TIME SERIES MODELING IN ECONOMICS

Workshop Proceedings No. 1

November 1987

edited by Robert M. Kunst

Contributions by (in alphabetical order):

Peter BRANDNER
Wolfgang A. FELS
Albert JAEGER
Robert M. KUNST
Klaus NEUSSER
Benedikt M. PÖTSCHER
Gerhard THURY
(affiliations see inside)

Preface

The present memorandum is a printed summary of speeches held at a workshop about Time Series Modeling and Economic Data at the Institute for Advanced Studies last November. It starts a memoranda series which has been designed especially for similar events which are to take place at the Institute irregularly but not infrequently in the future. The purpose of these workshops is to empossible the exchange of results of personal research on special topics. All but three of this volume's contributions were presented by members of this Institute but this should not be the rule for future workshops and presentation of work by "outside" researchers will be especially promoted.

The tentative character of workshop speeches implies that the contributions are quite heterogeneous. The spectrum ranges from "theoretical" to "empirical" and from provisional results to papers mature for publication. Some papers are on their way to being published in journals, so note the proviso on the title page. Let me express once more my thanks to the contributors without whose kind collaboration this volume (and the preceding workshop) would not have been possible.

Robert M. Kunst

April 1988

Contents

1 Institute for Advanced Studies, Vienna, Austria
2 Department of Economics, University of Vienna, Austria
3 Department of Econometrics, University of Technology, Vienna, Austria
4 Institute for Economic Research, Vienna, Austria

PRIOR SPECIFICATION AND FORECASTING PERFORMANCE: A BVAR MODEL OF AUSTRIA'S MONETARY SECTOR

Peter Brandner

Institute for Advanced Studies

Vienna

March 1988

PRIOR SPECIFICATION AND FORECASTING PERFORMANCE: A BVAR MODEL OF AUSTRIA'S MONETARY SECTOR

Peter Brandner

Institute for Advanced Studies

Vienna

March 1988

ABSTRACT

Bayesian vector autoregressive (BVAR) models have become an attractive alternative to structural econometric models or to time series techniques in order to forecast economic time series. Since the forecasting performance of BVAR models depends on the prior information, different procedures specifying the metaparameters are considered. This investigation focuses on those metaparameters which control the interaction among the variables of a model. Incorporating more economic theory rather than extracting the prior information from the data increases the predictive accuracy when the BVAR techniques are applied to monetary time series of the Austrian economy.

1. INTRODUCTION

The function of a model designed for forecasting purposes is to extract the relevant information from a set of data or, in other words, to filter the economic signal out of noisy time series. Unfortunately, forecasters are not only confronted with a limited amount of data but also face a high degree of uncertainty about the understanding of the economy's structure which, additionally, will perhaps change over time.

The standard approach to equation specification is to include only a few variables in line of a given economic theory while excluding the rest to keep the resulting econometric "structural" model identified. This practice has often been critizised among others by Sims (1980), Lucas and Sargent (1981) and Litterman (1986) both for theoretical as well as empirical reasons. Alternatively, Sims (1980) recommends the use of unrestricted vector autoregressive (VAR) models. Since unrestricted VAR models require each variable in each equation with the same lag specification, even in small models the available degrees of freedom are quickly exhausted. Such an overparameterized system will provide a good-in-sample fit but a bad out-of-sample forecasting performance because the parameters will fit the systematic relationship as well as random variations in the data. In view of these problems a lot of suggestions imposing restrictions on parameters have been made. Kling and Bessler (1985) described several types of restrictions to reduce the dimensionality of a VAR model.

This paper follows Litterman (1980, 1986) and Doan, Litterman, and Sims (1984) to apply Bayesian techniques directly to the estimation of the VAR coefficients. Litterman's approach combines prior and sample information via a shrinkage type of estimator. Basically, the procedure is to formulate the prior information as a set of dummy observations that is incorporated in the estimation process using Theil's (1971) mixed estimator.

However, the forecasting performance of such a BVAR model clearly depends on the choice of certain metaparameters which represent the prior. Special interest is put on those metaparameters that control the interaction among the variables. It is the purpose of this investigation to evaluate different strategies in specifying them. One procedure is to utilize statistical techniques and extract prior information from the data, the second procedure is to stress economic theory or rely on economic arguments in order to set the metaparameters.

Another aspect of this paper is the application of this approach to monetary time series. For a lot of reasons BVAR models seem to be well suited for forecasting financial variables especially when they are available on a weekly or monthly basis. In the Bayesian specification framework there are no limitations due to the degrees of freedom which arise in a classical framework. Since financial and monetary markets can hardly be characterized by a stable relationship over a longer period of time, one can account for important financial innovations or policy changes simply by dropping older observations. From a practical point of view, once a BVAR model is specified it allows one to process newly available information in a quite efficient way both for estimation as well as for forecasting. This gain comes up because the repeated construction of the model can be avoided (Neusser, 1986).

This paper is organized as follows: In section 2 the BVAR methodology along the line of Litterman (1986) is briefly reviewed. A model for the monetary sector of the Austrian economy is introduced in section 3 and different strategies for specifying metaparameters are discussed in section 4. In section 5 the forecasting performance of various prior specifications are compared. Finally, section 6 summerizes the results.

2. BAYESIAN VECTOR AUTOREGRESSIONS

The BVAR approach starts with a p-th order autoregressive representation of a m-dimensional time series vector Y giving the i-th equation the form:

$$(1) \qquad Y_{it} = c_{it}'\alpha_i + \sum_{j=1}^{m} \sum_{k=1}^{p} \beta_{ijk} Y_{it} + \epsilon_{it} \qquad i=1,2,\dots,m$$

where c_{it} is a vector of deterministic variables and can include the constant, trends, and dummies; ϵ_{it} is an error term distributed $N(0,\sigma_i^2)$. The prior is specified as a multivariate normal distribution for the coefficients β_{ijk} and can therefore be completely characterized by the mean and the variance-covariance matrix. No prior is given to the coefficients of the deterministic components, so they are estimated freely. The other parameters are all assumed to have means of zero except the coefficient on the first lag of the dependent variable which has a prior mean of one:

$$(2) \qquad \beta_{ijk} = \begin{cases} 1 & \text{for } i=j \text{ and } k=1 \\ 0 & \text{otherwise} \end{cases}$$

This specification of the prior means is justified on the observation that the behaviour of many economic time series can be approximated by a random walk around an unknown deterministic component (Nelson and Plosser (1982)). In a Bayesian framework one would expect an investigator to postulate some relationships among a model's variables derived from a particular economic theory. But as Litterman (1986) has already noted, the multivariate random walk specification has the ability to capture more accurately uncertain a priori information than the standard methods of restricting VAR representations or identifying equations of econometric models. Information about relationships among variables should be imposed in the estimation process.

Assuming the parameters to be uncorrelated with each other, the specification problem for the covariance matrix can be considerably reduced by setting all covariances equal to zero. Nevertheless, there still remain $(m \cdot m \cdot p)$ variances to be specified. A further reduction can be achieved by generating the variances as a function of only a few metaparameters. In particular, the standard deviation of the prior distribution for the coefficient on lag k of variable j in equation i is

$$(3) \qquad S_{ijk} = [\tau \cdot f(i,j) \cdot g(k)] \cdot s_i / s_j$$

where s_i is the estimated standard error of an unrestricted univariate autoregression on variable i. The scaling factor s_i/s_j is used because the standard deviations of lag coefficients on variables other than the dependent variable are not scale invariant. The parameter τ represents the overall tightness of the prior distribution, the function g(k) causes the prior standard deviations to decrease on higher lags, and the function f(i,j) controls the interaction among the variables. Since the prior information τ and g(k) treats each equation in the same manner it is the function f(i,j) (considered as weights) that allows to put more structural information on the prior. However, a smaller value of S_{ijk} ("tight" prior) gives more weight to the prior information while a larger value ("loose" prior) gives more weight to the sample information. From a Bayesian perspective, the standard approach to equation specification is to always select extreme values for S_{ijk}: excluding a variable from an equation is equivalent to setting S_{ijk} equal to zero (with a prior mean of zero), while including a variable is equivalent to setting S_{ijk} equal to infinity (without considering a prior mean).

After the prior information has been formulated as dummy observations, estimation is done by applying Theil's (1971) mixed estimator on an equation by equation basis. Although a gain in efficiency can be made using a seemingly unrelated regression procedure, such a procedure has not been attempted because of the computational burden (Litterman (1986)).

3. THE DATA SET

The purpose of the model is to forecast monetary time series which often show strong variations within a few months. A suitable model can be based on monthly data. This allows one to process newly available information quite efficiently and therefore to update the model as well as the forecasts.

The model originally built by Neusser (1986) includes three categories of variables:

- Variables that have to be forecast are currency in circulation (BGK), foreign reserves at the Austrian National Bank (FRC), demand deposits at banks (SI), time deposits at banks (TE), saving deposits at banks (SP), total central bank money (ZG), and the average bond yield in the Austrian secondary market (RS).
- Stochastic variables considered to characterize the rest of the economy or to be important in forecasting the monetary time series are the consumer price index (CPI), the index of industrial production (IP) which serves as a monthly proxy of real domestic product (in favour of a interpolation of quarterly GDP), the average yield on bonds in the Federal Republic of Germany secondary market (RSDM), and the nominal exchange rate of the Austrian Schilling vis-á-vis the Deutschmark (XDM). All stochastic variables enter each equation with 12 lags. Except for the interest rates each variable is logged to stabilize the variance.
- Deterministic variables include seasonal dummies for each month (S1 to S12) and the bank rate by the Austrian National Bank (RDISK) (contemporaneous and lagged) . This last variable is a policy variable because it is not market determined but set at irregular intervals by the National Bank.

Because of memory limitations on the personal computer a smaller model was adopted to accomplish the forecasting evaluation. This smaller model consists of the following stochastic variables: CPI, IP, KV, M2 (=BGK+SI+TE), RS, RSDM, SP, ZG as well as the deterministic variables S1 to S12 and RDISK. Furthermore, the lag length is reduced from 12 to 6.

4. PRIOR SPECIFICATION STRATEGIES

Before emphasis is put on the function f(i,j) which weighs the influence of variable i on variable j, the setup for the overall tightness τ and for the function g(k) has to be formulated. Three values for the parameter τ were used: 0.10, 0.18, and 0.25. Experience with BVAR models has shown that these values are reasonible choices. The function g(k) was choosen to decline geometrically in the lag length k

(4) $$g(k) = d^{k-1}, \qquad 0 < d < 1$$

Again, three values for d were selected: 0.95, 0.80 and 0.65. Together with the choices for τ, nine combinations can be considered. They provide a framework for the simulation experiment to analyze the importance of the f(i,j) specification. If the forecasting performance is mainly determined by the weighting function f(i,j), then a "good" weight specification should dominate a "poor" one regardless of the (τ,d) values. This investigation considers two strategies informulating f(i,j) by filling up a (mxm) matrix $W = (w_{ij})$ the elements of which represent the influence of variable i on variable j.

The first strategy utilizes statistical procedures in order to set up the matrix W. In particular a prior labelled *C* results from taking the correlation coefficients:

(5) $$w_{ij} = y_i'y_j \ / \ \surd(y_i'y_i)\cdot\surd(y_j'y_j)$$

This weight w_{ij} may serve as a simple measure of the linear influence of time series on each other. Calculating these coefficients gives the following matrix[1]:

(C)

	CPI	IP	KV	M2	RS	RSDM	SP	ZG
CPI	1.0	.74	.99	.97	.24	.16	.99	.98
IP	.74	1.0	.77	.76	.23	.06	.75	.74
KV	.99	.77	1.0	.98	.23	.14	.99	.99
M2	.97	.76	.98	1.0	.37	.28	.98	.98
RS	.24	.23	.23	.37	1.0	.71	.24	.25
RSDM	.16	.06	.14	.28	.71	1.0	.17	.19
SP	.99	.75	.99	.98	.24	.17	1.0	.99
ZG	.98	.74	.99	.98	.25	.19	.99	1.0

1) Because w_{ij} has to be nonnegative, the absolute value was taken

Note, however, that this is a symmetric specification. In order to incorporate approximations for causal relationships a second procedure, which leads to a non-symmetric prior labelled *L* was taken into account:

$$(6) \qquad w_{ij} = y_i'y_j^- \; / \; \sqrt{(y_i'y_i)} \cdot \sqrt{(y_j'y_j)}$$

where y_j^- denotes y_j lagged one period. These coefficients describe the correlation of a series i with a one period lagged series j. For the time series of the monetary model, the following matrix was obtained:

(L)

	CPI	IP	KV	M2	RS	RSDM	SP	ZG
CPI	1.0	.78	.98	.97	.33	.45	.98	.97
IP	.79	1.0	.81	.78	.33	.33	.80	.78
KV	.98	.80	1.0	.97	.35	.44	.98	.97
M2	.96	.78	.97	1.0	.43	.53	.96	.96
RS	.30	.33	.32	.41	1.0	.67	.31	.31
RSDM	.44	.34	.43	.53	.74	1.0	.44	.44
SP	.98	.79	.98	.97	.34	.45	1.0	.98
ZG	.97	.76	.97	.96	.34	.45	.98	1.0

The second strategy is to rely on economic intuition. Within this approach two priors are considered. The first specification labelled *P* stresses the small country assumption for the capital market assuming that the yield in the Austrian bond market is mainly determined by the yield in the German bond market. Both interest rates are modelled to be "more exogenous" relative to the rest of the model. In terms of f(i,j), this means that there is more influence from the interest rates to the other variables than the other way around. The next step is to translate those more qualitative considerations into a quantitative prior. Although one can introduce an additional metaparameter which controls the influence of the weighting function f(i,j) relative to a symmetric treatment of all variables, the present investigation proceeds in a more pragmatical way. By choosing specific values for the w_{ij}'s it is assumed that small variations of these w_{ij}'s (e.g. from 0.5 to 0.45 or to 0.55) would not affect the forecasting performance of a distinct prior relative to other priors. In particular, the W-matrix for the prior *P* was choosen to be

(P)

	CPI	IP	KV	M2	RS	RSDM	SP	ZG
CPI	1.0	.5	.5	.5	.2	.1	.5	.5
IP	.5	1.0	.5	.5	.2	.1	.5	.5
KV	.5	.5	1.0	.5	.2	.1	.5	.5
M2	.5	.5	.5	1.0	.2	.1	.5	.5
RS	.5	.5	.5	.5	1.0	.1	.5	.5
RSDM	.5	.5	.5	.5	.8	1.0	.5	.5
SP	.5	.5	.5	.5	.2	.1	1.0	.5
ZG	.5	.5	.5	.5	.2	.1	.5	1.0

For example, this matrix shows that RSDM influences RS with a weight of 0.8 and all other variables with a weight of 0.5. Specifying the opposite direction, only a small weight (0.1) is considered for the influence of all variables to RSDM.

Similar to the first specification, the second specification labelled *S* considered the German bond yield as the "most exogenous" variable of the system. Additionally, more structural information for the rest of the variables is put on the prior. Since the beginning of the 1980s Austrian exchange rate policy has been pegging the Austrian Schilling to the Deutschmark. Interest rates have to support the exchange rate goal. The "hard-currency" policy rules out an independent monetary policy, that means factors affecting demand rather than supply of liquidity will influence the development of monetary aggregates. These aspects lead to a specification in which - except for the German bond yield - industrial production and the yield in the Austrian bond market represent "core" variables while the rest represents "circle" variables. Core variables are seen to be of more importance and therefore have a higher weight. Again, following a more pragmatic way the following matrix W was used for the prior *S*.

(S)

	CPI	IP	KV	M2	RS	RSDM	SP	ZG
CPI	1.0	.2	.3	.3	.2	.1	.3	.3
IP	.7	1.0	.7	.7	.5	.1	.7	.7
KV	.3	.2	1.0	.3	.2	.1	.3	.3
M2	.3	.2	.3	1.0	.2	.1	.3	.3
RS	.7	.5	.7	.5	1.0	.1	.7	.7
RSDM	.2	.2	.2	.5	.8	1.0	.2	.2
SP	.3	.2	.3	.5	.2	.1	1.0	.3
ZG	.3	.2	.3	.5	.2	.1	.3	1.0

In addition the so-called "Minnesota" prior is used as a benchmark. This prior specification treats each variable in the same manner:

$$w_{ij} = \begin{cases} 1 & \text{if } i=j \\ w & \text{otherwise, } 0 \leq w \leq 1 \end{cases}$$

The priors *Y4*, *Y6*, and *Y8* refer to values 0.4, 0.6, 0.8 for w, respectively.

5. FORECAST EVALUATION

The comparsion of the forecasting performance relies on out-of-sample (ex-ante) forecasts which are performed over the period 1985:9 to 1987:9 in the following way: The model was estimated up to 1985:9 and k-step ahead predictions (k = 1, 3, 6, 12) are generated using the chain rule of forecasting. Then a new observation was added to the model. After the coefficients were updated via Kalman filtering, a new set of forecasts was generated. The last forecasts were made at 1987:8 and include only 1-step ahead predictions. This exercise provides 23 1-step ahead forecast errors, 21 3-step ahead forecast errors, 18 6-step ahead forecast errors, and 12 12-step ahead forecast errors.

Predictive accuracy is then evaluated with a measure suggested by Doan, Litterman, and Sims (1984):

(7a) $$_t\epsilon_{t+k} = {}_t\hat{Y}_{t+k} - Y_{t+k}$$

(7b) $$E_k = \sum_{s=1}^{T} ({}_s\hat{\epsilon}_{s+k}\ {}_s\hat{\epsilon}_{s+k}')$$

(7c) $$D_k = \text{k-step ahead log-determinant} = \log(|E_k|)$$

where the m-vector $_t\hat{Y}_{t+k}$ denotes the k-step ahead predictor at time t, $_t\hat{\epsilon}_{t+k}$ denotes the k-step ahead prediction errors at time t; E_k is (up to a factor T) an estimator of the covariance matrix of k-step ahead prediction errors. To compare different covariance matrices, the logarithm of the determinant of E_k was taken.

In general, the results show that the specification of the f(i,j) has a significant influence on the forecasting performance. Since it increases as the parameters τ and d decrease only the four most interesting combinations of the (τ,d) parameters are reported. However, the other five combinations show the same relative performance of the various f(i,j) specifications.

The results of *Table 1* are obtained from an estimation period starting 1980:1. The sample period coincides then with the period of the hard-currency exchange rate regime. It is assumed that earlier observations on data will contain less information useful in forcasting the late 1980s. The results demonstrate that significant improvements can be made by imposing the prior *(S)* on the system. Even the rather simple structured prior *(P)* outperforms the priors relying on statistical procedures. The prior *(C)* seems to work better than prior *(L)*. Furthermore, when the Minnesota prior is applied smaller weights give generally better forecasts.

In order to investigate whether the results are sensitive subject to the sample period, the same forecasting exercise was accomplished with the model estimated over the period starting 1974:1. The reason for this choice is the breakdown of the Bretton Woods system and, consequently, a renewal of the banking system and monetary policy in Austria. The results, reported in *Table 2*, are similar to those of *Table 1*. Again, those specifications that put more economic structure on the prior generally perform better although sometimes the best result is achieved from other priors (*Y4*, *C*). However, a larger sample period always reduces the influence of the prior on the posteriori distribution. Interestingly, a comparison of *Table 1* and *Table 2* shows that predictive accuracy is only slightly better when the larger sample period is used. This means that the data from 1970s provides less information useful in forecasting the late 1980s.

5. SUMMARY

This paper has argued that Bayesian vector autoregressions are well suited for forecasting monetary time series. Since the prior has to be formulated, different procedures specifying those metaparameters which control the interaction among the variables of a BVAR model were investigated. It was demonstrated that these metaparameters have a significant influence on the forecasting performance. The use of more informative priors which rely on economic arguments rather than on simple statistical procedures have increased predictive accuracy.

Table. 1

Negative k-step ahead log-determinants for estimation period starting from 1980:1

Prior:	d = 0.80							d = 0.65							
	(Y4)	(Y6)	(Y8)	(P)	(S)	(C)	(L)	(Y4)	(Y6)	(Y8)	(P)	(S)	(C)	(L)	
1-step ahead	65.879	65.670	65.503	65.971	66.119	65.567	65.441	66.059	65.885	65.728	66.185	66.295	65.774	65.656	τ=0.10
3-step ahead	57.978	57.491	57.079	58.191	58.418	58.429	57.174	58.488	58.112	57.721	58.718	58.813	57.992	57.757	
6-step ahead	54.447	54.666	54.081	55.599	56.234	54.784	54.343	56.196	55.415	54.784	56.218	56.790	55.345	54.951	
12-step ahead	57.749	57.136	56.296	58.939	59.247	57.916	57.030	58.302	57.908	57.247	59.547	59.745	58.438	57.761	
1-step ahead	65.928	65.674	65.396	65.969	66.177	65.558	65.378	66.268	66.065	65.889	66.356	66.605	65.915	65.764	τ=0.18
3-step ahead	57.486	57.137	56.853	57.796	58.038	57.116	56.881	58.224	57.853	57.548	58.492	58.745	57.705	57.491	
6-step ahead	54.847	54.458	54.126	55.206	55.632	54.703	54.285	55.503	54.998	54.628	55.742	56.288	55.132	54.728	
12-step ahead	57.667	56.887	56.157	58.672	59.094	57.854	56.917	58.645	57.967	57.138	59.480	59.962	58.522	57.883	

Each column refers to a different specification of f(i,j): (Y4), (Y6), (Y8) ... Minnesota prior;
(P) ... simple economic structure; (S) ... complicated economic structure; (C) ... correlation coefficients (contemporeanous); (L) ... correlation coefficients (lagged)
Double underlined values indicate the best forecast, underlined values the second best forecast

Table. 2

Negative k-step ahead log-determinants for estimation period starting from 1974:1

Prior:	d = 0.80							d = 0.65							
	(Y4)	(Y6)	(Y8)	(P)	(S)	(C)	(L)	(Y4)	(Y6)	(Y8)	(P)	(S)	(C)	(L)	
1-step ahead	66.228	66.251	66.248	66.365	66.406	66.426	66.356	66.174	66.133	66.090	66.277	66.390	66.211	66.133	τ=0.10
3-step ahead	59.049	58.936	58.970	59.252	59.219	58.949	58.884	59.173	58.986	58.791	59.333	59.398	58.896	58.828	
6-step ahead	58.595	58.312	57.925	58.573	58.421	57.953	58.049	58.797	58.459	58.122	58.666	58.623	57.894	58.113	
12-step ahead	60.488	59.869	59.337	60.837	60.772	59.844	59.853	60.824	60.320	59.916	61.220	61.100	60.322	60.374	
1-step ahead	66.263	66.177	66.027	66.436	66.418	66.517	66.314	66.248	66.158	66.090	66.403	66.539	66.401	66.244	τ=0.18
3-step ahead	58.785	58.608	58.451	59.078	58.901	58.712	58.536	58.930	58.701	58.551	59.109	59.211	58.713	58.600	
6-step ahead	58.318	57.943	57.550	58.603	58.493	57.906	57.703	58.517	58.222	58.041	58.655	58.693	58.072	58.112	
12-step ahead	59.773	59.110	58.652	60.518	60.689	59.252	58.996	60.504	60.051	59.705	61.249	61.352	60.340	60.196	

Each column refers to a different specification of f(i,j): (Y4), (Y6), (Y8) ... Minnesota prior; (P) ... simple economic structure; (S) ... complicated economic structure; (C) ... correlation coefficients (contemporeanous); (L) ... correlation coefficients (lagged)

Double underlined values indicate the best forecast, underlined values the second best forecast

REFERENCES

DOAN, T., LITTERMAN, R.B., SIMS, C.A. (1984), Forecasting and Conditional Projection Using Realistic Prior Distributions, Econometric Review, 3(1), 1-100.

KLING, J.L., BESSLER, D.A., (1985), A Comparsion of Multivariate Procedures for Economic Time Series, International Journal of Forecasting, 1, 5-24.

LITTERMAN, R.B. (1980), Techniques for Forecasting with Vector Autoregressions, Doctoral Thesis, University of Minnesota.

LITTERMAN, R.B. (1986), Forecasting with Bayesian Vector Autoregressions - Five Years of Experience, Journal of Business and Economic Statistics, 4, 25-38.

LUCAS, R.E., SARGENT, T.J. (1981), After Keynesian Macroeconomics, in: Lucas, R.E., Sargent, T.J. (eds.), Rational Expectations and Econometric Practice, Vol. 1, University of Minnesota Press, Minnesota, 295-319.

NELSON, R., PLOSSER C.I., (1982), Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications, Journal of Monetary Economics, 10, 139-162.

NEUSSER, K. (1986), Forecasting Monetary Time Series Using Bayesian Vector Autoregressions: A First Appraisal for Austria, mimeo.

SIMS, C.A. (1980), Macroeconomics and Reality, Econometrica, 48, 1-48.

THEIL, H. (1971), Principles of Econometrics, Wiley: New York.

EXPLORATION VORAUSEILENDER INDIKATOREN

Wolfgang A. Fels

Institute for Advanced Studies

Vienna

Wolfgang A. Fels

EXPLORATION VORAUSEILENDER INDIKATOREN

KAP 1

Ein Kriterium für vorauseilende Konjunkurindikatoren (LI) ist weniger die hohe Kreuzkorrelation oder eine erhebliche Phasenverschiebung bei deutlicher Kohärenz gegenüber einer Referenzreihe als vielmehr die möglichst verlässliche Vorwegnahme von konjunkturellen Wendepunkten. Um diesem qualitativem Kriterium gerecht zu werden treten paramentrische Schätzverfahren bei der Methodenwahl zur Bestimmung vorauseilender Sammelindikatoren in den Hintergrund. Das vom National Bureau of Economic Research (NBER)[1] entwickelte Verfahren, das das heute wohl weitverbreitetste dieser Art darstellt[2], orientiert sich an der Stabilität des Medians der Leads in Konjunkturhoch- und Tiefpunkten zwischen einer Referenzreihe und potentiellen Indikatoren. Notwendig ist hierfür einerseits eine Absplittung kurzer, auf unterschiedlichen Niveaus stattfindender Zyklen von einem längerfristigem Trend[3], und andererseits die Festlegung, d.h die Datierung von Konjunkturwendepunkten[4].

Ohne näher auf dieses Verfahren einzugehen sei hier nur angemerkt, daß letzgenannter Schritt über eine Serie sich überlappender gleitender arithmetischer Mittel erfolgt, ohne daß dadurch eine Wendepunktbestimmung ohne manuelles Eingreifen sichergestellt werden kann[5]. Die Nachteile solcher Filterungen und alternative Glätter werden in Kap. 2 diskutiert. Ein darauf aufbauendes Verfahren zur Bestimmmung LI wird in Kap. 3 vorgestellt und in Kap.4 und 5 für den Wachstumszyklus der österreichischen Industrieproduktion (IP) angewandt. Hiebei werden kurz- (Lead etwa ein halbes Jahr), mittel- (drei Quartale) und langfristige (ein Jahr) Indikatoren gebildet. Der eher inhaltlich als methodisch interessierte Leser seit insbesondere auf Kap.4 - der Datierung konjunktureller Leads einiger Zeitreihen - verwiesen.

---

1) Zur Dokumentation der NBER-LI sei auf Moore, G.H. (83) verwiesen.

2) Sowohl die Europäische Gemeinschaft (Klein,F., Nerb,G.,(85); Dryden,J., Reynard,B.;(85)) als auch österreichische Ökonomen (Breuss,F., (84); Breuss,F., Wüger,M.,(85); (86)) greifen und griffen auf diesen Ansatz zurück.

3) Für dieses "phase-average-trend" - Bestimmung (PAT) wird auf das NBER - Verfahren von Charlotte Boschan und Walter Ebanks (78) zurückgegriffen.

4) Die dafür verwendeten Prozeduren sind nach ihren Schöpfern G. Bry und C. Boschan (1971) benannt.

5) Das Ziel, eindeutige zweipolige Zyklen mit einer Mindestlänge von 15 Monaten zu diagnostizieren, ist durch diese Routine nicht gewährleistet.

KAP 2

Wie Mallows(80) und Velleman (80) reservieren wir den Begriff Filter für lineare sequenzielle Transformationen während mit Glättern (smoothing Algorithmus) nicht-lineare Ansätze zur Bestimmmung glatter Zeitreihen bezeichnet werden.

UNGERADSPANNIGE GLEITENDE MEDIANE:

Die einfachsten Glätter sind ungeradspannige gleitende Mediane mit Span v die durch die Zahl 'v' symbolisiert werden. Diese Selektoren bilden längere monotan ansteigende Sequenzen unverändert ab. Einbrüche bis zur Länge (v-1)/2 werden eingeebnet, Einbrüche bis zur Länge (v-1) werden zu Zwischenplateaus, Kuppeln zu Plateaus der Länge (v+1)/2 transformiert.

RESMOOTHING (R):

Kürzere Schwingungen können durch ungeradspannige Mediane umgekippen[6]. Dies kann jedoch durch weiteres Glätten mit v ausgebügelt werden. Mit der Anzahl der ungeradspannigen Glättungen konvergiert die Zeitreihe gegen eine mit **vR** (für resmoothed) bezeichnete Reihe.

GERADSPANNIGE GLEITENDE MEDIANE:

Geradspannige gleitende Mediane müssen immer zweimal hintereinander angewandt werden, wenn neue Werte für die gegebenen diskreten Zeitpunkte bestimmt werden sollen. Jede einzelne Glättung von Span v ergibt

$$z_{t+} = \text{med}\,\{y_{t-v/2+1},\ldots,y_t,\ldots,y_{t+v/2}\}$$

Wird als Folgeglättung ein 2-span Median, der dem arithmetischem Mittel entspricht, gewählt, so wird das Ergebnis

$$z_t = [\text{med}\{y_{t-v/2},\ldots,y_{t+v/2-1}\} + \text{med}\,\{y_{t-v/2+1},\ldots,y_{t+v/2}\}]\ /2$$

in der Regel als v2 bezeichnet. Da im Rahmen dieser Arbeit 2-Glättern immer geradspannige Mediane folgen werden wir uns lediglich auf die Bezeichnung v für v2 beschränken[7].

Im Vergleich zu ungeradspannigen gleitenden Medianen weisen geradspannige gleitende Mediane eine schwächere Plateaubildung, generell glättere Spuren und Transferfunktionen auf[8].

---

6) S. Polasek,W.,(82). Vgl. auch die IP-Wachstumsraten Mitte 1983 (Abb. 2).

7) Aus der Definition ergibt sich übrigens auch daß 2 dem Hanning-Glätter (H) bzw. dem gleitenden gewichteten Mittel mit den Gewichten 0.25,0.5,0.25 Span 3 (3gM) entspricht.

8) siehe Velleman,P.F. (80)

ENDWERTREGEL:

Durch Anwendung von gleitenden Mediane bleiben die ersten und letzten (v-1)/2 Daten unbestimmt. Von den in der Literatur vorgeschlagenen Regeln sollen zwei vorgestellt werden:

a) Span-Schrumpfung: Im Randbereich wird ein kürzerer unsymmetrischer Glätter verwendet.

b) Tukey-Regel: Ist f der letzte Zeitpunkt für den eine Glättung zf vorliegt so wird

$$z_{f+1} = \text{med}\,\{3z_f - 2z_{f-1},\ y_{f+1}, z_f\}$$

gesetzt, wobei y für die ursprüngliche, z für die geglättete Zeitreihe steht. In analoger Weise werden die Anfangswerte schrittweise bestimmt.

Regel a) verwenden wir für die unten dargestellte Bi- und Triweight-Schätzer; die Tukey-Regel fand bei gleitenden Medianen und dem Splitting Anwendung.

SPLITTING (S):

Zur Abrundung von Plateaus werden die Werte von beiden Kanten des Plateaus zum Zentrum hin nach der Endwertregel neu bestimmt.

BIWEIGHT (Bi)[9]:

Die Berechnung des Biweight-Schätzers (Bi) erfolgt in einem iterativen Prozess. In der k-ten Iteration ergeben sich die Gewichte (w_i), mit denen die v einzelnen Beobachtungen(x_i) in das Lagemaß eingehen aufgrund der Abweichungen, der Beobachtungen vom zuletzt berechneten Schätzer (x_tBi^{k-1}). Bei Span v ergibt sich[10])

$$x_tBi = \lim_{k\to} x_tBi^k = \frac{\Sigma_i \; w_{ik} \cdot x_i}{\Sigma_i \; w_{ik}} \quad , \quad i = t-{}^{(v-1)}/_2, \ldots, t, \ldots, t+{}^{(v-1)}/_2$$

$$\text{mit } w_i = \begin{cases} (1-((x_i - x_tBi^{k-1})/cS)\ ) & \text{falls } ((x_i - x_tBi^{k-1})/cS) \leq 1 \\ 0 & \text{falls } ((x_i - x_tBi^{k-1})/cS) \leq 0 \end{cases}$$

S steht hier für ein Streuungsmaß wie den Median der absoluten Abstände von x_tBi^{k-1} oder die Hälfte des Intervalls zwischen der ersten und dritten Quartilsobergrenze verwendet werden kann[11]).

c steht für einen zu wählenden Parameter - Tukey &. Mosteller empfehlen 6 oder 9 -, der die Krümmung der Reaktionsfunktion (Abb.2)

---

9) Im Rahmen dieser Arbeit soll, wenn von Biweight oder Triweight-Schätzern gesprochen wird, nur an Running Bi- oder Triweight- Schätzer gedacht werden.

10) Mosteller,M.; Tukey,J.; (77), S.205

11) In dieser Arbeit wurden Biweight und Triweight ausschließlich über die letztgenannte Variante, dem mittleren inneren Quartilsspann berechnet.

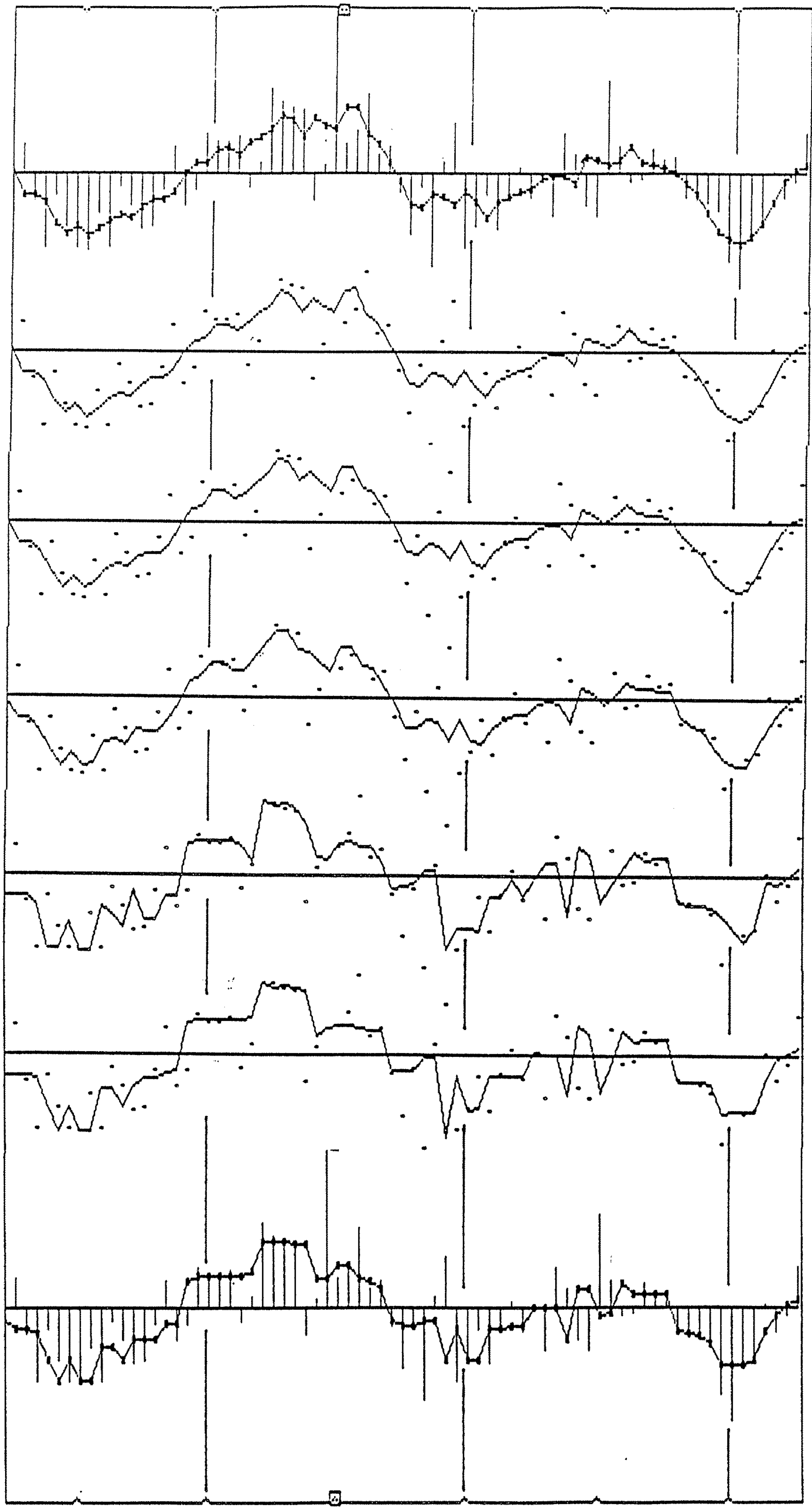


Abb.1 :
Industrie-
produktions-
Wachstumraten
(senkrechte
Linien)
und 5-Span
Glättungen
von 1977:07
bis 83:07.

oben: gl.Mittel.
unten:gl.Median.
dazwischen: Bi
mit sinkendem
c.

bestimmt. bei c -> konvergiert $x_t Bi^k$ unabhängig vom gewählten Startwert Bi^0 gegen das arithmetische Mittel, das die Schöpfer des Schätzers generell auch als Startwert empfehlen. Dieser Ratschlag erweist sich jedoch bei sinkendem c als Konvergenzfalle. Bei c -> 0 kann der Schätzer schließlich nur einen Wert - nämlich den Startwert selbst - wählen, wenn ein $x_t Bi^0$ aus dem x_i-Sample gewählt wurde. Durch die Verwendung des Medians als Selektor für $x_t Bi^0$ kann über die Wahl von c Bi stufenlos zwischen dem gleitenden Median und dem gleitenden Mittel ausgesteuert werden (vgl.Abb.1). c ist somit als Prior über die Normalität des zu glättenden Prozesses zu interprätieren. Es liegt nahe bei der ersten Glättung ein kleines c, bei wiederholten Glättungen ein wachsendes c zu wählen.

TRIWEIGHT (T):

Der Triweight-Schätzer ist ein für Analysen im Zeitbereich mit einer Lag-Gewichtung l_i erweiterter Biweight - Schätzer:

$$x_t T = \lim_{k->} x_t T^k = \frac{\Sigma_i w_{ik} l_i x_i}{\Sigma_i w_{ik} l_i} , \quad i = t-{}^{(v-1)}/_2,..,t,..,t+{}^{(v-1)}/_2$$

w_i sind wie oben definiert. Für Leads und Lags wird im Folgenden eine logarithmisch abfallende Lag-Gewichtung, die den aktuellen Zeitpunkt doppelt so hoch wie die maximalen Lags gewichtet[12)] gewählt.

VERGLEICH DER GLÄTTER

Während M durch maximale Gauss-Effizienz charakterisiert ist, zeichnet den Median Optimalität für Laplace-verteilte Populationen aus. Bei kontaminierten Verteilungen (insbesondere bei Langschwänzigkeit) büsst das arithmetische Mittel jedoch wesentlich dramatischer als der Median Effizienz ein. Dies wird besonders deutlich durch sogenannte Reaktionsfunktionen die den Zusammenhang zwischen einer Beobachtung des Sapels und dem daraus berechneten Lageparameters wiedergeben[13)]. Die Abbildung 2 zeigt Lageparameter mit Span 13 als Funktion der Beobachtung x aus dem Vektor (-6,-5,..,-1,x,1,..,5,6).

Gleitende Mediane wären bei diesem Sampel nur im Bereich (-1,1) sensitiv auf die Veränderungen von x. Ausserhalb dieses Intervalls bleibt der Median starr bei -1 bzw. +1. Das arithmetische Mittel steigt monoton linear während der Biweightschätzer bei größeren Abweichungen von x zum Bi der übrigen 12 Beobachtungen konvergiert. Bei kleinen Abweichungen reagiert dabei **Bi** etwas stärker als , da weiter weg liegende Beobachtungen beim Steigen des mittleren Wertes an

12) Bei maximal A Lags sei Gewicht des dt'ten Lags/Lead durch $l_{dt} = \exp({}^{dt}/_A .\ln(0.5))$ bestimmt. Damit gilt 3gM = H = 2.

13) Reaktionsfunktionen sind im wesentlichen proportional zu den bekannteren Psi-Funktionen Hubers. (s. Huber (81),p.45)

Gewicht verlieren. Bei der engeren Gewichtung mit c=6 (Bi(6)) steigt der Parameter zuerst stärker, schwingt sich aber dann auch früher auf 0 zurück als bei c=9.

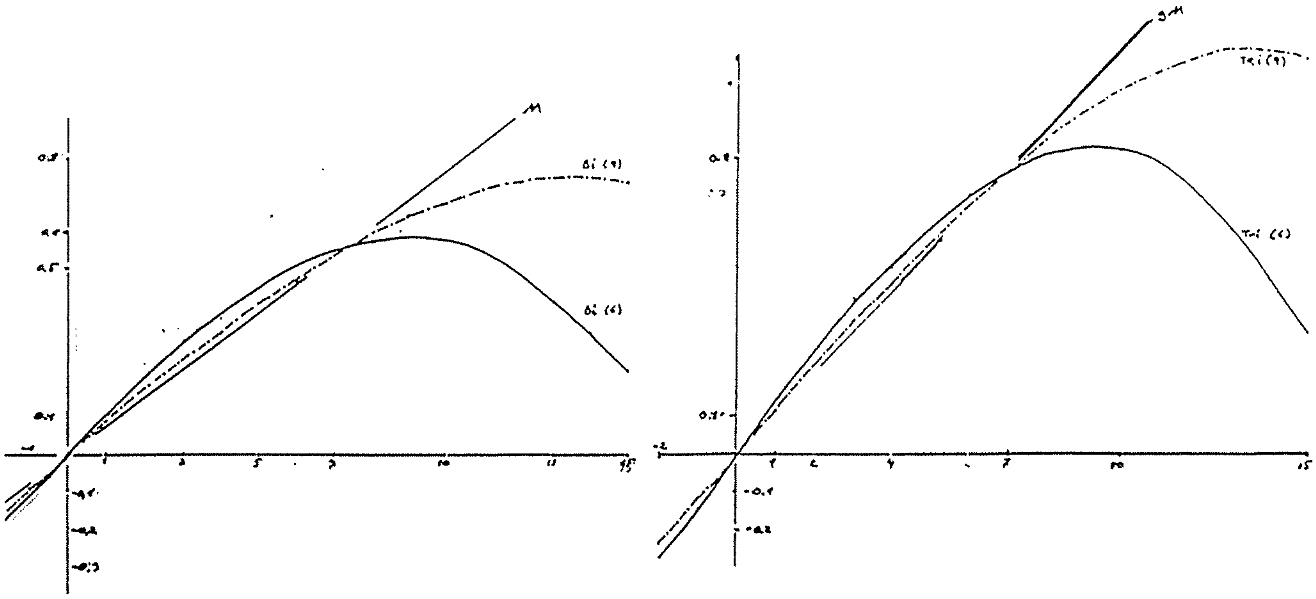


Abb.2 : Reaktionsfunktionen: Reaktion der Lageparameter auf Veränderungen eines Stichprobenwerts.
Links : ar.Mittel (M), Biweight mit c=9 und c=6.
Rechts: gewichtetes Mittel (gM), Triweight mit c=9 und c=6.

Duch die Lag-Gewichtung wird, wie der zweite Teil von Abb.2 zeigt, die mittlere Beobachtung des Sampels wesentlich stärker konserviert: Das gewichtete Mittel und der Triweight-Schätzer steigen mit ihr deutlich stärker an.

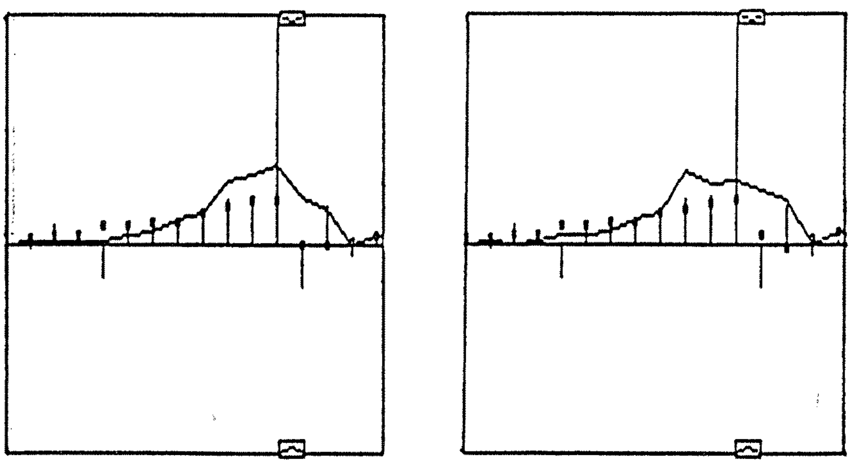


Abb.3 : Sensitivität auf Ausreißer (IP Dezember 79)
links : gM5 (Spur) und T5 (Knollen) .
rechts: M5 (Spur) und Bi5 (Knollen).

Abbildung 3 zeigt wie die diskutierten Glätter auf einen offensichtlichen Ausreißer in den IP-Wachstumsraten im Monat der Abweichung und der Nachbarschaft reagieren. Die klassische Lösung des Ausreißerproblems liegt im Ausschluß 'verdächtiger' Beobachtungen vor der Schätzung obwohl die Normaltheorie nicht adequat für solchermaßen 'gesäuberten' Samples ist. Robuste Schätzer wie der Biweight passen sich dagegen, da sie konzeptionell alle Beobachtungen spezifisch gewichten, automatisch an Ausreißer und Abweichungen von Verteilungsannahmen an. Dies ist insbesondere - wie in unserem Fall - für Parameterschätzungen bei Samples von 3, 5 oder 7 Beobachtungen, bei denen sich Verteilungstests erübrigen, wesentlich.

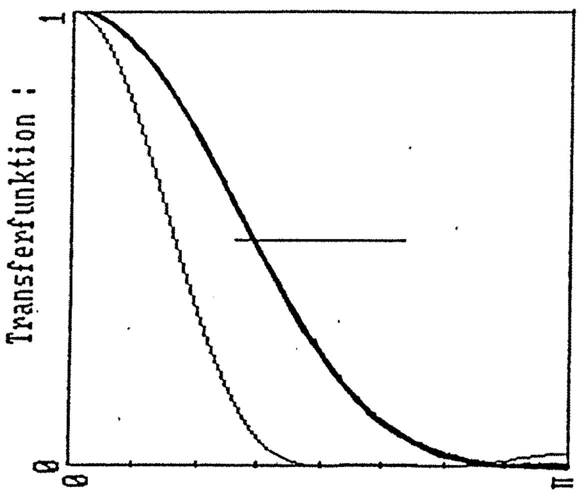


Abb.4 :Transferfunktionen der hier verwendeten gewichteten Mittel.
Dicke Linie : Span 3 (gM3)
Dünne Linie : Span 5 (gM5)

Ein wesentlicher Vorteil gleitender (M) und gewichteter(gM) Mittel gegenüber robusteren Glättern liegt in ihrer eindeutig bestimmbaren Transferfunktion. Abb. 4 zeigt jene der hier verwendeten gewichteten Mittel (gM) bei Span 3 und 5. Mit der Ausdehnung des Spans wird die spektrale Masse der Zeitreihe auf tiefere Frequenzen konzentriert, wobei hohe Frequenzen gedämpft erhalten blieben, wären sie nicht schon davor durch Glätter mit kürzerem Span völlig eliminiert worden.

Vellemans (80) Simulationen mit gleitenden Medianen zeigen, daß insbesondere ungeradspannige Glätter in regelmäßigen - vom Spann abhängenden - Abständen Einbrüche in den geschätzten Frequenz-Antwort-Funktionen aufweisen[14]. Unsere Untersuchungen von Transferfunktionen von Bi- und Triweight- Glättern, die vielmehr anhand verwendeter Daten die vorliegenden Konjunkturglättungen dokumentieren sollen als ein allgemeinens Ergebnis postulieren zu wollen - zeigen daß

evidenter Weise die Transferfunktionen der Bi-Glätter bei c->0 bzw. c -> gegen jene der gleitenden Mittel bzw. der gleitenden Mediane konvergieren.

Bi und Tri lediglich eine starke Dämpfung auf Frequenzen ausüben, die die entsprechenden M und gM Schätzern ausradieren.

14) Velleman Simulationsdesign erlaubt keine Schlüsse auf Schwingungen, die durch korrelierte Störungen erweitert werden. Zur Problematik Vellmans Untersuchung s. Härdler &. Tuen (86).

- Bi (c>0) auch jene Frequenzen stärker konservieren, bei denen die Transferfunktion der entsprechenden ungeradspannige gleitende Mediane regelmäßig einbricht.
- bei einem Spann, bei dem lineare Glätter höhere Frequenzen nicht auslöschen (d.h. v>3), Bi und Tri diese höhen Frequenzen deutlicher als M oder gM, aber auch deutlicher als gleitende Mediane drosseln.

Exemplarisch sei auf Abbildung 5 verwiesen.

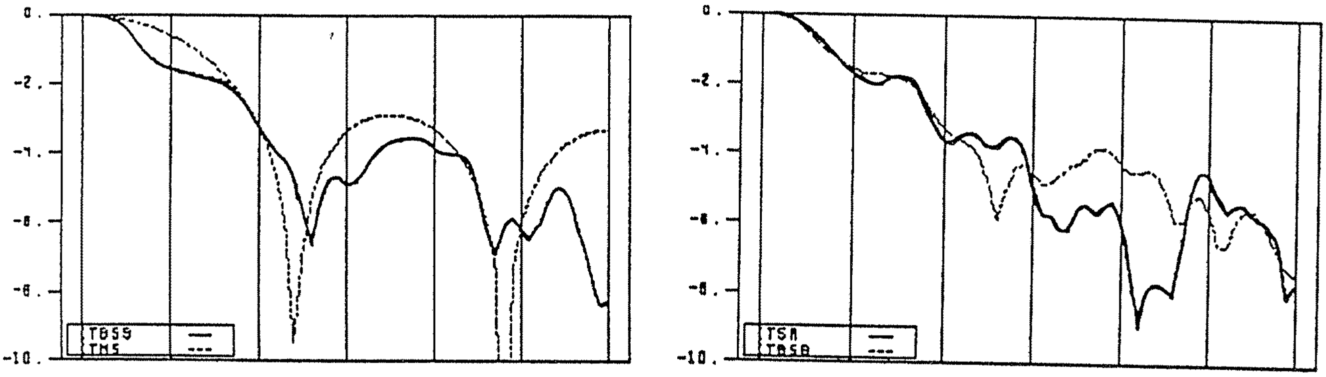


Abb. 5 : Log-Transferfunktionen.

links : Dünn : gl. Mittel Span5 (M5)
Dick : Bi5 mit c=9

rechts : Dünn : Bi5 mit c=6
Dick : gl. Median (5R)

KAP 3

Die Bestimmung von Leading Indicators erfolgt über ein mehrstufiges Verfahren. Im Folgenden seien die einzelnen Schritte genannt die zu den in den darauffolgenden Teilen wiedergegebenen Resultaten führten:

(1) Datenvorbereitung.

Nach eventueller Interpolation von vorliegenden Quartalsreihen zu Monatsreihen[15] werden einheitlich jährliche Wachstumsraten bestimmt[16].

(2) Datierung des konjunkturellen Vorlaufs einzelner Zeitreihen.

Zur Bestimmmung der Menge vorauseilenden Zeitreihen dient das Kriterium, in konjunkturellen Wendepunkten der Referenzreihe mit einem möglichst stabilen Lead gewünschter Länge vorauszueilen. Zu dieser Datierung sind drei Schritte erforderlich:

(i) Bereinigung von längerfristigen Trends.

Die Bereinigung von langfristigen Zyklen ist notwendig, da der Indikator der Referenzreihe nicht in langfristigen sondern primär

---

15) Die Monatswere der Reihe X (X_m) werden dabei durch Interpolation des bekannten Quartalswerts X_q aufgrund des Industrieproduktionsmonatsmusters gebildet: $X_m = X_q . IP_m/IP_q$

16) $X'_t = \ln (X_t/X_{t-12})$.

in konjunkturellen Schwankungen vorauseilen soll. Bei nicht-bereinigten Datenreihen würden konjunkturelle Schwankungen bei den folgenden Glättungen schließlich eher eingeebnet, wenn sie über langen Konjunkturauf- oder abschwüngen liegen als wenn sie mit Phasen ohne langfristige Trends fallen[17].

(ii) Glättung von kurzfristigen Schwankungen.
Die Glättungsprozedur greift auf ungeradspannige Triweigt-Schätzer(T) und geradspannige gleitende Mediane mit Splitting (S) zurück, wobei die Span-Ausdehnung monoton vergrößert wird[18]. Während wir letztere besonders wegen ihrer peak- und musterkonservierenden Eigenschaften schätzen, kommt den Triweight-Glättern, die relativ zu den Medianen glatte Schwingungen generieren, vor allem die Aufgabe der Abrundung jener Plateaus zu, die sich zwangsläufig aus der Glättung mit gleitenden Medianen ergeben. Wir beginnen mit 3T, und folgen mit 4S,5T,6S... Glättungen. Die Glättung wird solange fortgesetzt bis entweder die kürzeste Phase $\geq$ m, oder die längste Phase $\geq$ 2.m und der Median der Phasen > m ist.

Wir setzten hierbei m=12 Monate um Konjunkturzyklen mit einer Dauer von etwa 2 bis 4 Jahren zu bestimmen[19]. In der Regel wurde der Glättungsprozess nach etwa 5 oder 6 Glättungen abgeschlossen[20].

(iii) Bestimmung der Lead-Eigenschaften durch Zuordnung der Wendepunkte der konjunkturell geglätteten Reihen.

Die Zuordnung erfolgt

1) automatisch aufgrund Informationen der geglätteten Zeitreihen:

a) im Vergleich von Länge und Abfolge von Anstiegs- und Schrumpfungsphasen in der Umgebung der Extrema.

b) im Vergleich der Häufigkeiten und der Entfernung von Realisationen die kleiner (bzw. größer) als das zuzuordnende Minimum (bzw. Maximum) ist.

2) durch manuelle Kontrolle &. Korrektur.

a) mittels Vorinformationen über die Richtung des zeitlichen Zusammenhangs aufgrund der Betrachtung des bewegten Phasendiagramms[21].

---

17) Der längerfristige Trend, hier additiv zu kurz- und mittelfristigen Schwankungen modelliert, ergibt sich durch eine 3-Span Triweight-Glättung der durch einen 37-Span-Triweight geglätteten Zeitreihe wobei für die erste Glättung c=6, für die letztere c=9 gewählt wurde. Im Folgenden betrachten wir daher ausschließlich die transformierten Reihen Xneu = X - 3T(37T (X)). Dadurch wird nicht der Zeitpunkt sondern nur das Niveau der Extremwerte verschoben.

18) Siehe Kap. 2.

19) Daß die hier bestimmten mittleren Konjunkturzyklen kürzer als die des NBER (4 Jahre) sind ergibt sich aus der Länge einiger der zur Verfügung stehenden Zeitreihen.

20) Ergeben sich durch die Glättung Tableaus, so wird im Plateaubereich das Extremum der noch nicht geglätteten Reihe bestimmt und als Extremwerts des Plateaus für die spätere Datierung der konjunkturellen Wendepunkte registriert.

21) Dreht sich in einem üblichen 2D-Orthogormalsystem die Trajektorie im Uhrzeigersinn so läuft die auf der Abszisse aufgetragene Zeitreihe Y der auf der Ordinate aufgetragenen Reihe X voraus.

b) aufgrund apriori-Informationen über den zu erwartenden mittleren Lead infolge der Minimierung des mit Err% bezeichneten qualitativen Prognosefehlers[22] der bei Lead L wie folgt für die Referenzreihe X und den Indikator Y definiert wird:

$$Err\% \ (X,Y,L) = 100.p(sign(X_t - X_{t-1}) \ sign \ (Y_{t-L} - Y_{t-L-1}))$$

c) bei Unsicherheit über die Datierung einzelner Wendepunkte wird der Vergleich zu weniger stark geglätteten Versionen der Zeitreihen herangezogen[23].

3) Konsistenzprüfung:

Die zeitliche Reihung der Extrama einer Zeitreihen muß der der ihr zugeordenten Extrema der Referenzreihe entsprechen.

(3) Bestimmung der Sammelindikatoren.

Die 3 gewünschten Leading Indikators werden mittels Faktorenanalyse durch Extraktion der ersten Hauptkomponenten aus einer Menge von Zeitreihen mit gewünschtem mittleren Lead gebildet. Zeitreihen die weiter vorlaufen können gelaggt in kurzfristiger vorlaufende Indikatoren integriert werden.

In die endgültigen Linearkombinationen gehen allerdings nicht konjunkturell geglättete Einzelreihen sondern standardisierte 3R2S-Glättungen der jährlichen Wachstumsraten ein da eine möglichst frühe Verfügbarkeit der LI gewährleistet werden soll. Für jeden LI mehere Varianten zusammengestellt um im nächsten Schritt den optimalen auszuwählen.

(4) Überprüfung und Auswahl vorauslaufender Indikatoren .

Zur Auswahl der LI dienen folgende Kriterien und Statistiken:

i) Leadlänge: Median der Leads, mittlerer Lead.
ii) sabiler Lead: QS (Quartilsspann), (Standardabweichung) d. Leads.
iii) deutlicher Lead:Kreuzkorrellationen,Err%
iv) konjunktureller Lead: Phasenwinkel.

(5) Re-Trending

Die vom langfristigen Trend bereinigten Indikatoren müssen zuletzt dem langfristigen Trend der Referenzreihe angepasst werden[24].

22) Dieses rein qualitative Kriterium bietet ausschließlich bei konjunkturell glatten Zeitreihen einen Hinweis auf den mittleren Leads.

23) Zur Orientierung dienten in der Regel die 3T-5T (d.h. 3T4S5T) Glättungsvarianten.

24) Dieser Schritt stellt das Gegenstück zu dem in (2,ii) beschriebenem Detrending dar und sollte auf einen ebenso definierten langfristigem Trend aufbauen. Das Retrending ist insbesondere für die Präsentation von Grafiken wichtig, wenn der langfristige Trend von Indikator und Referenzreihe gegenläufig sind. Angezeigte Konjunkturwendepunkte können damit höchstens marginal verschoben werden.

KAP 4

Tabelle 1:

:::::::::::::::::::: Lead (>0) oder Lag (<0) ::::::::::

:::::::::::::::::::: Industrieproduktion ::::::::::

	Min	Q1	M	Q3	Max	Err%	QS
LKV Saldo langfr. KapVerkehr	2	15	20	21	24	20	3,0
IFH0 FinLage Haushalt vor 1 Jahr	4	8	14	16	16	40	4,0
IFH1 FinLage Haushalt in 1Jahr	1	2	13	15	23	39	6,2
M1 Geldmenge M1 real	-3	10	12	15	17	39	2,5
-R3CH -Schweizer 3-monats-Zins	3	7	10	12	17	24	2,5
-RSEC -Rendite Neuemissionen	0	0	7	8	3	35	4,0
ATE Auftragseingänge	-2	0	4	7	9	26	3,5
K$ Dollar-Schilling Kurs	-10	-3	4	7	14	40	5,0
TIEG EG Klima Industrie	1	2	4	6	12	24	2,0
TIVG Ifo Klima Verarb. Gewerbe	-8	0	4	5	8	23	2,5
TIGH Ifo Klima Grosshand.Ums.	-4	-1	3	8	15	38	4,5
TIVP Verkauspreise erwartet,Iv	-2	-1	3	6	10	28	3,5
ATEA Auftragseing. v. Ausland	-3	-1	2	4	5	20	2,5
IPGG Gr. Grundstoffe Indprod.	-2	1	2	12	18	32	5,2
UB Beschäftigte GesamtWirt.	-5	-1	2	6	15	38	3,5
UMS Umsätze Grosshandel	-3	-1	2	4	8	26	2,5
XR Exporte Real	-4	0	2	6	16	32	3,0
TIVA Auftragsbestände Ind. Iv	-5	-2	1	2	5	25	2,0
TPro Produktions-Beurteilung Iv	-5	-2	1	4	14	21	3,0
-TLag-Lager-Beurteilung Iv	-4	0	1	1	4	17	0,5
GAJA Arb.Stunden / Ind.Arbeiter	-9	-3	0	4	14	32	3,5
INV BruttoanlageInvests	-10	-2	0	1	18	26	1,5
TEXP Export Beurteilung Iv	-11	-1	0	2	13	21	1,5
TIVM Aufträge aus Ausland Iv	-6	-4	0	2	3	22	3,0
-BHA -Saldo Handelsbilanz	-14	-2	0	2	5	29	2,0
TATB Auftragsbestände Ind. Iv	-9	-2	-1	1	11	23	1,5
-TKap -freie Kapazitäten Iv	-23	-3	-1	0	15	22	1,5
VATI BNP - Beitrag Industie	-12	-4	-2	0	4	19	2,0
IPFI fertige InvestGueter	-17	-8	-3	0	7	36	4,0
PLC Verbraucherpreisindex	-19	-6	-3	3	7	38	4,2
TJob Beschäftigung Iv	-9	-8	-3	-1	1	23	3,5
UBI Beschäftigte Industrie	-10	-6	-4	-2	7	24	2,0
URI 1/Arbeitslosenrate	-17	-7	-4	0	6	33	3,5
BEUA Bruttoentgelt UnSelbstSt.	-17	-9	-8	-6	2	42	1,5

Die Resultate beziehen sich auf geglättete Zeitreihen bis Juli 86.

Min,Max Minimaler bzw. Maximaler Lead in Monaten.

Q1,Q3 Obergrenzen des 1. bzw. 3.Quartils der Leads.

M Median der Leads.

Err% $=100.p(sign(IndiZR_t - IndiZR_{t-1}) \neq sign(RefZR_{t+M} - RefZR_{t+M+1}))$

QS Mittlerer Quartilsspann der Leads = (Q3-Q1)/2.
Kriterium für die Stabilität der Leads.

\- Zusammenhang mit Referenzreihe negativ.

Iv Daten der österreichischen Industriellenvereinigung.

Die in Tabelle 1 wiedergegebenen mittleren konjunkturellen Vorläufe legen ein Kunjunkturzyklusmodell das wesentlichen Impulsen vom Kapital- und Geldmarkt sowie von der ausländischen Wirtschaftsentwicklung erhält nahe:

Entscheidungen des Geld - und Kapitalmarkts sind Investitions- und Produktionsentscheidungen vorgelagert. Der datierte Vorlauf des Kapitalverkehrsaldos LKV muß zwar als problematisch betrachtet werden[25)], deutlicher ist aber der Lead der realen Geldmenge (M1)[26)] und der untersuchten Zinssätze. Diese Leaddatierungen werden zudem deutlich von der Kreuzkorrelogramm gestützt. Freilich kann aus dem M1-Lead weder auf eine Kausalbeziehung[27)] noch auf die Effizenz der österreichischen Geldpolitik geschlossen werden[28)].
Den Haushaltserhebungen des IFES kommt unter den untersuchten Zeitreihen eine besondere Stellung zu. Explizit beziehen sich sich auf das Haushaltseinkommen im Befragungszeitpunkt im Vergleich zu dem vor einem Jahr (IFH0) bzw. auf die für nächste Jahr erwarteten Einkommensveränderungen (IFH1) wobei erstgenannte Zeitreihe den Erwartungen im Schnitt lediglich um einen Monat nachhinkt. Tatsächlich werden durch sie aber auch Konsum- und Sparbereitschaft, das erwartete effektive Arbeitsangebot aber auch allgemeine makroökonomische Einschätzungen der Tagespresse und die politische Propaganda erfasst[29)]. Im Rahmen dieser Arbeit wollen wir sie primär als Indikatoren für den Konsum[30)] und -sehr weit gefasst- für die wirtschaftliche Stimmung interprätieren. Sie eignen sich übrigens für die BIP- Vorhersage besser als für die der Industrieproduktion.
Daß wesentliche Konjunkturimpulse importiert werden verdeutlicht auch die Paralellität des Exports- (EX) und des IP-Wachstums, die isbesondere zwischen '72 und '82 frappierend[31)] ist. Der Lead zwischen unserer Referenzreihen und den Auftragseingängen aus dem Ausland (ATEA) ist stabiler als der zum gesamten Auftragsvolumen (ATE). Die Bedeutung eines Konjunkturimports wird zudem durch die

---

25) Bei extrem langen Verzögerung können Wendepunkte einer Zeitreihe den folgenden oder den bereits vergangenen der Referenzreihe zugeordnet werden. Die Analyse des Kreuzkorellogramms zwischen IP und LKV verweist ebenso wie die Err%-Funktion auf einen möglichen LKV-Lag von 10 Monaten. Da diese Statistik eher eine IP-LKV Sequenz, Err% dagegen eine LKV-IP Sequenz nahelegt wurden LKV in einigen LI-Kombinationen getestet. Es zeigte sich jedoch, daß die Aufnahme von LKV in langfristige KOnjunkturindikatoren diese nicht verbessern konnte.

26) Dieses für viele überraschende Ergebnis stimmt jedoch mit Breuss (84) überein, der ebenfalls bei der Untersuchung von Zeitreihen für die Bildung vorauseilender Indikatoren den Median der Leads von M1- zu IP-Wachstumswendepunkten mit 11 Monaten ausweist. Die für Höhe- und Tiefpunkte getrennte Datierung ergibt in ersteren einem mittleren Lead von 12 Monaten, in letzteren einen -wie bei den meisten vorauseilenden Zeitreihen - kürzeren Lead von 9 Monaten. (Breuss,F.(84),S.472)

27) Ohne näher auf die Problematik einzugehen, sei auf die von T. Sargents (82, Kap. XI,7) präsentierten Beispiele von Prozessen verwiesen, bei denen beliebige Phasenverschiebungen auftreten, obwohl keine Kausalbeziehung vorliegt, welche die Prognose der nachhinkende Variable durch die vorlaufende berechtigte.

28) Sofern letztere überhaupt existiert, steht sich doch im Schatten der Währungspolitik und reagiert damit primär auf internationale Wachstumsdifferenziale und die bundesdeutsche Wirtschaftspolitik.

29) vgl. Fels,W.(86)

30) Breuss,F.; Wüger,M., (85) registrieren einen mittleren Konjunkturlead zwischen den Zeitreihen und den Einzelhandelsumsätzen von 9 Monaten.

31) Ausserhalb dieses Bereichs vernebelt sich der Lead des Exports; in der letzten Phase scheint EX sogar hinter dem GDP herzulaggen.

Leads europäischer Konjunkturtests (etwa TIGH,TIEG) unterstrichen.
Im Vergleich zum Vorlauf der österreichischen Sekondärmarktrendite RSEC erweist sich zudem der Lead des schweizer Zinssatzes R3CH wesentlich stabilerer und längerer.
Im Konjunktureinbruch fallen dann etwa mit den Exporten die Inländischen Umsätze und die Beschäftigung sowie die Produktion der Industriegruppe Grundstoffe (IPGG) bevor die gesamte IP mit den Investitionen (INV) zurückgehen[32].
Als Kapazitätsanpassungsrunde kann die Anpassung der Lager (TLag,TKap)[33] und der durchschnittlich geleisteten Arbeitsstunden bezeichnet werden.
Die Beschäftigung in der Industrie, reagiert wie die Arbeitslosenrate erst mit einem Quartal Verzögerung auf IP- und INV-Einbrüche. Die Beschäftigungsanpassung der Gesamtwirtschaft folgt der in der Industrie (vgl. UB, UBI). Das Bruttoentgelt der Unselbstständigen (BEUA) reagiert schließlich auf die Veränderungen am Arbeitsmarkt[34].

KAP 5

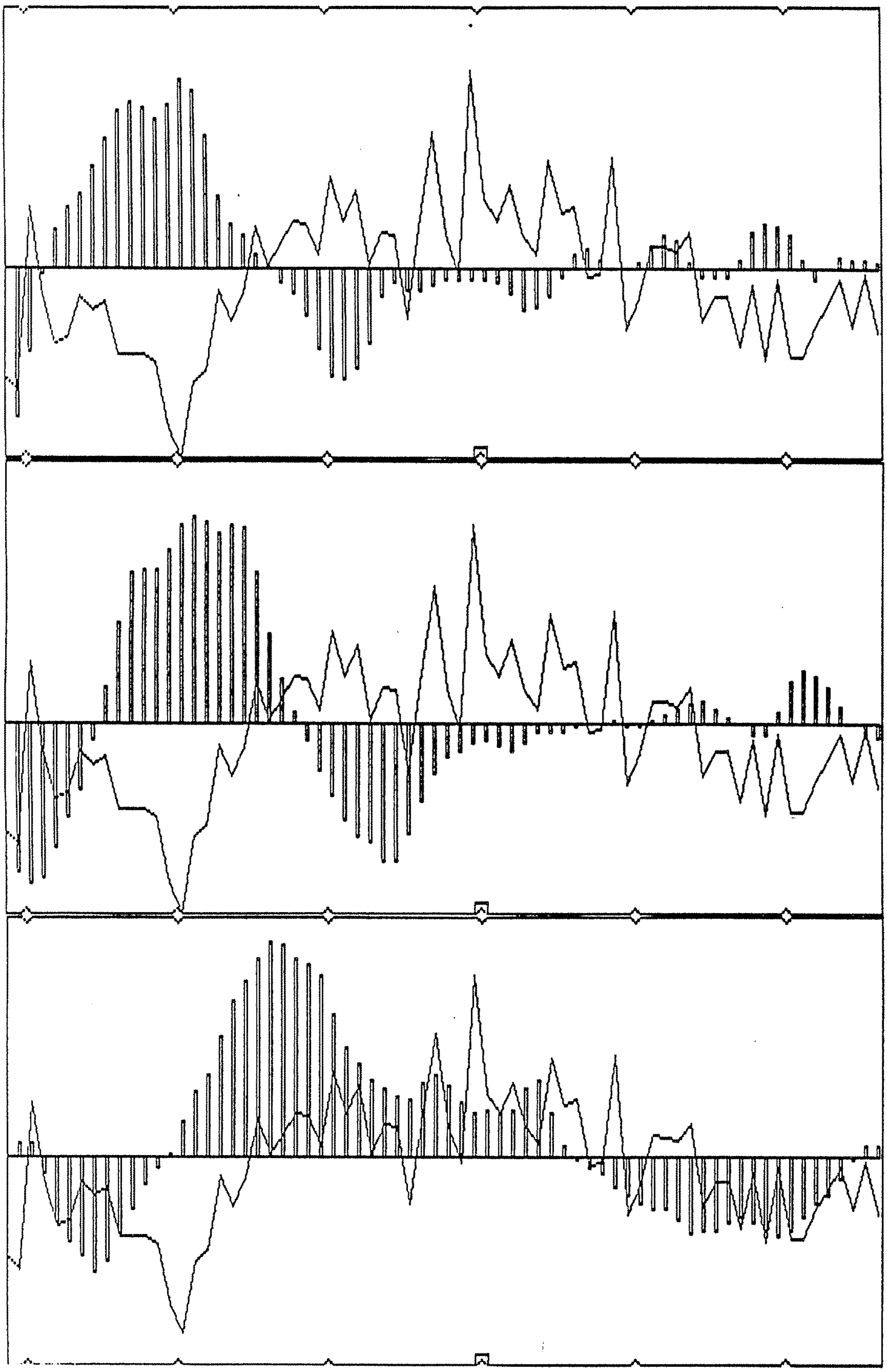
Tabelle 2 zeigt die Zusammenstellung jener LI, die aufgrund der angegebenen Kriterien sowie einer Untersuchung der Kreuzspektren zwischen der IP und möglichen Indikatoren aus der Menge möglicher Indikatoren ausgewählt wurden. Alle Indikatoren zeigen die dominanteste Kohärenz zu den Referenzreihen im tiefen konjunkturellen Bereich; bei manchen -insbesondere langfristige Indikatoren findet sich daneben noch eine hervorstechende Übereinstimmung im saisonellen Bereich ohne daß in diesem Bereich eine nennenswerte Phasenverschiebung auftritt.

Wesentlicher als jede Ex-post Analyse dürfte jedoch die Eignung in der tatsächlichen Prognose sein. Abb. 6 zeigt die Indikatoren und die IP-Wachstumsraten bis September 87 während alle bislang vorgestellten Analysen und Tests auf Daten bis zum September 86 basierten. Bedenkt man, daß zur Lead-Datierung der Zeitreihen mehrere Randmonate durch die Glättung verloren gingen und der letzte zuordbare Peak der IP der zum Jahresende 85 war, so muß die Periode seit Mitte 85 als echte Prognose bezeichnet werden. Es wurde bereits betont, daß die Indikatoren keine quantitativen Aussagen sondern Vorhersagen von

32) IPGG kommt damit eine besondere Stellung unter der gesamten Industie zu, die wie auch die Datierung von VATI- dem Beitrag der Industie um BIP - zeigt, gegenüber anderen Sektoren verzögert auf Wachstumsumschwünge reagiert. Nach der Untersucherung von Deistler und Schleicher (74) mit Daten von 54-70 leaden die Obergruppen Investitionsgüter (1.1 Monate), Bergbau und Grundstoffe (0.8 Monate) gegenüber dem Gesamtindex der IP. Besonders deutlich ist dabei der Vorlauf der vorallem exportorientierten Magnesitindustrie mit etwa 4.6 Monaten. Die Konsumgüterindustrie hinkt dagegen dem Gesamtindex um fast zwei Monate nach.

33) Dazu ist anzumerken, daß TLag und TKap, zwei ex-post Befragungen der österreichischen Industriellenvereinigung sind. Tatsächlich reagieren daher Kapazitäten etwa simultan zur Industrieproduktion, Lager aber etwa schon zwei Monate früher.

34) Die Beziehung von BEUA zu den Referenzreihen ist äusserst unstabil. Anzumerken ist, daß zwischen BEUA und GDP- (bzw. IP-) Wachstumsraten ein negativer Zusammenhang bei einem BEUA-Lead von 11 Monaten (bzw. etwa einem dreiviertel Jahr) mit Err% = 39 (bzw. Err% = 37).



<u>Abb.6 :</u> Balken : Vorlaufende Indikatoren IL (oben), IM (mitte), IK (unten)
durchgezogen : Referenzreihe (Standardisierte IP-Wachstumsraten)
Zeit : Dez. 82 und Aug. 87

Wachstumsumbrüchen erlauben sollten. Wir überlassen es dem Betrachter den Wert der LI bezüglich dieses Kriteriums einzuschätzen (s. Abb. 6).

Tabelle 2:

Zusammenstellung der vorauseilenden Indikatoren		kurz-, mittel- &.langfristige Indikatoren der Industrieproduktion		
		IK	IM	IL
Gewichte der Komponenten	bei Lag:			
ATE	0	,333		
IFH1<1>+IFH0	3		,453	
IHH1<1>+IFH0	0			,559
K$	0	,225		
M1	0			,531
M1	3		,438	
R3CH	0			-,637
R3CH	5		-,544	
R3CH	6	-,394		
RSEC	0		-,555	
TIEG	0	,589		
TIVG	0	,580		
Erfasste Varianz		,509	,705	,644

Qualitätskriterien	IK	IM	IL
(Leads)	4,75	9,82	13,28
(Leads)	2,14	3,31	3,12
Min	1	5	8
Q1	3	8	11
Lead Median M	5	9,5	13
Q3	6	11	14
Max	10	19	23
QuartilsSpann	1,5	1,5	1,5
Err% (M-3)	42	49	47
Err% (M-2)	33	43	41
Err% (M-1)	30	37	38
Err% (M)	28	35	37
Err% (M+1)	32	37	41
Err% (M+2)	36	45	48
Err% (M+3)	40	45	48
Max. Kreuzkorr. mit IP:	,744	,560	,638
bei Lead:	5	9	13

Erfasste Varianz : Anteil der im Indikator erfassten Varianz der verwendeten Reihen.

Err%(L) = 1oo.prob[sign(Ind(t-L)-Ind(t-L-1)) sign(IP(t)-IP(t-1))].

Datenquellen : Zur Zusammenstellung der vorauseilenden Indikatoren wie zum Vergleich mit den Referenzreihen (IP) dienten normierten 3R2-Glättungen der jährlichen Wachstumsraten.

LITERATURANGABEN

BOSCHAN, C., EBANKS, W., (78), 'The Phase-Average Trend, A Way of Measuring Economic Growth', in: Proceedings of the Business and Economic Statistics Section, American Statistical Association.

BRY,G., BOSCHAN, C.,(71), 'Progammed Turning Point Determination', in: Cyclical Analysis of Time-Series: Selected Procedres and Computer Programms, National Bureau of Economic Research.

BREUSS,F., (84), 'Konjunkturindikatoren für die österreichische Wirtschaft', WIFO-Monat, S.464-492.

BREUSS, F., WÜGER, M., (85), 'Consumers Sentiments as an Indicator for Consumtion Behaviour', Papier der 17. Ciret Konferenz, Wien.

BREUSS, F., WÜGER, M., (86), 'Consumer Sentiment as an Indicator for Consumption Behaviour in Austria', Empirica, Nr.13

DEISTLER, M., SCHLEICHER,S., (74), 'Strukturen der österreichischen Industrieproduktion, eine spektralanalytische Untersuchung', Empirica, S.72-97.

DRYDEN, J., REYNARD,B., (85), 'Cyclical Analysis and Leading Indicators - The OECD Experiment ', Papier der 17. Ciret- Konferenz, Wien.

HAERDLE,W., TUAN, P-H, (86), 'Some Theory of M-Smoothing of Time Series', Journal of Time Series Analysis, p. 191-204.

HUBER, P.J., (81), 'Robust Statistics', N.Y.

FELS,W., (86), 'Von Erwartungen zur Prognose', Diplomarbeit, Wien.

POLASEK, W., (82), 'Explorative Business-cycle Analysis Using Running Medians', Empirica, Nr.1, S.49-70

KLEIN,F.-J., NERB,G., (85), 'Performance of the EC composite Leading Indicator and its components', Papier der 17. Ciret Konferenz, Wien.

MOORE, G.H., (83), 'Business Cycles, Inflation and Forecasting', 2nd Edition, Cambridge,Mass.

MOSTELLER,F., TUKEY,J.W., (77), 'Data Analysis and Regression', Addison-Wesley.

TUKEY, J.W., (77), 'Exploratory Data Analysis', Addison-Wesley.

SARGENT, T.J, (82), 'Makroökonomik', Kap.XI,7, Oldenbourgh.

VELLEMAN, P.F.,(80),'Definition and Comparison of Robust Nonlinear Data Smootheng Algorithms', JASA, S.609-615.

SOFTWARE - QUELLEN

Für viele Teile dieser Arbeit lag keine Standartsoftware vor. Die wesentlichen Berechnungen wurden auf vom Autor erstellten PC-Turbo-Pascal Programmen durchgeführt. Die Programme 'DatOper' für Datenvorbereitungen und 'KonjAna' für Grafiken, Glättungen, Wendepunkt- und Leaddatierung stehen dem interessierten Leser zur Verfügung. Für einzelne Teile dieser Arbeit wurde daneben auf die Programmpakete 'RATS', 'IAS', und 'Statgraf' zurückgegriffen.

TESTING RICARDIAN EQUIVALENCE: ARE THE DATA INFORMATIVE?

Albert Jaeger*

Institute for Advanced Studies

Vienna

January 1988

TESTING RICARDIAN EQUIVALENCE: ARE THE DATA INFORMATIVE?

Albert Jaeger*

Institute for Advanced Studies

Vienna

January 1988

ABSTRACT

Recent empirical work on the Ricardian equivalence proposition has produced sharply conflicting evidence. This paper evaluates the claim that U.S. data for 1949-84 are not informative for discrimating between Ricardian and non-Ricardian restrictions on long-run consumption behavior. Evidence from cointegration tests supports the conclusion that U.S. data are not informative. Austrian data, on the contrary, are shown to contradict Ricardian equivalence.

* I thank Klaus Neusser for helpful discussion.

Abstract

Recent empirical work on the Ricardian equivalence proposition has produced sharply conflicting evidence. In this paper, I use a standard version of the permanent income hypothesis to derive a test for Ricardian equivalence that exploits long-run information in time series data. The test is based on the idea that "Ricardian consumers" will base their consumption decisions on an income concept which takes into account the intertemporal government budget constraint. Under Ricardian equivalence, consumption should be cointegrated with the modified income concept. It is possible, however, that time series data do not contain sufficient information for discriminating between Ricardian and non-Ricardian restrictions on long-run consumption behavior. The suggested test explicitly allows for this possibility.

Evidence from cointegration tests supports the conclusion that U.S. data spanning 1949-84 do not contain sufficient information to accept or reject the derived long-run implication of Ricardian equivalence. Austrian data, on the contrary, are shown to contradict Ricardian equivalence. The discriminatory power of the data in the Austrian case can be traced to a permanent shift in the mix between tax and bond financing of public expenditures around 1975.

I. INTRODUCTION

How changes in public deficits and debt affect private sector behavior is a perennial question in macroeconomics. The Ricardian equivalence proposition (REP), forcefully restated by Robert Barro (1974), holds that a change in the mix between tax and bond financing of a given public expenditure stream induces an offsetting change in private savings, leaving the economy's total savings unaffected.

Recent empirical tests of the REP using U.S. data rely on estimating consumption functions which include variables measuring public sector activities (taxes, transfers, expenditures, etc.) among the regressors. Restrictions on the parameters of these variables are employed to test whether consumers behave in a Ricardian or non-Ricardian manner. The empirical work following this approach has produced sharply conflicting evidence. Roger Kormendi (1983, p. 1007) reports "decisive rejection" of a specification assuming non-Ricardian behavior. Franco Modigliani and Arlie Sterling (1986, p. 1179), on the contrary, conclude that the data are "strikingly and unmistakenly consistent" with the assumption of non-Ricardian behavior. Other authors have adopted an agnostic attitude. James Poterba and Lawrence Summers (1987, p. 378), for example, point out that until recently U.S. economic history "has provided relatively few powerful tests" of the REP. A comprehensive survey of this research is contained in Douglas Bernheim (1987).

In this paper, I derive a novel test for Ricardian equivalence that exploits the long-run information in time series for private consumption, income, and budget deficits. I assume that a standard version of the permanent income hypothesis (PIH) describes consumption behavior. Under the PIH, consumption and a modified measure of disposable income which takes into account the government budget constraint should be cointegrated. Recent tests of the PIH have generally ignored this implication of infinitely forward looking consumption behavior. John Campbell's (1987) "rainy days" test of the PIH, for example, is based on the assumption that consumption and conventially defined disposable income are cointegrated, thereby effectively ignoring the government budget constraint. It can be shown, however, that these two implications will be observationally equivalent if no permanent shift in the mix between tax and bond financing took place in the time period used for testing. In this case, both measures of income will be cointegrated with consumption and the long-run information in the data will not be informative for testing the REP.

The finding that the data do not contain long-run information for testing Ricardian equivalence should go a long way towards explaining the disparate conclusions reached by different authors. I report results of cointegration tests for U.S. data covering the period 1949-84, supporting the thesis that the data are not informative. For illustration, I also report results for Austrian data that clearly reject the hypothesis that consumers take into account the government budget constraint. The discriminatory power of the data in the Austrian case can be traced to a permanent shift in the mix between tax and bond financing of public expenditures around 1975.

The paper is organized as follows: Section II discusses the theoretical framework. Section III reports the empirical findings for the U.S.A. and Austria. Finally, section IV contains a brief discussion of the results.

II. THEORETICAL CONSIDERATIONS

I assume that the standard life cycle-permanent income hypothesis with infinite horizon and rational expectations describes private consumption behavior

$$(1)\quad C_t = \Gamma(r/(1+r))\ [A_t + \sum_{i=0}^{\infty}(1/(1+r))^i\ E_t(Y_{t+i}-T_{t+i})].$$

In (1), C_t is real private consumption, A_t real non-human assets including real public debt, Y_t real gross labor income, T_t taxes net of transfers, r is the after-tax real interest rate taken to be constant and Γ is a proportionality factor. E_t denotes the mathematical expectations operator conditional on full public information at time t. The expression in brackets times r/(1+r) denotes permanent income defined as the annuity value of the sum of human and non-human wealth.

Decision rule (1) for private consumption is based on a set of rather restrictive assumptions. First, real interest rates are assumed constant and consumers can freely borrow and lend at this rate in a fictitious capital market subject only to their life time budget constraint. Second, the utility function is time-separable in private consumption and excludes public consumption as an argument. Third, the optimization problem of the representative consumer exhibits certainty equivalence, thereby excluding precautionary motives. And fourth, the planning horizon stretches into infinity and decisions are based on all relevant information about the future course of income and taxes.

The last assumption implies that private agents take into account the government budget constraint when deciding on consumption and saving

$$(2) \qquad B_t = \sum_{i=0}^{\infty}(1/(1+r))^i E_t T_{t+i} - \sum_{i=0}^{\infty}(1/(1+r))^i E_t G_{t+i}.$$

Here, B_t denotes real public debt and G_t real government consumption. This equation can be derived by iterating the government budget identity, $B_{t+1} := (1+r)[B_t+G_t-T_t]$, forward in time and imposing a solvency constraint.[1]

Inserting (2) in (1) results in

$$(3) \quad C_t = \Gamma(r/(1+r)) \, [(A_t-B_t) + \sum_{i=0}^{\infty}(1/(1+r))^i \, E_t(Y_{t+i}-G_{t+i})].$$

Equation (3) is the decision rule for consumption of private households under the standard version of the PIH used in this paper. Consumers will not consider the public debt as net private wealth and they deduct consumption of real resources by the public sector from gross income instead of taxes. In other words, Ricardian equivalence is assumed to hold.

Recent literature on testing the PIH, e.g. John Campbell (1987), has generally ignored the implication that consumers with infinite horizon and rational expectations will take into account the government budget constraint for calculating permanent income. They have relied on equation (1) or some "Euler equation" analogue of it to derive testable implications. In the next step, I show that equations (1) and (3) have rather different implications.

For the further derivations, I identify disposable income of private households (YD) as conventionally measured with the

1 The government budget identity excludes financing of the budget deficit by printing money. This assumption is probably not too restrictive for countries with low inflation rates.

following expression

$$(4) \quad YD_t = (r/(1+r))A_t + Y_t - T_t.$$

Similarily, a measure of "Ricardian disposable income" (YD^*) can be defined as

$$(4)' \quad YD^*_t = (r/(1+r))(A_t - B_t) + Y_t - G_t.$$

Multiplying (4) and (4)' by the proportionality constant Γ and subtracting these expressions from (1) and (3) respectively, results after some algebraic manipulations in the following equations

$$(5) \quad C_t - \Gamma YD_t = \Gamma \sum_{i=1}^{\infty} (1/(1+r))^i \, E_t[\Delta Y_{t+i} - \Delta T_{t+i}]$$

and

$$(5)' \quad C_t - \Gamma YD^*_t = \Gamma \sum_{i=1}^{\infty} (1/(1+r))^i \, E_t[\Delta Y_{t+i} - \Delta G_{t+i}].$$

An equation similar to (5) was first studied by Campbell (1987). To repeat, (5) is based on an unexplained failure of consumers to take into account the government budget constraint whereas (5)'is based on the full set of restrictions. As pointed out by Campbell, equations of this type possess a nice economic interpretation: The linear combination on the left hand side corresponds to a measure of savings (exactly so if Γ is 1) and is an optimal predictor of expected future changes in the income measures on the right hand side.

Using the concept of cointegration recently developed by Robert Engle and Clive Granger (1987), testable long-run implications can be derived from (5) and (5)'. Under the maintained hypothesis that gross labor income, net taxes, and government expenditures are stationary after first differencing, the sums on the right hand sides of (5) and (5)' will also be stationary. Given that consumption and the disposable income variables are non-stationary variables[2] in levels, (5) implies that C_t is cointegrated with YD_t and (5)' implies that C_t is cointegrated with YD^*_t.[3]

These results suggest the following test: Given the standard permanent income hypothesis is valid, C_t and YD^*_t should be cointegrated according to (5)'. But this result does not imply that C_t and YD_t can not be cointegrated. The Ricardian measure of disposable income, YD^*_t, is simply the sum of conventially measured disposable income, YD_t, and the government budget deficit. C_t and YD_t could as well be cointegrated if the deficit was stationary in the period under observation. In this case, the long-run information in the data will not be sufficiently powerful to accept or reject Ricardian equivalence. In summary, the cointegration tests could give rise to four possibilities:

(i) C_t is cointegrated with YD_t as well as YD^*_t. The data generated by this economy should be judged uninformative with respect to the validity of the REP.

2 If Γ is smaller than 1, consumption as well as disposable income will not be stationary after first differencing (see Campbell (1987, p. 1254). But the linear combination between the two variables will still be stationary.

3 Engle and Granger (1987) call a variable X_t integrated of order one if it is stationary after first differencing. Two variables X_t and Y_t are called cointegrated if they are (i) individually integrated of order one but (ii) a linear combination between the two variables is integrated of order zero. For a more exact definition see Engle and Granger (1987, pp. 252-53).

(ii) C_t cointegrates with YD_t but not with YD^*_t. The standard version of the PIH which implies Ricardian equivalence has to be rejected.

(iii) C_t cointegrates with YD^*_t but not with YD_t. PIH as well as REP can not be rejected.

(iv) C_t neither cointegrates with YD_t nor with YD^*_t. The assumed behavioral hypothesis for long run consumption behavior is rejected and we can infer nothing concerning the validity of the REP within the testing framework of this paper.

III. EMPIRICAL RESULTS

The empirical analysis is based on Austrian and U.S. data. The data for the USA are taken from Modigliani and Sterling (1986). These authors perform several adjustments of the offical data as decribed in the data appendix of their paper. The Ricardian measure of income employed in this paper was constructed by simply deducting the public deficit variable as listed by Modigliani and Sterling from the conventionally defined income variable. The Austrian data were constructed as follows: The income variable corresponds to the disposable income of private households. The Ricardo definition of income corresponds to the sum of convential disposable income and public savings as reported in the national income accounts.[4] This definition effectively pools private and public savings behavior. The consumption concept is total private consumption. The deflator of total private consumption was used to convert nominal into real magnitudes.

As a first step, in table 1 I report results of Dickey-Fuller tests for unit roots in the time series used for testing cointegration. The null-hypothesis of non-stationarity in levels of the three variables in lines (1) to (3) for both countries can not be rejected (results are not reported). Lines (1) to (3) contain test statistics for unit root tests after first differencing of the variables. A comparison of the statistics with the critical values listed at the bottom of table 1 reveals that all variables with the possible exception of the U.S. consumption series can be judged stationary after first differencing with high confidence.

In lines (5) and (6), Dickey-Fuller tests for the level as well as

4 Note that public savings as calculated in the national income accounts is the appropriate "public deficit" variable because it does not depend on public investment expenditures.

the first difference of the public deficit variable (D_t) are reported. If permanent shocks in this variable occured during the observation period, D_t should be non-stationary in levels. This is the case for the Austrian time series. The U.S. fiscal deficit appears to be stationary in levels according to the test statistics in line 5.

The results of the tests for cointegration in table 2 can be summarized as follows: For the Austrian data, consumption is cointegrated with conventially defined disposable income (equation 1) but not with the Ricardian definition of income (equation 2). The Dickey-Fuller tests reject the null hypothesis of no cointegration for convential disposable income at very low significance levels. In the cointegrating regression using the Ricardian income concept, the null hypothesis can not be rejected and the DW test statistic drops markedly. From the regressions for the U.S. data (equations 3 and 4) the conclusion is that cointegration seems to hold for both income concepts.[5] Here the test statistics for cointegration as well as the coefficient estimates are very similar across the equations demonstrating that the data do not contain enough information to discriminate between Ricardian and non-Ricardian behavior.

The results for the United States reflect the fact that the time series for the U.S. fiscal deficit was stationary over the period of observation. As more data on the U.S. experience with historically high peace-time deficits in the 80s become available, the data might well reject the REP as asserted by Poterba and Summers (1987), who evaluate data up to 1986 on an informal basis. The information content in the Austrian case can be traced to a

5 As pointed out by Anindya Banerjee et.al. (1986), OLS-estimates of cointegration vectors can exhibit considerable small-sample bias despite the "superconsistency"-result derived for this estimator by James Stock (1987). They suggest R^2 as an index for judging whether the bias is small. Informally, an estimate higher than .95 should be assurance that the bias is small. The reported R^2-statistics in table 2 are always higher than .95.

change in the mix between bond and tax financing around 1975. Although, the boosting of the public deficit was originally planned as a transitory measure to stabilize aggregate demand in the face of the first oil shock, it became permanent in the aftermath. As a result, public savings in percentages of private disposable income declined from 10.6 % (mean 1960-75) to 4.2 % (mean 1975-86). The permanent shock in the Ricardian measure of disposable income was obviously ignored by Austrian consumers who based their consumption plans on disposable income as conventially measured instead.

Table 1: Dickey-Fuller tests for unit roots*

		Country/Test Statistic			
		Austria 1960:86		USA 1949:84	
	Variable	DF	ADF	DF	ADF
(1)	ΔC_t	-6.00	-3.62	-2.84	-3.19
(2)	ΔYD_t	-4.51	-4.38	-5.47	-3.76
(3)	ΔYD_t^*	-3.73	-3.05	-5.31	-4.71
(4)	D_t	-1.73	-2.35	-3.16	-4.99
(5)	ΔD_t	-3.72	-4.53	-	-

* These tests are based on the regression $\Delta X_t = \mu + \alpha X_{t-1} + \sum_{i=1}^{n} \beta_i \Delta X_{t-i}$. The Dickey-Fuller test (DF) assumes n=0 and uses the t-statistic on α. The augmented Dickey-Fuller tests (ADF) were performed setting $n = 1$ and again use the t-statistic on α. Critical values from Fuller (1976) are: 1% -3.75; 5% -3.0; 10% -2.63; (sample size = 25).

Table 2: Cointegration tests*

Austria 1960:86

(1) $C_t = 5.746 + .892\ YD_t$ $\quad R^2 = .997$

$\qquad\quad (4.219) \quad (.010)$

DW = 1.59 DF = -4.02 ADF = -3.37

(2) $C_t = -27.541 + .910\ YD^*_t$ $\quad R^2 = .984$

$\qquad\quad (10.252) \quad (.023)$

DW = .45 DF = -2.02 ADF = -2.19

USA 1949:84

(3) $C_t = -.201 + .976\ YD_t$ $\quad R^2 = .990$

$\qquad\quad (.058) \quad (.017)$

DW = .911 DF = -3.36 ADF = -3.14

(4) $C_t = -.319 + .997\ YD^*_t$ $\quad R^2 = .969$

$\qquad\quad (.108) \quad (.031)$

DW = .818 DF = -2.86 ADF = -3.33

* The Engle and Granger (1987, p.269) critical values for the null hypothesis of no cointegration are Durbin-Watson (DW) 1 % .511, 5 % .386, 10 % .322; Dickey-Fuller (DF) 1 % -4.07, 5 % -3.37, 10 % -3.03; Augmented Dickey-Fuller (ADF) 1 % -3.77, 5 % -3.17, 10 % -2.84. These critical values are based on sample sizes of 100. Engle and Yoo (1987) present critical values for cointegration tests for sample sizes of 50. Their values are only slightly below those listed above.

IV. DISCUSSION

The empirical results in the preceding section suggest that the typical U.S. data set does not contain sufficient long-run information to accept or reject the REP whereas Austrian data clearly reject the proposition. Most of the research on this issue, summarized conveniently by Bernheim (1987, p. 278), has typically relied on regressing consumption on income, wealth, and variables like taxes, government expenditures, transfers, and public debt. If these regressions are performed in levels, the non-stationarity of the listed variables could make inference from the regression results hazardous. The usual asymptotic distribution theory can not be invoked and if consumption is cointegrated with a subset of the right-hand variables, the coefficients on the other non-stationary variables should be zero by construction. First differencing, as recommended by Kormendi (1983), would not be an appropriate remedy, however, as this operation destroys the long-run information in data. And this type of information is likely to be most helpful to distinguish between Ricardian and non-Ricardian behavior in view of the notorious difficulties to discriminate between macroeconomic hypotheses on the basis of noisy short-run fluctuations in time series. Given these considerations, the mixed evidence reached by different authors working with U.S. data should not come as a surprise.

Empirical evaluation of the REP faces some awkward questions as far as the quality and economic meaningfulness of the measures of public sector activities is concerned. Well kown non-sensical accounting conventions are discussed extensively by Robert Eisner (1986). Further research should therefore investigate whether the empirical results of the cointegration tests suggested in this paper are robust with respect to different measures of public sector activities.

A rather different issue is the question what can be learnt from rejections of the REP on the basis of cointegration tests. I have argued in this paper that Ricardian equivalence is a direct implication of the standard PIH. Therefore, rejections of the REP should lead to a reconsideration of the assumptions the PIH is derived from. Now, the standard version of the PIH as interpreted in literature is generally considered as being inconsistent with at least some features of the data (see e.g. Angus Deaton (1986)). Recent research concentrates on the separability assumption (e.g. John Muellbauer (1986)) or the assumptions of perfect capital markets (e.g. John Campbell and Gregory Mankiw (1987)) as the critical assumptions to be reconsidered. Rejection of the REP, however, points to the assumptions of an infinite planning horizon and rational expectations as the most likely culprits for the failure of the PIH. Finite horizons have always been considered as one of the possible stumbling blocks for Ricardian equivalence. Reconsideration of rational expectations as a working hypothesis is plausible because of the information assumptions necessary to derive the REP.

References

BANERJEE, Anindya, Juan J. DOLADO, David F. HENDRY and Gregor W. SMITH (1986): "Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence", Oxford Bulletin of Economics and Statistics 48, 252-277.

BARRO, Robert J. (1974): "Are Government Bonds Net Wealth?," Journal of Political Economy 82, 1095-117.

BERNHEIM, B. Douglas (1987): "Ricardian Equivalence: An Evaluation of Theory and Evidence", in: Stanley Fischer (ed.): Macroeconomics Annual 1987, 263- 304, MIT Press.

CAMPBELL, John Y. (1987): "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," Econometrica 55, 1249-73.

CAMPBELL, John Y. and MANKIW, Gregory N. (1987): "Permanent Income, Current Income and Consumption," mimeo.

DEATON, Angus S. (1986): "Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory", NBER Working Paper No. 1910.

EISNER, Robert (1986): How Real is the Federal Deficit?, Free Press, New York.

ENGLE, Robert F. and GRANGER, Clive W.J. (1987): "Co-Integration and Error Correction: Representation, Estimation and Testing," Econometrica 55, 251-276.

ENGLE, Robert F. and YOO, Byung Sam (1987): "Forecasting and Testing in Co-Integrated Systems," Journal of Econometrics 35, 143-159.

FULLER, Wayne A. (1976): Introduction to Statistical Time Series, Wiley, New York.

KORMENDI, Roger C. (1983), "Government Debt, Government Spending and Private Sector Behavior," American Economic Review 73, 994-1011.

MODIGLIANI, Franco and STERLING, Arlie (1986): "Government Debt, Government Spending and Private Sector Behavior: Comment," American Economic Review 76, 1168-79.

MUELLBAUER, John (1986): "Habits, Rationality and Myopia in the Life-Cycle Consumption Function", CEPR Discussion Paper no. 12.

POTERBA, James, and SUMMERS, Lawrence H. (1987): "Finite Lifetimes and the Effects of Budget Deficits on National Saving", Journal of Monetary Economics 20, 369-391.

STOCK, James H. (1987): "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors", Econometrica 55, 1035-56.

COINTEGRATION IN A MACRO-ECONOMIC SYSTEM[1]

Robert M. Kunst

Institute for Advanced Studies

Vienna

December 1987

1 Note: This publication makes part of a joint investigation on co-integrating structures within an Austrian macroeconomic system which is currently performed by the author together with Klaus Neusser, University of Vienna.

1. Preliminary remarks

In the past few years, the phenomenon of cointegration has attracted widespread attention. Many investigations in this area are still unpublished and are changing hands as mimeos and lots of problems connected with the idea are still unsolved. Fundamental properties are well summarized in the seminal paper by Engle and Granger(1987) and therefore the problem will be explained only in short here.

It is known that most economic variables - at least national accounts variables - are well represented individually by a staionary autoregression on first logarithmic differences. This means that the logs of the variables are "integrated" (of order one). Furthermore it is known that macroeconomic variables are interdependent which means that univariate time series models can hardly be an efficient strategy. However, the naive conclusion to use vector autoregressions (VAR) in differences is only correct in the special case that all linear combinations of the (trending) variables are also trending (i.e. integrated). If there is a linear combination which is wide-sense stationary, the variables are said to be "cointegrated" and the linear combination is sometimes called a "cointegrating vector".

Empirical investigations have shown that cointegration is not an exception but rather the rule within macroeconomic VAR systems. The present paper uses a variables set which is quite similar to a core set for the Austrian economy which already has been exploited successfully for VAR forecasting both by models in levels and in differences (see Kunst and Neusser(1986)). We will try to answer questions like the following ones:

Are there cointegrating vectors in the system ? If yes, how many ? Can they be given an economic interpretation ? Does forecasting benefit from incorporating cointegration into the model ?

2. The applied methods

The question concerning the existence of cointegrating relationships in a multivariate system may be investigated into by several procedures, e.g. Stock and Watson (1986), Phillips and Ouliaris (1987), Johansen (1987). The former two articles are concerned solely with the problem of performing tests about the number of relationships, whereas Johansen's paper provides for a unified approach for the testing problem as well as for the estimation of the cointegrating vectors or, rather, of an orthogonal basis for the r-dimensional cointegrating space if r denotes the number of independent relationships identified in the testing stage. This paper will be concerned primarily with Johansen's approach although the results of different tests will be also mentioned.

It seems worth contrasting the methods from the cointegration literature with an older technique based on canonical correlation analysis. The concept of canonical correlations was introduced by Hotelling(1936) and applied to the vector-autoregressive setting by Box and Tiao(1977). Similar to Johansen's method, but relying on stationary theory with near-unit roots (sometimes called "almost non-stationary" processes ANS, see Tjøstheim and Paulsen (1982)), their procedure identifies orthogonal vectors which transform the time series at hand into stationary (far-from-unit-root) series.

2.1 Johansen's method

Johansen(1987) starts from a vector-autoregressive process of order k and dimension p

(1) $Y_t = \Sigma\ \pi_i Y_{t-i} + e_t$

with non-singular (not necessarily diagonal) covariance matrix Ω. The process may be re-parametrized as a first-difference process

(2) $\Delta Y_t = \Sigma\ \Gamma_i \Delta Y_{t-i} + \Gamma_k Y_{t-k} + e_t \qquad \Gamma_i = -I + \pi_1 + \ldots + \pi_i$

If Γ_k is zero, (Y_t) is autoregressive as a first-difference process. This coincides with the case of Y_t depending on p independent trends in the terminology of Stock and Watson (1986). On the other hand, if Γ_k has full rank, Y_t is stationary. In all other cases, Γ_k has rank r between 0 and p, i.e. it may be represented as

(3) $-\Gamma_k = \alpha'\beta\ (= I - \pi_1 - \ldots - \pi_i)$

with the dimension of α and β equal to r X p and full rank r. The sign convention was chosen to recover the "impact matrix" in parentheses. β is not unique but the space spanned by its column vectors is and may be estimated together with the remaining parameters in the problem by the method of maximum likelihood assuming normally distributed errors. According to Johansen's derivation, this may be done by using moments and cross-moments matrices of the residuals from the regression of ΔY on the lagged differences

(4) $\Delta Y_t = \Sigma\, a_i \Delta Y_{t-i} + r_{0t}$

and of Y_{t-k} on the same regressors

(5) $Y_{t-k} = \Sigma\, a_i \Delta Y_{t-i} + r_{kt}$

The moment matrices are called S_{00}, S_{0k}, S_{kk}. Now, the eigenvectors of

(6) $S_{kk}^{-1} S_{0k}' S_{00}^{-1} S_{0k}$

corresponding to the smallest r eigenvalues give the required β basis and

(7) $\alpha = -S_{0k}\beta$

solves the corresponding α problem. Finally, the full model is given by

(8) $\Delta Y_t + \alpha'\beta Y_{t-k} = \Sigma\, \delta_i \Delta Y_{t-i}$

While these calculations are rather straightforward, two problems will arise in macro-economic practice which have not been covered fully by present theory. First, seasonality remaining in the data might cause roots on the unit circle different from one which will impair or even destroy the procedure. Second, the analysis is founded on polynomial matrices of order k while, in reality, the optimal order k may differ between elements rendering the procedure inefficient.

The number r which gives the number of cointegrating vectors or, likewise, the smaller dimension of α and β, is to be determined by a likelihood-ratio test of the null hypothesis of "at most r cointegrating vectors". The statistic to be used is

(9) $$-2\log Q = -T \sum_{i=r+1}^{p} \log(1-\tau_i)$$

comprising the p-r smallest eigenvalues τ_i of the above problem. The asymptotic distribution of (9) is tractable but complicated. Selected fractiles are given in Johansen(1987).

2.2 Box and Tiao's method

As stated above, Box and Tiao's method relies on the assumption of jointly stationary processes which, however, includes almost non-stationary cases. It is well known that discrimination between highly dependent stationary processes and pure random walks is impossible at conventional significance levels for smaller samples like the one under investigation (e.g. compare the simulation results in Stock and Watson(1986)). Conversely, the system which provided an example in Box and Tiao(1977) probably would be judged to be co-integrated by today's standards.

Given a correct forecast for the vector X_t from its past, Z_t, ideally the conditional expectation, the eigenvalues and eigenvectors of the matrix

(10) $(X'X)^{-1}(Z'Z)$

provide special information on the dynamic structure. In the univariate case, it is obvious that (10) is restricted to the interval (0,1) and that values close to one represent processes whose innovations variance is negligible as compared to the process variance. This condition is fulfilled for the near-unit root cases. Conversely, values close to zero represent random processes of low temporal dependence approaching white noise. In the multivariate case, the eigenvalues are restricted to the same interval. Additionally, the corresponding eigenvectors may be used to transform the vector (X_t) into a component vector (Y_t) consisting of recursively dependent components corresponding to the respective location of the eigenvalues.

Although Box and Tiao only used the case of first-order autoregressive forecasts the method directly extends to higher-order processes. However, if the lag order is increased, all eigenvalues are taken towards one since no AIC-like term is involved impeding parameter inflation relative to prediction accuracy. An alternative way of generalisation was followed by Velu et al.(1987) sticking to the interpretation of the AR(1) case as looking for the canonical correlations between X_t and X_{t-1}. Consequently, they look for the canonical correlations between X_t and X_{t-p} in the AR(p) model, adjusted for the observations in between. This method will not be used here.

What makes Box and Tiao's procedure interesting in the context of co-integration is that any component within (Y_t) corresponding to an eigenvalue far from unity, that is, a stationary component, necessarily gives a co-integrating relationship and makes the respective eigenvector a co-integrating vector. Thus, the Box-Tiao eigenvectors may be compared to the Johansen eigenvectors although, of course, only the latter ones are derived from genuine co-integration theory.

3. Data preparation

The procedures outlined in Section 1 demand for series that are almost or quite non-stationary individually but whose non-stationarity may be removed by first-order differencing. However, seasonally non-adjusted macroeconomic series tend to show seasonal near-unit roots commonly removed by fourth-order differencing. Although generalizing the outlined methods to the case of fourth-order differencing is straightforward in principle, distributional results are not yet available and, moreover, seasonal aberrations together with stochastic trends would complicate the matter considerably. For reasons not fully given in this paper, the following data set of quarterly Austrian series (1964-1985) was selected for the investigation:

GDP real gross domestic product
C real private consumption
IF real gross fixed investment
PGDP deflator of GDP
R interest rate on the secondary market
WAGE gross wages per employee

The latter two series are used in nominal terms. Except for R, all of these series have been logged to remove heteroskedasticity effects. Again except for R, the series seem to be trending and, except for R and IF, they show strong seasonal patterns. By performing the usual unit-root tests (see Dickey and Fuller(1976)), all series may be judged to be integrated of order one and to satisfy our assumptions - under the condition that some sort of seasonal adjustment is applied. Simple regression on four seasonal dummies seems enough to eliminate any indication of additional near-unit roots. It is, however, not enough to eliminate any indication of seasonal patterns. This procedure was selected since it is some kind of "soft" adjustment unlikely to destroy the relationships between individual series, which would happen, of course, if fourth differences were applied. The length of the time series does not allow to model the - obviously time-changing - seasonal patterns completely. Similar ratiocinations led Juselius(1987) in her report of an application of Johansen's technique to a similar step.

After extracting seasonal effects from the data, the question of integratedness should be revisited. Two tests were performed: the original test by Dickey and Fuller(1976) and the univariate version of Johansen(1987)'s canonical correlation analysis. Both statistics are similar and have equivalent asymptotic distributions. Both rely on partial correlations between differenced and level series, conditional on past differences. Both critically rely on the number of lags involved. In our case, the null hypothesis of first-order integratedness was rejected for the interest rate and accepted in the remaining cases. Further analysis showed that R contains a near-unit root and is somehow modeled more parsimoniously by differencing. Additionally, the tests are known to be biased against the null in smaller samples. Thus, R still may be viwed as a borderline case, and a model assuming stationary R will be compared to a model assuming integrated R.

4. Results

4.1. Results from Johansen's method

While running Johansen's procedure on the data the lag order K has to be fixed. This K corresponds to k-1 according to the first section. K gives the order of the autoregression within the differenced model while k corresponds to the order of the autoregressive model in levels (2.1). The shape of β as well as the identified number of cointegrating relations crucially depends on this parameter. The following list gives the identified β dimensions depending on the lag order based on a test significance level of 2.5 %.

lag order	# coint.
1	2
2	4
3	3
4	2
5	3
6	4
7	4
8	4

Our selection criterion for the lag order was the AIC (Akaike's Information Criterion) of the error correction representation of the final model (2.8). The AIC was applied to single equations rather than to the multivariate model in order to avoid the downward bias of multivariate information criteria. AIC reaches its minimum for K=1 in the case of the interest rate, for K=2 in the case of gross investment, and for K=4 in the remaining cases. This suggests K=4 for further research but may be also seen as indicating seasonality within the data. Of course, estimating a full model of order 6 to 8 would imply terrible over-parametrization effects, anyway.

A different criterion also suggests 4 lags within differences, namely the Q statistic of the final equations which tests for uncorrelatedness of the errors. For any k smaller than 4, residual correlation is significant, extremely so for prices and wages. This correlation disappears in the 4 lags specification. Of course, this is another indication of the presence of seasonality in the (adjusted) data.

After fixing the lag order at 4, the number of cointegrating vectors has to be identified. Although the automatic approach suggests 2 vectors, the decision of ignoring a third vector is close to the significance boundary. Therefore, both models - K=4 with 2 and 3 cointegrating vectors - have been tried. The resulting single equation standard errors are given in table 1. They are contrasted with the errors from a pure differenced model assuming no cointegration at all.

TABLE 1 : Standard errors in the final equations
(r = number of cointegrating vectors)

r =	2	3	0	1(5v)
GDP	.0122	.0120	.0135	.0132
C	.0195	.0182	.0204	.0210
IF	.0320	.0317	.0359	.0338
PGDP	.0112	.0111	.0122	.0112
RS	.2199	.2189	.2293	.2234
WAGE	.0156	.0155	.0156	.0157

The model with 3 vectors is slightly better than the other one for all series. However, the information criterion approach suggests the simpler model with 2 vectors since the marginal improvement cannot make up for the additional parameters used in estimating α and β.

It looks tempting to apply subset searches to the identified model in order to reduce the amount of insignificant coefficients. It is known that subset VAR models usually outperform unrestricted VAR models (see Kunst and Neusser 1986). Three special features of the cointegrated model, however, make this approach less attractive than in the classical model. First, it is unknown on which regression out of (2.4),(2.5),(2.8) elimination should rely. Second, the derivation of Johansen only works in the case of the same regressors entering into each of the system equations. Especially, if the set of regressors differs between equations, single equation OLS is no more maximum-likelihood. Third, and this is an empirical outcome of this investigation, subset selection relying on the final equation (2.8) is incompatible with the "non-stationary" regression (2.5) where the eliminated regressors are needed to establish co-integration. Consequently, performing the analysis on the reduced model renders a model without co-integration, i.e. a multivariate first-difference model. Even for very careful elimination only of the most insignificant parameters in (2.8), any co-integration within our model was destroyed.

As stated above, the 6 variables model (6v) is compared to a model assuming stationary interest rate R. In this case, R cannot partake in the formation of cointegrating vectors. Therefore, the 5 remaining variables were subjected to Johansen(1987)'s procedure. In the setting of the final equations (2.8), the stationary R is again included.

Again, lag order 4 is identified whereas the dimension of the cointegrating space now is reduced to one. The following vectors were given by the procedures:

	GDP	C	IF	PGDP	R	WAGE
5v	-61.89	7.79	21.50	31.12		-7.86
6v	52.23	-6.25	-22.98	-48.27	-1.28	23.59
	-7.03	-65.96	2.16	-41.97	-1.96	55.62

A full economic interpretation of these vectors is outside of the aims of this paper. Anyway, the first vector in (6v) tells that there is some sort of direct relation between the ratio of national income to prices and the ratio of investment to nominal wages. The second vector seems to link consumption to the real wage which also makes sense. The vector in (5v) is more difficult to interpret. It may be seen as a relation of the kind

$$GDP = a_1\, C/WAGE + a_2\, IF + a_3\, PGDP$$

with the sum of the coefficients approximately equal to one. The principal difficulty arises from a mixing of real and nominal variables without obvious theoretical sense. Note that the co-integrating 5v basis vector is not quite included in the cointegration space of the 6v version though it somehow resembles the first 6v vector.

If R is inserted into the final 5v system (2.8), standard errors deteriorate slightly relative to the 6v system (see Table 1, last column).

Since observations of all system variables are available after the end of the period used for the calculations, it is possible to compare some ex-ante forecasts based on the final equations with reality. Any evidence from this exercise is to be seen as preliminary. (6v) assumes integrated interest rate R which causes R to remain near the low level of 1985:4 if pure differencing is used. If 2 cointegrating vectors are inserted, R rises towards its long-run average. Of course, this rise also appears in the (5v) models with or without cointegrating restrictions. The use of co-integration in (5v) has little effects relative to pure (5v) differencing, except for higher investment IF. GDP growth during 1987 is predicted to a bit less than 2 %. The (6v) model with 2 vectors is the most inflationary but forecasts less growth in real aggregates GDP, C, and IF. Actual growth was low, which is in favor of (6v)-2, but inflation was also low, which is in favor of (5v) or of neglecting co-integration. Interestingly enough, actual R remained low, a property only mirrored by (6v) without cointegration and evidence in favor of the unit root hypothesis for the interest rate.

Of course, longer forecast horizons and more experiments will be necessary to decide on the predictive quality of the cointegrating system. With regard to forecast horizons, the Engle and Yoo (1987) get the interesting simulation result - based on an actually cointegrating system - that a model incorporating cointegration dominates a standard vector autoregression from 6 steps on but produces greater errors at shorter horizons. Their cointegrating model, however, uses a special method (Engle and Granger(1987)'s 2-step) whose performance might be different from the procedure used here.

4.2. Results from Box and Tiao's method

The Box and Tiao method relies on the assumption of stationary series which are allowed to be "almost non-stationary". The number of eigenvalues "significantly" (there are no statistical properties available) different from one should more or less correspond to the number of cointegrating vectors. The eigenvalues estimated are - in descending order - .99, .94, .88, .54, .26, .01. The last two are, obviously, different from 1. The first three are almost 1, and the value .54 represents a borderline case. This corresponds to the results from Johansen's procedure. The vectors corresponding to the two smallest eigenvalues are:

GDP	C	IF	PGDP	R	WAGE
.6484	-6.0017	.9528	-.2950	-.0415	2.0493
1.8822	-1.7886	-1.6659	-4.8373	-.0441	3.7350

The first vector is the less obvious. It links (real) consumption to (nominal) wages. The second one may be seen as establishing a link between real wages (wages by prices) and the remainder term if GDP is explained by C and IF, which mainly consists of exports and government policy. Note, however, that the data has been logged and that, therefore, GDP-C-IF is not exports minus imports plus public consumption plus statistical errors. In summary, the number of stationary combinations remains the same but the selected combinations differ widely between procedures. This could depend on the number of lags included (up to 5 with Johansen and only one with Box and Tiao). Increasing lag length to 5 - insignificant Q for each equation cannot be reached below this - gives eigenvalues closer to one as a whole: 1.00, .96, .93, .80, .69, .38. The eigenvectors corresponding to the smallest two values are:

GDP	C	IF	PGDP	R	WAGE
1.6165	-5.8975	-.2834	-3.8492	-.0604	4.4102
-2.8302	2.8546	-.8808	-4.5713	-.0286	3.3378

The first vector might be seen as to determine national income from the propensity to consume out of wage income or vice versa. The second vector relates the consumption quota (C by GDP) to the real wage, but negatively (!). This again might be seen as an indicator of foreign trade performance (strong exports resulting in low consumption quotas) determining real wage, which is somehow plausible in a small open economy like Austria.

Last not least, there is one obvious graphical result which even could favor Box and Tiao's method: the data transformation corresponding to the largest eigenvalue .99 really looks like a trend which is not the case in the other procedure.

Final note

The graphics in the appendix display the original series and the components from the 6v Johansen model as well as those from the Box and Tiao model with 5 lags.

References

BOX, G.E.P. and G.C. TIAO (1977): A canonical analysis of multiple time series; *Biometrika* 64,2,355-365.

DICKEY, D.A. and W.A. FULLER (1976): Distribution of the estimator for autoregressive time series with a unit root; *Journal of the American Statistical Association* 74,427-431.

ENGLE, R.F. and C.W.J. GRANGER (1987): Cointegration and error correction: Representation, estimation and testing; *Econometrica* 55,251-276.

ENGLE, R.F. and B.S. YOO (1987): Forecasting in co-integrated systems; *Journal of Econometrics* 35,143-159.

HOTELLING, H. (1936): Relations between two sets of variables; *Biometrika* 28,321-377.

JOHANSEN, S. (1987): Statistical Analysis of Cointegration Vectors; Paper presented at the Econometric Society European Meeting, Copenhagen.

JUSELIUS, K. (1987): Cointegration and Identification in a Vector Time Series Model. An Application to the Demand for Money in Denmark; Paper presented at the Econometric Society European Meeting, Copenhagen.

KUNST, R.M. and K. NEUSSER (1986): A Forecasting Comparison of Some VAR Techniques; *International Journal of Forecasting* 2,447-456.

PHILLIPS, P.C.B. and S. OULIARIS (1987): Testing for cointegration; Cowles Foundation discussion paper No. 809.

STOCK, J.H. and M.W. WATSON (1986): Testing for Common Trends; Harvard Institute of Economic Research, Harvard University, Discussion Paper No. 1222, March 1986.

TJØSTHEIM, D. and J. PAULSEN (1982): Empirical Identification of Multiple Time Series; *Journal of Time Series Analysis* 3,4,265-281.

VELU, R.P., D.W. WICHERN and G.C. REINSEL: A Note on Non-Stationarity and Canonical Analysis of Multiple Time Series Models; *Journal of Time Series Analysis* 8,4,479-487.

Figure 1: Original Series

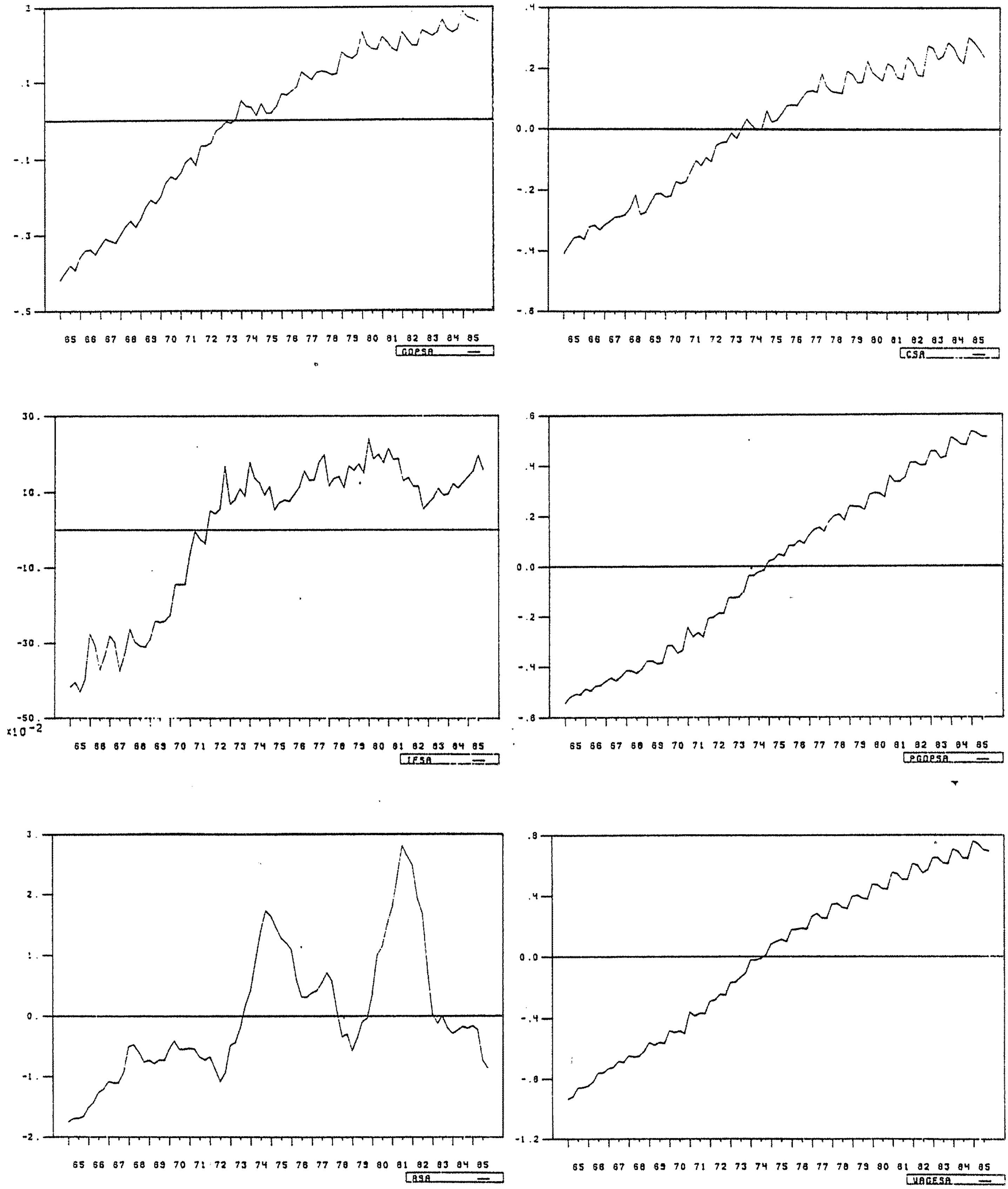


Figure 2: Components according to Johansen

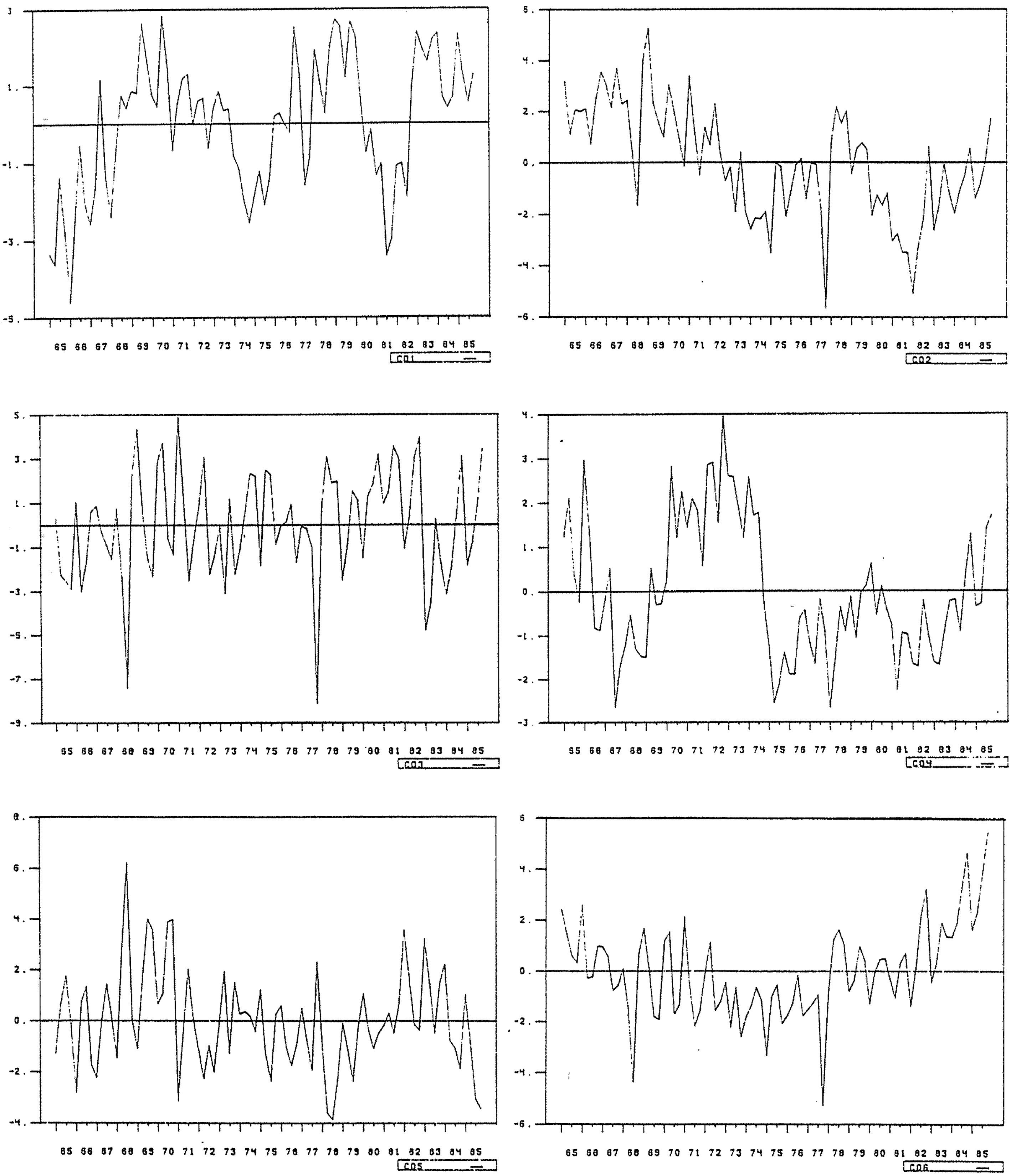
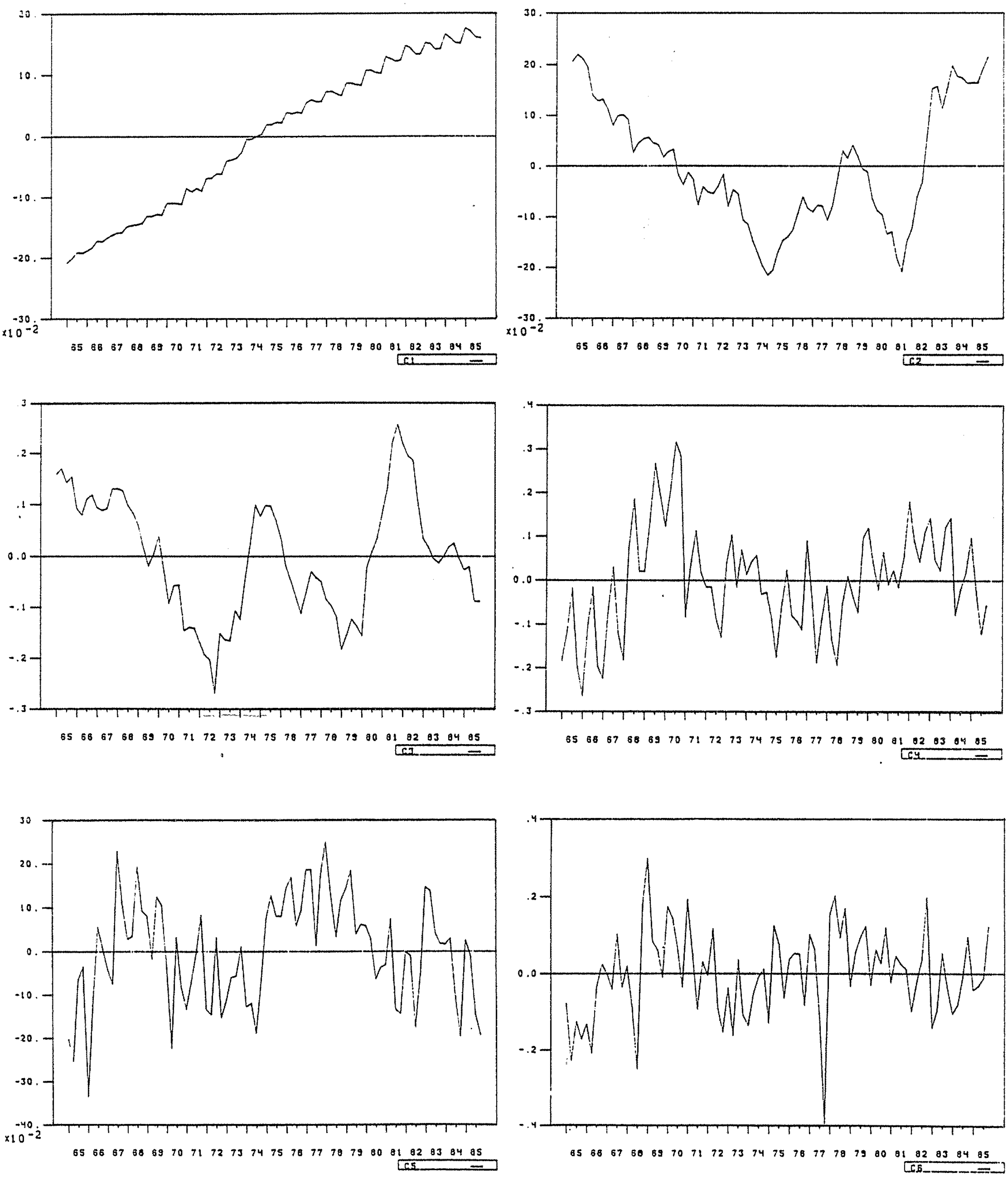


Figure 3: Components according to Box/Tiao



LIQUIDITY CONSTRAINTS VERSUS INTERTEMPORAL NON-SEPARABILITY OF AGGREGATE CONSUMER EXPENDITURES: AN EMPIRICAL TEST *

Klaus Neusser

University of Vienna

January 1988

* A preliminary version of this paper was presented at the Institute for Advanced Studies, Vienna. The author wants to thank Benno Pötscher, and Albert Jäger for helpful comments. The usual disclaimer, of course, applies.

1. Introduction

In his influencial paper Hall (1978) shows that the life cycle - permanent income hypothesis with rational expectations implies that consumption is a martingale. Several studies have investigated this proposition since then. Although consumption, or its logarithm, evolves fairly closely to a random walk, the hypothesis is often rejected on a formal statistical basis. Especially, Flavin (1981) found that current consumption expenditures react to current income in excess of the predictions based on the permanent income hypothesis with rational expectations. She interpreted this "excess sensitivity" as evidence that a sizeable fraction of households are subject to liquidity constraints. Campbell and Deaton (1987), and Campbell and Mankiw (1987)) present a further, but related troublesome piece of evidence. They show that consumption expenditures are "too smooth" with respect to innovations of an income process that is stationary in first differences.

Although the employed statistical procedures are still the subject of some controversies[1], the growing consensus emerges that the rational expectations permanent income hypothesis, at least in its time additive separable codification, is no longer a satisfactory working hypothesis (see Deaton (1986)). Against this background,

---

1 See the discussion in Mankiw and Shapiro (1985), Stock and Watson (1987), and Christiano, Eichenbaum, and Marshall (1987).

two promising, but competing lines of research[2] opened up. One tries to resolve the puzzles mentioned above in an explicitly formulated model of liquidity constrained households (see Hayashi (1985b)). The other gives up the assumption of intertemporal additive preferences in consumption expenditures. The rationale for changing the preference structure is the presumption that past expenditures will influence current and future marginal utilities either because of habit formation or because some goods, even though classified as a nondurable or a service, are better characterized as investment goods.[3]

This paper lies in the vein of the second approach. Along with Hayashi (1985a) it argues that goods are not valuable to the household by themselves but only through the services they provide. Therefore, a sharp distinction between expenditures and consumption is necessary.[4] Because only outlays are statistically observable, this approach requires the specification of a technology that relates expenditures to consumption. Following previous investigations by Eichenbaum, Hansen, and Richard (1984),

---

2 Other alternatives consist in relaxing the assumptions of a constant real interest rate or certainty equivalence (See Hall (1985) on the former, and (Blanchard and Mankiw (1987) and Neusser (1988) on the latter topic).

3 Hayashi (1985a) gives dental services or expenditures for education or recreation as examples. Another reason why goods considered to be nondurables have actually a large "durable" component is that they are used in fixed proportions with durables (Grossman and Laroque (1987)).

4 The literature addresses this problem by aggregating the services from durables and the outlays for nondurables into a composite consumption good; or by restricting the analysis to nondurables alone, possibly excluding expenditures on clothes and shoes as in Blinder and Deaton (1985). A test of these alternative specifications is provided in Eichenbaum and Hansen (1986).

Eichenbaum and Hansen (1987), Dunn and Singleton (1986), and Muellbauer (1986), this is accomplished by an intertemporal Gorman-Lancaster technology.[5] According to this specification current consumption is "produced" by a linear technology which is defined as a weighted sum of current and past expenditures. Then, retaining intertemporal additive preferences in terms of consumption - not in terms of expenditures - preserves the recursive nature of household's maximization problem.

Despite its simplicity, the intertemporal Gorman-Lancaster technology is rich enough to encompass habit formation as well. In this way, this line of research ties in with the important work that originates in Duesenberry's (1949) relative income hypothesis.[6] Habit formation is obtained when the marginal utility of consumption increases with past expenditures which are therefore seen as "negative inputs" to current consumption. The coefficients of the Gorman-Lancaster technology have therefore to be negative. On the other hand, durability of goods is implied when the marginal utility of consumption decreases with past expenditures. The coefficients of the Gorman-Lancaster technology are positive in this case.

After having developed the theoretical framework in section 2, the model is tested on time series data for Austria. This data set is particularily suited for this purpose. Austria can be considered

5 Browning (1987) and Zin (1987) propose other ways to formulate models with intertemporal nonseparability.

6 For a discussion of recent formulations of this hypothesis see Muellbauer (1986) and the references cited therein.

as a small open economy which has to take real interest rates from world capital markets. The real interest rate is therefore a weakly exogenous. This assumption then allows a simple test of the model and in addition leads to the identification of the signs of the Gorman-Lancaster technology so that it is possible to discriminate between habit formation and the "investment character" of goods. Furthermore, seasonally unadjusted data are available for all time series. In section 3, the paper proposes an explicit seasonal model that circumvents the important problems raised by Miron (1986) and Sargent (1987) in this context.

Following the work of Palm and Winder (1986) on The Netherlands, the model is tested in section 4 by investigating the restrictions it places on the joint vector autoregressive representation of real expenditures on nondurables and services and the real interest rate. The empirical findings for Austrian data can be summarized as follows. First, non-separable preferences are indeed an important element in consumer's choice. Second, the role of habit formation is rejected against durability. Third, it is not possible to reject the hypothesis that households are not subject to liquidity constraints.

2. The theoretical framework

Consider a representative household with preferences over sequences $\{C^*_t\}$ of consumption services. Assume that they are additive separable over time and can be represented by the utility functional:

$$(2.1) \qquad E_t \sum_{i=0}^{\infty} \beta^i \, U(C^*_{t+i})$$

where the subjective discount rate β is an element from the open interval (0,1). E_t denotes the mathematical expectations operator conditional on information available at time t. The period utility function U is assumed to have the usual smoothness and regularity properties.

Indexing bundles of goods by their date of purchase, the linear intertemporal Gorman-Lancaster technology relates current consumption services to current and past expenditures in the following way:

$$(2.2) \qquad C^*_t = A(L)\, C_t = \sum_{i=0}^{\infty} a_i \, C_{t-i}$$

where $\{C_t\}$ denotes sequences of purchases. The lag polynomial A(L) is normalized such that $a_0 = 1$ and is supposed to be invertible. Additionally, the sequence $\{a_i\}_{i\geq 0}$ is of exponential order less than β. The coefficients a_i, $i\geq1$, can have either sign a priori. They are positive, when the durability component is most important so that the marginal utility of current consumption decreases with past expenditures. They are negative, when habit formation is most important so that the marginal utility of current consumption increases with past expenditures (Muellbauer (1986)).

The assumption made above ensures that A(1) is finite. The coefficients of the lag polynomial therefore decline to zero. They may, however, follow a wide variety of patterns. For example, a hump-shaped pattern obtains when the full benefit is attained only after some time of learning. In this case, the contribution to current consumption first rises and declines only after some time has elapsed. A one-horse-shay type of technology, on the other hand, implies that a_i equals one for $0\leq i\leq T$ and zero for $i>T$ for some index T. In this case the polynomial A(L) is of finite order. The specification also encompasses a geometric decay pattern. This can be rationalized by taking C^*_t to be proportional to some capital stock which wears out at a fixed rate but can be replenished through goods purchases.

The household's period budget identitity is defined as:

$$A_{t+i+1} = R_{t+i} (A_{t+i} + Y_{t+i} - C_{t+i}) \qquad i\geq0 \tag{2.3}$$

with A_t given. Y_{t+i} denotes labor income of the agent as of time period t+i. A_{t+i} is the stock of assets valued in units of the consumption good held at the beginning of period t, and R_t is the real gross rate of return of this asset between dates t and t+1, measured in units of time t+1 consumption good per time t consumption good.

$\{R_t\}$ and $\{Y_t\}$ are random processes. The realization R_t becomes known to the agent only at the beginning of period t+1, so that the agent knows at time t when the consumption decision for this period has be to made only the values of R dated t-1 and earlier. The process $\{Y_t\}$ is assumed to be uncontrollable and its realizations dated t and earlier are known to the agent in period t. To rule out strategies of infinite consumption supported by infinite borrowing a suitable transversality condition is imposed.

The representative household's problem then consists in maximizing (2.1) subject to the constraints (2.2) and (2.3). Although the Gorman-Lancaster technology defined under (2.2) introduces non-separable intertemporal preferences in terms of expenditures, the maximization problem retains a recursive structure that allows the application of standard programming principles. As shown in appendix A, the Euler equations retain their familiar form, but are now expressed in terms of consumption rather than expenditures:

$$(2.4) \qquad E_t\,[\beta\, R_t\, U'(C^*_{t+1})/U'(C^*_t)] = 1 \qquad \text{for all } t$$

This equation states that ex-ante the agent equates the marginal rate of substitution with respect to consumption services, $U'(C^*_t)/[\beta U'(C^*_{t+1})]$, to the rate of transformation, R_t.

In the above form the Euler equation is, however, not easily econometrically tractable. Following the work of Hansen and Singleton (1982) a specific form of the utility function and of the joint distribution of the gross rate of return and consumption is postulated: $U(C^*_t) = [(C^*_t)^{1-\delta} - 1]/(1-\delta)$, $\delta>0$, and the conditional distribution of $(\ln C^*_{t+1}, \ln R_t)$ is normal. Using the property that the expected value of a log-normally distributed random variable Z is equal to $\exp[E\ln Z + \frac{1}{2}Var(\ln Z)]$, the Euler equation (2.4) becomes:

$$(2.5) \qquad E_t \ln C^*_{t+1} = \ln C^*_t + \delta^{-1} E_t \ln R_t + \tfrac{1}{2}\delta^{-1} V_t(\ln R_t - \delta \ln C^*_{t+1}) + \delta^{-1} \ln\beta$$

where V_t is the variance operator conditional on information at time t. According to the above equation the representative agent plans his expenditures in such a way that the service flow in period t+1 is equal to the current flow plus an adjustment positively related to the expected real gross rate of return and to the perceived risk measured by the conditional variance.

Instead of estimating equation (2.5) directly, the paper follows Palm and Winder (1986) in using the above relation to place restrictions on the bivariate process of consumption services and the interest rate. Let the law of motion of these two variables be described by the following autoregressive process:

$$(2.6) \quad \begin{bmatrix} 1-\mu_{11}(L) & -\mu_{12}(L) \\ -\mu_{21}(L) & 1-\mu_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta \ln C^*_{t+1} \\ \ln R_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \epsilon_{1t+1} \\ \epsilon_{2t+1} \end{bmatrix}$$

where $\mu_{ij}(L)$, $i,j\in\{1,2\}$, are homogenous polynomials in the lag operator L and where α_1 and α_2 are constants. The conditional as well as the unconditional mean of the error term $(\epsilon_{1t+1},\epsilon_{2t+1})'$ is zero. The Euler equation (2.5) then places restrictions on this bivariate process:

$$(2.7) \quad \begin{aligned} & [\mu_{11}(L) - \delta^{-1}\mu_{21}(L)] \ \Delta \ln C^*_{t+1} \\ + \ & [\mu_{12}(L) - \delta^{-1}\mu_{22}(L)] \ \ln R_t \\ + \ & \alpha_1 - \alpha_2/\delta - \delta^{-1}\ln\beta - (2\delta)^{-1} \ V_t[\ln R_t - \delta \ln C^*_{t+1}] = 0 \end{aligned}$$

Because this relation must hold for all realizations of the bivariate stochastic process, the following restrictions must hold:

$$(2.8a) \quad \mu_{11}(L) - \delta^{-1}\mu_{21}(L) \equiv 0$$

$$(2.8b) \quad \mu_{12}(L) - \delta^{-1}\mu_{22}(L) \equiv 0$$

$$(2.8c) \quad \alpha_1 - \delta^{-1}\alpha_2 - \delta^{-1}\ln\beta - \tfrac{1}{2}\delta^{-1} \ V_t[\ln R_t - \delta \ln C^*_{t+1}] \equiv 0$$

These restriction imply that the order of the polynomials $\mu_{11}(L)$ and $\mu_{21}(L)$ respectively $\mu_{12}(L)$ and $\mu_{22}(L)$ are equal. The coefficients of these polynomial are of equal sign and their values are proportional to the intertemporal elasticity of substitution $1/\delta$. Furthermnore, (2.8c) implies that the coefficients α_1 and α_2 are time varying whenever the conditional variance is not constant.

An interesting special case arises when the gross real interest rate is weakly exogenous with respect to consumption, i.e. $\mu_{21}(L)\equiv 0$. Then according to (2.8a) the AR-part of consumption services, $\mu_{11}(L)$, must also equal zero. This case is particularilly interesting for a small country like Austria which has to accept interest rates and prices determined in world markets.

3. The econometric model

The model is tested using seasonally unadjusted quarterly data on the Austrian economy. Real consumption expenditures are identified with expenditures on nondurables and services only. The analysis could easily be extended to include outlays on durables. The paper sticks, however, to the narrower concept to highlight the investment character of the goods included in this category . This implicitly implies that the services stemming from durables are regarded as completely separable from the services generated by nondurables and services.[7] The real interest rate is constructed by taking the yield on newly issued bonds[8] in period t and substracting the rate of change of the implicit deflator of consumer expenditures on non-durables and services that actually occured between t and t+1.

Both time series have a pronounced seasonal pattern that necessitates modifications of the bivariate autoregressive model (2.6). The seasonality of Austrian time series should, however, not considered as a nuisance but rather as a chance that allows the problems raised by Miron (1986) and Sargent (1987) in this context to be avoided.

The paper follows again Palm and Winder (1986) by postulating a

7 See the discussion of Eichenbaum and Hansen (1987) on this issue.

8 In the course of the investigation other interest rates have been tested. But it turned out that the choice is not crucial for the results.

simple seasonal model: the logarithm of consumption services[9] is the sum of a seasonal component and a systematic component; the latter being generated by the underlying model of consumer behavior. To economize on notation define c^*_t as $\ln C^*_t$ and r_t as $\ln(1+R_t)$. Then the seasonal model can written as:

$$\begin{aligned} c^*_{t+1} &= \tilde{c}^*_{t+1} + s_{1t+1} \\ r_t &= \tilde{r}_t + s_{2t+1} \end{aligned} \tag{3.1}$$

where ~ designates the systematic component. The seasonal components s_{1t+1} and s_{2t+1} sum on average to zero over four quarters. Therefore:

$$(1 + L + L^2 + L^3)\, s_{it+1} = \pi_{it+1} \qquad i \in \{1,2\} \tag{3.2}$$

where π_{it+1}, $i\in\{1,2\}$, are mutually independent, normally distributed white noise error terms. Taking the fourth difference of the vector autoregressive system (2.6) and using the seasonal model (3.1) gives:

$$\begin{bmatrix} 1-\mu_{11}(L) & -\mu_{12}(L) \\ -\mu_{21}(L) & 1-\mu_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta_4\Delta c^*_{t+1} - \Delta_4\Delta s_{1t+1} \\ \Delta_4 r_t - \Delta_4\, s_{2t+1} \end{bmatrix} = \begin{bmatrix} \Delta_4\epsilon_{1t+1} \\ \Delta_4\epsilon_{2t+1} \end{bmatrix} \tag{3.3}$$

where Δ and Δ_4 are defined by $1-L$ and $1-L^4$, respectively. The

9 The seasonality of expenditures is therefore implicitly attributed to seasonality in consumption services. Alternatively, seasonality can be modeled as preference shocks.

seasonal components can then be replaced by π_{it+1} using (3.2) to obtain:

$$\begin{bmatrix} 1-\mu_{11}(L) & -\mu_{12}(L) \\ -\mu_{21}(L) & 1-\mu_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta_4\Delta c^*_{t+1} \\ \Delta_4 r_t \end{bmatrix}$$

(3.4)

$$= \begin{bmatrix} \Delta_4\epsilon_{1t+1} + (1-\mu_{11}(L))(1-L)^2\pi_{1t+1} - \mu_{12}(L)(1-L)\pi_{2t+1} \\ \Delta_4\epsilon_{2t+1} - \mu_{21}(L)(1-L)^2\pi_{1t+1} + (1-\mu_{22}(L))(1-L)\pi_{2t+1} \end{bmatrix}$$

The result of these manipulations is a bivariate ARMA model for $(\Delta_4\Delta c_{t+1}, \Delta_4 r_t)'$. Because multivariate ARMA models are generally not identified, it is necessary to investigate the structure of the model (3.4) in detail to see whether the underlying economic theory provides enough identifying restrictions.

The MA part has the following diagonal structure:

$$\begin{bmatrix} \theta_{11}(L) & 0 \\ 0 & \theta_{22}(L) \end{bmatrix} \begin{bmatrix} v_{1t+1} \\ v_{2t+1} \end{bmatrix}$$

where $(v_{1t+1}, v_{2t+1})'$ is a bivariate, normally distributed white noise error term. The MA components $\theta_{ii}(L)\ v_{it+1} = [1+\theta_{ii1}L+\theta_{ii2}L^2+\ldots]v_{it+1}$, $i\in\{1,2\}$, are defined by the first and second row of the right hand side of the system (3.4). Under the assumption that the degrees of $\mu_{i1}(L)$ and $\mu_{i2}(L)$, $i\in\{1,2\}$, are less or equal to two and three, respectively, $\theta_{11}(L)$ and $\theta_{22}(L)$ are polynomials of degree less or equal to four. Teir degree is actually four, because their definition involves $_4\epsilon_{1t+1}$ and $\blacktriangle_4\epsilon_{2t+1}$, respectively. Then, the matrix consisting of the coefficients of the highest lags of the AR and MA part, $[\mu_p, \mathrm{diag}(\theta_{114}, \theta_{224})]$, has rank 2, because θ_{114} and θ_{224} are both different from zero. This is sufficient for identification of the model (3.4) (see Hsiao (1983,254)).[10] If on the other hand the degrees of $\mu_{i1}(L)$ and $\mu_{i2}(L)$, $i\in\{1,2\}$, are strictly greater than two or three respectively, so that $\theta_{11}(L)$ or $\theta_{22}(L)$ are of degree greater than four, identification is achieved by refering to the restrictions (2.8a) and (2.8b) which imply that the degrees of $\mu_{11}(L)$ and $\mu_{21}(L)$, and $\mu_{12}(L)$ and $\mu_{22}(L)$ are equal, respectively. In this case, the degrees of $\theta_{11}(L)$ and $\theta_{22}(L)$ are again equal and the above identification criterium still applies.

10 In addition, the polynomials representing the AR and MA part have to be relatively left prime.

4. Empirical results

Following a suggestion of Muellbauer (1986), appendix B shows that $A(L)\Delta_4\Delta\ln C_{t+1}$ is a good approximation of $\Delta_4\Delta\ln C^*_{t+1}$. The system (3.4) then becomes:

$$\begin{bmatrix} (1-\mu_{11}(L))A(L) & -\mu_{12}(L) \\ -\mu_{21}(L)A(L) & 1-\mu_{22}(L) \end{bmatrix} \begin{bmatrix} {}_4c_{t+1} \\ \Delta_4 r_t \end{bmatrix} = \begin{bmatrix} \theta_{11}(L) & 0 \\ 0 & \theta_{22}(L) \end{bmatrix} \begin{bmatrix} v_{1t+1} \\ v_{2t+1} \end{bmatrix} \qquad (4.1)$$

The system (4.1) is estimated on an equation by equation basis. To avoid the problem of changing seasonality, the paper reduced the estimation period to 1970:1 - 1986:4. It should, however, be noted that this choice is not affecting the conclusion in any significant way.

Table 1 reports the results[11] obtained for the real interest rate equation - the second equation in system (4.1). The first column gives the estimates of a MA(4) model. As one would expect when modelling financial variables the explanatory power is rather limited. Nearly 20 percent of the variance is explained. The Ljung-Box Q-Statistik indicates that no further autocorrelation is

---

11 Here and in all subsequent tables, the time indices are set in such a way that match the notation in (4.1).

present in the residuals. Looking at the MA-coefficients and their corresponding t-ratios, it is evident that only v_{2t-3} and to a minor extent v_{2t-1} represent significant contributions to the explanation $\Delta_4 r_t$. Despite the insignificance of v_{2t} and v_{2t-2}, the F-statistic F_{MA} for the null hypothesis that all MA-coefficients are simultanously zero takes on a value that rejects this hypothesis at the 1 percent significance level.

The inclusion lagged interest rates does not alter the picture very much (see equation (2)). The MA-coefficients remain highly significant as a group, but only v_{2t-3} is significant individually. The lagged dependent variables are neither significant as a group nor individually. The change in the MA-coefficients, however, suggests that there is some scope for these variables. Lagged fourth differences of consumption expenditure growth also enter in an insignificant way (see equation (3)). Reestimating the model but eliminating all MA-terms except for the seasonal one and including two respectively three lags of the two other variables confirms the results obtained previously (see equation (4)). The coefficients of the lagged dependent variable are just marginally significant at the 10 percent level. Lagged consumption expenditures still turn out to be unimportant.[12] This result suggests that $-\mu_{21}(L)A(L)$ is identically zero which implies that either the real interest rate is weakly exogenous with respect to consumption (i.e $\mu_{21}(L)\equiv 0$) or that habit formation and "durability" play no role in explaining expenditures on

12 This specification gives the most significant coefficients to lagged consumption.

nondurables and services (i.e $A(L)\equiv 1$). Comparing all four models the information criteria AIC and BIC just confirm again the results. They even suggest that the MA(4) is to be prefered to the other ones.

Table 2 reports the results for the consumption equation, the first equation in (4.1). Starting again with a MA(4) model, column (1) shows highly significant coefficients. The introduction of two lags of the dependent variable alters the picture radically. Although, taken as a group, the MA-coefficients still remain significant at the 1 percent level, only the last one has a significant t-ratio. In contrast to the model for the real interest rate, the lagged expenditure variables turn out to be highly significant. These results do not change when lagged interest rates are included (equation (3)) nor when the insignificant MA-coefficients are excluded (equation (4)).

The results of table 1 and 2 imply that only the assumption $A(L)\not\equiv 1$ is consistent with the theoretical framework. For suppose that $A(L)\equiv 1$, then table 1 implies that $\mu_{21}(L)\equiv 0$ which in turn by restriction (2.8a) implies that $\mu_{11}(L)\equiv 0$. The results of table 2, however, show that lagged expenditures are highly significant so that $\mu_{11}(L)\not\equiv 0$. Therefore, the simple intertemporal model with time additive preferences is rejected. Only a model with habit formation or "durability" (i.e. $A(L)\not\equiv 1$) allows a consistent interpretation of the empirical findings. Taking $\mu_{11}(L)$ and $\mu_{21}(L)$ equal zero, the consumption equation provides estimates for the coefficients of the Gorman-Lancaster technology. These estimates are given by minus the coefficients of the lagged expenditures. Table 2 then implies that a_1 and a_2 are positive, and according to appendix B upward biased. Consequently, habit formation is rejected against "durability". Eichenbaum and Hansen (1987) have obtained a similar conclusion with US data. In contrast, Muellbauer (1986) again for US data and Winder (1987) for The Netherlands found habit formation to the more important factor.

The paper next investigates Hall's (1978) orthogonality test by including lagged fourth differences of real net labor income growth rates[13] as additional regressors in the consumption equation. According to the theoretical model of section 2, the coefficients of these variables should not be significantly different from zero, because their informational content has already been incorporated into the household's expenditure plan. I, however, these variables enter in a significant way, the theoretical model is rejected. Although such a failure can have several causes, Flavin (1981) has interpreted such findings as evidence for liquidity constrained households. In contrast to Hall (1978) and Flavin (1981) this paper uses real net labor income, including social security payments, rather than real disposible income. The reason for this choice is twofold. First, capital market imperfections are more likely to be felt by wage earners who want to borrow against their human capital. Second, regressing growth rates of expenditures against growth rates of income leads to statistically correct inferences only when the two time series

13 The fourth difference is necessary to account for the seasonal component.

are not cointegrated.[14] Campbell (1987) proves that the permanent income hypothesis implies that consumption is cointegrated with disposible income and not with labor income. Jäger and Neusser (1987) corroberate this implication for the data at hand.

Adding four lags of the fourth difference of real net labor income growth, denoted by $\Delta_4\Delta \ln Y_{t-i}$, $i \in \{0,1,2,3\}$, in the consumption equation (4) of table 2 leads to the results reported in the first column of table 3. The inclusion of these new variables does not affect the results just reached: lagged real interest rates are neglegible and lagged expenditures remain highly significant. The contribution of lagged labor income is practically zero, according to the t-ratios and the F-statistic F_Y. Past labor income does therefore represent information already incorporated in the expenditure plan. This conclusion is also reflected in the deterioration of the information criteria.

The omission of the insignificant real interest rate variables corroberates this conclusion further (see equation (2)). The coefficients of labor income only become significant when the MA term is left out as in equation (3). This underlines the importance of a correct treatment of seasonality. In equation (4) and (5) lagged expenditures are excluded. Again, the coefficients of labor income remain insignificant, reinforcing the argument that liquidity constraints do not play an important role in

14 The concept of cointegration between time series has been initiated by work of Engle and Granger (1987). Stock and West (1987) analyze its consequences for the application of Hall's orthogonality test.

Austria. Interestingly, the exclusion of lagged expenditures lead to significant real interest rate coefficients. The Q-statistic, however, signals residual correlation. The model including lagged expenditures is therefore prefered.

Table 1: Estimates for ARMA models of the real interest rate

dependent variable: $\Delta_4 \ln(1+R_t)$

estimation period: 1970:1 - 1986:4

explanatory variable	equation (1)	(2)	(3)	(4)
$\Delta_4 \ln(1+R_{t-1})$		.120 (.44)		.045 (.34)
$\Delta_4 \ln(1+R_{t-2})$		.133 (.57)		.291 (2.24)
$\Delta_4 \ln(1+R_{t-3})$		.129 (.59)		.109 (.85)
v_{2t}	.091 (.78)	-.040 (.16)	.062 (.51)	
v_{2t-1}	.190 (1.63)	.071 (.34)	.267 (2.23)	
v_{2t-2}	.026 (.22)	-.129 (.58)	.045 (.36)	
v_{2t-3}	-.445 (3.76)	-.505 (3.93)	-.486 (3.90)	-.652 (5.88)
$\Delta_4 \Delta \ln C_t$			.082 (1.55)	.098 (1.92)
$\Delta_4 \Delta \ln C_{t-1}$			.035 (.70)	.030 (.62)
R^2	.176	.181	.192	.204
SEE	.0086	.0088	.0087	.0086
Q	6.095	5.932	6.520	6.732
AIC	-9.445	-9.363	-9.406	-9.421
BIC	-9.314	-9.135	-9.211	-9.225
F_{MA}	6.40***	5.33***	9.54***	34.52***
F_R		.24		2.20*
F_C			1.23	1.88

numbers in parenthesis are t-ratios
*, ** and *** indicate that the corresponding value of the test statistic is significant at 10, 5 and 1 percent, respectively.

Table 2: Estimates for ARMA models of expenditures

dependent variable: $\Delta_4 \Delta \ln C_{t+1}$

estimation: 1970:1 - 1986:4

explanatory variable	equation (1)	(2)	(3)	(4)
$\Delta_4 \ln(1+R_{t-1})$			-.300 (1.03)	-.296 (1.07)
$\Delta_4 \ln(1+R_{t-2})$			.093 (.33)	.223 (.80)
$\Delta_4 \ln(1+R_{t-3})$			.310 (1.20)	.294 (1.15)
v_{1t}	-.443 (4.00)	.107 (.52)	.009 (.04)	
v_{1t-1}	-.339 (2.97)	-.232 (1.63)	-.223 (1.47)	
v_{1t-2}	.429 (3.67)	-.079 (.46)	-.001 (.01)	
v_{1t-3}	-.486 (4.28)	-.530 (4.23)	-.550 (3.96)	-.510 (4.13)
$\Delta_4 \Delta \ln C_t$		-.681 (3.04)	-.518 (2.29)	-.490 (3.92)
$\Delta_4 \Delta \ln C_{t-1}$		-.323 (1.57)	-.197 (.94)	-.281 (2.29)
R^2	.434	.496	.515	.488
SEE	.0182	.0175	.0176	.0176
Q	19.222	10.337	12.812	19.142
AIC	-7.954	-8.011	-7.960	-7.994
BIC	-7.823	-7.816	-7.666	-7.799
F_{MA}	22.15***	11.44***	8.84***	17.03***
F_R			.83	1.13
F_C		5.15***	2.87*	8.14***

numbers in parenthesis are t-ratios
*, ** and *** indicate that the corresponding value of the test statistic is significant at 10, 5 and 1 percent, respectively.

Table 3: Testing for liquidity constraints

dependent variable: $\Delta_4\Delta\ln C_{t+1}$

estimation period: 1970:1-1986:4

explanatory variable	equation (1)	(2)	(3)	(4)	(5)
$\Delta_4\ln(1+R_{t-1})$	-.256 (.87)			.579 (1.83)	
$\Delta_4\ln(1+R_{t-2})$	.150 (.48)			-.665 (2.25)	
$\Delta_4\ln(1+R_{t-3})$	.273 (.99)			.225 (.75)	
$\Delta_4\Delta\ln C_t$	-.549 (4.06)	-.608 (5.11)	-.716 (6.08)		
$\Delta_4\Delta\ln C_{t-1}$	-.303 (2.32)	-.358 (3.01)	-.339 (2.88)		
$\Delta_4\Delta\ln Y_t$	.167 (1.05)	.181 (1.21)	.383 (2.66)	.012 (.07)	-.012 (.07)
$\Delta_4\Delta\ln Y_{t-1}$	-.023 (.14)	.029 (.19)	-.170 (1.07)	.050 (.30)	-.016 (.10)
$\Delta_4\Delta\ln Y_{t-2}$	.021 (.14)	.067 (.46)	.151 (.97)	.125 (.73)	.127 (.76)
$\Delta_4\Delta\ln Y_{t-3}$	-.015 (.10)	-.010 (.07)	-.135 (.94)	-.013 (.08)	.061 (.37)
v_{1t-3}	-.457 (3.55)	-.438 (3.29)		-.502 (4.03)	-.590 (5.38)
R^2	.498	.480	.431	.359	.262
SEE	.0180	.0179	.0186	.0200	.0210
Q	18.633	18.825	17.667	33.767*	34.286*
AIC	-7.897	-7.950	-7.890	-7.711	-7.659
BIC	-7.571	-7.722	-7.694	-7.450	-7.496
F_R	.68			3.19**	
F_C	8.59***	13.53***	18.60***		
F_Y	.31	.67	2.66**	.15	.18

numbers in parenthesis are t-ratios
*, ** and *** indicate that the corresponding value of the test statistic is significant at 10, 5 and 1 percent, respectively.

5. Conclusion

The paper has demonstrated that the introduction of intertemporally nonseparable preferences represents a viable way to resolve the problems of the conventional permanent income hypothesis. The framework relies on the distinction between expenditures and consumption. Because the latter are not directly observable a linear technology is formulated to relate unobservable consumption to observable current and past expenditures. Retaining time additive preferences in terms of consumption, the maximization problem leads to Euler equations which state that households plan expenditures in such a way as to hold the marginal utility of consumption constant. The setup also allows to distinguish between habit formation and "durablity" of goods classified as nondurables as a reason for nonseparability.

The empirical analysis used Austrian quarterly time series data of expenditures on nondurables and services to test the implications of the model. After taking care explicitly for seasonality, the following results are obtained. First, nonseparability is indeed an important element in explaining aggregate consumption behavior. Second, "durability" rather than habit formation accounts for this fact. Third, Hall's orthogonality test presents no evidence for the presence of liquidity constrained households.

It lies in the nature of this type of investigation that not all aspects can be adequatly incorporated in the analysis. An especially pressing one is time aggregation, at least according to Christiano, Eichenbaum, and Marshall (1987). Altough it may easily handled within this framework along the line of Hall (1985), the dynamic structure of the bivariate ARMA model is complicated a great deal. Proceeding in this direction is left for future investigations.

References

Benveniste, L.M., and J.A. Scheinkman (1979) On the Differentiability of the Value Function in Dynamic Models of Economics. Econometrica, Vol.47, 727-732.

Blanchard, Olivier J., and N. Gregory Mankiw (1987) Beyond Certainty Equivalence. Paper given at the American Economic Association Meeting 1987 in Chicago.

Blinder, Alan S., and Angus Deaton (1985) The Time Series Consumption Function Revisited. Brookings Paper on Economic Activity, 2:1985, 465-521.

Browning, Martin (1987) A Simple Non Additive Preference Structure for Models of Household Behavior over Time. McMasters University, Department of Economics, Working Paper No. 87-12.

Campbell, John Y. (1987) Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis. Econometrica, Vol.55, 1249-1273.

Campbell, John Y., and Angus S. Deaton (1987) Is Consumption Too Smooth? NBER Working Paper No.2134.

Campbell, John Y., and N. Gregory Mankiw (1987) Permanent Income, Current Income, and Consumption. Mimeo.

Christiano, Lawrence J., Martin Eichenbaum, and David Marshall (1987) The Permanent Income Hypothesis Revisited. NBER Working Paper No.2209.

Deaton, Angus S. (1986) Life-cycle Models of Consumption: Is the Evidence Consistent With the Theory? NBER Working Paper No.1910.

Dunn, Kenneth B. and Kenneth J. Singleton (1986) Modelling the Term Structure of Interest Rates under Non-Separable Utility and Durability of Goods. Journal of Financial Economics, Vol.17, 27-55.

Duesenberry, James S. (1949) Income, Saving, and the Theory of Consumer Behavior. Massachusetts: Harvard University Press.

Eichenbaum, Martin S., and Lars P. Hansen (1987) Estimating Models with Intertemporal Substitution using Aggregate Time Series Data. NBER Working Paper No.2181.

Eichenbaum, Martin S., Lars P. Hansen, and Scott Richard (1984) The Dynamic Equilibrium Pricing of Durable Consumption Goods. Mimeo, University of Chicago.

Engle, Robert F., and Clive W.J. Granger (1987) Cointegration and ErrorCorrection: Representation, Estimation and Testing. Econometrica, Vol.55, 251-276.

Flavin, Marjorie (1981) The Adjustment of Consumption to Changing Expectations about Future Income. Journal of Political Economy, Vol.89, 971-987.

Grossman, Sandford, and Guy Laroque (1987) Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods. NBER Working Paper No.2369.

Hall, Robert E. (1978) Stochastic Implications of the Life Cycle - Permanent Income Hypothesis: Theory and Evidence. Journal of Political Economy, Vol.86, 971-988.

Hall, Robert E. (1985) Real Interest Rate and Consumption. NBER Working Paper No.1694.

Hansen, Lars P., and Kenneth Singleton (1982) Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models. Econometrica, Vol.50, 1269-1286.

Hayashi, Fumio (1985a) The Permanent Income Hypothesis and Consumption Durability: Analysis Based on Japanese Panel Data. Quarterly Journal of Economics, Vol.100, 1083-1113.

Hayashi, Fumio (1985b) Tests for Liquidity Constraints: A Critical Survey. NBER Working Paper No.1720.

Hsiao, Cheng (1983) Identification. In: Zvi Griliches and Michael D. Intriligator (eds.) Handbooks of Econometrics, Volume I, New York: North-Holland, 223-283.

Jaeger, Albert, and Klaus Neusser (1987) Excess Consumption as a Predictor of Future Income Changes. University of Vienna, Working Paper No.8701.

Mankiw, N. Gregory, and Matthew Shapiro (1985) Trends, R Walks, and Test of the Permanent Income Hypothesis. Journal of Monetary Economics, Vol.16, 165-174.

Miron, J.A. (1986) Seasonal Fluctuations and the Life Cycle-Permanent Income Model of Consumption. Journal of Political Economy, Vol.94, 1258-1279.

Muellbauer, John (1986) Habits, Rationality, and Myopia in the LIfe-Cycle Consumption Function. CEPR Discussion Paper No.112.

Neusser, Klaus (1988) Consumption and Changing Income Uncertainty: An Empirical Investigation for Austria. In H. Frank, G. Plaschka, and D. Roessl (eds.) Umweltdynamik ("Changing Environment"), Springer-Verlag:Wien, 243-261.

Palm, Frans C., and Carlo C.A. Winder (1986) Incorporating Economic Theory into Structural Time Series Models with an Application to Consumption. Paper presented at the Econometric Society European Meeting in Budapest.

Sargent, Thomas J. (1987) Macroeconomic Theory. Second Edition. New York: Academic Press.

Stock, James H., and Kenneth D. West (1987) Integrated Regressors and Tests of the Permanent Income Hypothesis. NBER Working Paper No.2359.

Winder, Carlo C.A. (1987) Rational Habits in the Life Cycle Consumption Function. Free University of Amsterdam, Research Memorandum 1987-34.

Zin, Stanley E. (1987) Aggregate Consumption Behaviour in a Life Cycle Model with Non-Additive Recursive Utility. Queen's University, Discussion Paper No.693.

Appendix A: Derivation of the Euler equation (2.4)

Let B(L) denote the inverse of the polynomial A(L) defined in equation (2.2), such that $C_t = B(L)\, C^*_t = 1 + b_1 L + b_2 L^2 + \ldots$ The state variables at time t of the intertemporal optimization problem are real non-human wealth A_t and $X_t' = (C^*_{t-1}, C^*_{t-2}, \ldots)$. The transition equations are then given by:

$$A_{t+1} = R_t\, u_t$$

$$X_{t+1} = B\, X_t + e_1\, (A_t + Y_t - u_t)$$

where the control u_t and the vector e_1' are defined as gross savings, $A_t+Y_t-C_t$, and (1,0,0,...), respectively. The infinite dimensional matrix B is defined as:

$$B = \begin{bmatrix} -b_1 & -b_2 & -b_3 & . & . & . \\ 1 & 0 & 0 & . & . & . \\ 0 & 1 & 0 & . & . & . \\ 0 & 0 & 1 & . & . & . \\ . & . & . & . & . & . \end{bmatrix}$$

Because the optimization problem has a recursive structure, it is possible to write down Bellman's equation:

$$V(A_t,X_t) = \max_{u_t} \{ U(C^*_t) + \beta E_t V(A_{t+1},X_{t+1}) \}$$

$$= \max_{u_t} \{ U(A_t + Y_t - u_t - b_1C^*_{t-1} - b_2C^*_{t-2} - \ldots)$$

$$+ \beta E_t V(R_tu_t , BX_t+e_1(A_t+Y_t-u_t)) \}$$

where the maximization is subject to the budget constraint (2.3). The Euler equation is then given by:

$$-U'(C^*_t) + \beta E_t R_t V_A(A_{t+1},X_{t+1}) - \beta E_t V_1(A_{t+1},X_{t+1}) = 0$$

where V_A and V_i, i=1,2,3,..., denote the partial derivatives of the value function with respect to real non-human wealth and the i-th component of the vector X_t, respectively. According to Benveniste and Scheinkman (1979), these partial derivatives can be determined through the following expressions:

$$V_A(A_t,X_t) = U'(C^*_t) + \beta E_tV_1(A_{t+1},X_{t+1})$$

$$V_1(A_t,X_t) = -b_1U'(C^*_t) - \beta b_1 E_tV_1(A_{t+1},X_{t+1}) + \beta E_tV_2(A_{t+1},X_{t+1})$$

$$V_2(A_t,X_t) = -b_2U'(C^*_t) - \beta b_2 E_tV_1(A_{t+1},X_{t+1}) + \beta E_tV_3(A_{t+1},X_{t+1})$$

. . . .

$$V_i(A_t,X_t) = -b_iU'(C^*_t) - \beta b_i E_tV_1(A_{t+1},X_{t+1}) + \beta E_tV_{i+1}(A_{t+1},X_{t+1})$$

. . . .

After successive insertions one obtains:

$$\beta\ E_t\ B(\beta L^{-1})V_1(A_t,X_t) = -\ \beta b_1\ E_tU'(C^*_t) - \beta^2 b_2\ E_tU'(C^*_{t+1}) - \beta^3 b_3\ E_tU'(C^*_{t+2}$$
$$= U'(C^*_{t-1}) - E_t\ B(\beta L^{-1})U'(C^*_{t-1})$$

This implies that

$$\beta\ V_1(A_t,X_t) = E_t\ A(\beta L^{-1})U'(C^*_{t-1}) - U'(C^*_{t-1})$$

$V_A(A_t,X_t)$ is then given by:

$$V_A(A_t,X_t) = U'(C^*_t) + E_tA(\beta L^{-1})U'(C^*_t) - U'(C^*_t)$$
$$= E_tA(\beta L^{-1})U'(C^*_t)$$

Inserting these results into the Euler equation yields equation (2.4) in the text.

Appendix B: Approximating $\Delta \ln C^*_t$ by $A(L)\Delta \ln C_t$

Following a suggestion of Muellbauer (1986, the growth rate of consumption is approximated by distributed lag of growth rates in expenditures.

$$\Delta \ln C^*_t = \Delta \ln(C_t + a_1 C_{t-1} + a_2 C_{t-2} + \ldots)$$

$$= \Delta \ln[C_t(1 + a_1(C_{t-1}/C_t) + a_2(C_{t-2}/C_t) + \ldots]$$

$$= \Delta \ln C_t + \Delta \ln[A(1) - a_1(\Delta C_t/C_t) - a_2(\Delta_2 C_t/C_t) - \ldots]$$

$$= \Delta \ln C_t + \Delta \ln\{A(1) - [a_1 + a_2(1+L) + a_3(1+L+L^2) + \ldots](\Delta C_t/C_t)\}$$

$$\approx \Delta \ln C_t + \Delta \ln[\ 1 - \sum_{i=1}^{\infty} A_i \Delta \ln C_{t+1-i}\ ]$$

where $A_i = (a_i+a_{i+1}+a_{i+2}+\ldots)/A(1)$, $i\geq 1$, and $\Delta C_t/C_t$ has been approximated by $\Delta \ln C_t$. Denoting the average growth rate of expenditures by g, the first order Taylor expansion $\Delta \ln(1+x) = \Delta \ln(1+\hat{x}) + \Delta(x-\hat{x})/(1+\hat{x}) = \Delta x/(1+\hat{x})$ of the second term of the right hand side then yields:

$$\Delta \ln C^*_t \approx \Delta \ln C_t - (1+\hat{x})^{-1}\left[\sum_{i=1}^{\infty} A_i (1-L)L^{i-1}\right] \Delta \ln C_t$$

where $1+\hat{x}$ is defined by

$$1+\hat{x} = 1 - g \sum_{i=1}^{\infty} \sum_{i=1}^{\infty} a_i/A(1) = 1 - (g/A(1)) \sum_{i=1}^{\infty} i a_i$$

This expression is well defined, because the sequence $\{a_i\}_{i\geq 0}$ is of exponential order less than β. Then after further manipulations

$$\Delta \ln C^*_t \approx A(1)^{-1} \{ 1 + (1+\hat{x})^{-1} [a_1 L + a_2 L^2 + a_3 L^3 + \ldots]\} \Delta \ln C_t$$

In the ARMA representations the leading coefficient is normalized to one. For reasonable values of g and $a_i>0$, $i\geq 0$, $(1+\hat{x})^{-1}$ is a number greater then 1 so that the estimates of the Gorman-Lancaster technology are upward biased. The estimates are downward biased, if the a_i's are negative.

MODEL SELECTION UNDER NONSTATIONARITY:

AUTOREGRESSIVE MODELS AND STOCHASTIC LINEAR REGRESSION MODELS

B. M. Pötscher*
Department of Statistics
Yale University

October 1986

Revised March 1987

MODEL SELECTION UNDER NONSTATIONARITY:
AUTOREGRESSIVE MODELS AND STOCHASTIC LINEAR REGRESSION MODELS

B. M. Pötscher*
Department of Statistics
Yale University

October 1986

Revised March 1987

*Research supported by the Max Kade Foundation. I want to thank C.Z. Wei, University of Maryland, for drawing my attention to the papers by H. An, L. Gu and S. Wang and for helpful comments on the first version of this paper. Thanks are also due to J. Wan for the translation of the paper by S. Wang and H. An, and to B. Amato for typing the manuscript.

Author's address: Department of Statistics, Yale University,
2179 Yale Station, New Haven, CT 06520-2179.

Abstract

We give sufficient conditions for strong consistency of estimators for the order of general nonstationary autoregressive models based on the minimization of an information criterion à la Akaike's (1969) AIC. The case of a time dependent error variance is also covered by the analysis. Furthermore the more general case of regressor selection in stochastic regression models is treated.

Keywords: Model selection, order estimation, selection of regressors, strong consistency, autoregression, nonstationarity, non-ergodic models, information criteria.

AMS 1980 Subject Classification: Primary 62M10, 62J05, 60G10.
Secondary 62F12, 93E12.

1. Introduction

The statistical properties of order estimation and model selection procedures based on so-called information theoretic criteria like Akaike's AIC or its variants have been intensively studied in recent years. Most of the work has been done in the field of time series analysis. In the framework of stationary autoregressive models strong consistency of the order estimators obtained through minimization of certain variants of AIC has been discussed in Hannan and Quinn (1979) for the univariate case, and in Quinn (1980) for the multivariate case. Parallel results for stationary autoregressive moving average models are given in Hannan (1980), Hannan (1981). The latter two papers give also weaker conditions under which weak consistency holds. For an alternative approach using sequences of tests see Pötscher (1983, 1985). Weak consistency results for the closely related model selection problem for linear regression models with asymptotically stationary regressors and normal i.i.d. errors are given in Geweke and Meese (1981) (for a review of the literature on selection of regressors, see Amemiya (1980), Thompson (1978)). Contrarily, for nonstationary autoregressive models with i.i.d. errors weak consistency of the order estimators has been established independently by Paulsen (1984) and Tsay (1984), the former also treating the multivariate case. The nonstationarity considered in both papers arises from the fact that the characteristic polynomial is allowed to have roots not only outside but also on the unit circle. Another case of nonstationarity is considered in Paulsen and Tjøstheim (1985): here the autoregressive scheme has to be stable, i.e. all zeroes of the characteristic polynomial are outside the unit circle, but the error process is allowed to have a nonconstant variance.

The present paper gives strong consistency results for model selection procedures based on variants of AIC for general nonstationary stochastic linear regression models where the errors constitute a not necessarily stationary martingale difference sequence using recent results of Lai and Wei (1982a, b, 1983, 1985). These results are then applied to yield strong consistency results for orderestimation in nonstationary autoregressive models where the errors form a not necessarily stationary martingale difference sequence. The assumptions used in the present paper are weaker than each of the assumptions employed in Paulsen (1984), Paulsen and Tjøstheim (1985) and Tsay (1984), hence they provide a common framework for models exhibiting both kinds of nonstationarity. After the first version of this paper was written, C. Z. Wei brought the related papers by Wang and An (1984), An and Gu (1985) and Gu and An (1985) to my attention. The first one of these papers discusses model selection in stochastic linear regression models and autoregressive models by means of BIC, i.e. criterion (2.1) with $C(T) = \log T$ in our notation. In this paper Wang and An give conditions for "overconsistency" of the selected models, i.e. conditions under which the selected models contain all relevant regressors but possibly also redundant ones. The paper neither gives conditions for the full consistency property nor explores the domain of feasible rates for the penalty term $C(T)$ such that consistency results as is done in this paper. For a further discussion of the results of this paper by Wang and An (1984) and its relation to the present paper, see Sections 3 and 4. The two other papers, An and Gu (1985) and Gu and An (1985), deal essentially only with the stationary case. The paper is organized as follows: Section 2 gives consistency results in the general context of a linear regression model with stochastic regressors and martingale difference errors; in Section 3 the special case of autoregressive models is treated; Section 4 contains complementary remarks; all proofs are relegated to an appendix.

2. Model selection in linear regression models

The dependent variable of a linear regression is modelled as a real-valued stochastic process (y_t). The family of potential regressors under consideration is a family $\mathcal{Q} = ((z_{tk}) : k \in \mathcal{K})$ of real-valued stochastic processes defined on the same probability space $(\Omega, \mathfrak{F}, P)$ as y_t is. The set $\mathcal{K}$ is an arbitrary index set and the index t varies in $\mathbb{N}$. More specifically, the researcher has in mind a (nonvoid) set $\mathcal{M}$ of regression models M where each M is a finite subset of $\mathcal{K}$, i.e. under model M the regressors (z_{tk}) for $k \in M$ enter the regression equation for (y_t) (if M is void regression is on the nullspace). Important special cases are the case where the regressors are ordered in a natural way, e.g. $\mathcal{K}$ is $\mathbb{N}$ or an initial segment of $\mathbb{N}$ (in the natural order) and M runs through all initial segments of $\mathcal{K}$, or the case where $\mathcal{M}$ is the set of all finite subsets of $\mathcal{K}$. We shall be concerned with the problem of choosing a "minimal" and "true" model M from the set $\mathcal{M}$. The notion of a true model is here to be understood in the sense of the following definition and is defined relative to a given filtration $\mathfrak{F}_s$, $s \in \mathbb{N} \cup \{0\}$, of the σ-field $\mathfrak{F}$. <u>In this section, the filtration $\mathfrak{F}_s$ will be throughout assumed to have the property that z_{tk} is $\mathfrak{F}_{t-1}$-measurable for all $t \in \mathbb{N}$ and $k \in \mathcal{K}$.</u> The prototypical examples for this situation are the case of an autoregressive model where $\mathfrak{F}_{t-1}$ represents, for example, the σ-field generated by the past of the process (y_t), or the case of a general linear dynamic regression model. <u>For the rest of this section, we shall also assume that a true model in the sense of Definition 2.1 exists in $\mathcal{M}$.</u>

<u>Definition 2.1</u>: A model $M \in \mathcal{M}$ is called a true model for (y_t) relative to $(\mathfrak{F}_t)$ if there exist random variables β_k, $k \in M$, such that y_t can be

P-a.s. decomposed as $y_t = \Sigma_{k\in M} \beta_k z_{tk} + u_t$ such that for all $t \in \mathbb{N}$ hold:

(i) $\sum_{k\in M} \beta_k z_{tk}$ is P-a.s. measurable w.r.t. $\mathfrak{F}_{t-1}$,

(ii) u_t satisfies $E(u_t | \mathfrak{F}_{t-1}) = 0$,

(iii) u_t is measurable w.r.t. $\mathfrak{F}_t$.

Notice that conditions (i) and (ii) in Definition 2.1 determine u_t and $\Sigma\beta_k z_{tk}$ uniquely up to null sets. Similarly, if two models M and $\bar{M}$ satisfy conditions (i) and (ii), then the respective residuals u_t and $\bar{u}_t$ are P-a.s. equal; especially if M is a true model so is then $\bar{M}$. Notice that in a true model the error process u_t is a martingale difference sequence and that no integrability assumptions have been made for y_t or z_{tk} . Conditions (i) and (ii) state that in a decomposition $y_t = f_t + u_t$, where f_t is $\mathfrak{F}_{t-1}$-measurable and $E(u_t | \mathfrak{F}_{t-1}) = 0$, the component f_t is P-a.s. equal to a pointwise linear combination of all the regressors contained in model M and the coefficients of the linear combination do not depend on t (that the coefficients can always be chosen in an $\mathfrak{F}$-measurable way is easily seen). This reflects the notion of a linear model for y_t . Condition (iii) which makes the error process u_t then a martingale difference sequence and the $\mathfrak{F}_{t-1}$-measurability condition on z_{tk} are only necessary to make the asymptotics work. The reason for allowing the coefficients β_k to be random variables rather than constants is that from each single realization the following two models cannot be distinguished: $y_t = 1 + u_t$, $y_t = \beta x_t + u_t$ where (u_t) is i.i.d., integrable and independent of (x_t), $x_t \equiv x$ takes on only the values 1, -1 with positive probability, $\beta = x^{-1}$, $\mathfrak{F}_t$ is generated by $\{x, u_t, u_{t-1}, \ldots\}$. Of course in many cases β_k will not depend on ω .

The selection of a model given the first T observations of y_t and z_{tk} will be based on a minimization over $\mathcal{M}$ of one of the following criterion functions:

(2.1) $$\log \hat{\sigma}_T^2(M) + \text{size}(M)C(T)/T$$

or

(2.2) $$\hat{\sigma}_T^2(M) + \text{size}(M)C(T)/T$$

where $\hat{\sigma}_T^2(M)$ is the residual variance after fitting model M . Here C(T) denotes a (possibly stochastic) nonnegative real-valued function on the integers; further properties of C(T) will be specified later. The quantity size(M) which is assumed to be real-valued (and possibly stochastic) stands for any measure of the size or complexity of the model M as e.g. the number of parameters. If size(M) is chosen to be the number of parameters and C(T) = 2 then (2.1) reduces to Akaike's (1969) AIC criterion; if C(T) = log T (2.1) reduces to Schwarz's (1978) BIC criterion. Notice also that for the minimization of (2.1) it is irrelevant whether in (2.1) the term $\hat{\sigma}_T^2(M)$ is replaced by the residual sum of squares RSS(M) or not, since $\hat{\sigma}_T^2(M) = T^{-1}*RSS(M)$. The results of this paper also apply to criteria similar to (2.1) and (2.2) where the penalty term size(M)C(T)/T is replaced by C(M,T)/T ; see Section 4.

The following lemmata are the essential building blocks for the consistency result. The first lemma gives conditions which guarantee that the value of the criterion function (2.1) or (2.2) at an incorrect model is eventually larger than the corresponding value at a true model. We use the following notation and conventions: the quantity $\hat{\sigma}_T^2(M)$ is given as

$T^{-1}y'(I-Z_M(Z_M'Z_M)^+Z_M')y$ where $y = (y_1,\dots,y_T)'$. The matrix Z_M is $T \times m$, where $m = \text{card}(M)$, with its t^{th} row equal to $(z_{tk_1},\dots,z_{tk_m})$ and and $k_1,\dots,k_m$ enumerate M . (If M is void or M contains only the zero regressor, then $Z_M = 0 \in \mathbb{R}^T$.) The projection matrix on the column space of Z_M is denoted by $P_M = Z_M(Z_M'Z_M)^+Z_M'$ where A^+ denotes the Moore–Penrose inverse of A and the prime denotes transposition. Furthermore we use $f = (f_1,\dots,f_T)'$, $u = (u_1,\dots,u_T)'$ where f_t and u_t are defined via the decomposition: $y_t = f_t + u_t$ P-a.s., f_t is $\mathfrak{F}_{t-1}$-measurable and $E(u_t|\mathfrak{F}_{t-1}) = 0$. Such a decomposition exists in view of Definition 2.1 and of the assumed existence of a true model (if y_t is integrable of course such a decomposition always exists). We note that f_t and u_t do not depend on a particular model. As a convention we set $\log 0 = -\infty$, $\log(a/0) = \log \infty = \infty$ if $a > 0$. The symbol $\log^+ x$ is $\log x$ if $x \geq 1$ and zero if $0 \leq x < 1$. For $x \in \mathbb{R}^T$ we set $\|x\| = (x'x)^{1/2}$. Now consider the following four conditions with $M \in \mathcal{M}$:

(1) $\|f - P_Mf\|^2 > 0$ for large T a.s. and $\|f - P_Mf\|^{-2}\, u'u$ converges to zero a.s. as $T \to \infty$.

(2) $T^{-1}\|f - P_Mf\|^2 \to \infty$ a.s. as $T \to \infty$ and $\sup_{T\geq 1} T^{-1}\, u'u < \infty$ a.s..

(3) $\|f - P_Mf\|^2 \to \infty$ a.s., $\sup_{T\geq 1} T^{-1}\, u'u < \infty$ a.s., $\sup_{t\geq 1} E(u_t^2|\mathfrak{F}_{t-1}) < \infty$ a.s. and $[\log^+ \text{tr}(Z_M'Z_M)]^{1+\delta}/\|f - P_Mf\|^2$ converges to zero a.s. for some $\delta > 0$ as $T \to \infty$.

(4) $\|f - P_Mf\|^2 \to \infty$ a.s., $\sup_{t\geq 1} E(|u_t|^{\alpha}|\mathfrak{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$ and $\log^+ \text{tr}(Z_M'Z_M)/\|f - P_Mf\|^2$ converges to zero a.s. as $T \to \infty$.

<u>Lemma 2.1</u>: Let $M_1 \in \mathcal{M}$ be a true model and $M_2 \in \mathcal{M}$. If one of the assumptions (1) - (4) holds for M_2, then we have:

(a) $\hat{\sigma}^2_T(M_2) > 0$ for large T a.s., hence $\log(\hat{\sigma}^2_T(M_2)/\hat{\sigma}^2_T(M_1))$ is well-defined.

(b) $\hat{\sigma}^2_T(M_2) - \hat{\sigma}^2_T(M_1) > (\text{size}(M_1) - \text{size}(M_2))C(T)/T$ holds for large T a.s. if $C(T)/\|f - P_{M_2}f\|^2$ converges to zero a.s. as $T \to \infty$.

(c) $\log(\hat{\sigma}^2_T(M_2)/\hat{\sigma}^2_T(M_1)) > (\text{size}(M_1) - \text{size}(M_2))C(T)/T$ holds for large T a.s. if $C(T)/(T \log(1 + \|f - P_{M_2}f\|^2(u'u)^{-1}))$ converges to zero a.s. as $T \to \infty$.

We note that under (1) - (4) the quantity $\|f - P_{M_2}f\|^2$ is eventually positive, hence $C(T)/\|f - P_{M_2}f\|^2$ is well-defined, and similarly $B(T) = \log(1 + \|f - P_{M_2}f\|^2(u'u)^{-1})$ is well-defined and positive (possibly $+\infty$ if $u'u = 0$) for large T, almost surely. Under assumption (2) the condition that $C(T)/\|f - P_{M_2}f\|^2$ converges to zero a.s. is clearly satisfied if $C(T)/T$ is bounded a.s.; under the classical assumptions for the asymptotic theory in linear regression models $\|f - P_{M_2}f\|^2$ behaves like T and then the above condition reduces to $C(T)/T \to 0$ almost surely. Similarly under (1) and (2) the a.s. boundedness of $C(T)/T$ is sufficient for $C(T)/TB(T)$ to converge to zero almost surely. For later use we remark that under $\sup T^{-1}u'u < \infty$ a.s. the condition $C(T)/(T \log(1 + \|f - P_{M_2}f\|^2T^{-1})) \to 0$ a.s. implies $C(T)/TB(T) \to 0$ a.s., and the converse is true if $\liminf T^{-1}u'u > 0$ a.s. holds. Notice also that from the positivity of $\|f - P_{M_2}f\|^2$ it follows that M_2 is not a true model. The various conditions involving $\|f - P_{M_2}f\|^2$

are separation conditions, i.e., they tell us how well separated from the true models the incorrect models have to be in order that they can be detected as wrong models. Furthermore, since all the conditions on M_2 are only in terms of P_{M_2} and $tr(Z'_{M_2}Z_{M_2})$, they depend only on the space spanned by the model M_2, i.e., the image of P_{M_2}, and not on the special way this space is represented by the regressors in M_2 (cf. Lemma A.2 in the appendix). It is furthermore worth mentioning that for Lemma 2.1 to hold under condition (1) or (2) it is not necessary that the regressors z_{tk} are $\mathcal{F}_{t-1}$-measurable and that u_t is $\mathcal{F}_t$-measurable; only properties (i) and (ii) of Definition 2.1 are relevant. We note also that $\sup E(|u_t|^\alpha|\mathcal{F}_{t-1}) < \infty$ a.s. for $\alpha > 2$ implies $u'u = \sum_{t=1}^{T} E(u_t^2|\mathcal{F}_{t-1}) + o(T)$ (see Chow (1965), Lai and Wei (1982a)), hence $\sup T^{-1} u'u < \infty$ a.s. is automatically satisfied.

The next lemma gives conditions under which the value of the criterion function (2.1) or (2.2) at a true model of minimal size is eventually smaller than the corresponding value at a non-minimal true model. We introduce the conditions (5) and (6) for a model $M \in \mathcal{M}$:

(5) $\sup_{t\geq 1} E(u_t^2|\mathcal{F}_{t-1}) < \infty$ a.s., $C(T) \to \infty$ a.s. and $[\log^+ tr(Z'_M Z_M)]^{1+\delta}/C(T) \to 0$ a.s. for some $\delta > 0$ as $T \to \infty$.

Condition (6) is obtained from (5) by replacing $\sup E(u_t^2|\mathcal{F}_{t-1}) < \infty$ a.s. by the stronger assumption $\sup E(|u_t|^\alpha|\mathcal{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$ and by setting $\delta = 0$ in (5).

<u>Lemma 2.2</u>: Let $M_1 \in \mathcal{M}$ and $M_2 \in \mathcal{M}$ be true models with size(M_1) < size(M_2). If each one of M_1 and M_2 satisfies one of (5) or (6), then we have:

(a) If $\liminf_{T\to\infty} T^{-1}u'u > 0$ a.s. and $[\log^+ tr(Z'_{M_i}Z_{M_i})]^{1+\delta_i}/T \to 0$ a.s. for i=1,2 (for some $\delta_i > 0$ if (5) holds for M_i and for $\delta_i = 0$ if (6) holds for M_i), then $\hat{\sigma}^2_T(M_i) > 0$ for large T a.s., hence $\log(\hat{\sigma}^2_T(M_2)/\hat{\sigma}^2_T(M_1))$ is well-defined, and $\log(\hat{\sigma}^2_T(M_2)/\hat{\sigma}^2_T(M_1)) > (\text{size}(M_1) - \text{size}(M_2))\, C(T)/T$ holds for large T almost surely.

(b) $\hat{\sigma}^2_T(M_2) - \hat{\sigma}^2_T(M_1) > (\text{size}(M_1) - \text{size}(M_2))C(T)/T$ holds for large T almost surely.

Clearly, the a.s. boundedness of $C(T)/T$ is a sufficient condition for $[\log^+ tr(Z'_{M_i}Z_{M_i})]^{1+\delta_i}/T \to 0$ a.s. under (5) or (6). Notice that under $\sup E(|u_t|^{\alpha}|\mathfrak{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$ the condition $\liminf T^{-1} u'u > 0$ a.s. is equivalent to $\liminf_{T\to\infty} T^{-1}\sum_{t=1}^{T} E(u_t^2|\mathfrak{F}_{t-1}) > 0$ a.s. which is in turn implied by $\liminf_{t\to\infty} E(u_t^2|\mathfrak{F}_{t-1}) > 0$ a.s., a condition used in Lai and Wei (1982b, 1983). Furthermore the conditions in (5) and (6) depend on the models M_1, M_2 only through the space spanned by them (note, however, that models spanning the same space may have assigned different values of the complexity measure size(M)). We note that Lemma 2.2(b) holds under any alternative set of conditions which imply that $C(T)^{-1}u'P_{M_i}u$ goes to zero a.s. for i=1,2 and $C(T) > 0$ for large T a.s.; similarly part (a) of the lemma is true if these alternative conditions additionally imply that $\liminf T^{-1} u'u > 0$ a.s. and $\lim T^{-1}u'P_{M_i}u = 0$ almost surely. (The $\mathfrak{F}_{t-1}$-measurability of z_{tk} and the $\mathfrak{F}_t$-measurability of u_t need not be satisfied under these alternative conditions.) This will be of importance in Section 3. A simple consequence of Lemmata 2.1 and 2.2 is Corollary 2.3. Denote by

$\hat{M}(T,1)$ and $\hat{M}(T,2)$, respectively, an arbitrary model which minimizes the criterion function (2.1) or (2.2), respectively, over $\mathcal{M}$. Under a minimal true model we understand a true model which has minimal size among all true models.

Corollary 2.3: Let $\mathcal{M}$ be finite.

(a) Assume that for each $M \in \mathcal{M}$ which is not a true model one of the conditions (1) - (4) holds. Then firstly $\hat{M}(T,1)$ is a true model for large T a.s. if $C(T)/(T \log(1 + \|f - P_M f\|^2 (u'u)^{-1}))$ goes to zero a.s. as $T \to \infty$ for all $M \in \mathcal{M}$ which are not true models. Secondly, $\hat{M}(T,2)$ is a true model for large T a.s. if $C(T)/\|f - P_M f\|^2$ goes to zero a.s. as $T \to \infty$ for all $M \in \mathcal{M}$ which are not true models.

(b) Assume that for each true model $M \in \mathcal{M}$ one of (5) or (6) holds. Then firstly $\hat{M}(T,1)$ is not a nonminimal true model for large T a.s. if $\liminf_{T\to\infty} T^{-1}u'u > 0$ a.s. and if for all true $M \in \mathcal{M}$ we have

$[\log^+ \mathrm{tr}(Z_M' Z_M)]^{1+\delta}/T \to 0$ a.s. for some $\delta = \delta(M)$ with $\delta(M) > 0$ if (5) holds for M and with $\delta(M) = 0$ if (6) holds for M. Secondly, $\hat{M}(T,2)$ is not a nonminimal true model for large T almost surely.

Combining parts (a) and (b) of Corollary 2.3 one obtains rather general conditions under which the selected models $\hat{M}(T,1)$ and $\hat{M}(T,2)$ are eventually minimal and correct. Two of the most important special cases of Corollary 2.3 are more explicitly given in the following theorem.

Theorem 2.4: Let $\mathcal{M}$ be finite and assume that $\sup_{t\geq 1} E(|u_t|^\alpha|\mathfrak{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$, and $C(T) \to \infty$ a.s. as $T \to \infty$ hold. Assume that for all models M_1, M_2 in $\mathcal{M}$ where M_1 is true and M_2 is not true the following holds:

(i) $\|f - P_{M_2}f\|^2 \to \infty$ a.s., as $T \to \infty$,

(ii) $\log^+ tr(Z'_{M_2}Z_{M_2})/\|f - P_{M_2}f\|^2 \to 0$ a.s. as $T \to \infty$,

(iii) $\log^+ tr(Z'_{M_1}Z_{M_1})/C(T) \to 0$ a.s. as $T \to \infty$.

Then:

(a) $\hat{M}(T,2)$ is a minimal true model for large T a.s. if $C(T)/\|f - P_{M_2}f\|^2 \to 0$ a.s. for all $M_2 \in \mathcal{M}$ which are not true models. Furthermore $\hat{M}(T,1)$ is a minimal true model for large T a.s. if $\liminf_{T\to\infty} T^{-1}u'u > 0$ a.s., if $\log^+ tr(Z'_{M_1}Z_{M_1})/T \to 0$ a.s. for all true $M_1 \in \mathcal{M}$ and if $C(T)/(T\log(1 + \|f - P_{M_2}f\|^2 T^{-1})) \to 0$ a.s. for all $M_2 \in \mathcal{M}$ which are not true models.

(b) The same conclusion as in (a) holds if (i) is replaced by $T^{-1}\|f - P_{M_2}f\|^2 \to \infty$ a.s. and (ii) is dropped.

Notice that if at least one model $M_2 \in \mathcal{M}$ exists which is not true, then the other conditions in Theorem 2.4 already imply $C(T) \to \infty$ almost surely: condition (i) then implies $\lim \|f\|^2 = \infty$ a.s., i.e. $\lim \|Z_{M_1}\beta\|^2 = \infty$ a.s. for some vector β as in Definition 2.1. This implies $\lim \lambda_{max}(Z'_{M_1}Z_{M_1}) = \infty$ a.s., hence $tr(Z'_{M_1}Z_{M_1}) \to \infty$ a.s. which then gives $C(T) \to \infty$ a.s. using (iii).

Remark 1: (i) The conditions involving $\|f - P_M f\|^2$ and $tr(Z'_M Z_M)$ in the results above can be expressed in terms of eigenvalues. For example the condition $\lim\|f - P_{M_2}f\|^2 = \infty$ a.s. for models M_2 which are not true models is equivalent to $\lim \lambda_{min}((Z_{M_2}:f)'(Z_{M_2}:f)) = \infty$ a.s. in view of (1.6) in Lai and Wei (1982b) which simply means that f and Z_{M_2} are asymptotically not multicollinear. The other condition in (3) or (4) involving $\|f - P_{M_2}f\|^2$ balances the growth rates of the largest eigenvalue of $Z'_{M_2}Z_{M_2}$ and of the smallest eigenvalue of $(Z_{M_2}:f)'(Z_{M_2}:f)$. The conditions on $C(T)$ in Lemma 2.1 give an upper bound for the growth rate of $C(T)$ in terms of the growth of $\lambda_{min}((Z_{M_2}:f)'(Z_{M_2}:f))$. Similarly the conditions on $C(T)$ in (5) or (6) specify a minimal rate of divergence of $C(T)$ in terms of $\lambda_{max}(Z'_{M_1}Z_{M_1})$, M_1 a true model.

(ii) In the special case when all models $M \in \mathcal{M}$ are submodels of an overall true model (not necessarily a member of $\mathcal{M}$) , i.e. Z_M is a selection of columns from a matrix X corresponding to the overall model, and when the coefficients β in the overall model can be chosen so as not to depend on ω , then the condition $\lim\lambda_{min}(X'X) = \infty$ a.s. is sufficient for $\lim \|f - P_M f\|^2 = \infty$ a.s. when M is not a true model: clearly $\|f - P_M f\|^2 \geq \|f - P_{M*}f\|^2$ where Z_{M*} is a matrix containing all columns of X except one for which the corresponding coefficient $\beta_k \neq 0$, and is such that Z_{M*} contains all columns of Z_M . This choice is possible since M is not a true model and β is nonrandom. But then $\|f - P_{M*}f\|^2 \geq \beta_k^2 K^{-1}\lambda_{min}(X'X)$ where K is the number of columns of X . Given $\lim \lambda_{min}(X'X) = \infty$ a.s. one similarly shows that $\lambda_{min}^{-1}(X'X) \log \lambda_{max}(X'X) \to 0$ a.s. implies the condition in (4) involving $\|f - P_M f\|^2$ and $\log^+ tr(Z'_M Z_M)$ and that the conditions for $C(T)$ in Lemma 2.1 are now satisfied if $C(T)/\lambda_{min}(X'X) \to 0$ a.s. and $C(T)/(T \log (1 + \lambda_{min}(X'X)(u'u)^{-1})) \to 0$ a.s., respectively; the second condition on $C(T)$ in (5) or (6) is implied by $\log \lambda_{max}(X'X)/C(T) \to 0$ a.s..

(Of course an overall true model can always be constructed; however the sufficient conditions in terms of $X'X$ may then be overly restrictive.)

The following example shows how Theorem 2.4 works in the framework of asymptotically stationary processes.

<u>Example</u>: Let z_{tk}, $1 \leq k \leq K$, $K \in \mathbb{N}$, be jointly asymptotically stationary processes in the sense that $T^{-1} \sum_{t=1}^{T} z_{tk} z_{tj}$ converges a.s. to some real-valued (possibly stochastic) quantity q_{kj}. We assume that the processes are not multicollinear in the sense that the matrix $Q = (q_{kj})$, $1 \leq j, k \leq K$, is a.s. positive definite. The process y_t is generated as $y_t = \sum_{k=1}^{K} \beta_k z_{tk} + u_t$ where $\beta = (\beta_1, \ldots, \beta_K) \in \mathbb{R}^K$ is nonrandom, and where u_t is a martingale difference sequence w.r.t. a filtration $\mathfrak{F}_t$ such that z_{tk} is $\mathfrak{F}_{t-1}$-measurable (typically the filtration will be generated by current and past errors, current and past regressors as well as the regressors leading by one). We assume that $\sup E(|u_t|^{\alpha} | \mathfrak{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$ and that $\liminf T^{-1} u'u > 0$ a.s.. If we choose $\mathcal{M}$ as the set of all models M_ℓ, $0 \leq \ell \leq K$, where M_ℓ contains all regressors with indices running from 1 to ℓ, and M_0 is the void model, we face the situation of model selection where we have an a priori ordering of the regressors expressed through their enumeration. If we choose $\mathcal{M}$ as the set of all possible models (including the void model) with regressors from $(z_{tk}, 1 \leq k \leq K)$, then we have the subset selection problem. In the first case we want to end up with the model M_{ℓ_0} where ℓ_0 is the maximum of $\{k : \beta_k \neq 0\}$ and $\ell_0 = 0$ if this set is empty; in the second case the desired model contains only the regressors with indices from the set $\{k : \beta_k \neq 0\}$. We shall now verify the conditions of Theorem 2.4 and show that in

both cases we shall eventually pick the desired model if either one of the model selection criteria (2.1) or (2.2) is used with size(M) = number of regressors in M (size(M) = 0 if M is void) and C(T) is such that log T/C(T) and C(T)/T converge to zero. First of all a model M is true iff it contains all regressors with indices from $\{k : \beta_k \neq 0\}$. One half of this statement is trivial; the other one follows from the fact that

$\liminf T^{-1}\|f - P_M f\|^2 \geq \liminf T^{-1} \|f - P_{M*}f\|^2 \geq \beta_{k_1}^2 K^{-1} \lambda_{min}(Q) > 0$ where M* contains all variables z_{tk} except for $k = k_1$ where $\beta_{k_1} \neq 0$ and z_{tk_1} is not contained in M (if $y_t = u_t$ then there are only true models and the above claim is trivial). Hence we have also shown that condition (i) in Theorem 2.4 is satisfied and that $\|f - P_M f\|^2$ grows at least linearly for M not a true model. Next for an arbitrary model M we have

$\lim T^{-1} tr(Z_M' Z_M) = \sum q_{kk}$ where the summation is over all k such that $k \in M$, hence $\log^+ tr(Z_M' Z_M) = O(\log T)$. But then the conditions in (ii), (iii) and (a) of Theorem 2.4 are satisfied which gives the desired conclusion. Obviously in this example the analogous result holds for any other set of models $\mathcal{M}$ too.

Remark 2: Finally we note that the assumption $\liminf T^{-1} u'u > 0$ a.s., which means that the noise does not die out eventually, has only been used for the "underestimation" part and for the criterion function (2.1). Clearly an (asymptotically) vanishing noise should make model selection easier and therefore such a condition like $\liminf T^{-1} u'u > 0$ a.s. is unnatural. As the above results show we can indeed avoid this condition for the criterion function (2.2). For the criterion function (2.1) the "overestimation" part also works without this condition since the systematic bias in $\hat{\sigma}_T^2(M_2)$, M_2 not a true model, dominates the residual error variance $\hat{\sigma}_T^2(M_1)$ of a true model anyway. However, in the "underestimation" part we have to differentiate between different true models M_1 and M_2 on the basis of $\log \hat{\sigma}_T^2(M_1)$, $\log \hat{\sigma}_T^2(M_2)$ and their respective sizes if we use criterion (2.1). Now if the noise is for example not existent, i.e. $u'u = 0$ a.s., then the first term in (2.1) equals minus infinity regardless of the size of M_1, M_2 and hence we cannot discriminate between a minimal and a non-minimal true model in this case. This is of course an extreme case and the condition $\liminf T^{-1} u'u > 0$ can be relaxed somewhat. An inspection of the proof of Lemma 2.2 shows that the following is true: Let M_1 and M_2 be true models with size(M_1) < size(M_2) and assume that $u'u > 0$ for large T a.s., $(u'u)^{-1}u'P_{M_i}u$ goes to zero a.s. and that $T/C(T)$ is bounded almost surely (actually $TC(T)^{-1}(u'u)^{-1}u'P_{M_i}u \to 0$ a.s. and $C(T) > 0$ eventually a.s. suffice). Then $\hat{\sigma}_T^2(M_i) > 0$ for large T a.s. and $\log(\hat{\sigma}_T^2(M_2)/\hat{\sigma}_T^2(M_1)) > (\text{size}(M_1) - \text{size}(M_2))\, C(T)/T$ holds for large T almost surely. In this respect it is also of interest to note that $\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1) > (\text{size}(M_1) - \text{size}(M_2))\, C(T)/T$ for large T a.s. holds if only $u'u/C(T) \to 0$ a.s. and $C(T) > 0$ for large T a.s. (of course in many cases this is a much weaker result than Lemma 2.2b). It should also be noted that for these alternative versions of Lemma 2.2 to hold it is not necessary that the regressors are $\mathcal{F}_{t-1}$-measurable and that u_t is $\mathcal{F}_t$-measurable.

3. Autoregressive Models

In this section we apply the general results obtained in Section 2 to the orderestimation problem in general autoregressive models. We shall first discuss nonexplosive models and purely explosive ones and then general autoregressive models. The result for nonexplosive models will then be sharpened in the case of stable autoregressive models where the errors satisfy stronger conditions.

Let us now fix the notation. The process y_t is for $t \geq 1$ assumed to be generated by

$$y_t = \beta_1 y_{t-1} + \cdots + \beta_{p_0} y_{t-p_0} + u_t \tag{3.1}$$

where $\beta = (\beta_1,\ldots,\beta_{p_0})'$ is the parameter vector (assumed not to depend on $\omega \in \Omega$) satisfying $\beta_{p_0} \neq 0$. Let $\mathfrak{F}_0$ denote the σ-field generated by the starting values $y_0,\ldots,y_{1-p_0}$ (if $p_0 = 0$, i.e. $y_t = u_t$, then $\mathfrak{F}_0$ is set equal to the trivial σ-field) and let $\mathfrak{F}_t$ be the σ-field generated by the starting values and $u_1,\ldots,u_t$. The error process is assumed to be a martingale difference with respect to the filtration $\{\mathfrak{F}_t\}$, i.e. $E(u_t|\mathfrak{F}_{t-1}) = 0$. These assumptions will be maintained throughout this section, except where otherwise noted.

The order p_0 of the autoregressive process will in general not be known and hence has to be estimated. One way to do this is to minimize as usual one of the criterion functions (2.1) or (2.2) over all autoregressive models M_p of order p , $0 \leq p \leq P$, where it is assumed that $p_0 \leq P$ and $P \geq 1$ is a prespecified constant. In more detail, autoregressive models of order p are fitted to the data $y_{P+1},\ldots,y_T$ by least squares and the residual variance $\hat{\sigma}^2_T(p)$ is calculated which is then used to calculate (2.1) and (2.2). We set $y = (y_{P+1},\ldots,y_T)'$ and Z_p is the $(T-P) \times p$ matrix whose i^{th} row is given by $(y_{P-1+i},\ldots,y_{P-p+i})$. If $p = 0$ we put $Z_p = 0 \in \mathbb{R}^{T-P}$. This conforms with the notation of the previous section if we take into account that the sample period used for the calculation of $\hat{\sigma}^2_T(p)$ is now $P + 1 \leq t \leq T$, i.e. $\hat{\sigma}^2_T(p) = (T-P)^{-1} y'(I-Z_p(Z_p'Z_p)^+Z_p')y$. Let $\hat{p}_T(1)$ and $\hat{p}_T(2)$, respectively, denote a minimizer of $\log \hat{\sigma}^2_T(p) + pC(T)/T$ and of $\hat{\sigma}^2_T(p) + pC(T)/T$ over $0 \leq p \leq P$, respectively. Then we have the following consistency result for nonexplosive processes.

<u>Theorem 3.1</u>: Assume that the characteristic polynomial of (3.1), i.e. $1 - \beta_1 z - \cdots - \beta_{p_0} z^{p_0}$, has all its zeroes outside or on the unit circle in the complex plane. Let $\sup_{t \geq 1} E(|u_t|^\alpha | \mathcal{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$ and $\liminf_{T\to\infty} T^{-1} \sum_{t=1}^{T} E(u_t^2 | \mathcal{F}_{t-1}) > 0$ a.s. hold. Then $\hat{p}_T(1) \geq p_0$ and $\hat{p}_T(2) \geq p_0$ hold for large T a.s. if $C(T)/T \to 0$ a.s. as $T \to \infty$. Furthermore $\hat{p}_T(1) \leq p_0$, $\hat{p}_T(2) \leq p_0$ hold for large T a.s. if $C(T)/\log T \to \infty$ a.s. as $T \to \infty$.

Of course combining both parts of the theorem shows that a growth rate of $C(T)$ between $\log T$ and T gives consistent estimators. For purely explosive models Theorem 3.2 provides the analogous result. Note that the assumptions of Theorem 3.2 and $\beta_{p_0} \neq 0$ imply that (3.1) is purely explosive in the sense of Lai and Wei (1983). Of course $p_0 \geq 1$ holds in Theorem 3.2 below.

<u>Theorem 3.2</u>. Assume that the characteristic polynomial of (3.1) has all its zeroes inside the unit circle in the complex plane. Assume $\sup_{t \geq 1} E(|u_t|^\alpha | \mathcal{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$ and $\liminf_{t\to\infty} E(u_t^2 | \mathcal{F}_{t-1}) > 0$ a.s. hold. Then $\hat{p}_T(1) \geq p_0$ holds for large T a.s. if $C(T)/T^2 \to 0$ a.s. as $T \to \infty$.

Similarly $\hat{p}_T(2) \geq p_0$ holds for large T a.s. if $C(T)/e^{aT} \to 0$ a.s. as $T \to \infty$ for some $0 < a = a(\omega) < -2 \log m$ where m is the maximum of the moduli of the zeroes of the characteristic polynomial. If $\liminf_{T\to\infty} C(T)/T > 0$ a.s. then $\hat{p}_T(1) \leq p_0$ and $\hat{p}_T(2) \leq p_0$ for large T a.s. .

The slightly stronger assumption on the conditional variance of the error process in Theorem 3.2 compared to Theorem 3.1 is needed to ensure the validity of Theorem 2 in Lai and Wei (1983) which gives an exponential rate for the eigenvalues of $Z'_{p_0}Z_{p_0}$. The general case, i.e. the case where there are no restrictions on the location of the zeroes of the characteristic polynomials, is treated in the next theorem.

<u>Theorem 3.3</u>: Assume $\sup_{t\geq 1} E(|u_t|^\alpha|\mathcal{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$ and $\liminf_{t\to\infty} E(u_t^2|\mathcal{F}_{t-1}) > 0$ a.s. hold. Then $\hat{p}_T(1) \geq p_0$ and $\hat{p}_T(2) \geq p_0$ for large T a.s. if $C(T)/T \to 0$ a.s. as $T \to \infty$. Furthermore $\hat{p}_T(1) \leq p_0$ and $\hat{p}_T(2) \leq p_0$ hold for large T a.s. if $\liminf_{T\to\infty} C(T)/T > 0$ almost surely.

<u>Remark 1</u>: It is interesting to note that (2.1) and (2.2) lead to different feasible rates for $C(T)$ in explosive models. Notice also that Theorem 3.3 does not give a common feasible rate for $C(T)$ such that both parts of that theorem are satisfied and consistency is ensured. Of course such a rate for $C(T)$ may nevertheless exist since Theorem 3.3 gives sufficient conditions only. A proof of such a result, however, would essentially have to produce a

convergence rate for the least squares estimator in a general autoregressive model which seems to be very difficult.

Remark 2: The proof of the overestimation part, i.e. the first half of Theorem 3.3, is based on the method of proof used in Wang and An (1984). Lemma A.1 is implicit in the proof of Theorem 1 of Lai and Wei (1983) and appears also in Wang and An (1984). In the context of general autoregressive models Wang and An (1984) prove only that $\hat{p}_T(1) \geq p_0$ eventually a.s. for $C(T) = \log T$ (they actually treat the corresponding problem of subset selection of autoregressive parameters, see also Remark 3 below).

Remark 3: The results of this section can be extended to model selection in autoregressions where models are chosen from a more general set of models $\mathcal{M}$, to some extent. Theorem 3.1 carries over completely to this case; however the full strength e.g. of the overestimation part of Theorem 3.2 does not necessarily go through. This is so because estimating $\|f - P_M f\|^2$ for wrong models M from below by $\lambda_{min}(Z'_{\bar{M}} Z_{\bar{M}})$, where $\bar{M}$ is the smallest true model containing M, may give only a linear growth rate since the model $\bar{M}$ is not necessarily purely explosive in the sense of Lai and Wei (1983) although M_{p_0} is. Another potential difficulty is that the proof of Theorem 3.3 relies on the consistency of least squares estimators in general autoregressive models as established in Lai and Wei (1983). If a subset autoregression is estimated by least squares, then it is not evident that the consistency result carries over to this case (note, however, that Wang and An (1984) make use of this "result" in their proof).

Remark 4: (i) The conditions $\liminf T^{-1} \sum E(u_t^2 | \mathcal{F}_{t-1}) > 0$ a.s. and $\liminf E(u_t^2 | \mathcal{F}_{t-1}) > 0$ a.s. are, in the context of autoregressive models, used to ensure that the design matrix Z_p does not degenerate and that the lower bounds for $\lambda_{min}(Z'_p Z_p)$ given in Lai and Wei (1983, 1985) hold. Hence these conditions are essential for the overestimation parts of the theorems of this section. As discussed in Remark 2 of the previous section these conditions are also used to prove the underestimation parts for the estimator $\hat{p}_T(1)$. In contrast to that the underestimation result for $\hat{p}_T(2)$ in Theorem 3.1 and in Theorems 3.2, 3.3 (in the latter two theorems under the slightly stronger condition $C(T)/T \to \infty$ a.s.) hold without any of these two conditions as can be seen from the proofs. Furthermore $\hat{p}_T(2) \leq p_0$ for large T a.s. holds under the single condition $u'u/C(T) \to 0$ a.s. (and $C(T) > 0$ eventually a.s.) without any further assumptions on u_t . However, without such further conditions also models with $p < p_0$ might be true models.

(ii) Under the assumptions of the theorems of this section p_0 is of course uniquely determined. Clearly this is already true under weaker conditions on u_t .

Finally we show that in the stable case the condition $C(T)/\log T \to \infty$ can be weakened to $C(T)/\log \log T \to \infty$ if the errors are essentially a stationary and ergodic martingale difference sequence. A similar result was proved in Hannan and Quinn (1979) in the stationary case (and where $\hat{\sigma}^2_T(p)$ was obtained from the Yule-Walker equations). We have not been able to prove a similar result e.g. for nonexplosive models although it might be possible. Note that in the theorem below the error process has to be a martingale difference sequence only w.r.t. the smaller σ-fields $\mathfrak{F}_t$.

Theorem 3.4: Assume that the characteristic polynomial of (3.1) has all its zeroes outside the unit circle in the complex plane, that u_t is a strictly stationary ergodic martingale difference sequence w.r.t. the σ-fields $\mathfrak{F}_t$ where $\mathfrak{F}_t$ is generated by $\{u_s : 1 \leq s \leq t\}$ and that the fourth moment of u_t is finite and positive. Then $\hat{p}_T(1) \geq p_0$, $\hat{p}_T(2) \geq p_0$ for large T a.s. if $C(T)/T \to 0$, and $\hat{p}_T(1) \leq p_0$, $\hat{p}_T(2) \leq p_0$ for large T a.s. if $C(T)/\log\log T \to \infty$ a.s. as $T \to \infty$.

The $\log \log T$ bound for $C(T)$ in Theorem 3.3 can be slightly weakened to $\liminf C(T)/\log\log T > c$ for a suitable constant c which, however, may depend on (y_t). This bound is then sharp, in general; compare the discussion in Hannan and Quinn (1979), p. 193. We note that Akaike's AIC, i.e. (2.1) with $C(T) = 2$ and $\text{size}(M_p) = p$ does not satisfy all the conditions in the above theorems. Actually, AIC does not even give weakly consistent order estimators in nonexplosive models, see Tsay (1984). The criterion BIC, i.e. (2.1) with $C(T) = \log T$ and $\text{size}(M_p) = p$ does not satisfy all the conditions in Theorem 3.1. A close inspection of its proof, however, reveals that Theorem 3.1 is still valid if $C(T)/\log T \to \infty$ is replaced by $\liminf C(T)/\log T > c$. The constant c, however, depends on (y_t). Thus if BIC would be modified to BIC' by setting $C(T) = c' \log T$, $c' > c$, then we could be sure that BIC' gives consistent estimates. This is of course of no great practical interest since c is unknown. Furthermore BIC gives consistent estimators under the conditions of Theorem 3.3 of course. Finally we note that all results in this section obviously remain true if the upper bound P increases with the sample size (possibly depending on ω) slowly enough. We do not know how fast P is allowed to increase. For some results in this direction for stationary autoregressions, see e.g. An, Chen and Hannan (1982).

4. Complementary Remarks

The "overconsistency" result for stochastic linear regression models selected by BIC as given in Wang and An (1984) can be derived under weaker assumptions as is seen from an inspection of their proof: conditions (2.6) and (2.8) can be dropped without any loss since — in the notation of that paper — $S(J_k)^{-1}$ can be estimated from below by $(s's)^{-1}$ and hence Lemma 2.1 in that paper is not necessary for the proof of their Theorem 2.1 (notice that their assumption (2.1) implies $(s's)^{-1} > MT^{-1}$, $M = M(\omega) > 0$).

The results of the present paper easily carry over to more general criteria of the form (2.1) or (2.2) where the penalty term size(M)C(T)/T is replaced by $C(M,T)/T$. Lemma 2.1 carries over where now the conditions for $C(T)$ in (b) and (c) of this lemma have to be satisfied by $\Delta C(M_1,M_2,T) = C(M_1,T) - C(M_2,T)$. For Lemma 2.2 to carry over the penalty term has to be such that at least for true models M and M' always either $\Delta C(M,M',T) \geq 0$ for large T a.s. or ≤ 0 for large T a.s. holds. This gives then the required ordering of the models according to their "size". Lemma 2.2 is then true if all conditions for $C(T)$ are satisfied by $\Delta C(M_1,M_2,T)$ and $\Delta C(M_1,M_2,T) \leq 0$ for large T a.s. holds. These results are of some importance for an analysis of model selection criteria such as Mallow's (1973) C_p .

Appendix

Proof of Lemma 2.1:

We start from the basic identity

(A.1) $T(\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1)) = \|f - P_{M_2}f\|^2 + 2f'(I-P_{M_2})u - u'P_{M_2}u + u'P_{M_1}u$.

The second term on the r.h.s. is now in absolute value not larger than $2\|f - P_{M_2}f\| (u'u)^{1/2}$, the third term not larger than $u'u$ and the fourth term is nonnegative. Hence since $\|f - P_{M_2}f\| > 0$ eventually under (1) - (4),

$T\|f - P_{M_2}f\|^{-2}(\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1)) \geq 1 - 2\|f - P_{M_2}f\|^{-1}(u'u)^{1/2} - \|f - P_{M_2}f\|^{-2} u'u$.

Now under (1) or (2) the r.h.s. converges to one; hence for every $0 < \varepsilon < 1$ we have for large T almost surely

(A.2) $\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1) > \varepsilon \|f - P_{M_2}f\|^2/T$.

Under (3) (or (4) respectively) the second term on the r.h.s. of (A.1) is $O(\|f - P_{M_2}f\| [\max(1, \log^+(\|f - P_{M_2}f\|), \log^+(tr(Z_{M_2}'Z_{M_2}))]^{(1+\delta)/2})$ for every $\delta > 0$ (or $\delta = 0$ respectively) and the term $u'P_{M_2}u$ on the r.h.s. of (A.1) is $O(\{\max[1, \log^+(tr(Z_{M_2}'Z_{M_2}))]\}^{1+\delta})$ for every $\delta > 0$ (or $\delta = 0$ respectively). This follows from Lai and Wei (1982b), Theorems 4 and 3 (in Theorem 3, eq. (2.2) a printing error occurs: the term in braces on the r.h.s. of (2.2) should read $\max(1, \log^+(\Sigma\Sigma z_{ij}^2)))$. But then clearly (A.2) also holds under (3) or (4). Now (A.2) proves that $\hat{\sigma}_T^2(M_2) > 0$ eventually and that $\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1) > (size(M_1) - size(M_2))C(T)/T$ under (b).

Since $\hat{\sigma}_T^2(M_1) = T^{-1}(u'u - u'P_{M_1}u) \leq T^{-1}u'u$ we obtain from (A.2) under either set of assumptions that almost surely for large T

(A.3) $[\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1)]\, \hat{\sigma}_T^{-2}(M_1) \geq \varepsilon\|f - P_{M_2}f\|^2 (u'u)^{-1}$

Notice that (A.3) trivially holds if $\hat{\sigma}_T^2(M_1) = 0$ or $u'u = 0$ because $\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1) > 0$ as shown above. Now for large T almost surely the r.h.s. of (A.3) is positive. But then $\log(\hat{\sigma}_T^2(M_2)/\hat{\sigma}_T^2(M_1)) = \log(1 + [\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1)]\ \hat{\sigma}_T^{-2}(M_1)) \geq \log(1 + \varepsilon\|f - P_{M_2}f\|^2(u'u)^{-1})$ for large T a.s. and hence (c) holds, since $\log(1 + x)/\log(1 + \varepsilon x)$ is bounded on the interval $(0,\infty)$.

Proof of Lemma 2.2:

Clearly for $i = 1,2$ we have $T\hat{\sigma}_T^2(M_i) = u'u - u'P_{M_i}u$. From Lai and Wei (1982b), Theorem 3, we get $u'P_{M_i}u = O(\{\max[1, \log^+(tr(Z_{M_i}'Z_{M_i}))]\}^{1+\delta})$ for all $\delta > 0$ if M_i satisfies (5) and for $\delta = 0$ if M_i satisfies (6). Now since $C(T) \to \infty$ and $[\log^+ tr(Z_{M_i}'Z_{M_i})]^{1+\bar{\delta}_i}/C(T) \to 0$ for some $\bar{\delta}_i > 0$ under (5) and for $\bar{\delta}_i = 0$ under (6) we get $\lim C(T)^{-1}u'P_{M_i}u = 0$ a.s. . Since size(M_1) < size(M_2) this proves (b) because $TC(T)^{-1}(\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1)) = C(T)^{-1}(u'P_{M_1}u - u'P_{M_2}u)$ which goes to zero as just shown. Now under the assumptions of (a) we clearly have $\liminf \hat{\sigma}_T^2(M_i) = \liminf T^{-1}u'u > 0$ a.s. since $\lim T^{-1}u'P_{M_i}u = 0$ follows from $[\log^+tr(Z_{M_i}'Z_{M_i})]^{1+\delta_i}/T \to 0$ and Theorem 3 in Lai and Wei (1982b). This clearly implies $\hat{\sigma}_T^2(M_i) > 0$ for large T a.s.; $\lim(\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1))\ \hat{\sigma}_T^{-2}(M_1) = 0$ a.s. also follows. Finally $\log(\hat{\sigma}_T^2(M_2)/\hat{\sigma}_T^2(M_1)) = \log(1 + (\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1))\ \hat{\sigma}_T^{-2}(M_1))$ and $(\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1))\ \hat{\sigma}_T^{-2}(M_1)$ goes to zero as just shown. Hence

$$TC(T)^{-1}\log(\hat{\sigma}_T^2(M_2)/\hat{\sigma}_T^2(M_1)) = (1+\xi_T)^{-1}\ TC(T)^{-1}(\hat{\sigma}_T^2(M_2) - \hat{\sigma}_T^2(M_1))\ \hat{\sigma}_T^{-2}(M_1)$$

where ξ_T is a mean value going to zero almost surely. But then (a) follows since the r.h.s. of the last equation goes to zero a.s. by what has already been established.

Proof of Corollary 2.3:

Since $\mathcal{M}$ is finite $\hat{M}(T,1)$ and $\hat{M}(T,2)$ exist. We give the proof only for $\hat{M}(T,1)$; the proof for $\hat{M}(T,2)$ is completely analogous. Assume (a) is not true. Since $\mathcal{M}$ is finite we would have $\hat{M}(T,1) = M_2$ infinitely often on a set of positive probability where M_2 is a fixed model which is not true. Choose a fixed true model $M_1 \in \mathcal{M}$. Applying Lemma 2.1 to M_1 and M_2 we see that the value of the criterion function (2.1) at M_2 is eventually almost surely larger than the value at M_1 which leads to a contradiction. To prove (b) assume that it would not be true. Then similar as above $\hat{M}(T,1) = M_2$ infinitely often with positive probability where M_2 is now a fixed nonminimal true model. By finiteness of $\mathcal{M}$ a true minimal model exists, say M_1 . Now from Lemma 2.2 we conclude that the value of (2.1) at M_2 is eventually and almost surely larger than the value at M_1 which leads to a contradiction.

Proof of Theorem 3.1:

First we show that an autoregressive model M_p with $p < p_0$ is not true by showing that $\liminf_{T\to\infty} (T-P)^{-1} \|f - P_{M_p}f\|^2 > 0$ almost surely. Of course it suffices to prove this for $p = p_0 - 1$. Now $(T-P)^{-1} \|f - P_{M_{p_0-1}}f\|^2 = (T-P)^{-1} \|Z_{p_0}\beta - P_{M_{p_0-1}}Z_{p_0}\beta\|^2 = (T-P)^{-1}\| v\, \beta_{p_0} - P_{M_{p_0-1}} v\, \beta_{p_0}\|^2$ where v denotes the last column of Z_{p_0} . The last expression equals $\beta^2_{p_0}(T-P)^{-1} \|v - P_{M_{p_0-1}}v\|^2 \geq p_0^{-1}\beta^2_{p_0}(T-P)^{-1} \lambda_{min}(Z'_{p_0}Z_{p_0})$, the inequality following from (1.6) in Lai and Wei (1982b). Since $\beta_{p_0} \neq 0$ and $\liminf (T-P)^{-1} \lambda_{min}(Z'_{p_0}Z_{p_0}) > 0$ a.s. by Theorem 3 and Example 3 in Lai and Wei (1985), we arrive at the desired conclusion. Furthermore for $p < p_0$ we have $\log^+tr(Z'_pZ_p) \leq \log^+tr(Z'_{p_0}Z_{p_0}) \leq \log^+p_0 + \log^+ \lambda_{max}(Z'_{p_0}Z_{p_0})$. It follows now from Corollary 1 in Lai and Wei (1985) that $\lambda_{max}(Z'_{p_0}Z_{p_0}) =$

$= O(T)$ if all zeroes of the characteristic polynomial are outside the unit circle and $\lambda_{max}(Z'_{p_o}Z_{p_o}) = O(T^{2\rho} \log\log T)$ otherwise where ρ is the sum of the multiplicity of all of the zeroes on the unit circle. In any case $\log^+ \lambda_{max}(Z'_{p_o}Z_{p_o}) = O(\log T)$. Since $\|f - P_{M_p}f\|^2$ increases at least as T as just shown above, we have $\log^+ tr(Z'_{p_o}Z_{p_o})/\|f - P_{M_p}f\|^2 \to 0$ for $p < p_o$. This shows that assumption (4) is satisfied and applying Corollary 2.3 we obtain the first half of Theorem 3.1 taking into account the remarks after Lemma 2.1. On the other hand it is obvious that M_p is a true model if $p \geq p_o$. Furthermore $\log^+ tr(Z'_pZ_p) \leq \log^+ p + \log^+\lambda_{max}(Z'_pZ_p) = O(\log T)$ again by Corollary 1 in Lai and Wei (1985). Hence $\log^+ tr(Z'_pZ_p)/T \to 0$ and $\log^+ tr(Z'_pZ_p)/C(T) \to 0$ since $C(T)/\log T \to \infty$ by assumption. This establishes condition (6) and the conditions in Corollary 2.3. The theorem then follows from Corollary 2.3.

Proof of Theorem 3.2:

For $p < p_o$ we have similarly as before $\|f - P_{M_p}f\|^2 \geq \|f - P_{M_{p_o-1}}f\|^2 \geq p_o^{-1}\beta^2_{p_o}\lambda_{min}(Z'_{p_o}Z_{p_o}) \geq p_o^{-1}\beta^2_{p_o}e^{aT}$ for large T a.s. and for $0 < a < -2 \log m$ using Theorem 2 in Lai and Wei (1983) (to make this theorem applicable the origin of time has to be shifted). This shows that condition (2) is satisfied. Furthermore $C(T)/e^{aT} \to 0$ a.s. clearly implies $C(T)/\|f - P_{M_p}f\|^2 \to 0$ a.s. and $C(T)/T^2 \to 0$ a.s. implies $C(T)/T \log(1 + \|f - P_{M_p}f\|^2(u'u)^{-1}) \to 0$ a.s. since certainly $\sup T^{-1}u'u < \infty$. The first half of the theorem then follows from Corollary 2.3. Now for $p \geq p_o$ the model M_p is certainly true and $C(T)^{-1}u'P_{M_p}u \to 0$ a.s. by Lemma A.1 and since $\liminf C(T)/T > 0$ a.s.; this gives $TC(T)^{-1}(\hat{\sigma}^2_T(p_2) - \hat{\sigma}^2_T(p_1)) = C(T)^{-1}(u'P_{M_{p_1}}u - u'P_{M_{p_2}}u) \to 0$ a.s. for $p_2 > p_1 \geq p_o$, hence $\hat{\sigma}^2_T(p_2) + p_2\, C(T)/T > \hat{\sigma}^2_T(p_1) + p_1\, C(T)/T$ eventually and almost surely. But then

$\hat{p}_T(2) \leq p_0$ for large T a.s. follows. To prove $\hat{p}_T(1) \leq p_0$ for large T a.s. observe that $\liminf T^{-1}u'u > 0$ a.s. holds and Lemma A.1 implies $\liminf \hat{\sigma}^2_T(p_1) > 0$ a.s. and $(\hat{\sigma}^2_T(p_2) - \hat{\sigma}^2_T(p_1))\, \hat{\sigma}^{-2}_T(p_1) \to 0$ a.s. . Proceeding as in the proof of Lemma 2.2 we get $\log \hat{\sigma}^2_T(p_2) + p_2\, C(T)/T > \log \hat{\sigma}^2_T(p_1) + p_1\, C(T)/T$ for large T a.s. from which the result follows. (We note that if $C(T)$ satisfies $C(T)/T \to \infty$ a.s. then $\hat{p}_T(2) \leq p_0$ eventually a.s. can be proved without Lemma A.1 by directly verifying (6) using Corollary 1 in Lai an and Wei (1985)).

<u>Proof of Theorem 3.3</u>:

For $p < p_0$ we have $\liminf (T-P)^{-1} \| f - P_{M_p} f \|^2 > 0$ by a similar argument as in the proof of Theorem 3.1, hence M_p is true iff $p \geq p_0$. The proof of the second part of the theorem is identical to the proof of the corresponding part of Theorem 3.2. The first part is proved as follows: for $p < p_0$ we have $\hat{\sigma}^2_T(p) - \hat{\sigma}^2_T(p_0) \geq \hat{\sigma}^2_T(p_0-1) - \hat{\sigma}^2_T(p_0) = (T-P)^{-1}\hat{\beta}^2_{p_0} \| v - P_{M_{p_0}} v \|^2$ where v is the last column of Z_{p_0} and $\hat{\beta}_{p_0}$ is the least squares estimator for β_{p_0} based on model M_{p_0} . This follows from a standard formula used in stepwise regression and (3.3) in Lai and Wei (1982b). Hence $\hat{\sigma}^2_T(p) - \hat{\sigma}^2_T(p_0) \geq (T-P)^{-1} p_0^{-1} \hat{\beta}^2_{p_0}\, \lambda_{min}(Z'_{p_0} Z_{p_0})$ and the r.h.s. is larger than $(p_0-p)\, C(T)/T$ for large T a.s. since $\hat{\beta}^2_{p_0} \to \beta^2_{p_0} > 0$ by Theorem 1 in Lai and Wei (1983), $\liminf T^{-1}\lambda_{min}(Z'_{p_0} Z_{p_0}) > 0$ a.s. by Theorem 3 in Lai and Wei (1985) and since $C(T)/T \to 0$ a.s. by assumption. This shows also that $\liminf \hat{\sigma}^2_T(p) > 0$ a.s.; furthermore since $\hat{\sigma}^2_T(p_0) \leq T^{-1}u'u$ and $\sup T^{-1}u'u < \infty$ we obtain similar as in the proof of Lemma 2.1 that $[\hat{\sigma}^2_T(p) - \hat{\sigma}^2_T(p_0)]\, \hat{\sigma}^{-2}_T(p_0) \geq c(\omega) > 0$ from which we conclude that $\log(\hat{\sigma}^2_T(p)/\hat{\sigma}^2_T(p_0)) > (p_0-p)\, C(T)/T$ for large T a.s. holds. The result $\hat{p}_T(1) \geq p_0$, $\hat{p}_T(2) \geq p_0$ then follows along the lines of the proof of Corollary 2.3.

Proof of Theorem 3.4:

Since all conditions and claims depend on the finite dimensional distributions only, we can assume without loss of generality that the underlying probability space is rich enough to allow u_t to be extended to a strictly stationary and ergodic process for all $t \in Z$. Necessarily u_t is then a martingale difference sequence with respect to $\mathfrak{F}_t$ where $\mathfrak{F}_t$ is the σ-field generated by $\{u_s : -\infty < s \leq t\}$ (this follows from stationarity and the martingale convergence theorem). The process y_t can then be decomposed as $y_t = y_t^* + y_t^{**}$ where y_t^* is the stationary solution of (3.1), i.e., $y_t^* = (1 - \beta_1 L - \cdots - \beta_{p_0} L^{p_0})^{-1} u_t$, L the shift operator, and y_t^{**} is a finite sum of expressions of the form $p(t)\lambda^t$ where $p(t)$ is a polynomial with random coefficients and $|\lambda| < 1$. Hence y_t^{**} is exponentially decaying. First of all it is now clear that $\liminf (\hat{\sigma}_T^2(p) - \hat{\sigma}_T^2(p_0))$ is almost surely positive for $p < p_0$ and $\hat{\sigma}_T^2(p_0) \to \sigma^2 = E(u_t^2) > 0$. Hence as in the proof of Lemma 2.1 we have that $\hat{p}_T(1)$ and $\hat{p}_T(2)$ are not less than p_0 for large T. For $p \geq p_0$ we have now $T(\hat{\sigma}_T^2(p_0) - \hat{\sigma}_T^2(p)) = u' Z_p(Z_p' Z_p)^+ Z_p' u - u' Z_{p_0}(Z_{p_0}' Z_{p_0})^+ Z_{p_0}' u$. To complete the proof of the theorem it is enough to show that both terms on the r.h.s. are at most of order $\log \log T$. Since of course $T^{-1} Z_p' Z_p$ and $T^{-1} Z_{p_0}' Z_{p_0}$ converge to nonsingular matrices Q_p and Q_{p_0}, respectively, it is enough to show that a typical entry of $u' Z_p$ (or $u' Z_{p_0}$ respectively) is at most of order $(T \log \log T)^{1/2}$. Now a typical entry is of the form $\sum_{i=P+1}^{T} u_t y_{t-i}$ where $1 \leq i \leq p$. Clearly $\sum u_t y_t^{**}$ is of smaller order than $(T \log \log T)^{1/2}$ by its nature and by Kronecker's lemma. Hence it suffices to consider $\sum u_t y_{t-i}^*$. But $u_t y_{t-i}^*$ is a stationary, ergodic and square integrable martingale difference sequence w.r.t. $\mathfrak{F}_t'$ and therefore satisfies the law of the iterated logarithm, see e.g. Stout (1970).

This shows that $TC(T)^{-1}(\hat{\sigma}_T^2(p) - \hat{\sigma}_T^2(p_0)) \to 0$ for $C(T)$ as specified in the theorem and $p \geq p_0$. Now proceeding as in the proof of Lemma 2.2 completes the proof.

Lemma A.1: In the notation of Section 3 let $\sup_{t \geq 1} E(|u_t|^\alpha | \mathcal{F}_{t-1}) < \infty$ a.s. for some $\alpha > 2$ and $\liminf_{t\to\infty} E(u_t^2 | \mathcal{F}_{t-1}) > 0$ a.s. hold and let y_t be generated by (3.1). Then $\lim_{T\to\infty} T^{-1} u' P_{M_p} u = 0$ a.s. for $p \geq p_0$ holds, where $u = (u_{p+1},\dots,u_T)'$.

Proof: The case $p = p_0 = 0$ is trivial. For $p \geq 1$ it is shown in the proof of Theorem 1 in Lai and Wei (1983) that there is a matrix A_T which is eventually a.s. nonsingular such that $A_T Z_p' Z_p A_T'$ converges to an a.s. nonsingular matrix. Furthermore, it is shown that $A_T Z_p' u = o(T^{1/2})$ (note that the cases $r = 0$ or $s = 0$ in the notation of Lai and Wei (1983) are covered by their arguments). The lemma follows now from $u' P_{M_p} u = u' Z_p A_T'(A_T Z_p' Z_p A_T')^{-1} A_T Z_p' u$.

Lemma A.2:

Let A,B be real matrices with countable infinitely many rows and k_1 , respectively k_2 , many columns. For $T \geq 1$ let A_T, B_T denote the submatrices consisting of the first T rows. If the columnspaces of A_T and B_T coincide for all $T \geq 1$, then $c_1 tr(A_T' A_T) \leq tr(B_T' B_T) \leq c_2 tr(A_T' A_T)$ holds for all $T \geq 1$ where c_1 and c_2 are positive real numbers.

Proof: From the assumptions we immediately see that A and B span the same space. Hence we can find matrices M and N such that $AM = B$ and $BN = A$ hold. But then $A_T M = B_T$ and $B_T N = A_T$. We arrive at $tr(B_T' B_T) = tr(M' A_T' A_T M) \leq \lambda_{max}(MM')\, tr(A_T' A_T) = c_2\, tr(A_T' A_T)$ and $tr(A_T' A_T) \leq \lambda_{max}(NN')\, tr(B_T' B_T) = c_1^{-1}\, tr(B_T' B_T)$. Clearly $c_1 > 0$ and $c_2 > 0$, if not $A = 0$ and $B = 0$. If $A = 0$ and $B = 0$, then the result trivially holds.

References:

Akaike, H. (1969): Fitting autoregressive models for prediction. Ann. Inst. Stat. Math. 21, 243-247.

Amemiya, T. (1980): Selection of regressors. Int. Econ. Rev. 21, 331-354.

An, H.Z., Chen, Z.G. and Hannan, E.J. (1982): Autocorrelation, autoregression and autoregressive approximation. Ann. Statist. 10, 926-936.

An. H. and Gu, L. (1985): On the selection of regression variables. Acta Mathematicae Applicatae Sinica 2, 27-36.

Chow, Y.S. (1965): Local convergence of martingales and the law of large numbers. Ann. Math. Statist. 36, 552-558.

Geweke, J. and Meese, R. (1981): Estimating regression models of finite but unknown order. Int. Econ. Rev. 22, 55-70.

Gu, L. and An, H. (1985): Statistical analysis of subset AR models. Acta Mathematicae Applicatae Sinica 8, 433-445 (in Chinese).

Hannan, E.J. (1980): The estimation of the order of an ARMA process. Ann. Statist. 8, 1071-1081.

Hannan, E.J. (1981): Estimating the dimension of a linear system. J. Multivariate Anal. 11, 458-473.

Hannan, E.J. and Quinn, B.G. (1979): The determination of the order of an autoregression. J. Roy. Statist. Soc. Ser. B 41, 190-195.

Lai, T.L. and Wei, C.Z. (1982a): Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. Ann. Statist. 10, 154-166.

Lai, T.L. and Wei, C.Z. (1982b): Asymptotic properties of projections with applications to stochastic regression problems. J. Multivariate Anal. 12, 346-370.

Lai, T.L. and Wei, C.Z. (1983): Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters. J. Multivariate Anal. 13, 1-23.

Lai, T.L. and Wei, C.Z. (1985): Asymptotic properties of multivariate weighted sums with application to stochastic regression in linear dynamic systems. In Multivariate Analysis VI, P.R. Krishnaiah (ed.), Elsevier Science Publishers B.V., 375-393.

Mallows, C.L. (1973): Some comments on C_p . Technometrics 15, 661-675.

Paulsen, J. (1984): Order determination of multivariate autoregressive time series with unit roots. J. Time Series Anal. 5, 115-127.

Paulsen, J. and Tjøstheim, D. (1985): Least squares estimates and order determination procedures for autoregressive processes with a time dependent variance. J. Time Series Anal. 6, 117-133.

Pötscher, B.M. (1983): Order estimation in ARMA-models by Lagrangian multiplier tests. Ann. Statist. 11, 872-885.

Pötscher, B.M. (1985): The behaviour of the Lagrangian multiplier test in testing the orders of an ARMA-model. Metrika 32, 129-150.

Quinn, B.G. (1980): Order determination for multivariate autoregression. J. Roy. Statist. Soc. Ser. B 42, 182-185.

Schwarz, G. (1978): Estimating the dimension of a model. Ann. Statist. 6, 461-464.

Thompson, M.L. (1978): Selection of variables in multiple regression I, II. Int. Statist. Rev. 46, 1-20 and 129-146.

Tsay, R.S. (1984): Order selection in nonstationary autoregressive models. Ann. Statist. 12, 1425-1433.

Stout, W.F. (1970): The Hartman-Wintner law of the iterated logarithm for martingales. Ann. Math. Statist. 41, 2158-2160.

Wang, S. and An, H. (1984): Consistency of selection of variables in stochastic regression. J. of Engineering Mathematics 1, 13-22 (in Chinese).

INTERVENTION ANALYSIS OF CONSUMER EXPENDITURE IN AUSTRIA

Gerhard Thury

Austrian Institute for Economic Research

Vienna

December 1987

INTERVENTION ANALYSIS OF CONSUMER EXPENDITURE IN AUSTRIA

Gerhard Thury

1. Introduction

Intervention analysis was first introduced to the profession in an article by BOX and TIAO (1975). This model type constitutes an extension of the class of ARIMA models. In applied work, it was namely soon discovered that pure ARIMA models often are a too restrictive approach for adequately modelling economic time series. These series are often heavily affected by policy interventions. Ignoring the effects of such measures means to invite difficulties. Substantial improvements of estimated ARIMA models can be achieved by modelling explicitly the effects of known policy interventions.

First attempts to apply intervention analysis to model Austrian consumer expenditure can be found in THURY (1982) and BRANDNER (1986). Consumer expenditure offers itself definitely for this type of analysis, because it has been affected strongly by different fiscal policy measures during the last twenty years. Our primary motif for replicating this analysis lies in the fact that the above cited studies have serious short-comings. Above all, it was attempted in both studies to estimate the effects of fiscal policy measures, which affect only specific consumption categories, from analysis of a series for total consumer expenditure. We suspect that serious distortions of the estimated effects might be the consequence of this way of proceeding. We broaden therefore, in the present study, the scope of the investigation. Besides total consumer expenditure, we analyse here explicitly also its two components, namely expenditure on nondurables and purchases of durables. Additionally, the latter categoty is further divided into purchases of cars and of other durables. We believe that this detailed analysis should provide better estimates for the effects of the various fiscal policy measures, especially as far as their shape and their magnitude is concerned.

2. The Theoretical Background of Intervention Analysis

The general form of an intervention model is given by

$$Y_t = f(x_t;\delta,w) + N_t. \tag{1}$$

Thus, an observed time series, Y_t, is thought to consist of two components. The function $f(\cdot)$ is meant to capture the effects of a known intervention. Here, x_{it} denotes a vector of exogenous intervention indicators, i.e. variables assuming the values 0 or 1. δ and w are the parameter vectors of the accompanying rational polynomial $w(B)/\delta(B)$, which is usually chosen to model the effects of an intervention x_{it}. The disturbance, N_t, represents the portion of the series which does not follow a purely deterministic model. It is modelled by an ARIMA process, i.e.

$$\varphi(B)N_t = t(B)a_t, \quad (2)$$

where $\varphi(B) = (1-B)^d Ø(B) = (1-B)^d(1-Ø_1B- \dots -Ø_pB^p)$ and $t(B) = (1-t_1B- \dots -t_qB^q)$ with all the roots of $Ø(B)$ and $t(B)$ lying outside the unit circle. For seasonal time series, multiplicative seasonal operators are introduced additionally on both sides of eq. (2).

2.1 Model Identification

2.1.1 The Process Noise

BOX and JENKINS (1970) suggest to employ the autocorrelation and the partial autocorrelation function for the identification of an ARIMA model. In the practical application of this suggestion several problems may occur. First, most economic time series are nonstationary, thus requiring some form of differencing. Unforunately, the exact order of this differencing is often hard to determine. The danger of overdifferencing is always present. Second, the ACF of a pure AR or MA process is easy to interpret. The ACF of a mixed process, however, often exhibits no clear-cut pattern, making model identification hazardeous.

Concentrated research efforts in recent years have succeeded in reducing these problems substantially. Especially, the development of the extended autocorrelation function was a great step forward. The EACF was introduced by TSAY and TIAO (1982,1983). The basic idea behind the EACF approach is as follows. Consistent estimates of the autoregressive parameters in a stationary or nonstationary ARMA model can be obtained by means of iterated regressions. Based on these estimates, the observed series can then be transformed into a process, which asymptotically follows a pure MA model. The order of this MA model can then be determined very easily from an inspection of the ACF. Thus, one of the main advantages of the EACF approach is that it can handle nonstationary time series directly without requiring differencing in advance.

Apart from this methodological problems, it will be generally difficult to identify the model for N_t adequately during the first stage of the iterative model building procedure, N_t represents a unobservable residual series from eq. (1). Therefore, N_t is often tentatively modelled based on observations prior to the occurence of exogenous effects. This tentatively identified model is then modified after these exogenous effects have been identified and estimated.

2.1.2 The Intervention Model

The exogenous portion of eq. (1), $f(\cdot)$, is generally postulated using knowledge of the system under study, and subquently tested and possibly modified during the estimation-diagnostic checking stage of a model building process. Frequently, the effects of a set of exogenous variables, x_{it}, can be represented by a dynamic model of the form

$$f(x_t;\delta,w) = \sum_{j=1}^{k} [w_j(B)/\delta_j(B)]\, x_{jt}. \quad (3)$$

The polynomials $w_j(B) = (w_{0j}+w_{1j}B+ \ldots +w_{r_j j}B^{r_j})$ and $\delta_j(B) = 1-\delta_{1j}B- \ldots -\delta_{s_j j}B^{s_j}$ are of degrees r_j and s_j, respectively. We assume that $w_j(B)$ has roots outside, and $\delta_j(B)$, outside or on, the unit circle.

2.2 Model Estimation and Diagnostic Checking

Once an intervention model is tentatively identified, the parameters can be estimated by maximizing the corresponding likelihood function. This function may be reasonably approximated by a conditional likelihood function due to BOX and JENKINS (1970). The SCA system (see LIU and HUDAK, 1985), which is used for estimation purposes in this study, additionally offers an approximation due to HILLMER and TIAO (1979). With n observations $Y_1, \ldots, Y_n$, where Y_t is assumed to be stationary, both approaches compute the likelihood function on the basis of the stochastic structure of n-p observations,

$$Y_t = \sum_{h=1}^{k} \left(\frac{w_h(B)}{\delta_h(B)}\right) x_{th} + \sum_{i=1}^{p} Ø_i Y_{t-i} - \sum_{j=1}^{q} t_j a_{t-j} + a_t, \quad t=p+1,\ldots,n \quad (4)$$

where $Y_1, \ldots, Y_p$ are regarded as fixed. The two methods differ in that the "conditional" likelihood approach assumes $a_p = \ldots = a_{p-q+1} = 0$, while the "exact" likelihood approach computes estimates for these values. A GAUSS-MARQUARDT nonlinear least squares algorithm is then used in both approaches to derive parameter estimates. In practice, one usually employs the conditional approach in the intial phases of the iterative modelling process and switches to the exact likelihood function in order to obtain parameter estimates of the final model version.

Once the parameters are estimated, various diagnostic checks should be performed on the estimated residuals

$$a_t = Y_t - \sum_{h=1}^{k} \left(\frac{w_h(B)}{\delta_h(B)}\right) x_{th} - \sum_{i=1}^{p} Ø_i Y_{t-i} + \sum_{j=1}^{q} t_j a_{t-j}, \quad t=p+1,\ldots,n \quad (5)$$

to determine the adequacy of the fitted model. Useful methods to aid diagnostic checking include plotting the residual series and studying the sample ACF of the residual series in order to see whether it is consonant with that of a white noise process.

2. Empirical Results

Consumer expenditure in Austria was heavily affected by a variety of fiscal policy measures during the last 20 years. A list of the most important of these measures is provided in table 1.

Table 1
Fiscal Policy Measures Affecting Consumer Expenditure

We see immèdiately that most of these fiscal policy measures affected only the purchases of durables, and here especially those of cars. But although these categories constitute only relatively small fractions of total consumer expenditure (in 1985, purchases of durables amounted to roughly 12 percent and purchases of cars to less than 4 percent of total consumer expenditure in constant prices), the listed fiscal policy measures distorted even the growth pattern of total consumer expenditure substantially on some occasions. For the specific consumption categories themselves, the observed effects were of a considerable order of magnitude.

Before we can start our analysis, we have to develop ideas about the possible shape of the consumers' reactions to the above mentioned fiscal policy measures. In the present context, the solution of this problem is relatively simple. In all cases, these measures were changes of sales tax system (either introducing a new tax or raising the rates of an existing one), which were known several months in advance. Thus, we would expect to observe increased purchases immediately before the introduction of a particular policy measure, and a corresponding restraint from buying after the introduction. Figure 1 shows a graph of this reaction pattern.

Figure 1
Suggested Reaction of Consumer Expenditure to Fiscal Policy Measures

In an algebraic formulation this can be written as

$$\left[w_0 + \frac{w_1 B}{1-\delta B} + \frac{w_2 B^2}{1-B}\right] P_t^{(T)},$$

where

$$P_t^{(T)} = \begin{cases} 0 & t \neq T \\ 1 & t = T. \end{cases}$$

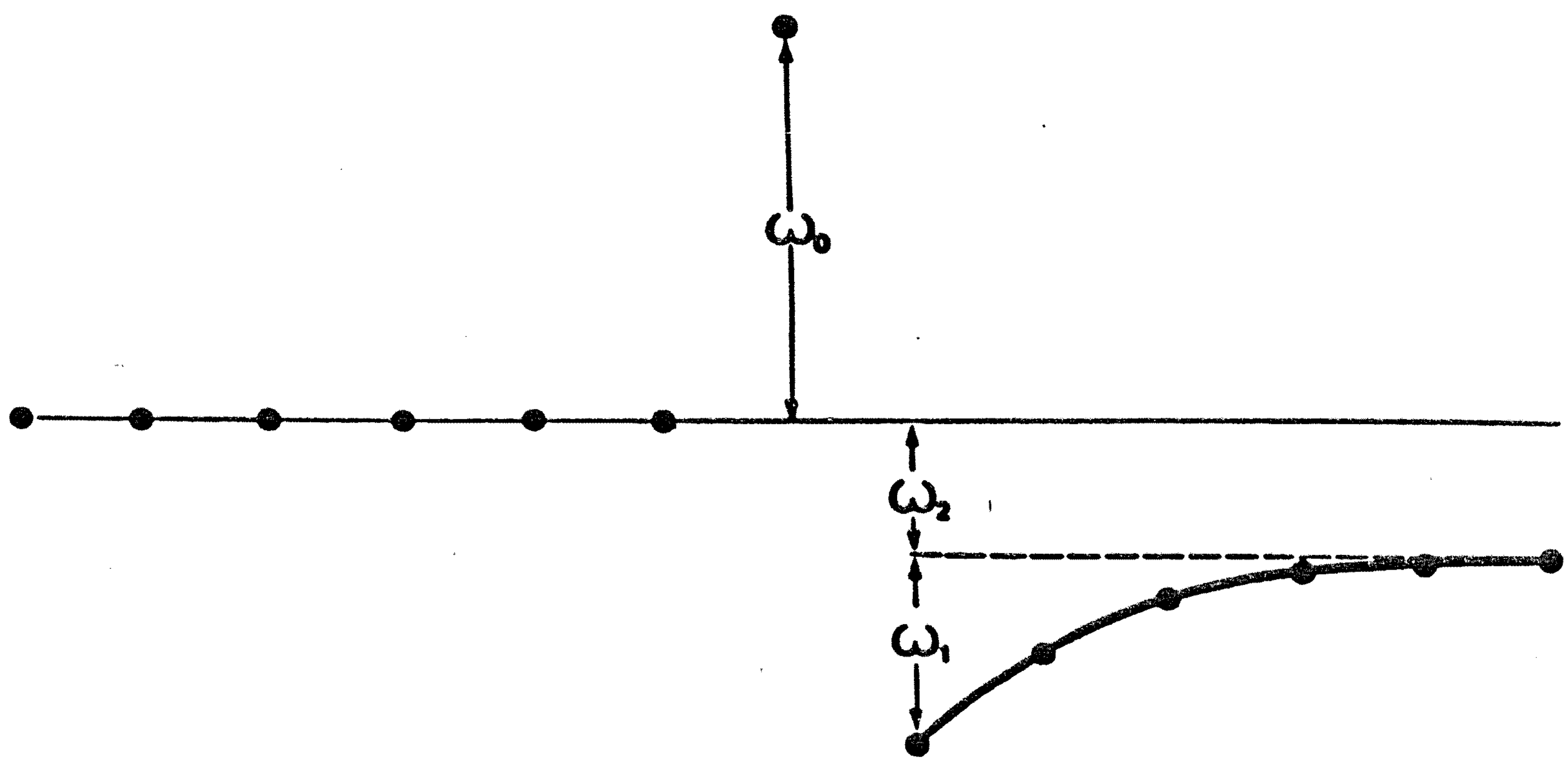
The parameter w_0 is expected to be positive, measuring the amount of advance purchasing which is induced by the impending tax increase. The decline in consumer expenditure immediately after the policy intervention will be measured by the sum w_1, w_2, with $w_1, w_2 < 0$, and the permanent effect will be given by w_2. This

Table 1
List of Fiscal Policy Measures Affecting Consumer Expenditure

Date	Policy Measure	Intervention	Expected Effects on Purchases of: Nondurables	Cars	Other Durables
Sep 1,1968	Introduction of surtax on purchases of new cars	x_{1t}	-	↑in 1968:3 ↓from 1968:4 on	↓in 1968:3
Dec 31,1970	Abolition of surtax	x_{2t}	-	↑from 1971:1 on	-
Jan 1,1973	Introduction of VAT	x_{3t}	↑in 1972:4 ↓from 1973:1 on	↑in 1972:4 ↓from 1973:1 on	↑in 1972:4 ↓from 1973:1 on
Jan 1,1975	Reduction in tariffs vis-a-vis Common Market	x_{4t}	-	↑in 1975:1	↑in 1975:1
Jan 1,1978	Introduction of special VAT rate on 'luxury goods'	x_{5t}	-	↑in 1977:4 ↓from 1978:1 on	↑in 1977:4 ↓from 1978:1 on
Jan 1,1984	Increase in VAT	x_{6t}	↑in 1983:4 ↓from 1984:1 on	↑in 1983:4 ↓from 1984:1 on	↑in 1983:4 ↓from 1984:1 on
Oct 1,1985	Cathalytic Converter Regulation	x_{7t}	-	↓in 1985:3 ↑from 1985:4 on	-

↑ symbolizes an increase in purchases
↓ denotes a reduction in purchases

FIGURE 1
SUGGESTED REACTION OF CONSUMER
EXPENDITURE TO FISCAL POLICY MEASURES



interpretation follows from the identity

$$S_t^{(T)} = \frac{1}{(1-B)} P_t^{(T)},$$

where $S^{(T)}$ denotes a step function, i.e.

$$S_t^{(T)} = \begin{cases} 0 & t < T \\ 1 & t \geq T. \end{cases}$$

3.1 Identification and Estimation of Intervention Models

Having settled the question of the possible shape of the reaction pattern, we can now turn to the identification and estimation of intervention models for the different categories of consumer expenditure. We start with total consumer expenditure, because the aggregate model can serve as helpful baseline, with which the results for the different consumption categories might be compared.

3.1.1 Total Consumer Expenditure

We analyse in this study the official consumption series in constant prices from the WIFO databank, covering the period 1954:1 to 1986:2. As first step, we have to identify an ARMA model for the noise process N_t. Looking at the original series of total consumer expenditure in constant prices, we note that the variance of this series seems to increase with its level. Therefore, we take the natural logarithm of this series in order to stabilize its variance. Investigating the ACF of the transformed series reveals that it is highly nonstationary. If we intended to employ the ACF in the identification process, some form of differencing would be necessary. By employing the above mentioned extended ACF, we can avoid this somewhat arbitrary choice of difference operators. The EACF for total consumption is given in table 2.

Table 2
The Extended ACF Table for Total Consumption

Theoretically, only observations prior to the occurence of the first intervention should be used for the identification of the noise model. But, since one of the most important interventions occurred already at the end of 1968, this way of proceeding would reduce the number of observations drastically. We calculated the EACF tables for both sample sizes. Fortunately, it turned out that they are not very different from one another. Apparently, what is gained by omitting the distorted observations, is lost again by the drastic reduction in the degrees of freedom. Therefore, similarly as BRANDNER (1986), we use all available observations

Table 2
The Extended ACF Table for Total Consumer Expenditure

Q / P	0	1	2	3	4	5	6	7	8
0	.90	.87	.85	.91	.81	.78	.77	.82	.73
1	-.45	-.09	-.37	.95	-.43	-.08	-.36	.92	-.42
2	-.53	.06	-.16	.95	-.50	.05	-.15	.93	-.49
3	-.37	-.25	-.35	.95	-.34	-.24	-.34	.92	-.33
4	.27	.15	.15	-.17	-.04	-.08	-.11	.09	.03
5	-.43	-.05	.28	-.51	.22	.04	-.12	.04	.08
6	-.49	-.08	.06	-.47	.26	.09	-.12	.04	.06
7	.47	-.17	-.18	-.45	-.06	-.08	.08	-.05	.02
8	.40	-.24	-.46	-.43	.24	-.05	.00	.00	.01

for the identification of an ARMA model for N_t.

The first line in an EACF table is the well known ACF, the second line contains the values of the first EACF, and so on. This table provides useful information on the maximum orders of p and q in an ARMA model. The order of p and q can be determined tentatively by finding a position (p_0,q_0) in the table, so that all values are insignificant in the triangular region where $i=p_0+k$ and $j>q_0+k$, $k=0,1,2,...$ This region is indicated in table 2. The vertex of the triangle of insignificant values in the EACF table points to an ARMA(4,4) model. The significant value of the 5-th EACF in the 6-th row of table 2 provides some justification to consider an ARMA(5,5) model elsewell. After the order (p,q) of an ARMA model is tentatively specified, TSAY and TIAO (1983) recommend that one should check the least squares estimates of the AR parameters in the q-th iterated AR(p) regression for supplementary information on the structure of AR portion of the model. Additionally, these

Table 3
Least Squares Estimates of Iterated AR(5) Regressions
for Total Consumption

AR estimates can be used to transform the original process into a pure MA series, the ACF of which is much simpler to interpret than that of a mixed process. Table 3 shows the LS estimates of the iterated AR(5) regressions. This table contains the results for several iterations, namely for j=0,1, ... ,9. These etsimates are reasonably stable for j>5, and suggest that the AR polynomial is of the form $(1-\emptyset B)(1-\Phi B^4)$ with both $\emptyset$ and Φ being close to 1. This follows from line 6 of table 3, i.e.

$$(1 - 0.99B - 0.01B^2 - 0.99B^4 + 0.99B^5) \approx (1-B) - B^4(1-B) = (1-B)(1-B^4).$$

Thus, the AR estimates of the iterated regressions provide justification for the use of regular and seasonal differences. We see that the EACF table of the undifferenced data can directly identify the order (p,q) of a seasonal ARMA model. Additionally, the LS estimates of the iterated AR regressions even spot the multiplicative form of the AR polynomial and thus help to resolve the ambiguity of differencing. Now, the differenced series $(1-B)(1-B^4)\ln\text{CONSUMPTION})_t$ is a pure MA process, the model of which can be identified from an inspection of the ACF. A graph of

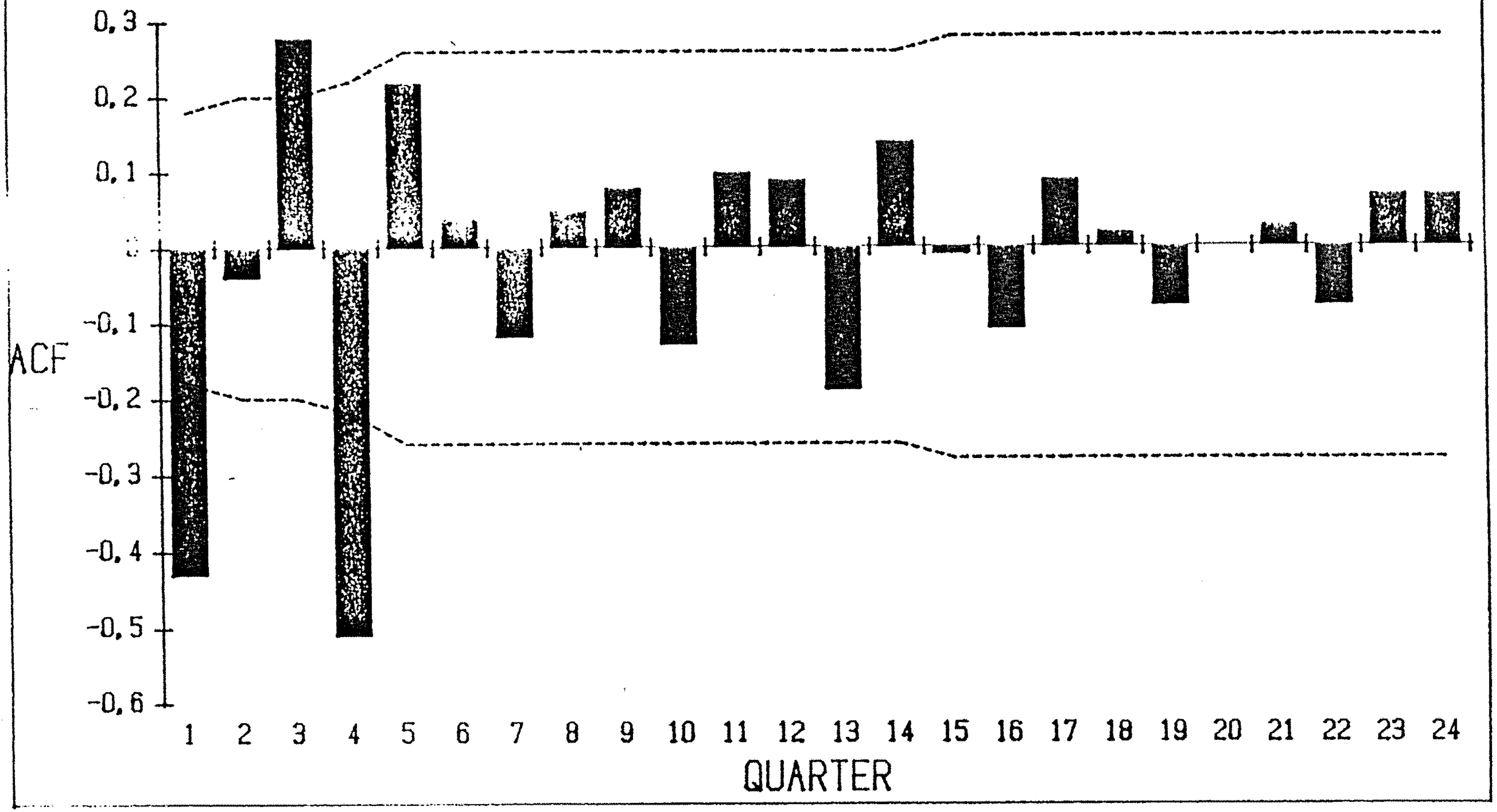
Figure 2
Autocorrelation Function of $(1-B)(1-B^4)\ln(\text{CONSUMPTION})_t$

this ACF together with its 95 percent confidence intervals is shown in figure 2. Already visual inspection reveals that this series is not white noise. The LJUNG-BOX Q statistics for the first 24 sample autocorrelations turns out as Q(24) = 97.7, which must be compared to a 5 percent critical value of 36.4 in a chi-square table. Thus, the null hypothesis of no autocorrelation

Table 3
Least Squares Estimates of Iterated AR(5) Regressions for Total Consumer Expenditure

j \ p	1	2	3	4	5
0	.78	.01	.01	.96	-.77
1	.95	.00	.00	.97	-.92
2	.97	.00	.00	.97	-.94
3	.82	-.00	-.01	.98	-.78
4	1.12	-.01	.00	1.00	-1.10
5	.99	.01	.00	.99	-.99
6	.97	.01	.00	.99	-.96
7	1.03	.01	.00	.99	-1.03
8	1.03	.01	.00	.99	-1.02
9	.95	.00	.00	.98	-.94

FIGURE 2
AUTOCORRELATION FUNCTION FOR
(1-B) (1-B**4) ln (CONSUMPTION)



must be rejected. The shape of the ACF suggests a model of the following form for total consumer expenditure in constant prices:

$$(1-B)(1-B^4)\ln(\text{CONSUMPTION})_t = (1-t_1B)(1-t_4B^4)a_t, \qquad (6)$$

where the sequence (a_t) is iid with mean 0 and variance σ^2.

In the estimation and diagnostic checking stage of eq. (6), it turned out that a model, in which the regular autoregressive operator $(1-\emptyset B)$ was not replaced beforehand by the difference operator $(1-B)$ but estimated from the data instead, performed slightly better. This result is in line with the findings of ANDERSON and GOIJER (1980). Parameter estimates for eq. (6), which are derived by a GAUSS-MARQUARDT nonlinear least squares algorithm employing the 'exact' likelihood approach of HILLMER and TIAO (1979), are given in column 2 of table 4. At a first glance, this

Table 4
Parameter Estimates for an ARIMA and an Intervention Model of Total Consumer Expenditure

relatively simple ARIMA model for total consumer expenditure looks quite adequate. The residual ACF for this model is given as figure 3. None of the first 24 sample autocorrelations is

Figure 3
Residual Autocorrelation Function for Consumption (ARIMA)

significant. The LJUNG-BOX Q statistic (5 percent critical value for 21 degrees of freedom is 32.7) seems to corroborate this visual impression. A detailed analysis of this residual series, however, reveals that errors of a substantial order of magnitude occur at the dates of the different fiscal policy measures, which are listed in table 1. Additionally, we observe large errors in the first or second quarters in various years which, apparently, are caused by the date of Easter. From this evidence, we must conclude that a pure ARIMA model provides no adequate description of total consumer expenditure.

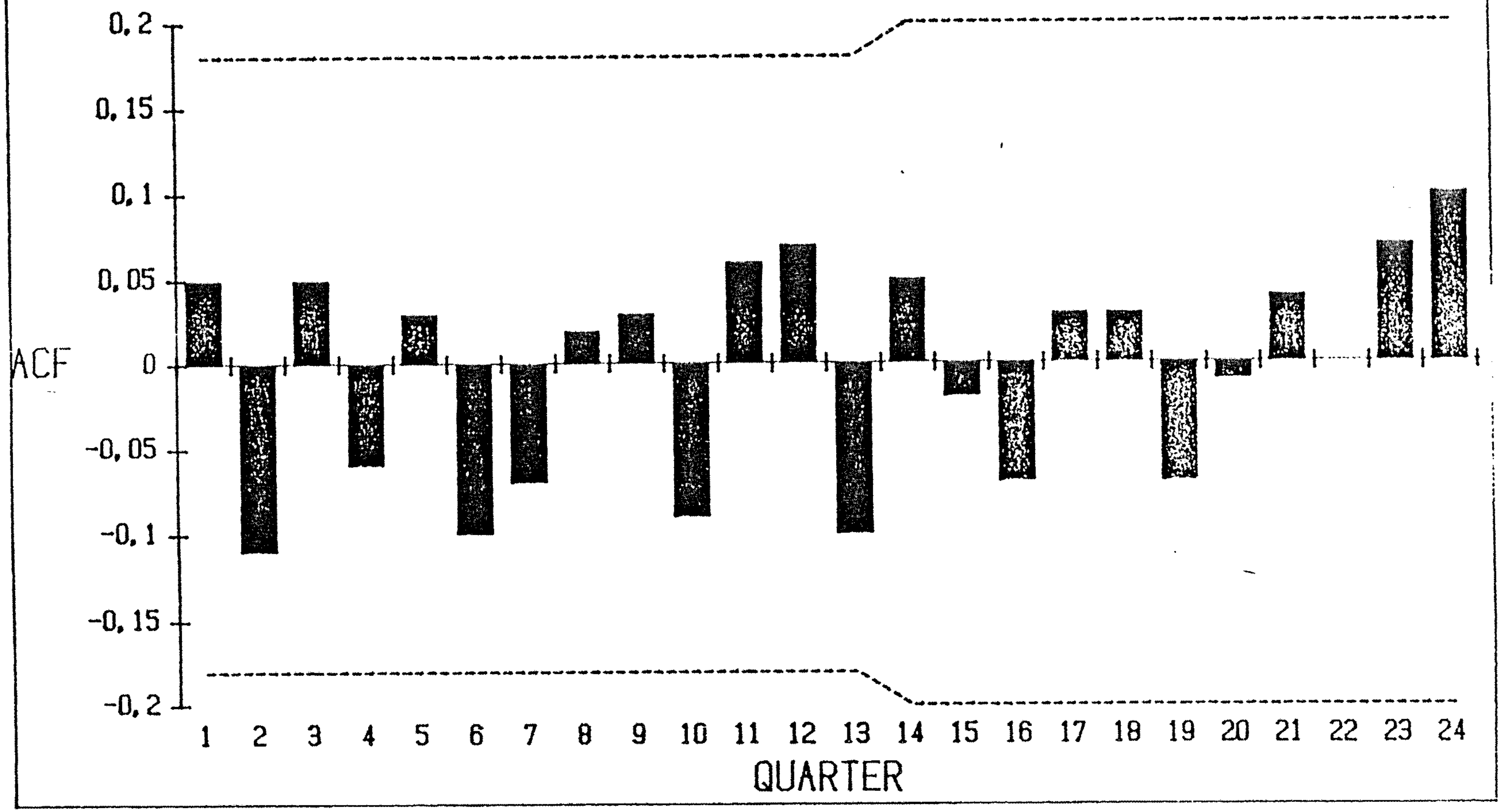
We try now to correct the deficiency of the pure ARIMA model by introducing relevant exogenous variables. In order to capture the effects of fiscal policy measures, we use the intervention variables of table 1. An idea about the possible shape of these reactions is depicted in figure 1. For modelling the effect of calendar variations, i.e. the effect of the changing date of Easter in our case, we follow an approach suggested in LIU (1983). It is well known that Easter leads to increased expenditure on certain categories of nondurables. A model for a holiday effect can be written as

$$f(H_{1t};\alpha_1) = \alpha_1 H_{1t}, \qquad (7)$$

Table 4
Parameter Estimates for an ARIMA and an Intervention Model
of Total Consumer Expenditure

Parameter	ARIMA Model	Intervention Model
α		.2375 (.0339)
w_{10}		.0344 (.0110)
w_{11}		-.0235 (.0110)
w_{50}		.0589 (.0122)
w_{51}		-.0392 (.0119)
δ_3		.7266 (.1917)
w_{60}		.0440 (.0119)
Ø	.9834 (.0080)	.9752 (.0119)
t_1	.6041 (.0715)	.4722 (.0809)
t_4	.6101 (.0713)	.3396 (.0864)
R^2	.9970	.9980
see	.0210	.0151
DW	1.9500	1.9500
Q(21)	13.4000	23.6000

FIGURE 3
RESIDUAL AUTOCORRELATION FUNCTION
FOR CONSUMPTION (ARIMA)



if the effect due to a specific holiday is constant over the years. Here, the artificial variable H_{1t} represents the proportion of the holiday effect falling into the t-th month. If the holiday effect increases (decreases) linearly over the years, then the following model might capture the effect more adequately:

$$f(H_{1t},H_{2t};\alpha_1,\alpha_2) = \alpha_1 H_{1t} + \alpha_2 H_{2t}, \qquad (8)$$

where $H_{2t} = H_{1t}*T$, with T denoting a linear time trend. The earliest and latest dates on which Easter can fall are March 22 and April 25, respectively. Thus, for quarterly data, Easter can lie either in the first or second quarter of a year. The dates of Easter and the proportions of the Easter effect falling into the first or second quarter are given in table 5. In the construction

Table 5
Dates of Easter and H_{1t} Proportions: 1954 - 1986

of this table it is assumed that the extra consumer expenditure due to Easter is distributed uniformly during a ten day period beginning on Saturday before Palm Sunday and ending on Easter Monday.

Parameter estimates of this extended model version for total consumer expenditure are given in column 3 of table 4. We see that the adequate modelling of fiscal policy measures and calendar effects results in an almost 30 percent reduction of the standard error of estimate. A comparison of the estimates for the noise model parameters Ø, t_1, t_4 in the two models is also very illuminating. The autoregressive parameter p is hardly different in the two models. The estimates of the moving average parameters, however, deviate substantially. Especially, the estimates of the seasonal moving average parameter t_4 bear no resemblance (0.61 for the pure ARIMA model and 0.34 for the intervention model). Ignoring the effects of exogenous variables apparently results in heavily upward biased estimates for the seasonal moving average parameters. This will have extremely negative consequences when such a model is used as input for a model-based seasonal adjustment procedure. The ACF of the residual series is shown in figure 4. At first sight, it is somewhat surprising that the

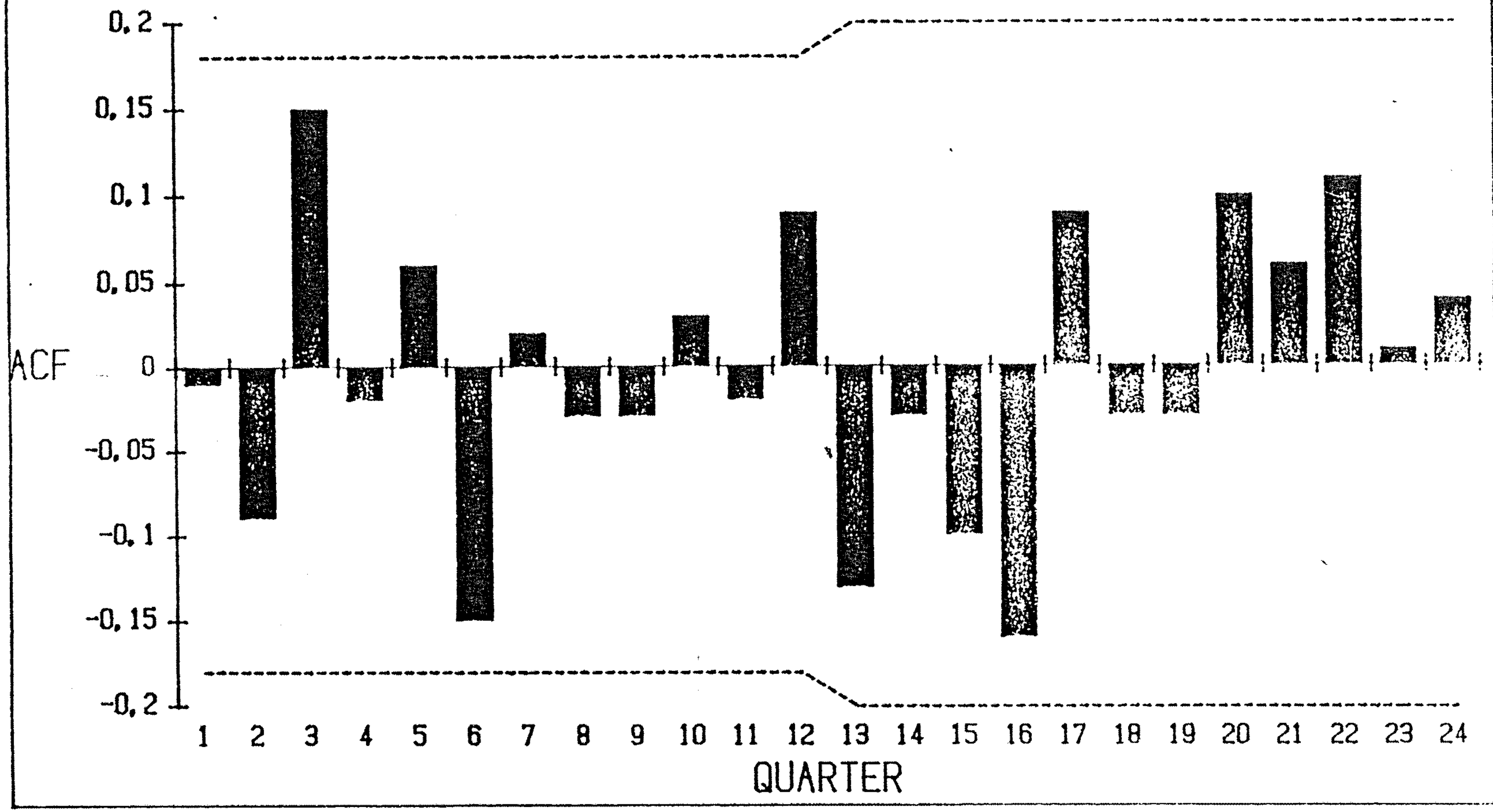
Figure 4
Residual Autocorrelation Function for Consumption (INTERVENTION)

residuals from the intervention model exhibit higher autocorrelation than those from the pure ARIMA model. The higher, although still not significant value of the LJUNG-BOX Q statistic seems also to indicate this. Some reflection, however, reveals that this fact is simply a consequence of the large number of outliers, which are present in the residual series of the pure ARIMA model. A series, which is heavily dominated by erratic outliers, will not exhibit a high degree of autocorrelation. This

Table 5
Dates of Easter and H_{1t} Proportions: 1954 - 1986

Year	Date of Easter	Proportion of Easter Effect (H_{1t}) in Quarter 1	Quarter 2
1954	April 18 and 19	0.0	1.0
1955	April 10 and 11	0.0	1.0
1956	April 1 and 2	0.8	0.2
1957	April 21 and 22	0.0	1.0
1958	April 6 and 7	0.3	0.7
1959	March 29 and 30	1.0	0.0
1960	April 17 and 18	0.0	1.0
1961	April 2 and 3	0.7	0.3
1962	April 22 and 23	0.0	1.0
1963	April 14 and 15	0.0	1.0
1964	March 29 and 30	1.0	0.0
1965	April 18 and 19	0.0	1.0
1966	April 10 and 11	0.0	1.0
1967	March 26 and 27	1.0	0.0
1968	April 14 and 15	0.0	1.0
1969	April 6 and 7	0.3	0.7
1970	March 29 and 30	1.0	0.0
1971	April 11 and 12	0.0	1.0
1972	April 2 and 3	0.7	0.3
1973	April 22 and 23	0.0	1.0
1974	April 14 and 15	0.0	1.0
1975	March 30 and 31	1.0	0.0
1976	April 18 and 19	0.0	1.0
1977	April 10 and 11	0.0	1.0
1978	March 26 and 27	1.0	0.0
1979	April 15 and 16	0.0	1.0
1980	April 6 and 7	0.3	0.7
1981	April 19 and 20	0.0	1.0
1982	April 11 and 12	0.0	1.0
1983	April 3 and 4	0.6	0.4
1984	April 22 and 23	0.0	1.0
1985	April 7 and 8	0.2	0.8
1986	March 30 and 31	1.0	0.0

FIGURE 4
RESIDUAL AUTOCORRELATION FUNCTION
FOR CONSUMPTION (INTERVENTION)



finding has serious implications for diagnostic checking. If one relies in the diagnostic checking stage exclusively on an inspection of the ACF or on portmanteau statistics such as the LJUNG-BOX Q, a completely inadequate model might be accepted in situations, where the residual series is dominated by the existence of erratic outliers. A detailed analysis of the residual series itself will help to avoid this danger.

Let us now turn to a closer examination of the deterministic part of our model. We see that not all of the fiscal policy measures listed in table 1 give, in fact, rise to significant parameter estimates. Such estimates were obtained for only three fiscal policy measures, namely

(i) the introduction of a surtax on purchases of new cars in September 1968;
(ii) the introduction of a special VAT rate on 'luxury goods' in January 1978;
(iii) the increase in VAT in January 1984.

Only in case (ii), we found an effect similar in shape to that depicted in figure 1. Contrary to an earlier study (see THURY, 1982), we could not establish a permanent negative effect for this policy measure. This then found permanent negative effect seems to have been an artefact caused by the relatively small number of observations after the introduction of the luxury tax in that earlier study. In case (i), we spot advance purchasing in the quarter immediately before the introduction of the surtax, and a corresponding decline in the quarter following the introduction. In case (iii), finally, we find only evidence that the impending increase of the VAT rates induced consumers to shift expenditure in a moderate degree to the quarter before the introduction of the tax change. A corresponding decline in consumer expenditure after the tax change was not to establish. All in all, we would say nevertheless that the obtained results come up to expectations. That we find only a limited number of significant interventions and that only one of these follows the expected reaction pattern is a consequence of the fact that all fiscal policy measures affected only specific components of consumer expenditure, which constitute only small fractions of total consumption.

The effect of the changing date of Easter is measured by the coefficient a in table 4. We could work either with a constant or a time-varying holiday model. A combination of the two as given in eq. (8) was rejected by the data. The time-varying formulation proved superior. The variable T is defined as decreasing linear trend, so that we observe an Easter effect which loses in importance over the years.

Table 6, finally, provides information on the forecasting behaviour of the pure ARIMA model and our preferred intervention model. Predictions derived by HOLT-WINTERS exponential smoothing, which is known to perform relatively good in macroeconomic forecasting also (see THURY, 1986) are included in this table as a standard of reference. The usual forecast accuracy measures for the log-changes of 1-step and 5-step ahead predictions of total consumer expenditure for the period 1977:4 to 1986:2 are given.

Table 6
Forecast Accuracy Measures
Total Consumer Expenditure

We note that, for 1-step ahead forecasts, the intervention model clearly outperforms the pure ARIMA model and HOLT-WINTERS exponential smoothing. For 5-step ahead forecasts, the observed differences between the different models are rather small. This is simply the consequence of the fact that we decided to ignore the intervention portion of the model when making 5-step ahead forecasts. Since we intend to simulate the practical forecasting situation as closely as possible, it would be illegitimate to include intervention variables for fiscal policy measures, which were not even known to be adopted at the time when these forecasts had to be made. It seems even doubtful whether it is really legitimate to do this for 1-step ahead forecasts. This depends on the type of the impending intervention. If a similar intervention was encountered in the past already, the existing parameter estimates could be used to evaluate tentatively the consequences of an impending re-use. Intervention models will be of no help for forecasting, however, when one is facing a completely new situation.

All in all, we can say that our estimated intervention model provides valuable information on the effects of major past fiscal policy measures. Whether this information can be used to evaluate the effects of impending future measures depends on the character of these measures. A similar intervention model for total consumer expenditure is presented in table 2 of BRANDNER (1986). The differences in explanatory power are small. Our slightly better explanatory power might be due to the more detailed modelling of the intervention part. Substantial differences can be found, however, between the ARIMA parts of the two models. The ARIMA part of our model follows a relatively simple multiplicative seasonal model, the properties of which are well established in the relevant literature (see, for example, BOX and JENKINS, 1970), while BRANDNER's noise model is of a unconventional, highly complex type.

3.1.2 Expenditure on Nondurables

Next, we turn to an analysis of the expenditure on nondurables, which constitute the by far most important component of total consumer expenditure (88 percent in 1985). The EACF table for expenditure on nondurables, calculated for the sample period 1954:1 to 1986:2 is shown as table 7. We see immediately that

Table 7
The Extended ACF Table for Expenditure on Nondurables

there exists hardly any difference between the tables 2 and 7. This is exactly what we would have expected taking into account the overwhelming share of expenditure of nondurables in total consumer expenditure. Thus, we believe that an ARMA(5,5) model

Table 6
Forecast Accuracy Measures
Total Consumer Expenditure

Forecast Accuracy Measure	ARIMA	INTERVENTION	HOLT-WINTERS
		1-step ahead	
Correlation coefficient	.9688	.9936	.9701
RMSE	.0333	.0172	.0334
MAE	.0245	.0145	.0251
Regression coefficient	.9025	.9241	.8919
THEIL's inequality coefficient	.0723	.0194	.0727
U^M	.0000	.0000	.0000
U^R	.1514	.3430	.1900
U^D	.8486	.6570	.8100
		5-step ahead	
Correlation coefficient	.9672	.9732	.9716
RMSE	.0350	.0294	.0342
MAE	.0266	.0210	.0266
Regression coefficient	.8861	.9432	.8717
THEIL's inequality coefficient	.0799	.0563	.0764
U^M	.0000	.0000	.0000
U^R	.1933	.0610	.2675
U^D	.8067	.9390	7325

Table 7
The Extended ACF Table for Expenditure on Nondurables

Q / P	0	1	2	3	4	5	6	7	8
0	.90	.86	.85	.91	.81	.78	.77	.83	.73
1	-.40	-.18	-.33	.95	-.38	-.18	-.32	.92	-.37
2	-.54	.08	-.15	.95	-.51	.08	-.15	.93	-.50
3	-.41	-.15	-.40	.95	-.39	-.14	-.38	.92	-.38
4	.19	.13	.19	-.15	-.03	-.11	-.12	.11	.05
5	-.48	-.04	.34	-.52	.26	-.02	-.10	.03	.05
6	-.53	-.07	.14	-.44	.23	-.04	-.12	.03	.08
7	.46	-.19	-.27	-.40	-.12	-.16	.02	-.02	.07
8	.45	-.29	-.29	-.36	.18	-.20	-.12	.02	.00

should again be fully adequate for modelling this series. The LS estimates of the iterated AR(5) regressions are given in table 8.

Table 8
Least Squares Estimates of Iterated AR(5) Regressions for Expenditure on Nondurables

These estimates are reasonably stable for $j \geq 5$, and point to the existence of multiplicative ARMA polynomial of the form $(1-\varnothing B)(1-\Phi B4)$ with both $\varnothing$ and Φ being again close to 1.

The ACF of $(1-B)(1-B^4)\ln(\text{NONDURABLES})_t$ is depicted in figure 5. It points to a model of the form

$$(1-B)(1-B^4)\ln(\text{NONDURABLES})_t = (1-t_1B)(1-t_4B^4)a_t. \qquad (9)$$

In the estimation stage, we discovered - what is hardly surprising, however - that a model with an AR operator $(1-\varnothing B)$ instead of the difference operator $(1-B)$ performs slightly better. Parameter estimates for this model are given in column 2 of

Figure 5
Autocorrelation Function for $(1-B)(1-B^4)\ln(\text{NONDURABLES})_t$

table 9. Inspection of the residual series reveals large errors in the first and second quarters of certain years. This fact is clear evidence for the presence of an Easter effect. Influences of fiscal policy measures, however, could not be spotted. When we

Table 9
Parameter Estimates for an ARIMA and an Intervention Model of Expenditure on Nondurables

started this analysis, we thought it possible that the introduction of VAT in January 1973 might have also influenced expenditure on nondurables. This hypothesis is not supported by the data, however. There is some indication for the existence of such an effect in the residuals, but it is obviously not strong enough to produce significant parameter estimates. This is not further surprising because the possibility for advance purchases of nondurables is rather limited by definition.

Parameter estimates for a model with an Easter effect included are given in the third column of table 9. The modelling of this effect results in a moderate reduction of the standard error of estimate, and a corresponding slight improvement of the forecasting performance. The residual ACF for this model is shown in figure 6. Neither this graph nor the LJUNG-BOX Q statistic point to any

Table 8
Least Squares Estimates of Iterated AR(5) Regressions
for Expenditure on Nondurables

j \ p	1	2	3	4	5
0	.80	.00	.01	.97	-.78
1	.98	-.01	.01	.97	-.95
2	1.00	-.01	.01	.97	-.97
3	.84	-.01	.00	.97	-.81
4	1.10	-.01	.00	.99	-1.08
5	.98	.01	.00	.99	-.97
6	.98	.01	.00	.99	-.98
7	1.03	.01	.00	.99	-1.02
8	1.03	.01	.00	.98	-1.02
9	1.03	.01	.00	.98	-.96

FIGURE 5
AUTOCORRELATION FUNCTION FOR
(1-B) (1-B**4) ln (NONDURABLES)

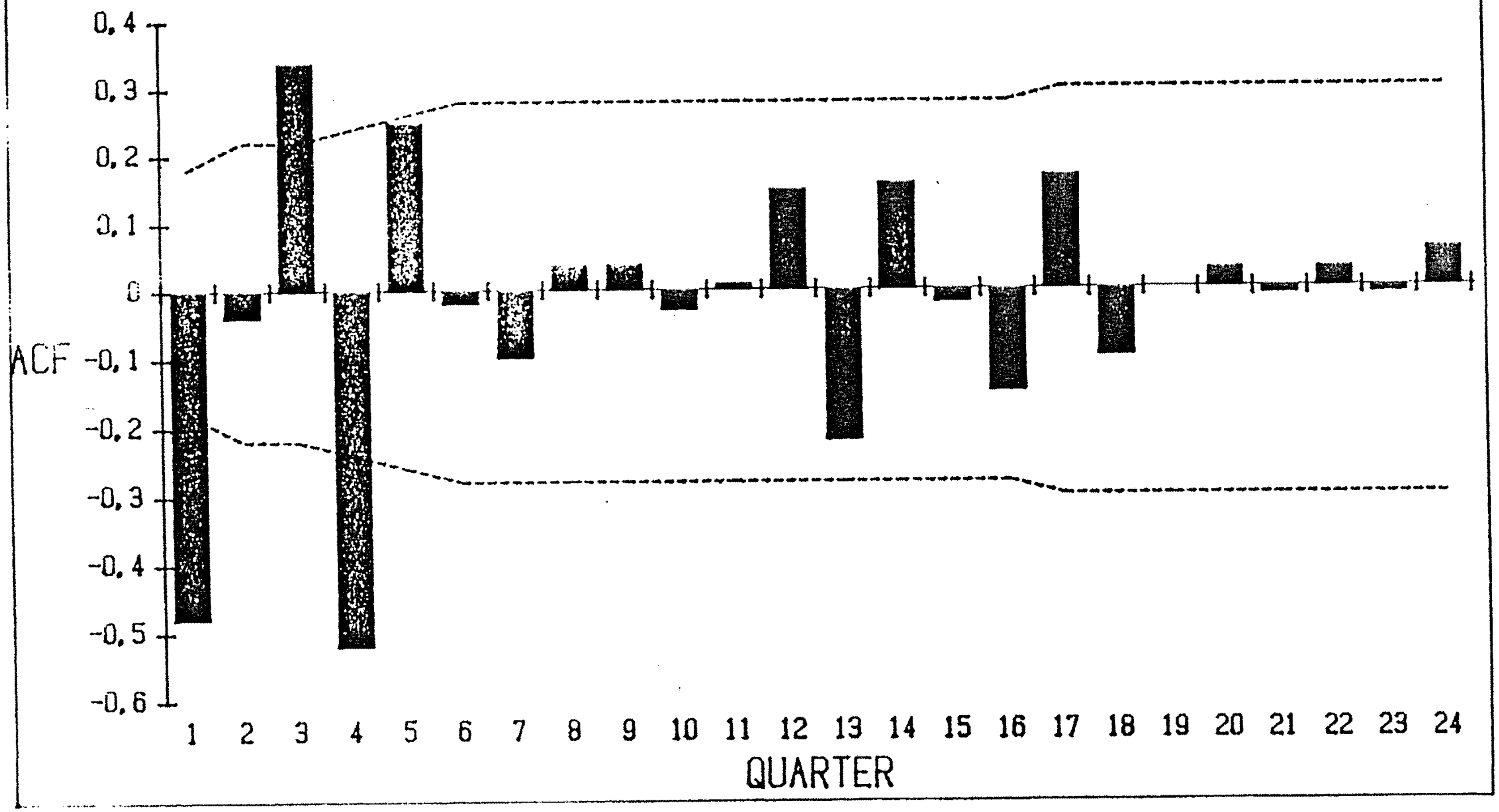


Table 9
Parameter Estimates for an ARIMA and an Intervention Model of Expenditure on Nondurables

Parameter	ARIMA Model	Intervention Model
α		.2473 (.0394)
Ø	.9858 (.0066)	.9824 (.0086)
t_1	.6746 (.0665)	.6098 (.0717)
t_4	.5653 (.0739)	.4417 (.0796)
R^2	.9970	.9980
see	.0176	.0156
DW	1.9700	1.9700
Q(21)	15.3000	19.5000

Figure 6
Residual Autocorrelation Function for Nondurables (EASTER)

model inadequacy. We regard the relatively large autocorrelation for lag 6 as statistical artefact and, therefore, do not attempt to eliminate it by including an own MA parameter.

3.1.3 Purchases of Durables

We proceed now to an analysis of those components of total consumer expenditure, for which the estimation of intervention models should be greater relevance. We begin with an analysis of purchases of durables as a whole, and then split up this component later on into purchases of cars and of other durables.

The EACF for purchases of durables is given in table 10. Somewhat surprisingly, the structure of this table is not different from that of total consumer expenditure. Again, it seems to indicate

Table 10
The Extended ACF Table for Purchases of Durables

that an ARMA(5,5) model is appropriate to characterize this series. Apparently, all consumption categories are nonstationary and exhibit a pronounced seasonal pattern. These two factors seem to be so strong that they veil existing differences in the behaviour of the various components. LS estimates of the iterated AR(5) regressions for purchases of durables are shown in table 11.

Table 11
Least Squares Estimates of Iterated AR(5) Regressions for Purchases of Durables

These estimates indicate that the AR polynomial might be of the form $(1-\emptyset B)(1-\Phi B)$ with both $\emptyset$ and Φ close to 1. Contrary to total consumption or to expenditure on nondurables, however, these estimates are far less stable. Primarily two factors might be responsible for that. First, purchases of durables are per se a more volatile series than total consumer expenditure. Second, this category was heavily affected by numerous fiscal policy measures during the last 20 years. It is impossible to identify a noise model for this series employing only undistorted observations, because we would end up with 50 observations coming from the late fifties or early sixties. Employing all 130 observations seems to be the smaller evil in this situation. A graph of the ACF for $(1-B)(1-B^4)\ln(\text{DURABLES})_t$ is depicted in figure 7. This ACF looks

Figure 7
Autocorrelation Function for $(1-B)(1-B^4)\ln(\text{DURABLES})_t$

quite well-behaved, but we have certain doubts whether this is not

FIGURE 6
RESIDUAL AUTOCORRELATION FUNCTION
FOR NONDURABLES (EASTER)

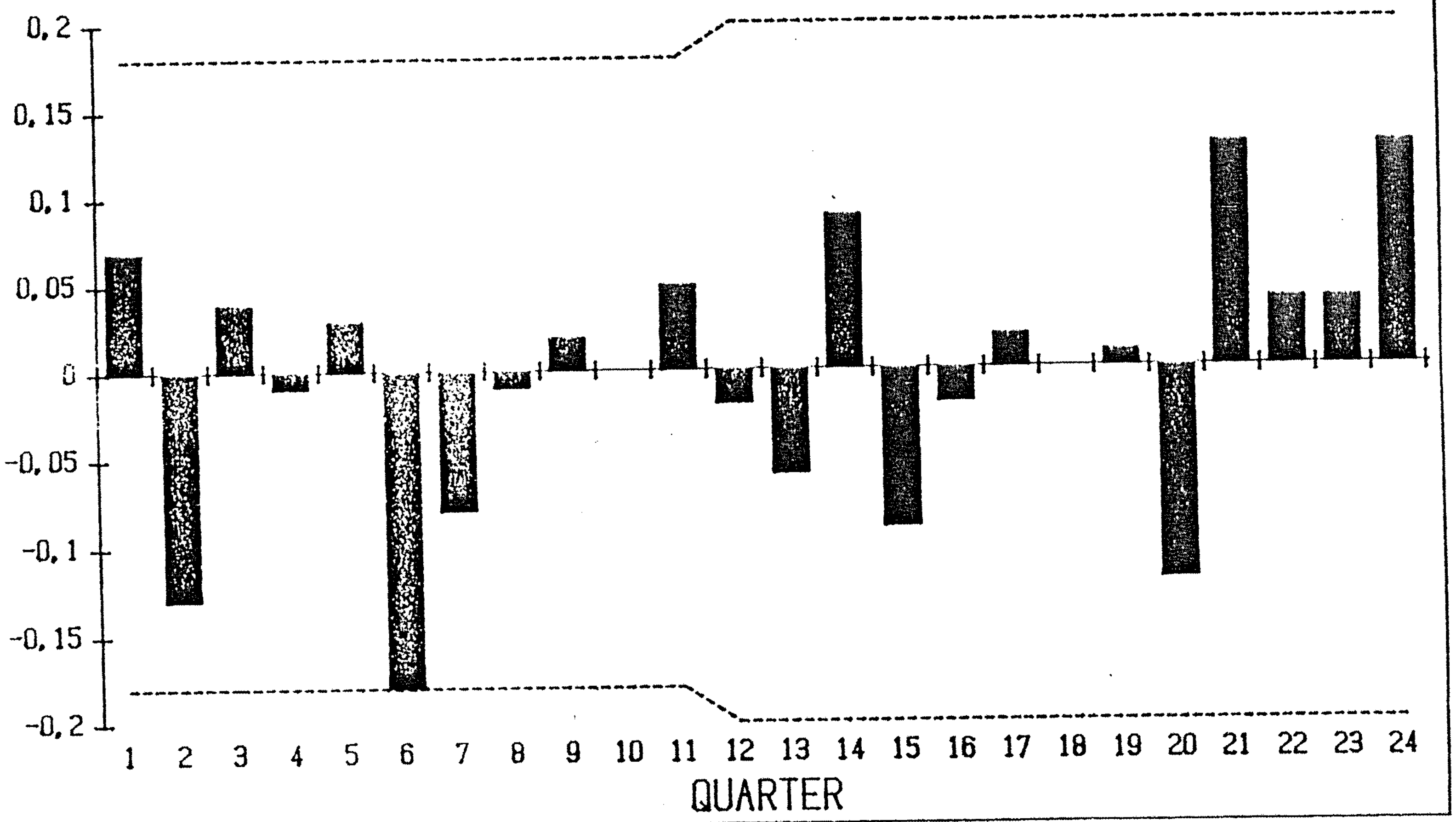


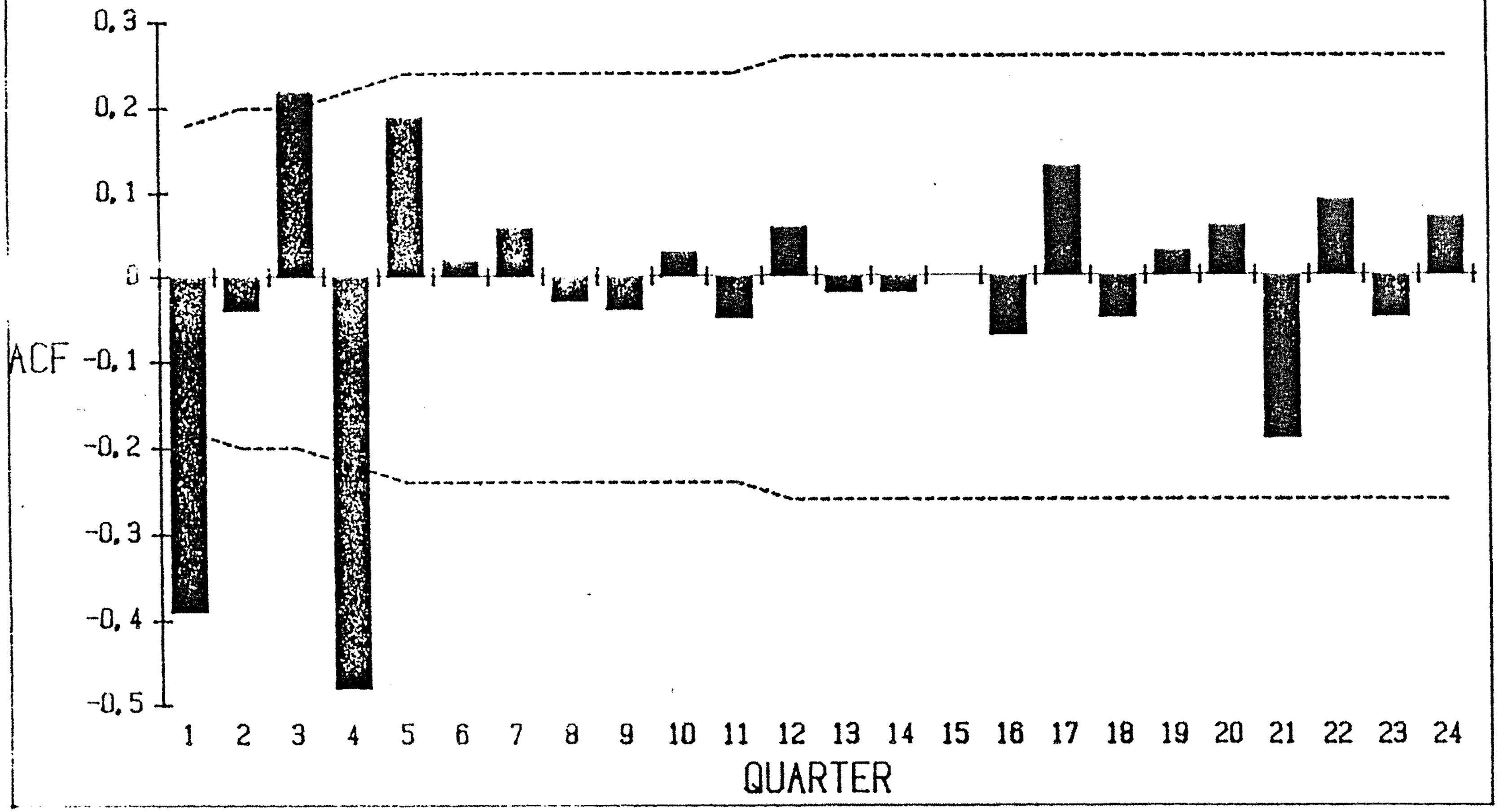
Table 10
The Extended ACF Table for Purchases of Durables

Q / P	0	1	2	3	4	5	6	7	8
0	.89	.89	.83	.87	.78	.78	.73	.78	.69
1	-.65	.35	-.55	.88	-.60	.35	-.53	.86	-.59
2	-.22	-.49	-.10	.81	-.19	-.44	-.09	.79	-.19
3	-.18	-.59	-.14	.81	-.15	-.52	-.12	.79	-.13
4	.36	.18	.14	-.23	-.02	-.04	-.02	-.08	-.13
5	-.35	-.08	.17	-.38	.19	.02	.05	-.03	-.04
6	-.48	-.11	-.03	-.44	.22	-.03	.07	-.03	-.04
7	.37	-.11	-.17	-.47	.24	.25	.18	-.01	-.04
8	.40	-.10	-.42	-.53	.18	-.08	.14	-.03	-.06

Table 11
Least Squares Estimates of Iterated AR(5) Regressions
for Purchases of Durables

j \ p	1	2	3	4	5
0	.51	.09	.01	.85	-.50
1	.77	.06	-.03	.87	-.69
2	.85	.07	-.04	.86	-.75
3	.37	.04	-.06	.91	-.30
4	2.33	-.05	-.09	1.06	-2.16
5	.97	.02	-.02	.99	-.96
6	.99	.00	-.01	.99	-.98
7	.79	-.01	-.03	1.00	-.77
8	.93	-.01	-.02	1.01	-.92
9	1.15	-.02	-.02	1.02	-1.13

FIGURE 7
AUTOCORRELATION FUNCTION FOR
(1-B) (1-B**4) ln(DURABLES)



also only a consequence of the large number of exogenous shocks, which affected our data. It points to a noise model of the form

$$(1-B)(1-B^4)\ln(\text{DURABLES})_t = (1-t_1B)(1-t_4B4)a_t. \qquad (10)$$

In the estimation stage, we found it advantageous to replace the difference operator (1-B) by an AR operator (1-ØB) and to add a constant term to the model. The parameter estimates for this pure ARIMA model are given in column 2 of table 12. A graph of the

Table 12
Parameter Estimates for an ARIMA and an Intervention Model
for Purchases of Durables

residual ACF for this model is shown in figure 8. Neither this graph nor the corresponding LJUNG-BOX Q statistic point to any model inadequacy. But, we have seen above that both tools are severly biased in favour of accepting an inadequate model under the present circumstances. Inspection of the residual series

Figure 8
Residual Autocorrelation Function for Durables (ARIMA)

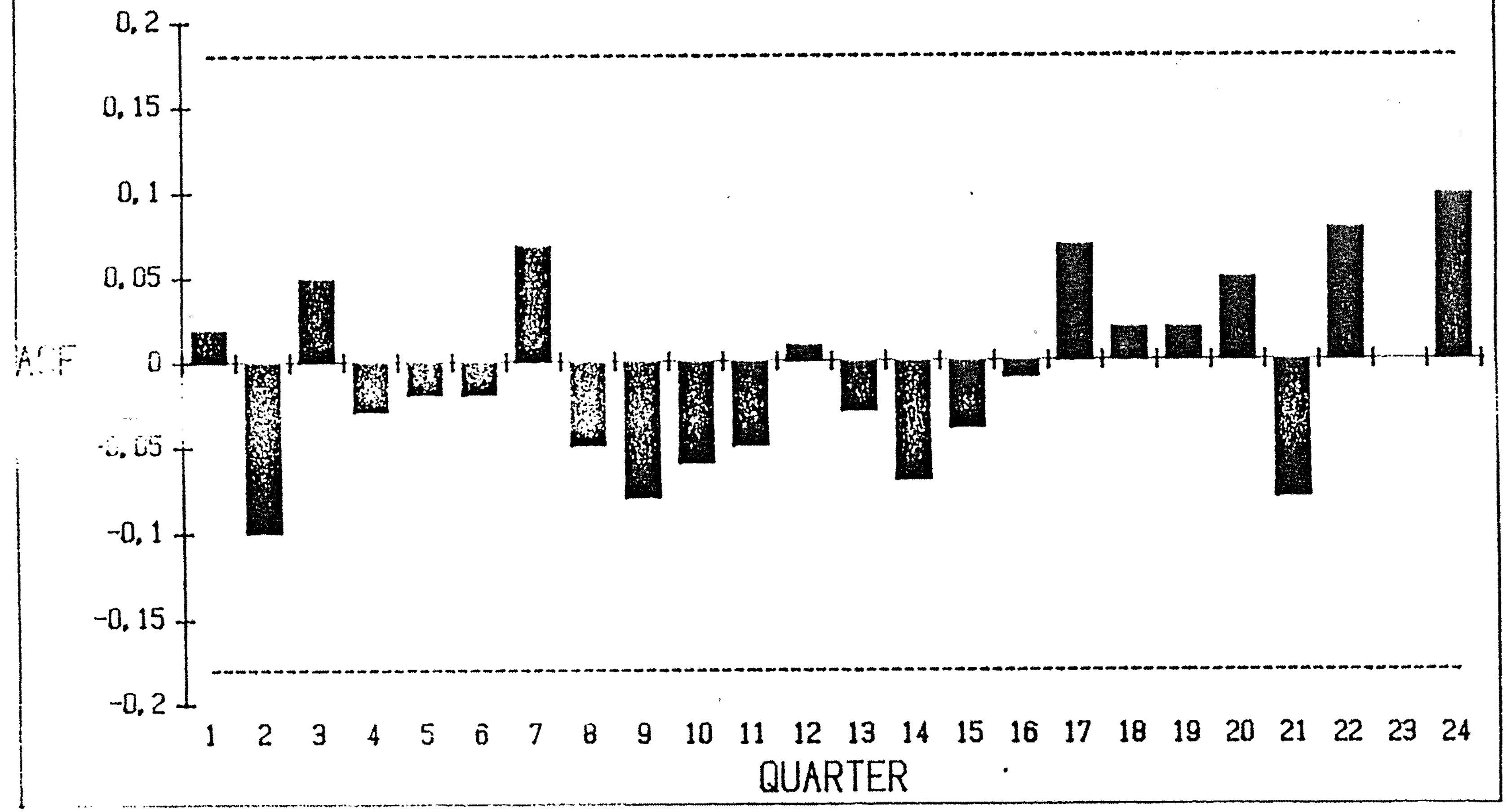
reveals then immediately that a substantial number of excessively large errors is present in this series.

Inclusion of the intervention variables of table 1 results in significant parameter estimates in almost all cases; the two exceptions being the reduction in tariffs vis-a-vis the Common Market and the cathalytic converter regulation. We note that the introduction (captured by w_{10}, w_{11}, δ_1) and abolition (w_2, δ_2) of a surtax on purchases of new cars, the introduction of VAT (w_3), the introduction of a special VAT rate for 'luxury goods' (w_{50}, w_{51}, δ_3) and the increase in VAT (w_{60}, w_{61}, δ_4) all had substantial effects on the purchases of durables. All these effects are of the expected sign and, with the exception of the introduction of VAT in January 1973, of the shape as depicted in figure 1. Some of the δ-parameters are very close to unity, thus pointing to the existence of a permanent effect. We believe, however, that one should be careful with such a conclusion. We have the impression that these high estimates for the d's are caused by the structure of our data. We observe namely interventions following each other in relatively short intervals of time. So, with quarterly data, we have often only 10 to 12 data points in order to establish the shape of a certain intervention. Consequently, it should be surprise that all these data points are affected and, therefore, signal a permanent effect. A similar problem may arise with interventions taking place shortly before the end of the sample period. This fact was pointed out already by BRANDNER (1986). The inclusion of intervention variables leads to a significant improvement in the explanatory power of the model for durables. The standard error of estimate declines by more than 40 percent from 0.078 to 0.046. Significant consequences for the

Table 12
Parameter Estimates for an ARIMA and an Intervention Model of Expenditure on Durables

Parameter	ARIMA Model	Intervention Model
CONST	.0454 (.0138)	.0523 (.0146)
w_{10}		.0150 (.0407)
w_{11}		-.1808 (.0426)
δ_1		.7903 (.1404)
w_{21}		.1082 (.0391)
δ_2		.9848 (.1042)
w_{30}		.1576 (.0352)
w_{50}		.2841 (.0396)
w_{51}		-.2706 (.0413)
δ_3		.5061 (.1248)
w_{60}		.1234 (.0362)
w_{61}		-.1144 (.0393)
δ_4		.9715 (.1066)
Ø	.9155 (.0303)	.8696 (.0349)
t_1	.4957 (.0876)	.3563 (.0965)
t_4	.7402 (.0667)	.4782 (.0845)
R^2	.9850	.9950
see	.0782	.0462
DW	1.9800	1.9800
Q(21)	10.0000	35.4000

FIGURE 8
RESIDUAL AUTOCORRELATION FUNCTION
FOR DURABLES (ARIMA)



estimates of the ARIMA model parameters can be observed also. These estimates, and here especially those of the seasonal parameter, are substantially different between the two models. The residual ACF for our intervention model is shown in figure 9. This

Figure 9
Residual Autocorrelation Function for Durables (INTERVENTION)

graph proves that our doubts concerning the seemingly well-behaved shape of the ACF in figure 7 are not unjustified. Apparently, there exist problems with modelling purchases of durables. Both, figure 9 and the relatively large LJUNG-BOX Q statistic point in this direction. Especially, the significant autocorrelations at lags 11 and 22 are somewhat suspicious, and should not be regarded as statistical artefacts. They might be caused by the existence of certain echo effects in the purchases of durables, which are due to the fiscal policy measures. Especially, for the purchases of cars such echo effects are observed. Simply including a lag 11 MA parameter does not seem to be an adequate solution here. It would cure only the symptom and not the problem. Finding an adequate model for these echo effects would be more appropriate.

Table 13, finally, provides a set of forecast accuracy measures for the pure ARIMA model, the intervention model, and HOLT-WINTERS exponential smoothing as benchmark. These measures are calculated

Table 13
Forecast Accuracy Measures
Purchases of Durables

for predictions of the log-changes of purchases of durables covering the period 1977:4 to 1986:2. We see that partial consideration of impending fiscal policy measures would improve the resulting forecasting performance immensely. There exist, however, practical limitations for doing this, which were discussed already above. Finally, it is interesting to note that we observe still significant differences in the quality of 5-step ahead predictions between the pure ARIMA and the intervention model. This fact might be the consequence of the better estimates for the MA parameters in the intervention model.

3.1.3.1 Car Purchases

We come now to an analysis of that consumption category which, definitely, was affected most by fiscal policy measures, namely car purchases. Car purchases constitute approximately one third of the expenditure on durables. From total consumer expenditure, they form consequently only a very small fraction (less than 4 percent in 1985). However, the effects of some fiscal policy measures were so strong, that they could be felt even in total consumer expenditure.

Table 14 presents the EACF for car purchases. The observations on car purchases from the end of 1968 onward are so heavily distorted

FIGURE 9

RESIDUAL AUTOCORRELATION FUNCTION FOR DURABLES (INTERVENTION)

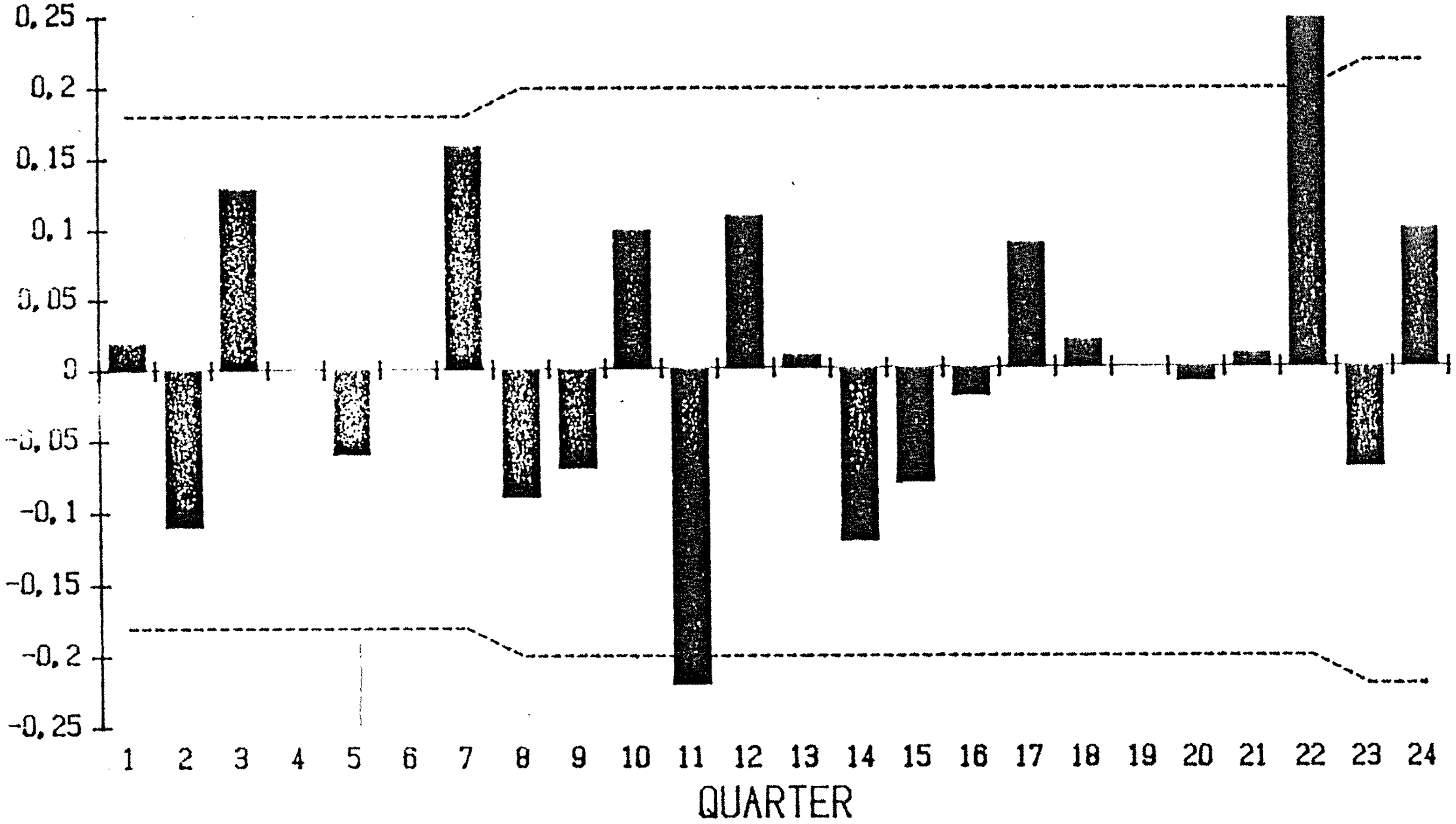


Table 13
Forecast Accuracy Measures
Purchases of Durables

Forecast Accuracy Measure	ARIMA	INTERVENTION	HOLT-WINTERS
		1-step ahead	
Correlation coefficient	.7573	.9752	.7527
RMSE	.1709	.0576	.1729
MAE	.1199	.0466	.1232
Regression coefficient	.7029	.9025	.6965
THEIL's inequality coefficient	.5290	.0600	.5410
U^M	.0001	.0003	.0001
U^R	.1937	.1847	.1988
U^D	.8062	.8150	.8011
		5-step ahead	
Correlation coefficient	.7994	.8662	.7810
RMSE	.1564	.1177	.1703
MAE	.1154	.0747	.1296
Regression coefficient	.7360	.9647	.6803
THEIL's inequality coefficient	.4432	.2509	.5250
U^M	.0004	.0007	.0005
U^R	.1854	.0040	.2566
U^D	.8142	.9953	.7429

by fiscal policy measures, that they can not be used for the identification of the noise model. Willy-nilly, we had to rely for

Table 14
The Extended ACF Table for Car Purchases

this purpose on the first 58 observations only. Consequently, this table will provide only very vague information on the possible form of the underlying noise model. It points to the existence of an ARMA(4,4) model. Inspection of the LS estimates of the iterated

Table 15
Least Squares Estimates of Iterated AR(4) Regressions for Car Purchases

AR(4) regressions in table 15 provides some scanty evidence that the AR polynomial contains a factor $(1-B^4)$. Consequently, the ACF of $(1-B^4)\ln(CARS)_t$ is shown in figure 10. Based on the also not

Figure 10
Autocorrelation Function for $(1-B^4)\ln(CARS)_t$

too reliable information provided by this graph, we tentatively entertain the model

$$(1-\emptyset B)(1-B^4)\ln(CARS)_t = (1-t_1B)(1-t_4B^4)a_t \qquad (11)$$

for further investigation. Parameter estimates are given in the column 2 of table 16. It is hardly any surprise that the obtained

Table 16
Parameter Estimates of an ARIMA and an Intervention Model for Car Purchases

pure ARIMA model seems to be of rather low quality. A standard error of estimate of almost 25 percent renders any hope to apply this model for forecasting purposes illusory. Here again, we find strong evidence that the residual autocorrelation function, which is shown in figure 11, and the LJUNG-BOX Q statistic are inadequate for diagnostic checking in certain situations. Closer

Figure 11
Residual Autocorrelation Function for Cars (ARIMA)

inspection of the residual series itself reveals immediately that it is dominated by an excessive number of very large errors. We employ again our set of intervention variables in order to get rid

Table 14
The Extended ACF Table for Car Purchases

Q / P	0	1	2	3	4	5	6	7	8
0	.82	.75	.68	.65	.53	.50	.46	.45	.34
1	-.38	-.13	-.28	.81	-.33	-.12	-.26	.74	-.35
2	-.48	.04	-.11	.79	-.40	.03	-.09	.64	-.35
3	-.42	-.16	-.26	.75	-.35	-.18	-.19	.59	-.28
4	.18	.14	.16	.35	.01	-.06	.02	.01	-.13
5	-.41	.02	.18	-.33	.21	-.15	.23	-.21	-.01
6	-.31	.04	.06	-.26	-.05	.09	.14	-.22	-.03
7	.41	-.10	-.11	-.26	-.05	.14	.12	-.04	-.11
8	.38	-.17	-.26	-.30	.06	-.07	.08	-.05	-.04

Table 15
Least Squares Estimates of Iterated AR(4) Regressions
for Car Purchases

j \ p	1	2	3	4
0	.38	.06	-.05	.43
1	-.90	.73	.07	.81
2	-.09	-.80	.66	.97
3	.04	-.02	-.63	1.41
4	.29	.09	-.06	.57
5	-.70	.55	.14	.86
6	-.15	-.56	.52	1.05
7	.04	.01	-.49	1.34
8	.19	.11	-.06	.71
9	-.39	.32	.09	.92

FIGURE 10
AUTOCORRELATION FUNCTION
FOR (1-B**4) ln(CARS)

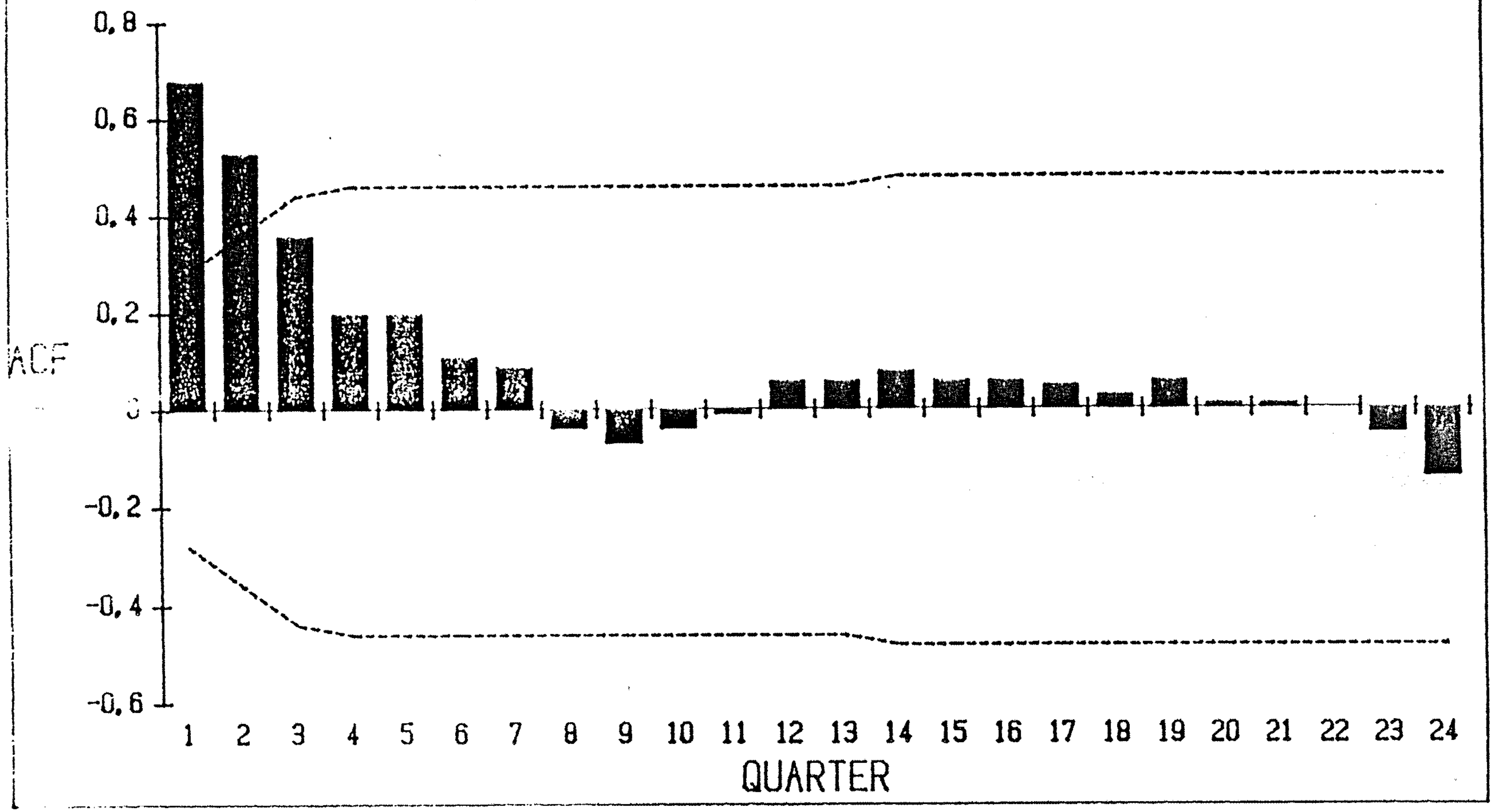
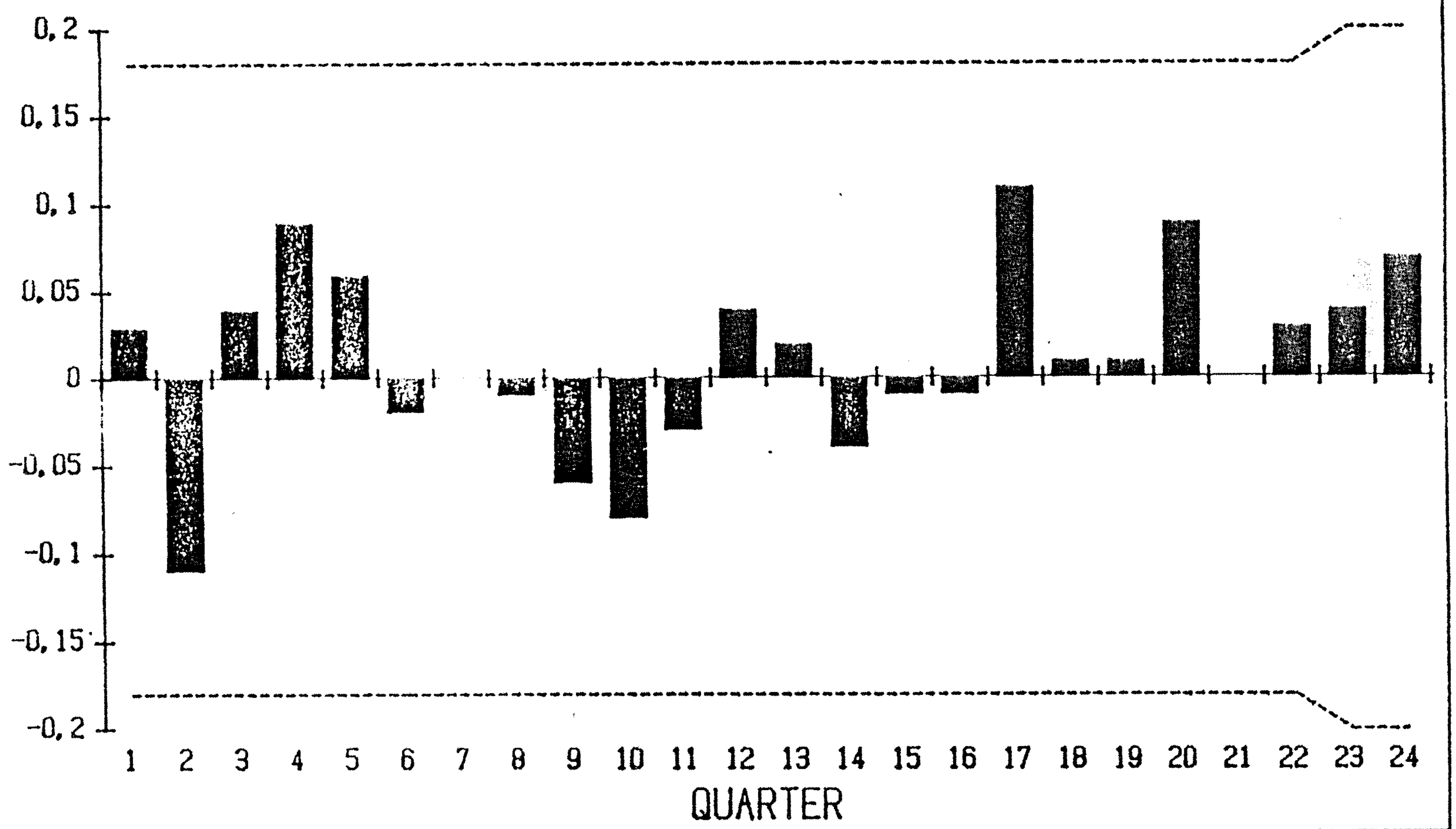


Table 16
Parameter Estimates for an ARIMA and an Intervention Model of Purchases of Cars

Parameter	ARIMA Model	Intervention Model
CONST	.0708 (.0085)	.0791 (.0189)
w_{10}		.5245 (.0789)
w_{11}		-1.4188 (.0887)
δ_1		.4999 (.0517)
w_{21}		.2732 (.0798)
δ_2		.9081 (.1122)
w_{30}		.5207 (.0694)
w_{40}		.1494 (.0703)
w_{50}		.6719 (.0823)
w_{51}		-.7939 (.0903)
δ_3		.5285 (.0906)
w_{60}		.1590 (.0744)
w_{61}		-.2657 (.0812)
δ_4		.9651 (.0942)
w_{70}		.1756 (.0641)
w_{71}		-.2323 (.0833)
Ø	.8274 (.0169)	.8039 (.0314)
t_1	.5256 (.0779)	.2104 (.1002)
t_4	.9359 (.0319)	.4965 (.0817)
R^2	.9350	.9900
see	.2413	.0930
DW	1.9700	2.0000
Q(21)	10.0000	31.9000

FIGURE 11
RESIDUAL AUTOCORRELATION FUNCTION
FOR CARS (ARIMA)



of the most disturbing errors at least. Here, all the intervention variables given in table 1 enter the relation with highly significant coefficients. With one exception, all policy measures have the expected effects, and the shape of these effects is in most cases of the form illustrated in figure 1. Only the cathalytic converter regulation had an effect, which is surely not in accordance with the intentions of the legislator. Although a premium for buying cars with cathalytic converter was offered additionally, it produced a substantial increase in purchases of conventional cars at the time immediately before the introduction of the new car-tax scheme followed by a corresponding decline of car sales in the subsequent quarter.

The introduction of intervention variables for the various fiscal policy measures improves the explanatory power of our equation substantially. It leads to a decline of more than 60 percent in the standard error of estimate. It must be admitted, however, that the remaining error of this intervention model is still large. The inclusion of intervention variables has also dramatic effects for the estimates of the MA parameters in the noise model. The regular MA parameter t_1 declines from 0.53 to 0.21, and the seasonal MA parameter t_4 from 0.94 to 0.50. The residual ACF is given in

Figure 12
Residual Autocorrelation Function for Cars (INTERVENTION)

figure 12. Again we find some evidence of the above mentioned echo effects.

Table 17, finally, provides information about the forecasting performance of our time series models for car purchases. It corroborates our opinion that the pure ARIMA model is completely useless for forecasting purposes. Taking account of impending

Table 17
Forecast Accuracy
Car Purchases

fiscal policy measures would improve the situation considerably. The more estimates of MA parameters alone already give rise to a significant improvement in the forecasting performance as can be seen from a comparison of the 5-step ahead predictions.

3.1.3.2 Purchases of Other Durables

We come now to the last consumption category, which we intend to analyse in this paper, namely purchases of other durables. The EACF for this series is shown in table 18. The autocorrelations

Table 18
The Extended ACF Table for Purchases of Other Durables

FIGURE 12
RESIDUAL AUTOCORRELATION FUNCTION
FOR CARS (INTERVENTION)

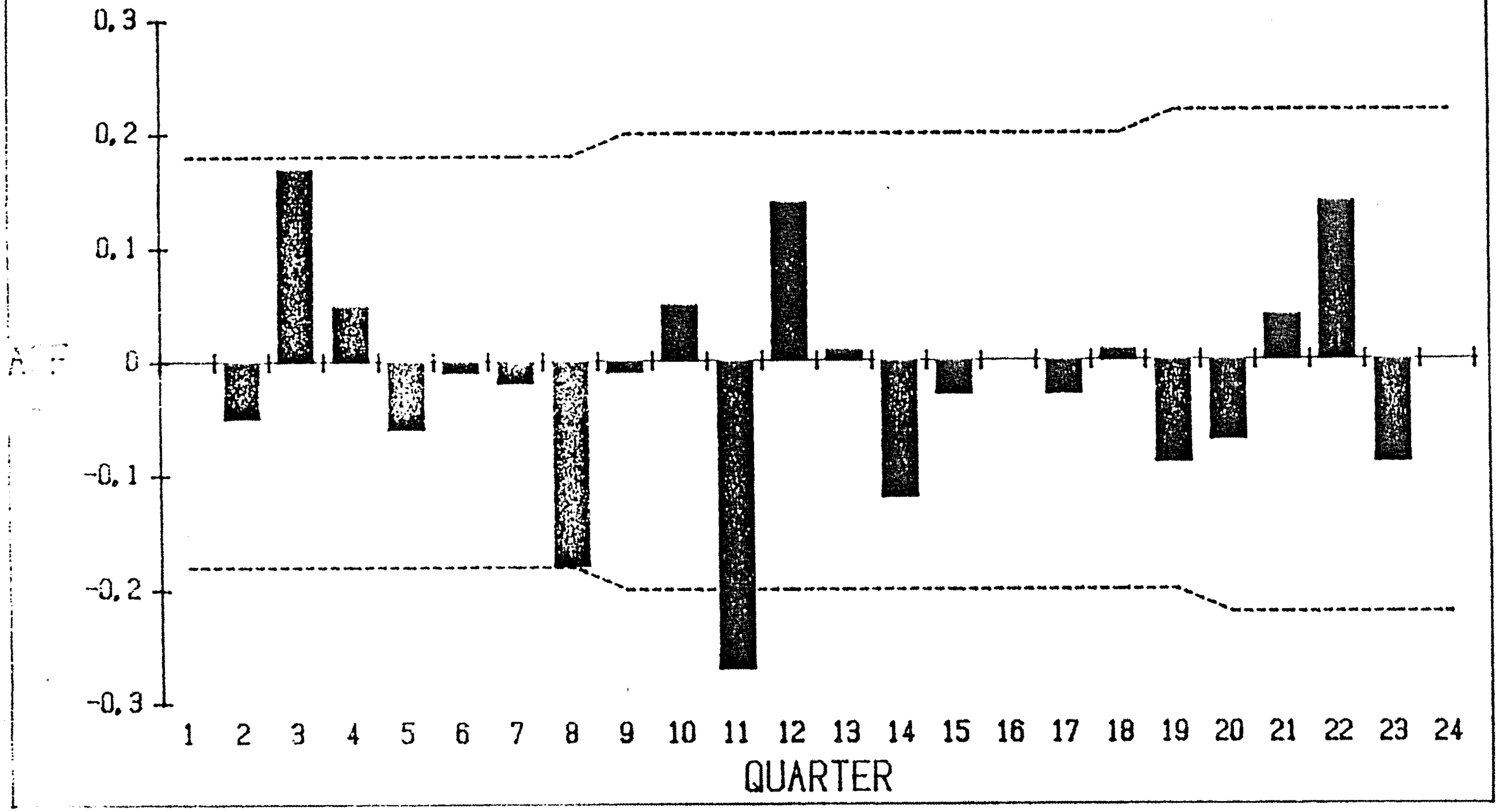


Table 17
Forecast Accuracy Measures
Car Purchases

Forecast Accuracy Measure	ARIMA	INTERVENTION	HOLT-WINTERS
		1-step ahead	
Correlation coefficient	-.1029	.9480	.1253
RMSE	.4266	.0959	.4294
MAE	.3102	.0799	.2829
Regression coefficient	-.2769	1.0436	.2016
THEIL's inequality coefficient	1.2141	.0001	1.2298
U^M	.0003	.0692	.0002
U^R	.1853	.0142	.2000
U^D	.8144	.9166	.7998
		5-step ahead	
Correlation coefficient	-.1419	.5338	.2889
RMSE	.4251	.3316	.4242
MAE	.3405	.1812	.2962
Regression coefficient	-.4266	.7952	.3514
THEIL's inequality coefficient	1.2056	.7336	1.2003
U^M	.0010	.0002	.0003
U^R	.1868	.0257	.2367
U^D	.8122	.9741	.7630

Table 18
The Extended ACF Table for Purchases of Other Durables

Q / P	0	1	2	3	4	5	6	7	8
0	.83	.81	.78	.89	.73	.72	.69	.80	.65
1	-.51	.03	-.38	.95	-.49	.03	-.36	.93	-.47
2	-.47	-.04	-.14	.95	-.45	-.03	-.13	.93	-.44
3	-.30	-.34	-.28	.95	-.28	-.32	-.28	.93	-.27
4	.42	.30	.20	-.13	.01	-.04	.00	-.02	-.06
5	-.39	-.01	.21	-.46	.19	-.05	.07	.00	-.07
6	-.39	.00	.02	-.47	.09	.11	.07	.00	.05
7	.37	-.05	-.19	-.47	.17	.10	.17	.01	-.02
8	.43	-.02	-.41	-.50	.21	-.13	.15	-.06	-.05

of this table indicate that an ARMA(4,4) or ARMA(5,5) model should be appropriate. The LS estimates of the iterated AR(5) regressions in table 19 supply evidence that the AR polynomial could be of the

Table 19
Least Squares Estimates of Iterated AR(5) Regressions for Purchases of Other Durables

multiplicative form (1-ØB)(1-ΦB) with both Ø and Φ close to 1. The ACF for $(1-B)(1-B^4)\ln(\text{OTHER DURABLES})_t$ is depicted in figure 13.

Figure 13
Autocorrelation Function for $(1-B)(1-B^4)\ln(\text{OTHER DURABLES})_t$

This ACF exhibits a pattern, which was encountered frequently in this study already. A model of the form

$$(1-B)(1-B^4)\ln(\text{OTHER DURABLES})_t = (1-t_1B)(1-t_4B^4)a_t \qquad (12)$$

should provide an adequate representation of this series. Parameter estimates are given in table 20. In the estimation stage, we detected that it was better to replace the difference

Table 20
Parameter Estimates for an ARIMA and an Intervention Model for Purchases of Other Durables

operator (1-B) by the AR operator (1-ØB) and to add a constant term to the equation. The residual ACF for eq. (12) is given in figure 14. Neither this graph nor the LJUNG-BOX Q statistic point

Figure 14
Residual Autocorrelation Function for Other Durables (ARIMA)

to the existence of autocorrelation in this residual series. Closer inspection of this series, however, reveals that a number of unduly large residuals exist. Consequently, some of our intervention variables (namely, those which are supposed to capture the introduction of VAT and the various changes in the VAT rates) give significant parameter estimates. Of special interest in this context is the - unfortunately insignificant - negative coefficient of the intervention variable x_1, standing for the introduction of the surtax on car purchases (w_{10} in table 20). It provides some weak evidence that this policy measure induced some substitution between car purchases and expenditure on other durables. The introduction of intervention variables leads to a decline in the standard error of estimate by almost 20 percent. The residual ACF for the intervention model is shown in figure 15.

Table 19
Least Squares Estimates of Iterated AR(5) Regressions
for Purchases of Other Durables

j \ p	1	2	3	4	5
0	.67	.00	.00	.97	-.67
1	.88	.00	.00	.98	-.86
2	.88	.00	.00	.98	-.86
3	.63	-.01	-.02	.99	-.62
4	1.23	-.01	.00	1.02	-1.22
5	.91	.00	.00	1.01	-.91
6	1.05	-.01	.00	1.01	-1.05
7	.94	-.01	-.01	1.01	-.94
8	.99	-.01	-.01	1.01	-.94
9	1.12	-.01	-.01	1.01	-1.12

FIGURE 13
AUTOCORRELATION FUNCTION FOR
(1-B) (1-B**4) ln (OTHER DURABLES)

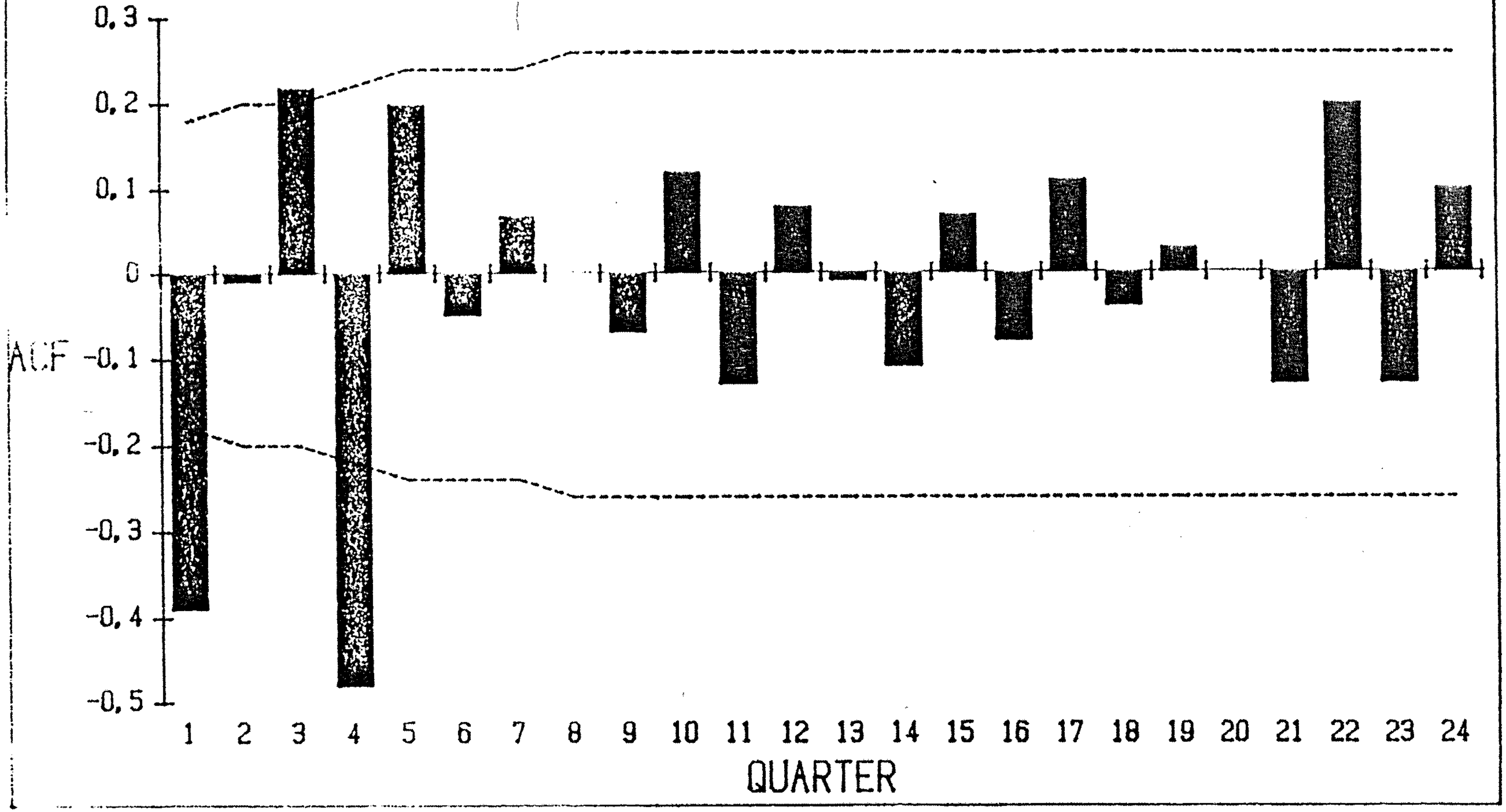


Table 20
Parameter Estimates for an ARIMA and an Intervention Model of Expenditure on Other Durables

Parameter	ARIMA Model	Intervention Model
CONST	.0374 (.0154)	.0406 (.0149)
w_{10}		-.0445 (.0303)
w_{30}		.0370 (.0323)
w_{31}		-.0750 (.0322)
w_{50}		.1556 (.0325)
w_{51}		-.0730 (.0322)
w_{60}		.1170 (.0307)
Ø	.9245 (.0287)	.9035 (.0331)
t_1	.4293 (.0888)	.2930 (.0953)
t_4	.6247 (.0715)	.5102 (.0809)
R^2	.9930	.9950
see	.0501	.0413
DW	1.9900	1.9800
Q(21)	12.2000	20.0000

FIGURE 14
RESIDUAL AUTOCORRELATION FUNCTION
FOR OTHER DURABLES (ARIMA)

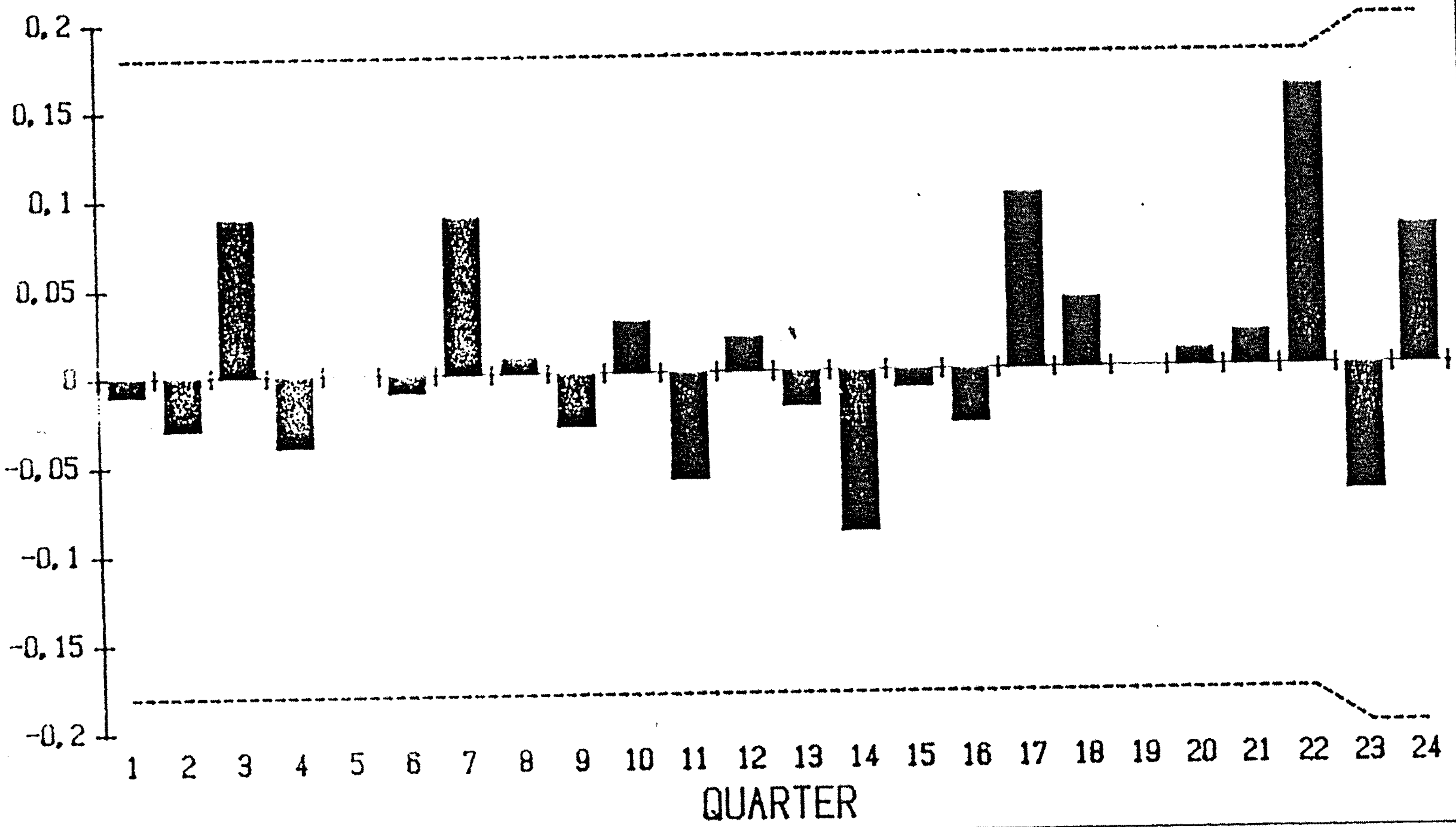


Figure 15
Residual Autocorrelation Function
for Other Durables (INTERVENTION)

Neither this graph nor the LJUNG-BOX Q statistic signal any porblems with residual autocorrelation.

Table 21, finally, provides information on the forecasting performance of the pure ARIMA and the intervention model, with HOLT-WINTERS exponential smoothing again serving as baseline

Table 21
Forecast Accuracy
Other Durables

These measures are calculated from predicted and realized log-changes of purchases of other durables covering the period from 1977:4 to 1986:2. We note again that a correct anticipation of impending fiscal policy measures could reduce the committed forecast errors substantially. The better estimates of the MA parameters in the intervention model lead to a higher forecast accuracy even in situations, where impending interventions are not taken into account.

3.2 Numerical Evaluation of the Effects of Fiscal Policy Measures

The above intervention models can be used to quantify the effects of the different fiscal policy measures. We shall concentrate here on an analysis of the two most important measures, namely the introduction and abolition of a 10 percent surtax on purchases of new cars and the introduction of a 30 percent VAT rate for socalled 'luxury goods'. Apart from expenditure on nondurables, substantial effects of these two measures were found in all of our estimated intervention models. Therefore, we have a number of sources from which we can derive numerical estimates for the order of magnitude of these effects:

(i) the intervention model for total consumer expenditure;
(ii) the intervention model for durables;
(iii) the intervention models for cars and for other durables.

A comparison of these effects will provide additional information on the quality of our estimated models. Above all, we want to check whether the effects derived from different models bear some similarity. Of course, it is not to be expected that they are exactly the same, but the observed discrepancies should stay within certain limits. If this is not the case, the usefulness of our estimated intervention models must be questioned.

We begin with an analysis of the effects of the surtax on car purchases. Estimates for the numerical order of magnitude of these effects in billions of AS are given in table 22. We note that there exist substantial discrepancies between the estimates of the

FIGURE 15

RESIDUAL AUTOCORRELATION FUNCTION FOR OTHER DURABLES (INTERVENTION)

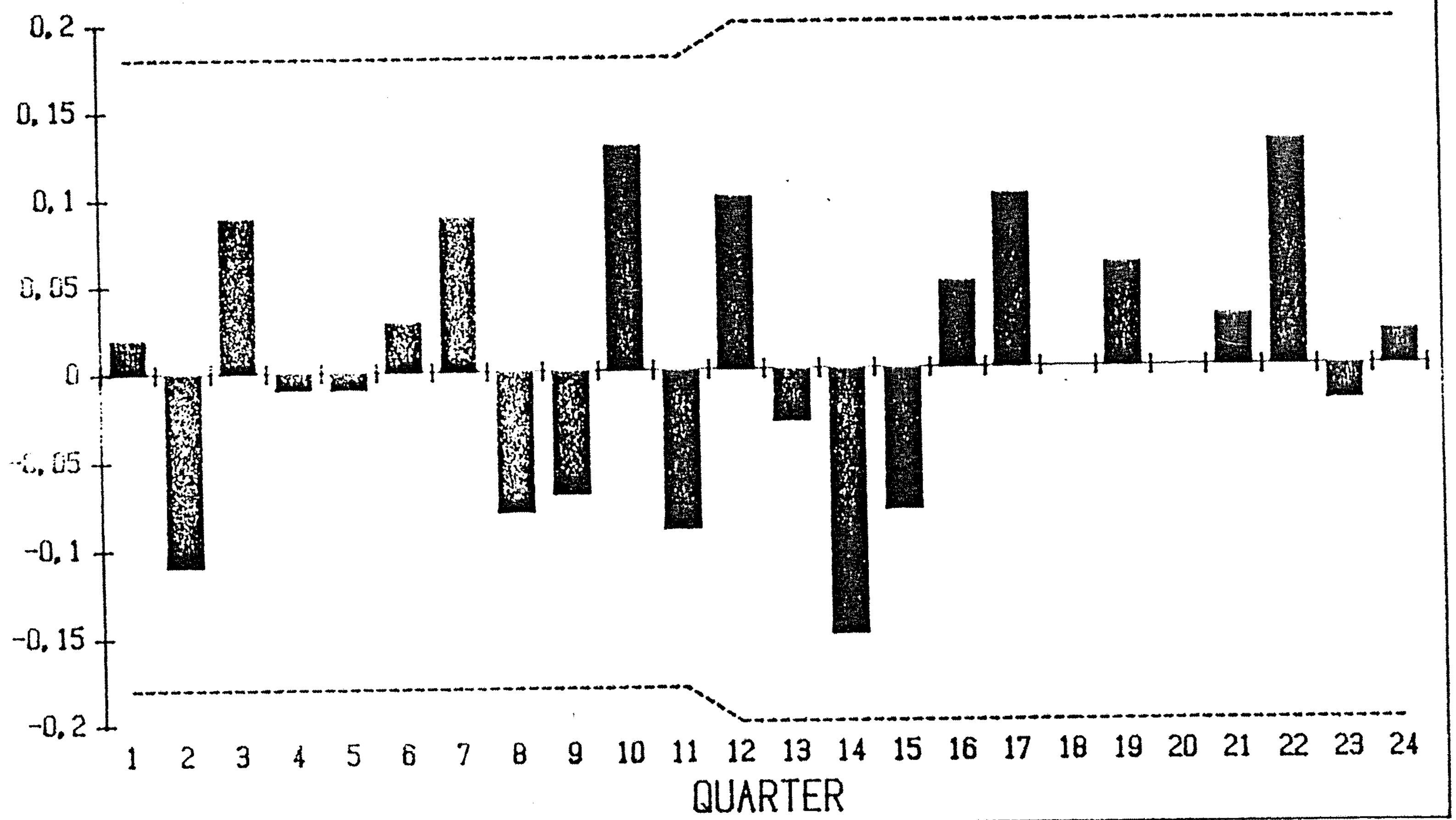


Table 21
Forecast Accuracy Measures
Purchases of Other Durables

Forecast Accuracy Measure	ARIMA	INTERVENTION	HOLT-WINTERS
		1-step ahead	
Correlation coefficient	.9555	.9909	.9606
RMSE	.0935	.0453	.0862
MAE	.0712	.0373	.0646
Regression coefficient	.9093	.9365	.9345
THEIL's inequality coefficient	.0961	.0226	.0817
U^M	.0000	.0004	.0001
U^R	.0946	.2003	.0554
U^D	.9054	.7993	.9445
		5-step ahead	
Correlation coefficient	.9598	.9741	.9531
RMSE	.0917	.0686	.1004
MAE	.0721	.0536	.0792
Regression coefficient	.8915	.9746	.8731
THEIL's inequality coefficient	.0925	.0518	.1108
U^M	.0002	.0004	.0005
U^R	.1474	.0125	.1731
U^D	.8524	.9871	.8264

Table 22
Effects of the Surtax on Purchases of New Cars

amount of advance purchasing derived from different sources. Above all, we observe significant differences between the estimates derived from intervention models for durables and for cars on the one side, and those obtained from the intervention model for total consumer expenditure on the other. We believe that this discrepancy is the consequence of the widely diverging orders of magnitude of the variables under study. Car purchases constitute only a minor fraction of total consumer expenditure. Therefore, it will be very hazardeous to derive the size of the effect of a fiscal policy measure, which affects only car purchases, from a model for total consumer expenditure. The slightest bias in the parameter estimates will produce huge errors. Simply because even very small errors in total consumer expenditure will constitute substantial errors for car purchases, when imputed completely to that consumption category. Consequently, effects of fiscal policy measures should be calculated only from models, which try to explain the specific variable being directly affected by the measure. The estimates for the amount of advance purchasing derived from the intervention models for durables and for cars lie much closer together. Moreover, the observed discrepancies have an obvious explanation. Apparently, they result from a substitution between car purchases and purchases of other durables. Thus, these two sets of estimates are in fact fully consistent. Advance purchasing in an order of magnitude of 1.2 billions of AS seems to be a reasonable estimate. Increased purchases in 1968:3 are followed by a corresponding decline in 1968:4. The restraint in buying during the whole year 1969 could, at first sight, be interpreted as evidence for changed consumer attitudes. However, the reaction to the abolition of this surtax contradicts this interpretation. Rather, it proves that consumers just postponed car purchases, because they expected an abolition of the surtax in the near future. Adding up all distributed effects, we find that the overall effect is practically zero. The introduction and abolition of this surtax only induced shifts in car purchases between single years, but had no permanent effect on consumer behaviour.

Next, we turn to a more detailed analysis of the effects of the introduction of a special VAT rate for socalled 'luxury goods' in January 1978. Estimates of the numerical size of these effects are presented in table 23. We note immediately that this policy measure had by far the strongest effects of all interventions which are analysed in this paper. It caused a substantial amount

Table 23
Effects of the Introduction of a Special VAT Rate on 'Luxury Goods'

of advance purchasing immediately before its adoption and a corresponding negative effect, which loses gradually in importance. Summing these negative effects, we see that there is

Table 22
Effects of the Surcharge on Purchases of New Cars

	Effect Calculated from Equation for Durables	Cars	Total Consumption
Introduction: Sept. 1, 1968			
1968:3	+0.8268	+1.2002	+2.6060
1968:4	-1.6319	-1.3193	-1.9293
1969:1	-0.8226	-0.8750	
1969:2	-0.8723	-0.8039	
1969:3	-0.6979	-0.3046	
1969:4	-0.7145	-0.1461	
Abolition: Jan. 1, 1971			
1971:1	+0.8305	+0.6606	
1971:2	+1.0135	+0.7995	
1971:3	+0.9791	+0.5436	
1971:4	+1.2101	+0.4474	

Table 23

Effects of Special Value-Added Tax Rates on 'Luxury Goods'

	Effect Calculated from Equation for				
	Durables 1	Cars 2	Other Durables 3	2 + 3 4	Total Consumption 5
1977:4	+5.6804	+3.5349	+2.2781	+5.8130	+7.3657
1978:1	-2.9954	-2.5116	-0.5730	-3.0846	-3.8195
1978:2	-1.8396	-1.7825		-1.7825	-2.9594
1978:3	-0.9100	-0.7686		-0.7686	-2.2569
1978:4	-0.5672	-0.3636		-0.3636	-1.8242
1979:1	-0.2194	-0.2715		-0.2715	-1.1046
1979:2	-0.1253	-0.1628		-0.1628	-0.8595
1979:3	-0.0601	-0.0627		-0.0627	-0.6485
1979:4	-0.0360	-0.0307		-0.0307	-0.5274

not much difference between the overall positive and negative effects. Thus, we can conclude again that the adopted fiscal policy measure caused only a shift of consumer expenditure between 1977 and 1978, but no permanent decline.

Comparing the effects, which are derived from different sources, we see that relying on the intervention model for total consumer expenditure apparently results in a drastic overestimation of the numerical size of the effect. This result is consistent with our findings about car purchases, because purchases of durables also constitute only a small portion of total consumer expenditure. The effects calculated from the intervention models for durables and for its two components, on the other side, lie surprisingly close together. Their order of magnitude, however, is considerably larger than the 4.5 billions of AS reported in PUWEIN, STANKOVSKY, and WÜGER (1984). This difference is obviously due to the method used by these authors. They calculate the effect as difference between the seasonal adjusted series and its trend-cycle component. Using X-11 and not a model-based approach for seasonal adjustment, this procedure must necessarily lead to a downward biased estimate of the fiscal policy measure, because the effect will be reflected already partly in the estimate of the trend-cycle component.

4. Conclusions

Our time series analysis of consumer expenditure reveals that this aggregate is heavily affected by various fiscal policy measures. We find that these fiscal policy measures produce only shifts in consumer expenditure between different years, but no permanent changes in consumer attitude. One should proceed very carefully when trying to estimate the effects of fiscal policy measures. It might be very problematic to attempt to recover the effects of a certain measure, which hits only a specific and, moreover, relatively small component of total consumer expenditure, from an analysis of the global aggregate. The effects of important fiscal policy measures, like the introduction of a surtax on purchases of new cars or the introduction of a special VAT rate for 'luxury goods' for example, can be spotted in total consumer expenditure also. But, unfortunately, it turned out that the derived effects were extremely biased. Much more reliable estimates of these effects were obtained from equations explaining particular consumption categories. Additionally, we documented that the correct anticipation of impending fiscal policy measures would improve the forecast accuracy of time series models enormously. Whether this can be actually done in practice, depends on the specific situation. The estimation of intervention models might provide considerable help in situations where fiscal policy measures, which have been adopted in the past already, are going to be reused again. Finally, we discovered that the estimates of the moving average parameters of ARIMA models are severely upward biased for series, which are heavily affected by policy interventions. This is especially true for the estimates of seasonal moving average parameters. This fact will have strong implications in situations, where one intends to employ a model-based approach for the seasonal adjustment of such series. Eliminating the effects of policy interventions before the seasonal adjustment will greatly improve the resulting outcome.

5. References

ANDERSON, O.D. and J.G. de GOIJER, Distinguishing Certain Stationary Time Series Models from their Nonstationary Approximations and Improved BOX-JENKINS Forecasting, in: O.D. ANDERSON (ed.), Analysing Time Series, North-Holland, Amsterdam, 1980, pp. 21 -42.

BOX, G.E.P. and J.M. JENKINS, Time Series Analysis: Forecasting and Control, Holden Day, San Francisco, 1970.

BOX, G.E.P. and G.C. TIAO, Intervention Analysis with Applications to Economic and Environmental Problems, JASA, 70(349), March 1975, pp. 70 - 79.

BRANDNER, P., Interventionsanalyse im Rahmen des Zeitreihenmodellbaus - eine Darstellung anhand des österreichischen Konsums, Seminararbeit am Institut für höhere Studien in Wien, Wien, 1986, (mimeo).

HILLMER, S.C. and G.C. TIAO, Likelihood Function for Stationary Multiple Autoregressive Moving Average Models, JASA, 74(367), September 1979, pp. 652 - 660.

LIU, L.-M., Identification of Time Series Models in the Presence of Calendar Variation, Working Paper No. 84-02, Department of Quantitative Methods, University of Illinois at Chicago, Chicago, 1983.

LIU, L.-M. and G.B. HUDAK, The SCA Statistical System Reference Manual, Scientific Computing Associates, P.O. Box 625, DeKalb, Illinois 60302, 1985.

PUWEIN, W., J. STANKOVSKY, and M. WÜGER, Vorkäufe im Jahr 1983, WIFO Monatsberichte, 4/1984, pp. 209 - 216.

THURY, G., Macroeconomic Forecasting in Austria. An Agnostic Viewpoint, Empirica 2'86, forthcoming.

THURY, G., Modelling Private Consumer Expenditure in Austria by Intervention Analysis, in: O.D. ANDERSON (ed.), Time Series Analysis: Theory and Practice 1, North-Holland, Amsterdam, 1982, pp. 307 - 321.

TSAY, R.S. and G.C. TIAO, Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models, Technical Report No. 683, University of Wisconsin at Madison, Department of Statistics, Madison, 1982.

TSAY, R.S. and G.C. TIAO, Identification of Multiplicative ARMA Models for Seasonal Time Series, Technical Report No. 7, Statistics Research Center, Graduate School of Business, University of Chicago, Chicago, 1983.

REPORTS IN THIS SERIES

The following titles can be obtained from WT Druck und Verlag, Wirtschaft und Technik Druck- und Verlagsgesellschaft mbH & Co KG, Neubaugasse 1, A-1071 Vienna, Austria (Tel. 0222/93 33 75).

Price per volume AS 180,- or foreign currency equivalent.