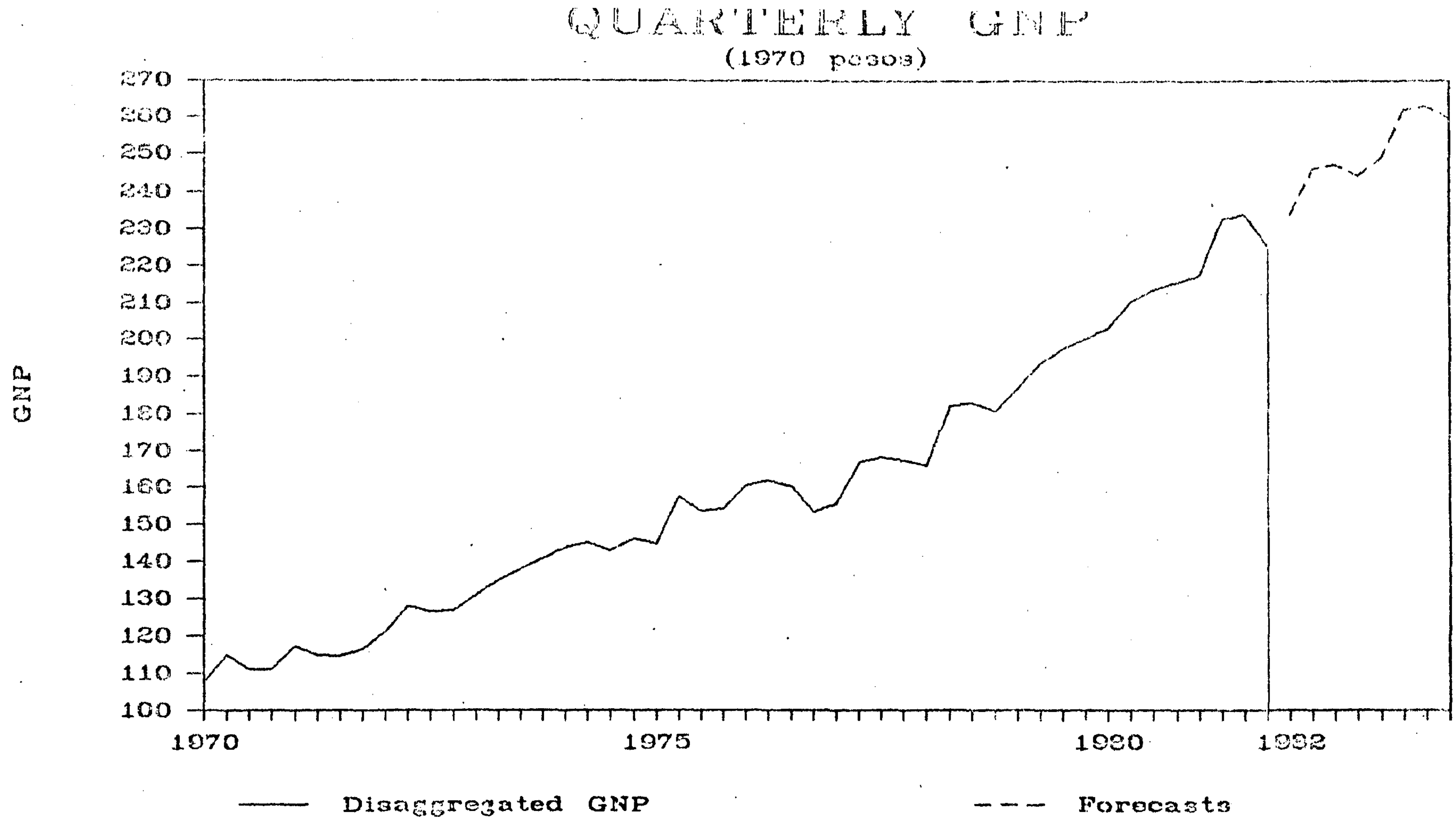

QUARTERLY GNP
(1970 pesos)
GNP
270
260
250
240
230
220
210
200
190
180
170
160
150
140
130
120
110
100
1970
1975
1980
1982
Disaggregated GNP
Forecasts

FIGURE 1

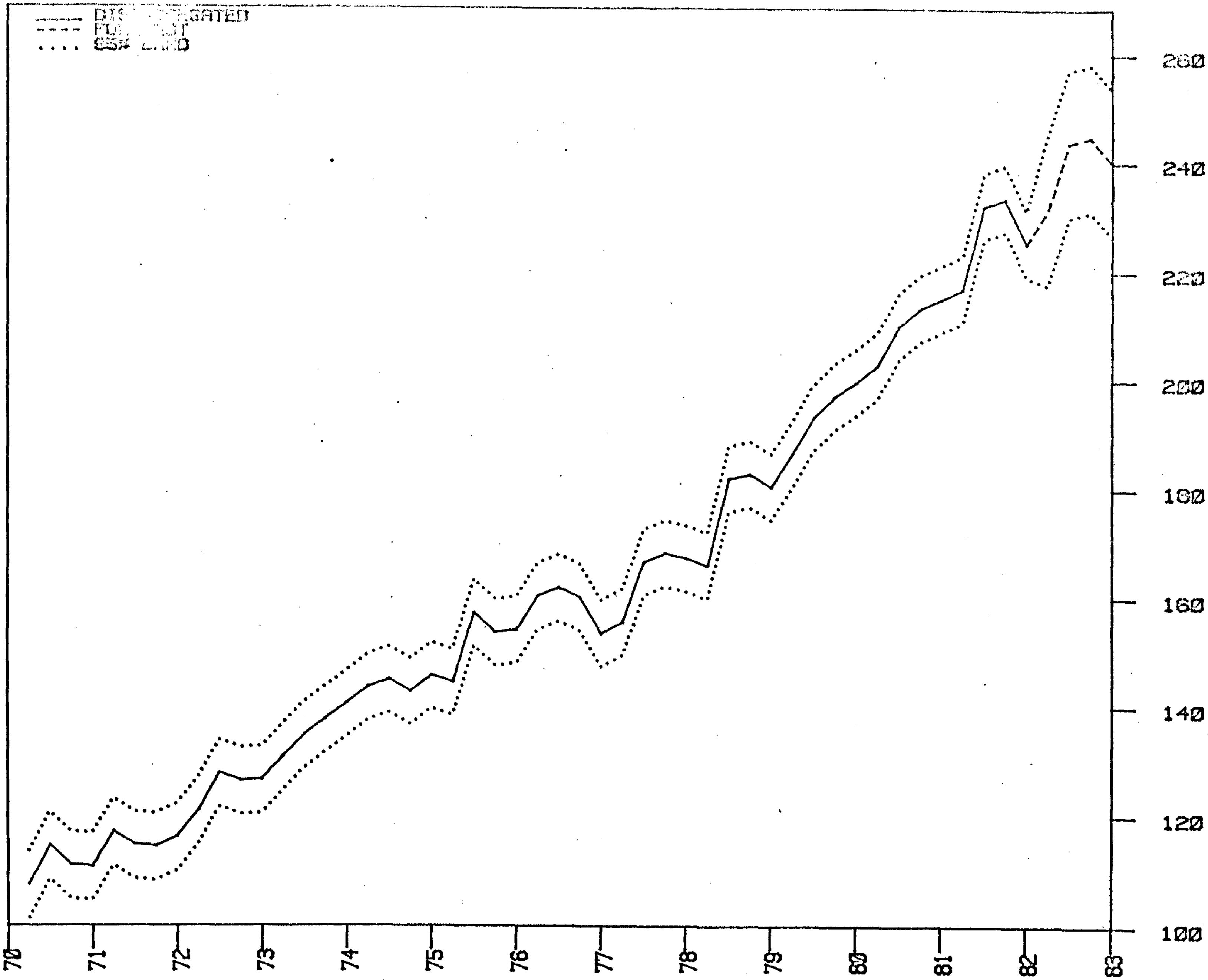
DISAGGREGATED GNP

VARIANCE RATIO = 100

PRIOR MEAN FOR NEXT YEAR = ZERO GROWTH

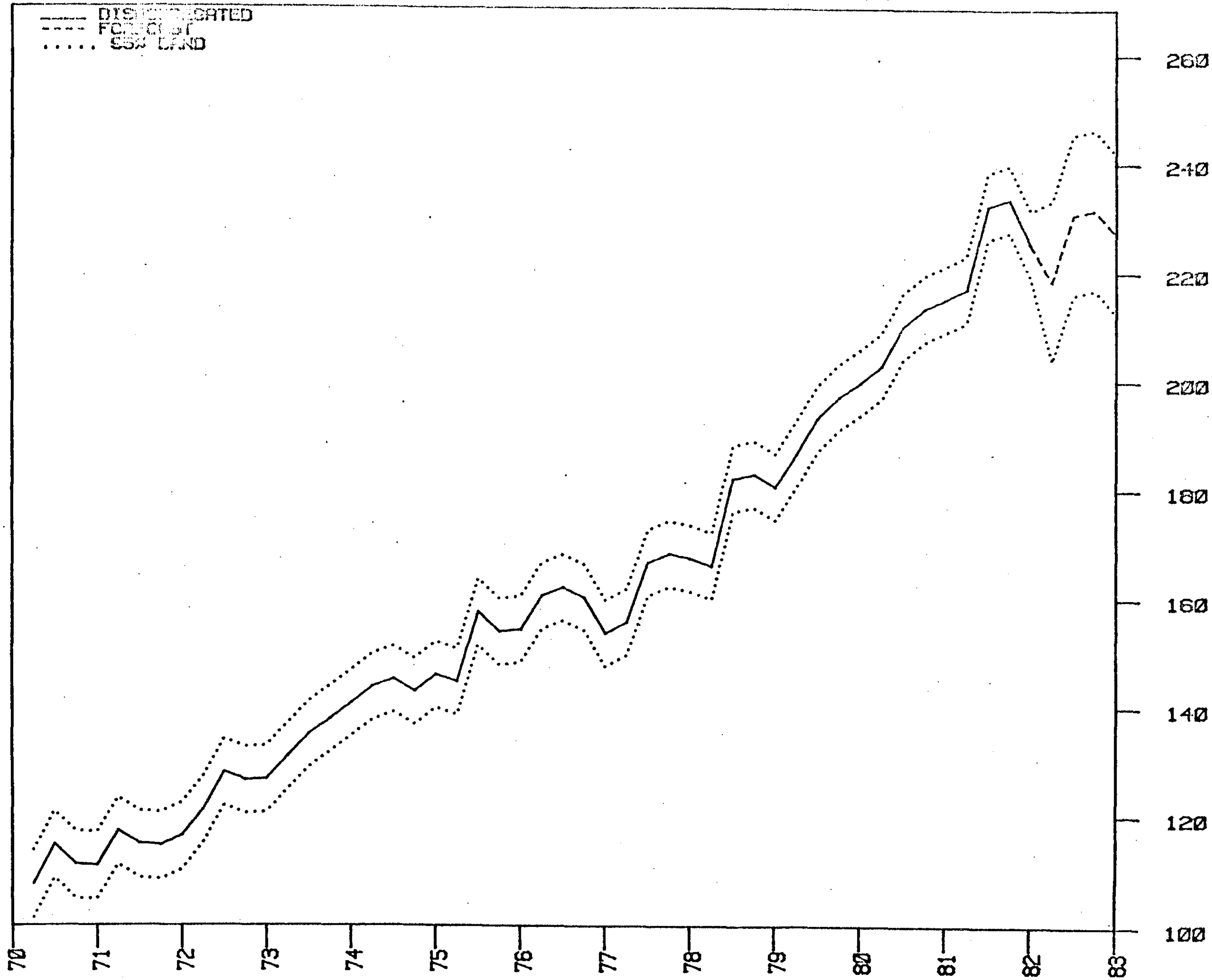
S.A.I.

D.G.S.T.G.



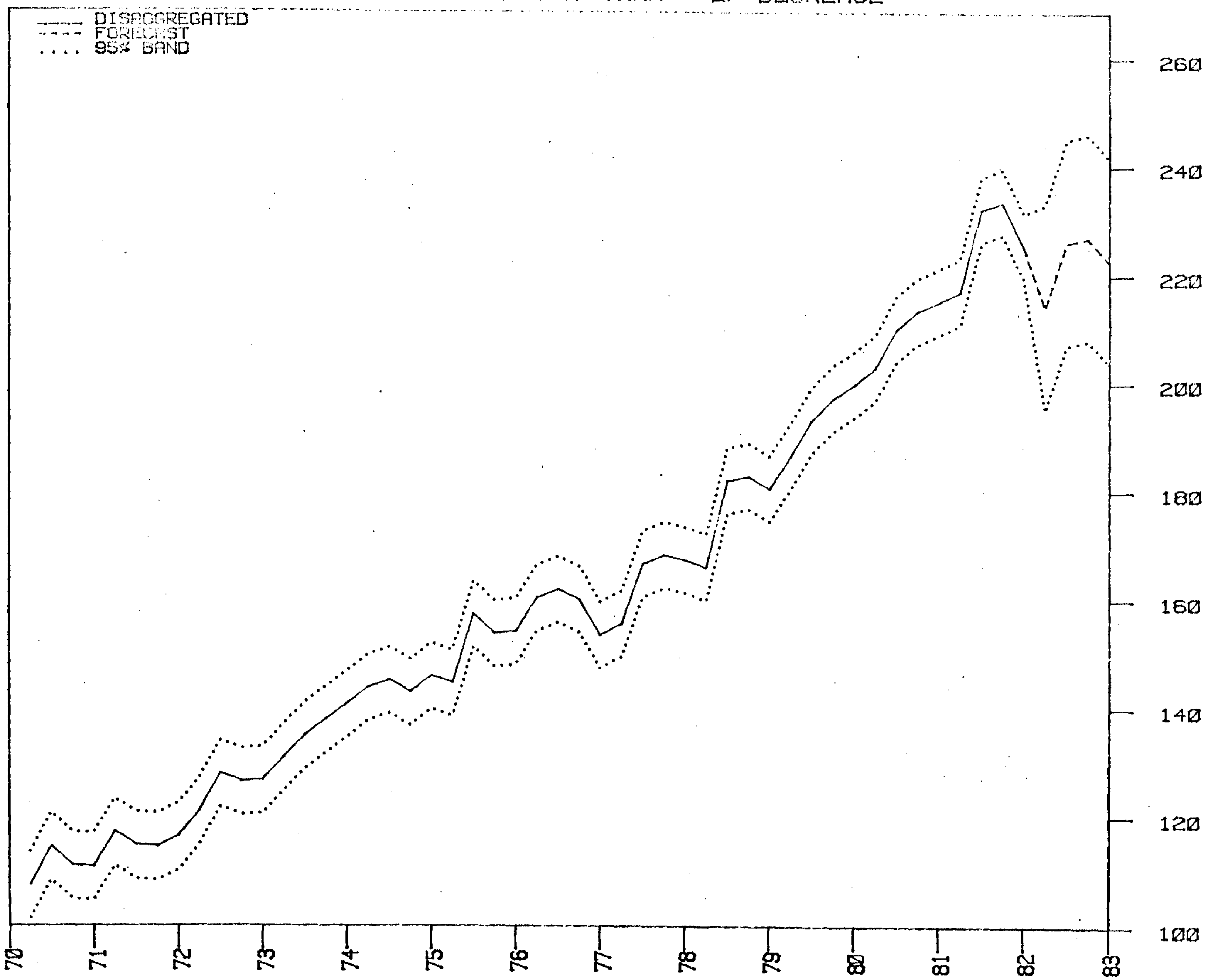
DISAGGREGATED GNP
VARIANCE RATIO = 0.01
PRIOR MEAN FOR NEXT YEAR = ZERO GROWTH

S.A.I.
D.G.S.T.G.



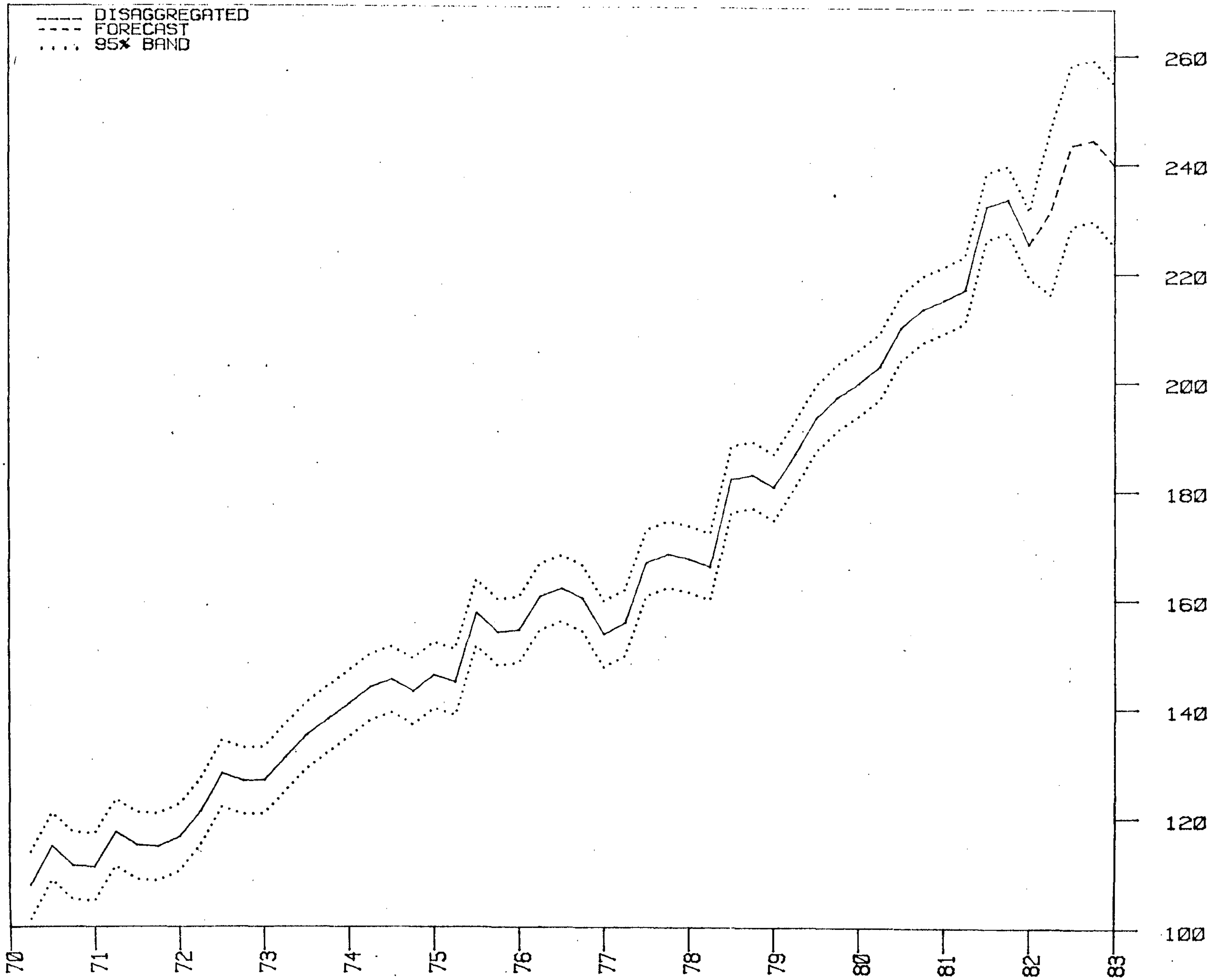
S.A.I.
D.G.S.T.G.

DISAGGREGATED GNP
VARIANCE RATIO = Ø.Ø1
PRIOR MEAN FOR NEXT YEAR = 2% DECREASE



S.A.I.
D.G.S.T.G.

DISAGGREGATED GNP
VARIANCE RATIO = 100
PRIOR MEAN FOR NEXT YEAR • 2% DECREASE



Serie Documentos de Trabajo 1985

No. I Bhaduri, Amit, "The Race in Arms: its Mathematical Commonsense".

No. II Garber, Peter M., and Vittorio U. Grilli, "The Belmont-Morgan Syndicate as an Optimal Investment Banking Contract".

No. III Ros, Jaime, "Trade, Growth and the Pattern of Specialization".

No. IV Nadal, Alejandro, "El Sistema de Precios de Producción y la Teoría Clásica del Mercado".

No. V Alberro, José Luis, "Values and Prices in Joint Production: Discovering Inner-Unproductivities".

No. VI De Urquijo Hernández, Luis Alfredo, "Las Políticas de Ajuste en el Sector Externo: Análisis de un Modelo Computable de Equilibrio General para la Economía Mexicana".

No. VII Castañeda Sabido, Alejandro I., "La Proposición de Inefectividad de la Nueva Macroeconomía Clásica: Un Estudio Crítico".

No. VIII De Alba, Enrique y Ricardo Samaniego, "Estimación de la Demanda de Gasolinas y Diesel y el Impacto de sus Precios sobre los Ingresos del Sector Público".

No. IX De Alba, Enrique y Yolanda Mendoza, "Disaggregation and Forecasting: A Bayesian Analysis"

Serie Documentos de Trabajo 1984

No. I Alberro, José Luis, "Introduction and Benefit of Technological Change under Oligopoly"

No. II Serra-Puche, Jaime y Ortíz, Guillermo, "A Note on the Burden of the Mexican Foreign Debt"

No. III Bhaduri, Amit, "The Indebted Growth Process"

No. IV Easterly, William, "Devaluation in a Dollarized Economy"

No. V Unger, Kurt, "Las Empresas Extranjeras en el Comercio Exterior de Manufacturas Modernas en México"

No. VI De Alba, Enrique y Mendoza, Yolanda, "El Uso de Modelos Log-Lineales para el Análisis del Consumo Residencial de Energía"

No. VII García Alba, Pascual, "Especificación de un Sistema de Demanda y su Aplicación a México"

No. VIII Nadal, Alejandro y Salas Páez, Carlos, "La Teoría Económica de la Sociedad Descentralizada", (Equilibrio General y Agentes Individuales).

No. IX Samaniego Breach, Ricardo, "The Evolution of Total Factor Productivity in the Manufacturing Sector in Mexico, 1963-1981"

No. X Fernández, Arturo M., "Evasión Fiscal y Respuesta a la Imposición: Teoría y Evidencia para México"

No. XI Ize, Alain, "Conflicting Income Claims and Keynesian Unemployment"

Serie Documentos de Trabajo 1983

No. I Bhaduri, Amit, "Multimarket Classification of Unemployment"

No. II Ize, Alain y Salas, Javier, "Price and Output in the Mexican Economy: Empirical Testing of Alternative Hypotheses"

No. III Alberro, José Luis, "Inventory Valuation, Realization Problems and Aggregate Demand"

No. IV Sachs, Jeffrey, "Theoretical Issues in International Borrowing"

No. V Ize, Alain y Ortíz, Guillermo, "Political Risk, Asset Substitution and Exchange Rate Dynamics: The Mexican Financial Crisis of 1982"

No. VI Lustig, Nora, "Políticas de Consumo Alimentario: Una Comparación de los Efectos en Equilibrio Parcial y Equilibrio General"

No. VII Seade, Jesús, "Shifting Oligopolistic Equilibria: Profit-Raising Cost Increases and the Effects of Excise Tax"

No. VIII Jarque, Carlos M., "A Clustering Procedure for the Estimation of Econometric Models with Systematic Parameter Variation"

No. IX Nadal, Alejandro, "la Construcción del Concepto de Mercancía en la Teoría Económica"

No. X Cárdenas, Enrique, "Some Issues on Mexico's Nineteenth Century Depression"

No. XI Nadal, Alejandro, "Dinero y Valor de Uso: La Noción de Riqueza en la Génesis de la Economía Política"

No. XII Blanco, Herminio y Garber, Peter M., "Recurrent Devaluation and Speculative Attacks on the Mexican Peso"

El Centro de Estudios Económicos de El Colegio de México, ha creado la serie "Documentos de Trabajo" para difundir investigaciones que contribuyen a la discusión de importantes problemas teóricos y empíricos aunque estén en versión preliminar. Con esta publicación se pretende estimular el análisis de las ideas aquí expuestas y la comunicación con sus autores. El contenido de los trabajos es responsabilidad exclusiva de los autores.

Editor: José Luis Alberro

Serie Documentos de Trabajo 1982

No. I Ize, Alain, "Disequilibrium Theories, Imperfect Competition and Income Distribution:

No. II Levy, Santiago, "Un Modelo de Simulación de Precios para la Economía Mexicana"

No. III Persky, Joseph and Tam, Mo-Yin S., "On the Theory of Optimal Convergence"

No. IV Kehoe, Timothy J., Serra-Puche, Jaime y Solís, Leopoldo, "A General Equilibrium Model of Domestic Commerce in Mexico"

No. V "Guerrero, Víctor M., "Medición de los Efectos Inflacionarios Causados por Algunas Decisiones Gubernamentales: Teoría y Aplicaciones de Análisis de Intervención"

No. VI Gibson, Bill, Lustig, Nora and Taylor, Lance, "Terms of Trade and Class Conflict in a Computable General Equilibrium Model for Mexico"

No. VII Dávila, Enrique, "The Price System in Cantillon's Feudal Mercantile Model"

No. VIII Ize, Alain, "A Dynamic Model of Financial Intermediation in a Semi-Industrialized Economy"

No. IX Seade, Jesús, "On Utilitarianism and Horizontal Equity: When is the Equality of Incomes as such Desirable?"

No. X Cárdenas, Enrique, "La Industrialización en México Durante la Gran Recesión: Política Pública y Respuesta Privada"