SEGMENTED REGRESSION MODELLING WITH AN APPLICATION TO GERMAN EXCHANGE RATE DATA

by

EDWARD ANDREW SUSKO

B.A., The University of Windsor, 1990

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

THE FACULTY OF GRADUATE STUDIES
THE DEPARTMENT OF STATISTICS

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1991

Abstract

Segmented regression models are the topic of this thesis. These are regression models in which the mean response is thought to be linear in the explanatory variables within regions of a particular explanatory variable. A criterion for estimating the number of segments in a segmented model is given and the consistency of this estimator is established under rather general conditions.

There have been many studies on modeling and forecasting foreign exchange rates using various models, notably the random walk model, the forward rate model, monetary models and vector autoregressions, see, for example, Meese and Rogoff (1983) and Baillie and McMahon (1989). The general conclusions have been that most of the models cannot outperform the random walk model by a significant margin. The observation that the dependence of the exchange rate on the key macroeconomic indicators is time varying, nonstationary and nonlinear leads to consideration of nonlinear models. In this thesis segmented models are fitted to German exchange rate data using least squares and forecasting results obtained from these models are compared with forecasting results from widely used models in exchange rate prediction. The segmented models tend to perform better than models that have been established in the literature, notably, the random walk model.

Table of Contents

List of Tables

List of Figures

Section 1

Exchange Rates

1.1 Introduction

Exchange rates are an important source of variation in financial decisions. However, attempts to model and predict exchange rates beyond 1971 have in general not performed better than that of a random walk, Meese and Rogoff (1983). It is the purpose of at least part of this thesis to provide a model which helps to explain some the variation contained in exchange rates.

Fixed currency exchange rates have been the norm for many years (until 1971 we had been more or less under a system of fixed exchange rates). They have also been one of the goals that the international community has tried to achieve. When exchange rates are fixed or reasonably predictable it is believed that trading in global markets is less tentative and that economic activity is more efficient. One of the problems with trying to maintain fixed exchange rates, however, is that government policy will be restricted, sometimes to the detriment of domestic considerations. As a result of this problem exchange rates have been allowed to float.

Prior to the great depression the exchange rate market operated according to the gold standard. Under the gold standard exchange rates were fixed. Currencies were defined in terms of gold. A loss or gain in domestic money supply corresponded directly with a loss or gain in gold. This would then be accompanied by an appropriate change in interest rates, GDP and prices. These changes would in turn be accompanied by appropriate changes in foreign investment and trade which would adjust the exchange rate to its appropriate price. As mentioned above domestic policy was restricted by such a system. Furthermore the inflation rate was dependent upon gold discoveries. The gold Standard collapsed during the chaos of the great depression of the 1930's.

In the intrests of economic efficiency the industrialized nations met in Bretton Woods, New Hampshire in 1944 to set up a system which would fix exchange rates. The United States pegged its currency to gold and the other nations pegged their currency to that of the United

States. The International Monetary Fund was set up to police the situation. Once again one of the major problems with such a system was that countries were forced to adopt monetary policies that may not have been in their best domestic interests. Furthermore, devaluations were permitted for countries only after long balance of payment deficits. This made such devaluations easy to predict and speculators were hence able to increase the magnitude of the devaluation. Revaluations expected from balance of payment surplus nations were not readily received and so the United States was forced to accept a balance of payments deficit. All of these problems combined to cause the fall of the Bretton Woods system in 1971.

In general exchange rate modeling attempts have been based on the theoretical relationships between the exchange rate and known (or approximately known) indicators of exchange rate movements. One favoured method used in exchange rate modeling is to look at the relationship between the spot rate and the forward rate. The spot rate is the price of one currency in terms of another currency. The forward rate is the price at which one currency can be purchased in terms of another currency at some prespecified time in the future. The usual setup is:

$$S_{t+1} = \rho + F_t + \epsilon_{t+1},$$

where S_t is the spot rate at time t, F_t is the forward rate at time t and ρ is a risk premium. Often ρ is taken to be equal to zero.

Another approach to exchange rate modeling is to assume the approximate satisfaction of theoretical money demand retationships. These type of models are called monetary models and are quite common in the economic literature. See for instance Bilson (1978, 1979), Frenkel (1976), Frankel (1979) and Hooper and Morton (1982). The models that fall under this category are usually based on the money demand equation. For a given country the aggregate supply of money is M/P where M is the nominal supply of money and P is the price level. The aggregate demand for money is of the form $m_d(Y, R, \ldots)$, where Y is real income and R is the prevailing interest rate. A widely used form for m_d is:

$$m_d = KY^{\alpha}e^{-\beta R},$$

where K is a country specific constant and α, β are independent of the country we are talking about. So at equilibrium,

$$M/P = KY^{\alpha}e^{-\beta R}$$

or,

$$P = \frac{M}{KY^{\alpha}e^{-\beta R}}.$$

Let quantity* denote the relevant quantities for some other country. Now the theory of purchasing power parity assumes that the rate of exchange between two countries is directly related to the relative price levels of those two countries. So

$$S = P/P^{*} = M/M^{*}\frac{K^{*}Y^{*\alpha}e^{-\beta R^{*}}}{KY^{\alpha}e^{-\beta R}},$$

where S is the exchange rate. Letting lower case letter=ln(upper case letter), we have

$$s = (m - m^{*}) + (k^{*} - k) + \alpha(y^{*} - y) + \beta(R - R^{*}).$$

For a further more detailed discussion of monetary models see Baillie and McMahon(1989).

In the previous two paragraphs an attempt has been made to introduce some of the prevalent models used in forecasting and empirical research. Again however these models have not been extremely successful in prediction or accurate modeling. There are several problems with the above approaches. The forward rate models seem viable since they take into account expectations concerning future market conditions and these expectations may be more important in a forecasting sense than the actual future state of affairs or predictions thereof. One problem is that the forward rates are constrained by the interest rate parity theory to be dependent primarily on the previous exchange rate and interest rates. Trade balance differentials and other economic indicators of exchange rate movements are restricted in their effect on the forward rate. To explain, interest rate parity theory implies that

$$F_t = \frac{(1 + R_d/4)}{(1 + R_f/4)}S_t,$$

where S_t is the rate of exchange for foreign currency in terms of domestic currency, F_t is the corresponding forward rate for one quarter ahead, R_d is the domestic annual interest rate and R_f is the foreign interest rate. Thus if as the forward rate model implies, $S_{t+1} \approx F_t$, then

$$s_{t+1} \approx s_t + ln(1 + R_d/4) - ln(1 + R_f/4)$$

Note the similarity between this model and the monetary models stated above. For R small a Taylor series approximation suggests that $ln(1+R) \approx R$. Hence the model is approximately a special case of the above mentioned monetary models with some coefficients set equal to zero and a random walk component.

The monetary models also have some difficulties. For one, there is no direct attempt to include more than two countries in the models. It seems sensible that the exchange rate is determined by the explanatory variables of more than two nations. Another problem is that as certain explanatory variables approach inordinately high or low values governments, central banks and foreign investors will take note and take action. The implication is that the values of parameters may not be stable across all regions of the explanatory variables. It is primarily these last two criticisms that the models to be suggested intend to deal with. Ideally, the more explanatory variables we include the more information we gain. However due to the limited number of observations some type of trade-off must be made. By taking principal components with respect to explanatory variables across countries it is hoped that some information can be gained about the relationships between the exchange rate and the explanatory variables of several countries as opposed to just two. By considering a model that is segmented with respect to the values of an explanatory variable it is hoped that the problem of instability of the parameter values may be dealt with.

1.2 Determination of Exchange Rates

There are several approaches to exchange rate determination. One idea is that of Purchasing Power Parity. Under this theory the exchange rate between any two countries is believed to reflect the relative price levels of those two countries. An implication of this theory is that inflation rates are the major determinants of exchange rates. However, empirically such relationships have been anything but clear in the 1970's and 1980's. Still this theory was widely accepted pre 1971 for explaining long range behaviour of exchange rates. It would seem that this theory is useful in explaining exchange rate behaviour during fixed exchange rate periods.

Another approach is to consider the exchange rate as an asset value. Under this approach the exchange rate one period ahead is determined by the present exchange rate plus some expected changes in the exchange rate market. The problem then arises as to how to account for expected changes. One thing to do is look at the forward rates, as mentioned in the introduction, this approach is useful in that it takes expectations about future events into account and often these expectations can have more relevance in exchange rate prediction than the actual occurrence of future events. These expectations will be dependent upon monetary sources, but it should also be said that these expectations will be dependent on other things, such as news. A case in point is the upward movements in the exchange rates throughout many of the industrialized nations (notably Germany) in 1980. It is hypothesized that these changes were (Issard, 1983) at least partly a result of the news that Ronald Reagan could be expected to win the upcoming presidential election. With his support for tight monetary policy and his support for the stimulation of U.S. competetiveness it was expected that the dollar would increase in value. The point is that the exchange rate increased irrespective of whether U.S. monetary policy actually was tight, the determining factor was the expectation as opposed to the actual occurence of the event. It should be added that asset model aproaches are interlinked in some sense with monetary models. Expectations about future conditions will be dependent upon news about political situations and the like but they will also be dependent on what the current or past state of affairs is like with respect to key economic variables. Furthermore it has been mentioned that at least theoretically the forward rate approach can be approximated by monetary approach with constrained variables and a random walk component.

The final approach I will consider with respect to exchange rate determination, and the approach that forms the basis for my models, is that of monetary models. The idea is that the exchange rate is the price of one country's currency in terms of another country's currency and hence that the laws of supply and demand apply. Thus, this approach looks at economic variables which are considered important in the determination of the supply of and demand for a country's money by foreign interests. Some key factors in such a determination would be:

1. Trade Balance and Current Account - A country sells its exports in it's home currency. So holding all other variables constant an increase in the demand for a country's currency would correspond to an increase in exports.

2. Foreign Investment - An increase in investment in a country would be accompanied by an increase in demand for that country's currency in order to finance the investment.

3. Interest Rates - An increase in interest rates attracts foreign investment.

4. Money Supply - An increase in a country's domestic money supply implies an increase the money supplied to foreign investors.

5. Inflation rates - rising inflation rates make a country's currency less favourable to foreign investors.

Such factors are important in the determination of exchange rates but certain problems and ideas need to be kept in mind when using these models. Changes in such indicators would not usually correspond to simultaneous changes in money supply and demand this makes it necessary to investigate the use of explanatory variables at different lags. Volatility may confuse relationships, large changes in economic indicators may cause concern as to the health of a nation's economy thus resulting in the reverse of or at least tempering of the effect on the exchange rate. Many of these variables are endogenous and thus again expected effects may be tempered. Another concern should be policy changes, they may change the relationships between these variables and certainly affect expectations about these variables and exchange rates (the "Lucas Critique", Frenkel 1983). It is important to consider changes between countries with respect to these variables. For instance, holding all other variables constant an increase in one country's interest rates should have no effect on exchange rates if the interest rates of all other nations increase at the same rate.

It is this approach with the above mentioned considerations taken into account that leads to the segmented regression models that I consider. I try to take the important determinants of supply and demand for foreign currency into account as explanatory variables. That is, the variables mentioned above (1-5) and subsets of these variables are used to determine the segmented regression relationship. Since international trade is greatest between the larger industrial nations it is reasonable to restrict attention to the exchange rate and explanatory variables corresponding to these nations. Thus the possible explanatory variables were chosen to be the differentials of the important determinants between G-7 nations.

Suppose we can assume that government policies of the major trading partners are more or less *rational* over the years and that interest rates and money supplies digest market information relatively efficiently. Then since policy decisions will often be determined or acompanied by hi-lo values of these variables it seems reasonable to consider seperate regression segments with respect to some segment explanatory random variable. For example governments are certain to act in the case of large trade balance deficits or large interest rate differentials. What emerges is a model where the exchange rate is dependent upon certain explanatory variables and where this relationship is different at extreme values of the explanatory variables.

The modeling considered is inherently long run modeling. Changes in these determinants and changes in the exchange rates will not be simultaneous, there will be some time lag. Furthermore, this time lag may be of variable length with repect to, say, monthly periods. Hence monthly models may not be appropriate in that the time lag chosen as 'best' may not be constant with respect to time. If the time periods concerned are increased to say quarterly periods then the models should be more robust with respect to this assumption of constant time lag. For this reason it was decided to concentrate on long run modeling.

Section 2

Data Analysis

2.1 Introduction

In section 1 it was mentioned that policy considerations may have effects on the relationships between exchange rates and certain economic indicators. Thus a segmented regression model could be appropriate. In this section the results of fitting various segmented models to the German exchange rate movements post 1971 are discussed.

The general form of a segmented time series regression model can be stated as

$$Y_t = \mathbf{x}'_{t-1}\beta_i + \epsilon_t, \ if \ x_{td} \in (\tau_{i-1}, \tau_i], \ i = 1, \ldots, l+1,$$

where Y_t is the exchange rate at time t, $\mathbf{x}_t$ is a vector of explanatory variables mentioned above at time $t-1$, ϵ_t is an error term, x_{td} is one of the components of $\mathbf{x}_t$ (the segmentation variable), $-\infty = \tau_0 < \tau_1 < \cdots < \tau_{l+1} = \infty$ and $\{\mathbf{x}_t\}$ and $\{\epsilon_t\}$ are independent series.

2.2 Explanatory Variables

To use a segmented model it was neccesary to make certain decisions about explanatory variables. Some important theoretical determinants of exchange rates were discussed in section 1 and are listed below

1. Interest Rate Differentials
2. Trade Balance Differentials
3. Money Supply Differentials
4. Inflation Rate Differentials

Some things should be mentioned at this point. It is the differential that is important. In section 1 it was pointed out, for instance, that if a country increases its supply of money that this will in general make its currency less attractive. This will not be true if every other country also increases its money supply by an appropriate amount. Hence the importance of differentials. However if differentials between all countries are considered there will be hundreds of explanatory variables. In view of this it was decided to consider differentials between important trading partners and economic powers. In particular it was decided to consider the differentials between Germany and the other G-7 nations. Preliminary simultaneous time series plots and scatter plots of the exchange rate vs. inflation rates suggested that inflation rate differentials would not be very useful as explanatory variables and so it was decided not to consider them. Further the effects of the inflation rate differentials are expected to be imbedded in the interest rate differentials. Thus the explanatory variables to be considered were interest rate differentials, money supply differentials, and trade balance differentials all between Germany and the G-7 nations. It was important to consider these variables with some time lag since it may be that the effect of an explanatory variable on the exchange rate will not be felt immediately. Thus the explanatory variables were considered with and without time lags. Finally some sort of standardization, to be mentioned shortly, was necessary.

In view of the fact that exchange rates were relatively fixed via the Breton Woods Monetary System until 1971 and that the explanatory variables are theoretically more valuable as long term predictor than as short term predictors, it was decided to use quarterly data from 1971 through to the second quarter of 1990. The decision to use quarterly data as opposed to monthly data was partly because of some of the concerns mentioned in the previous section and partly because of some practical problems related to the availability and/or reliability of certain economic data on a monthly basis. To repeat at least one of the concerns mentioned in the previous section it can be expected that the lag used in the explanatory variables is more likely to be variable with respect to time in the monthly model than in the corresponding quarterly model. To avoid this problem without adding parameters a useful approach is to consider quarterly data. For each G-7 country the following data were obtained from the international financial statistics published by the IMF (International Monetary Fund).

1. Exchange Rates - These were end of period spot rates. They were expressed as Marks/U.S. Dollar.

2. Money supply - This was taken to be M1 money. It was calculated as demand deposits plus currency in circulation. For each country money supply was given in that country's currency.

3. Interest Rates - This was the discount rate/Bank rate for the country of interest.

4. Trade Balance - This was calculated as merchandise imports minus merchandise exports. This quantity was given in U.S. Dollars.

5. C.P.I - Consumer Price Index. This quantity was used as the index for inflation of a particular country.

Several things should be mentioned with respect to the data. Some standardization is necessary. In particular one wants to consider real values. In order to do so it was necessary to adjust for the rate of inflation. Thus to get real money supply the nominal money supply was multiplied by $(100/CPI_{country})$. To get real trade balance nominal trade balance was multiplied by $(100/CPI_{US})$. To get real bilateral exchange rates, exchange rates were multiplied by $CPI_{US}/CPI_{Germany})$, since the exchange rate can be regarded as the price in Deutsche Marks for one U.S. Dollar. Since money supplies were given in home currencies a further adjustment was necessary to obtain money supply differentials. Each country's money supply was standardized, i.e. (quantity-mean(quantity)/s.e.(quantity)).

Now even in the presently described situation where only differentials between G-7 nations are considered there are still 3x6=18 explanatory variables. Since there are 78 observations this would have been too many explanatory variables for a reasonable segmented model. Thus it was decided to consider as possible explanatory variables the principal components for each differential which explained most of the within differential variability. For example the principal components from (Germany-Canada interest rate differential, Germany-France interest rate differential,...,Germany-U.S. interest rate differential). Another possible approach was to consider the principal components as defined in Tsay (1990). These principal components are optimal in a predictive sense. Models using these principal components were explored.

2.3 Plot Suggestions

The next step in the analysis was to consider the plots of the various explanatory variables vs. the exchange rates. Of course the explanatory variables are related to each other and hence two or even three dimensional plots will not give the complete picture. The plots are shown on pages 39-54. Time series plots are given on pages 39-44. Through each of the plots the real

German exchange rate movements are traced with a solid line. The corresponding scale is given on the left hand side of the plot. Each plot also gives the sample path of some explanatory variable with a broken line. The corresponding scale being given on the right hand side of the plot. Notice the volatility of the exchange rate particularly post 1980. The period in the mid '80's is particularly volitile. This 'hill' is not seen in many of the other time series plots although there is some indication of it in for instance the time series plot of standardized money supply differential with respect to Germany and the U.S.. Also, notice the time series plots of the trade balance differentials of the G-7 nations given on page 40. These plots hint at the interdependence of the exchange rates and the chosen explanatory variables. The plots of the trade balance differentials often exhibit 'hills' and 'valleys' after that same behavior is observed in the exchange rates.

The next type of plots to be observed are the two dimensional scatter plots given on pages 45-50. These plots are labeled lag plots in reference to the fact that they plot the explanatory variable listed at the top of the page vs. exchange rates at six different time lags. These plots must be observed with caution of course since the explanatory variables are certain to be interrelated. For instance money supply differentials are definitely going to be dependent upon prevailing interest rates to a certain extent. The plots are not altogether encouraging. For example the relationships between money supply differentials, trade balance differentials and exchange rates are not particularly evident in the corresponding scatter plots of the exchange rates vs. the first principal components. However, the plot of the exchange rates vs. the first principal component of the interest rate differentials does suggest a relationship, in fact it appears from the plots that a segmented approach with respect to this explanatory variable may be the way to go. For a given explanatory principal component notice the relative homogeneity of the plots of the exchange rates vs. explanatory principal components across different time lags. In view of this apparent similarity with respect to lags it was decided that it may not be neccesary to fit a lot of models with many different lag combinations. Experimentation with a few lags, to be mentioned below, confirmed that this was indeed the situation. Also, notice that the scatter plots of the exchange rates vs. the interest differentials seem to hint at the existence of at least two segments. The plot of the exchange rate vs. the exchange rate lagged one quarter is given on page 53. Clearly there is a very good linear relationship. Hence it was decided to include the exchange rate lagged one quarter in all further modeling attempts. The plot of the forward rate vs the spot rate one period forward is given on page 54. Again there is a clear linear relationship although the variability is somewhat greater than that in the plot of the exchange rate vs. the exchange rate lagged one quarter. Still the plot does lend credence to the forward rate models.

It has been mentioned that two dimensional plots may not be sufficient in exploring the possible relationships in the data. On pages 51-52 several three dimensional plots are given. On page 51 three dimensional plots of each of the first explanatory principal components with the exchange rates lagged one period and at present are given. The latter five plots do not give much suggestion to possible relationships along the direction of the explanatory variable but the plot concerning the principal component for the interest differentials does suggest a possible segmented relationship with possibly three segments. The plot on page 52 rotates the three dimensional plot of the first principal component for the interest rate differentials. As in the two dimensional plot a segmented relationship appears possible and it appears that a segmented model with possibly two or three segments may be appropriate.

In table 3, some forecasting results are quoted. One stunning aspect of this table is with respect to the models with Interest Differential PC1 as segmentation variable and Money Supply PC1 as an additional segmentation variable. There is quite a large dicrepency between the model with two segments and the model with three segments. The model with three segments performs substantially better than the model with two segments. Thus an immediate question is to what extent the number of segments could have been predicted via exploratory plots. One such plot is given on page 55. $Y_t - \hat{\beta}_i Y_{t-1}, \quad i = 1 \ldots l$) versus lag-1 interest rate differential and money supply differential was plotted, where l was any plausible number of segments and $\hat{\beta}_i$ the estimated coefficients. The plots were then rotated to get images from different angles. The plot shown is the one corresponding to $l = 3$, and it does appear in this plot that three segments are appropriate.

2.4 Results from Model Fitting

In this subsection the fitting of various models is discussed. The general form of the models can be stated as

$$Y_t = \mathbf{x}'_{t-1}\beta_i + \epsilon_t, \ if \ x_{td} \in (\tau_{i-1}, \tau_i], \ i = 1, \ldots, l+1,$$

where Y_t is the exchange rate at time t, $\mathbf{x}_t$ is a vector of explanatory variables at time $t-1$, ϵ_t is an error term, x_{td} is one of the components of $\mathbf{x}_t$ (the segmentation variable) and $-\infty = \tau_0 < \tau_1 < \cdots < \tau_{l+1} = \infty$. Note that the special case where $l = 0, \tau_0 = -\infty, \tau_1 = \infty$ is the familiar linear regression setup. The situation where $l > 0$ is that of segmented regression, namely a linear regression setup is assumed on 'segments' of a particular explanatory variable. The types of models that were fitted can be categorized. The first type is the segmented models with standard principal components. These models are segmented models with various combinations of the explanatory variables mentioned in section 2.2 each model having one explanatory variable determining the segmentation. The second type of models is the segmented models with Tsay type principal components. These models are the same in spirit as the previous type of models but the principal components adopted were taken according to Tsay (1990) in order to maximize the lag 1 autocorrelation of a linear combination of the variables of interest. There are two other types of models, these are competing models in the sense that they are the favoured models in the economic literature. These models have been described in the first section where the general problems in exchange rate determination were discussed. The first of the two would be monetary models. The form and theoretical background behind these models was discussed in the first section. For previous analysis of these models using monthly data pre 1980's see Meese and Rogoff (1983). The other type of competing model is the forward rate model. Again, some of the merits and pitfalls of using this type of model have been discussed in section 1.

Segmented regression is a form of nonlinear regression and criterion for testing model assumptions and parameter values are not well developed and would involve assumptions about independent identically distributed errors. As a practical issue the first thing we are compelled to consider is the appropriate number of segments. This issue is discussed in more detail in the next section as it pertains to large samples. In the present situation we do not have a large sample. For small samples this issue has not been resolved and often the best policy is to experiment with different numbers of segments perhaps as suggested in the plots. Furthermore small samples place restrictions on the ability to estimate large numbers of parameters and thus some further retrictions must be placed upon the number of segments. However criteria are needed to adress the issue of model appropriateness. The primary criteria considered were mean squared error (MSE), and the sum of out of sample one step ahead forecasting errors for a five year period (SSFE) from the second quarter of 1985 to the second quarter of 1990. The quantities from the random walk model were used as the yardstick to measure performance. Some residual analysis was done as well. Scatter plots of the residuals acf plots and time series plots of the residuals were examined.

The results for the segmented models with standard principal components are listed on pages 31-34. A variety of combinations of explanatory variables, segmentation variables and lags were experimented with. In most cases the reduction in mse over the random walk was similar for models having the same explanatory and segmentation variables but with different lags. This tends to agree with the plot suggestions. It does not appear that choosing the correct lag for the explanatory variables is of utmost importance. This may be a result of the fact that quarterly as opposed to monthly data were used. The models in which the segmentation variable was the principal component from the interest differentials had significant reductions in MSE over those models which used a different segmentation variable. Again this was indicated to a certain extent by the plots. From the results corresponding to MSE it was decided to use the models which had the principal component from the interest rate differentials as the segmentation variable to evaluate forecasting ability. As is seen on page 34 the reduction in one step ahead sum of squared forecasting error (SSFE) over the random walk model for a five year period (from the second quarter of 1985 to the second quarter of 1990) was as large as 42 percent in the case where the principal components from the interest differentials and money supply differentials were the explanatory variables. This result is quite impressive in view of the fact that it has been widely accepted that large reductions in forecasting ability over the random walk will not usually be obtained.

The results for the segmented models using Tsay type principal components are listed on pages 36-38. The results are similar to the results obtained using the standard principal components. In fact plots of the principal components suggested that the principal components obtained in this fashion were very similar to the principal components obtained the standard way. The results using standard principal components tended to be somewhat better than the results using Tsay type principal components but it seems clear from the similarity of the results that the appropriateness of the principal components taken was not that much of an issue.

The results from fitting forward rate models are given on page 30. The forecasts came from the models

$$S_{t+1} = F_t + \epsilon_t,$$

and

$$S_{t+1} = \rho + F_t + \epsilon_t.$$

The scatter plot of of the exchange rate vs. the forward rate in conjunction with the interest rate parity theory suggested that it would be appropriate to ignore the risk premium ρ. So analyses were performed with and without the risk premium. When S_t, F_t were expressed in nominal terms there was a small reduction in forcasting error over the random walk. As mentioned before large gains cannot reasonably be expected from these types of forecasts since $F_t \approx S_t$ due to arbitrage concerns. When models with adjustments for inflation were included there was no reduction in forecasting errror over the random walk model.

The results of fitting the monetary models are listed on page 35. Some discussion of these models was given in the first section. For a more complete discussion of these models see Meese and Rogoff (1983). The underlying equation for these models is

$$s = a_0 + a_1(m - m^*) + a_2(y - y^*) + a_3(r_s - r_s*) + a_4(\pi_e - \pi_e^*) + a_5 TB + a_6 TB^* + u,$$

where s is the logarithm of the price of dollars in term of foreign currency, m-m* the logarithm of foreign to U.S. money supply, y-y* the logarithm of foreign to U.S. real income, $r_s - r_s^*$ the interest rate differential, and $\pi_e - \pi_e^*$ is the expected inflation differential. TB and TB* represent the foreign and U.S. cumulative trade balances. The respective bank rates were taken as the short term interest rates. The expected inflation rate was taken to be the inflation rate of the previous period. In detail, $\pi_e = (CPI_t - CPI_{t-1})/CPI_{t-1}$. From this general formulation come the three models that were analyzed. The flexible price (Frenkel-Bilson) monetary model in which $a_4 = a_5 = a_6 = 0$. The sticky price (Dornbusch-Frankel) monetary model in which $a_5 = a_6 = 0$ and the sticky price (Hooper-Morton) asset model in which none of the coefficients are zero. See Bilson (1978, 1979), Frenkel (1976), Dornbusch (1976), Frankel (1979, 1981) for further discussion of these models. These models were fit with no lag in the explanatory variables. The results were poor, none of the models fared better in terms of MSE or SSFE in comparison to the random walk model. The other type of model with which comparisons were made were vector autoregression (VAR) type models as described in Meese and Rogoff (1983). These models performed considerably better. They can be described by

$$s_t = a_{i1}s_{t-1} + a_{i2}s_{t-2} + \ldots a_{in}s_{t-n} + B'_{i1}X_{t-1} + B'_{i2}X_{t-2} + \ldots B'_{in}X_{t-n} + u_{it},$$

where X_{t-j} is the vector of the explanatory variables in the equation above lagged j periods. The results are listed on page 35. The best result was from the Dornbusch-Frankel model with two lags. This model achieved a reduction of 2.26 percent over the random walk model in SSFE. None of the other models reported a reduction over the random walk model. It should be mentioned that these results should be interpreted with caution when comparing to the segmented models. Since the results quoted for the segmented models were in terms of the exchange rate as opposed to the logarithm of the exchange rate the sums of squares were calculated in the same fashion for the monetary models. This tended to give more dramatic results than would have been obtained if the sums of squares had been computed directly using the logarithmic values. For instance the above mentioned 2.26 percent reduction would be a reduction of 12.26 percent when the ln transformation was not taken in computing SSFE. Still the results are striking in contrast to the superior results obtained from the segmented models.

A question that arises is, to what extent can the nonlinear srtucture noticed in the complete data set be seen in truncated versions of the data? One would hope, for instance, that the number of segments would remain constant. In order to attempt to answer this question several plots have been included which give a visual comparison of the data prior to 1985 with the full data set. These are given on pages 56 and 57.

2.5 Conclusions

The monetary models in general did poorly. There was an exception as mentioned above but even this model still performed poorly in comparison to the segmented model and the best reductions were comparable to the best reductions a decade ago as reported in Meese and Rogoff (1983).

The forward models gave slight reductions in SSFE over the random walk models. This can be expected since the interest rate parity theory implies that $F_t \approx S_t$. Hence large gains over a random walk model should never be expected.

The segmented models do seem to do better than the random walk model. Certainly more testing using the exchange rates of other nations, following up with further testing on the German exchange rates of the future and other variations on the theme are neccesary to make any statments with a degree of certainty. The main reason that this type of modeling was initiated was the concern about the effect of policy decisions, speculation and central bank actions in times when economic indicators are taking on extreme or unusual values. It appears that this may be a valid concern. The segmented models with the principal component from the interest differentials as the segmentation variable performed significantly better than many other models. It does appear that there is some segmentation with respect to this variable. Perhaps this can be attributed to market efficiency in some manner. When differentials are large central banks and speculators will be forced to pay attention but when the differentials are small it may be that only the more informed, aware and less risk averse partipants act.

Section 3

Estimating the Number of Segments

3.1 Introduction

Segmented models may be useful in many situations. For example Yeh et al.(1983) discuss the the idea of an 'anaerobic threshold'. It is hypothesized that if a person has his workload steadily increased through some form of exercise there comes a point where the muscles cannot get enough oxygen and what were anaerobic metabolic processes become aerobic processes. This point is referred to as the 'anaerobic threshold'. In this situation two segments are what is suggested by the subject oriented theory. So it would be natural for the modeler to fit a model with two segments. However in some situations it may be suspected that a segmented model should be adopted but the appropriate number of segments may not be known. For instance in the exchange rate problem it is suspected that a segmented model is appropriate due to policy changes. It is not, however, clear how many segments will be necessary beforehand. One immediate approach to this problem is to graphically attempt to determine how many thresholds seem to be appropriate. This is worthwhile as a first step and in the case of a single explanatory variable but may not be appropriate in the multivariate case. In the multivariate case the interrelationships of the explanatory variables may confuse such an approach. Further this approach lacks objectivity, some sort of automated rule is desired. In this section I discuss a consistent procedure for identifying the number of segments.

3.2 A Criterion for Estimating the Number of Segments

Consider the following segmented linear regression model.

$$Y_t = \mathbf{x}_t'\beta_i + \epsilon_t, \ if \ x_{td} \in (\tau_{i-1}, \tau_i], \ i = 1, \ldots, l+1,$$

where $\epsilon_t = \sum_0^\infty \psi_i z_{t-i}$, $\sum_0^\infty |\psi_i| < \infty$, ,with the $\{z_t\}$ iid, mean zero and variance σ^2 and independent of $\{\mathbf{x}_t\}$, $\mathbf{x}_t = (1, x_{t1}, \ldots, x_{tp})'$ and $-\infty = \tau_0 < \tau_1 < \cdots < \tau_{l+1} = \infty$. Further we assume that there exists $\delta > 3/2$, $k > 0 \ni |\psi_i| \leq k/i^\delta \ \forall i$. Note that this implies that $\{\epsilon_t\}$ is a stationary ergodic process.

Consider the following regression setup.

$$Y_t = \mathbf{x}_t'\beta + \epsilon_t.$$

Recall that the least squares estimate of β is given by

$$\beta = (X_n'X_n)^{-1}X_n\mathbf{Y_n},$$

and the sum of squared error is given by

$$S_n = \mathbf{Y_n}'(I - H_n)\mathbf{Y_n},$$

where $H_n = X_n(X_n'X_n)^{-1}X_n'$. The situation here in the segmented regression case is completely analogous.

Let

$$I_n(\alpha,\beta) := diag(\mathbf{1}_{(x_{1d}\in(\alpha,\beta])}, \ldots, \mathbf{1}_{(x_{nd}\in(\alpha,\beta])}), \forall\ \alpha < \beta,$$

$$X_n := \begin{pmatrix} \mathbf{x}_1' \\ \vdots \\ \mathbf{x}_n' \end{pmatrix},\ \mathbf{Y}_n := \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \tilde{\epsilon}_n := \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}, X_n(\alpha,\beta) = I_n(\alpha,\beta)X_n,$$

and

$$H_n(\alpha,\beta) := X_n(\alpha,\beta)[X_n(\alpha,\beta)'X_n(\alpha,\beta)]^{-}X_n(\alpha,\beta)',$$

where in general A^{-} will denote a generalized inverse for any matrix A, and $\mathbf{1}_{(.)}$ is the indicator function,

$$S_n(\alpha,\beta) := \mathbf{Y}_n'(I_n(\alpha,\beta) - H_n(\alpha,\beta))\mathbf{Y}_n,\quad S_n(\tau_1,\ldots,\tau_l) := \sum_{j=1}^{l+1} S_n(\tau_{i-1},\tau_i),$$

and

$$T_n(\alpha,\beta) := \tilde{\epsilon}_n'H_n(\alpha,\beta))\tilde{\epsilon}_n, \hat{\mathbf{Y}}_n(\alpha,\beta) := H_n(\alpha,\beta)\mathbf{Y}_n.$$

Then, in terms of true parameters, our model can be rewritten in the vector form,

$$\mathbf{Y}_n = \sum_{i=1}^{l_0+1} X_n(\tau_{i-1},\tau_i)\tilde{\beta}_i + \tilde{\epsilon}_n.$$

The estimation of all the parameters is done primarily in two steps. First we estimate l, the number of thresholds $\tau_1, \ldots, \tau_{l^0}$. This is done by minimizing the modified Schwarz' criterion

$$MIC(l) := \ln[S(\hat{\tau}_1, \ldots, \hat{\tau}_l)/(n-l)] + l\frac{c_0(\ln n)^{2+\delta_0}}{n},$$

for some constants $c_0 > 0, \delta_0 > 0$, where for any fixed l, $\hat{\tau}_1, \ldots, \hat{\tau}_l$ are the least squares estimates which minimize $S_n(\tau_1, \ldots, \tau_l)$ subject to $-\infty = \tau_0 < \tau_1 < \cdots < \tau_{l+1} = \infty$. With the available estimates, $\hat{l}$, and, $\hat{\tau}_i$, $i = 1, \ldots, \hat{l}$, we then estimate the other regression parameters $\{\tilde{\beta}_i\}$ and the residual variance σ_0^2 by the ordinary least squares estimates,

$$\hat{\tilde{\beta}}_i = [X_n(\hat{\tau}_{i-1}, \hat{\tau}_i)' X_n(\hat{\tau}_{i-1}, \hat{\tau}_i)]^- X_n(\hat{\tau}_{i-1}, \hat{\tau}_i)\mathbf{Y}_n, \ i = 1, \ldots, \hat{l}+1,$$

and

$$\hat{\sigma}_0^2 = S_n(\hat{\tau}_1, \ldots, \hat{\tau}_l)/(n-\hat{l}).$$

Under some regularity conditions essential to the identifiability of the regression parameters, we shall see below that the ordinary least squares estimates $\hat{\tilde{\beta}}_j$ will be unique with probability approaching 1, for $j = 1, \ldots, \hat{l}+1$, as $n \to \infty$.

After estimation of the $\hat{\tilde{\beta}}_j$ is completed we can then use the estimated residuals $\hat{\epsilon}_t = Y_t - \mathbf{x}_t'\hat{\beta}_i, if x_{td} \in (\hat{\tau}_{i-1}, \hat{\tau}_i]$ to estimate the model for the $\epsilon_t's$.

3.3 Consistent Estimation of the Number of Segments

Consider the segmented linear regression model discussed in the previous section. Let $\hat{l}$ minimize $MIC(l)$. To identify the number of thresholds l, and hence the number of segments consistently, assume:

Condition 1

The sequence $\{\mathbf{x}_t\}$ is strictly stationary and ergodic and has positive definite matrices, $E\{\mathbf{x}_t\mathbf{x}_t'$ $\mathbf{1}_{(x_{td}\in(\tau_i-\delta,\tau_i))}\}$ and $E\{\mathbf{x}_t\mathbf{x}_t'\mathbf{1}_{(x_{td}\in(\tau_i,\tau_i+\delta))}\}$ within a small δ-neighbourhood of each of the true thresholds $\tau_1, \ldots, \tau_l$, for any t or

Condition 2

if the covariates are not random, $(1/n)\sum_{t=1}^n \mathbf{x}_t\mathbf{x}_t'\mathbf{1}_{[x_{td}\in(\tau_i-\delta,\tau_i]]}$ and $\frac{1}{n}\sum_{t=1}^n \mathbf{x}_t\mathbf{x}_t'$ $\mathbf{1}_{[x_{td}\in(\tau_i,\tau_i+\delta]]}$ converge to positive definite real matrices for $\delta \in (0, \min_{1\le j\le l}(\tau_{j+1}-\tau_j)/4)$.

Under *Condition 1*, the design matrix $X_n(\alpha, \beta)$ has full column rank a.s. as $n \to \infty$ for every open interval (α, β) in the small neighborhood of the true thresholds $\tau_i, i = 1, \ldots, l$ for which x_d has positive probability density. And $X_n(\alpha, \beta)$ will have full column rank for large n and for every open interval (α, β) in the small neighborhood of $\tau_i^0, i = 1, \ldots, l^0$ for which *Condition 2* is satisfied. So $\hat{\hat{\beta}}_i$ will be unique with probability tending to 1 as $n \to \infty$, for $i = 1, \ldots, \hat{l}$, provided that $\hat{l}$ converges to l, the true number of thresholds in probability.

When the segmented regression model reduces to the segmented polynomial regressions or functional segmented regressions as discussed by Feder (1975), then a similar condition to either *Condition 1* or *Condition 2* is essential for identifying the segmented model parameters. For details, see Feder (1975). In particular, for the segmented polynomial regression model, Condition 1 is automatically satisfied if the key covariate x_d has positive density within a small neighborhood of each of the thresholds.

In addition, we need to place some restriction on the distribution of the iid errors $\{z_t\}$. And this is the so-called *local exponential boundedness condition.* A random variable Z is said to be *locally exponentially bounded* if there exist constants c_0 and T_0 in $(0, \infty)$ such that

$$E(e^{uZ}) \leq e^{c_0 u^2}, \ \forall \ |u| \leq T_0. \tag{3.1}$$

Many of the commonly used distributions such as the normal, the symmetrized Poisson and exponential distributions have such a property.

In reality, the sample size n is always finite and hence the number of thresholds that can be effectively identified is always bounded. So we will assume throughout that there always exists an upper bound L of the true nummber of thresholds. Another simplification we gain in the nonlinear minimization of $S(\tau_1, \ldots, \tau_l)$ is obtained by limiting the possible values of $\tau_1 < \ldots < \tau_l$ to the finite discrete set, $\{x_{1d}, \ldots, x_{nd}\}$. This restriction induces no loss of generality.

Theorem 1 *Consider the segmented linear regression model with X_n independent of $\tilde{\epsilon}_n$. Suppose $\{z_t\}$ are iid with a locally exponentially bounded distribution having mean zero and variance σ^2. Assume for the true number of thresholds, l, that $l \leq L$ for some specified upper bound $L > 0$ and that one of Conditions 1 or 2 is satisfied. Then $\hat{l}$ converges to l in probability as $n \to \infty$.*

The proof of the theorem will be given after a series of related lemmas.

Lemma 1 a Suppose $\sum_{i=1}^{\infty}|a_i| < \infty$ and $|a_i| \leq k/i^{\delta}$ for some $k > 0,\ \delta > 3/2$. Then $\sum_{i=1}^{\infty}(\sum_{l=1}^{\infty}|a_{l+i}|)^2 < \infty$

Proof: By assumption $|a_i| \leq k/i^{\delta}$ for some $k > 0,\ \delta > 3/2$. Therefore,

$$\sum_{i=1}^{\infty}(\sum_{l=1}^{\infty}|a_{l+i}|)^2 \leq k_2 \sum_{i=1}^{\infty}(\sum_{l=1}^{\infty}\frac{1}{(i+l)^{\delta}})^2.$$

Now,

$$\begin{aligned}
\sum_{l=1}^{\infty}\frac{1}{(i+l)^{\delta}} &= \sum_{j=i+1}^{\infty} 1/j^{\delta} \\
&= \sum_{j=i+1}^{\infty} 1/j^{\delta}\int_{j-1}^{j} dt \\
&= \sum_{j=i+1}^{\infty} \int_{j-1}^{j} 1/j^{\delta} dt \\
&= \sum_{j=i+1}^{\infty} \int_{j-1}^{j} \min_{j-1\leq t\leq j} 1/t^{\delta} dt \\
&\leq \sum_{j=i+1}^{\infty} \int_{j-1}^{j} 1/t^{\delta} dt \\
&= \int_{i}^{\infty} 1/t^{\delta} dt \\
&= \frac{1}{(\delta-1)i^{\delta-1}}.
\end{aligned}$$

So,

$$\sum_{i=1}^{\infty}(\sum_{l=1}^{\infty}|a_{l+i}|)^2 \leq k_3 \sum_{i=1}^{\infty} 1/i^{2(\delta-1)}.$$

By assumption, $\delta > 3/2$, so $2(\delta-1) > 1$, and hence

$$\sum_{i=1}^{\infty}(\sum_{l=1}^{\infty}|a_{l+i}|)^2 < \infty.$$

Lemma 1 *Let $\{z_t\}$ be i.i.d. locally exponentially bounded random variables, i.e. assume for some $T_0 > 0$ and $0 < c_0 < \infty$, $E(e^{tZ_1}) \leq e^{c_0 t^2}$ for $|t| \leq T_0$. Let $\epsilon_t = \sum_0^{\infty}\psi_i z_{t-i}$ where we assume that there exists $\delta > 3/2,\ k > 0 \ni |\psi_i| \leq k/i^{\delta}\ \forall i.$. Let $S_k = \sum_{i=1}^{k} a_i\epsilon_i$, where the $a_i^{\prime}s$ are constants. Then there exists $0 < c < \infty$, $0 < t_0 < T_0$ such that for any $x \geq 0$,*

$$P\{|S_k| \geq x\} \leq 2e^{-t_0 x + c t_0^2 \sum_{i=1}^{k} a_i^2}.$$

Proof It follows from Markov's inequality that for $0 < t_0 < T$ satisfying $|t_0 a_i| \leq T$ for all $i \leq k$,

$$P\{S_k \geq x\} = P\{e^{t_0 S_k} \geq e^{t_0 x}\} \leq e^{-t_0 x} E(e^{t_0 S_k})$$

$$S_k = \sum_1^k a_i \epsilon_i = \sum_1^k a_i \sum_0^\infty \psi_j z_{i-j} = A(k) + B(k)$$

where,

$$A(k) = \sum_{i=0}^{k-1} z_{k-i} \sum_{j=0}^{i} a_{k-j} \psi_{i-j}$$

$$B(k) = \sum_{i=0}^{\infty} z_{-i} \sum_{j=1}^{k} a_j \psi_{j+i}.$$

$$E(e^{(B(k))}) \leq e^{c_0 \sum_{i=0}^\infty (\sum_{l=1}^k a_l \psi_{l+i})^2} \leq e^{c_0 ||a||^2 \sum_{i=0}^\infty (\sum_{l=1}^k |\psi_{l+i}|^2)}.$$

From the previous lemma there exists M_1 such that

$$\sum_{i=0}^{\infty} (\sum_{l=1}^{k} |\psi_{l+i}|)^2 \leq \sum_{i=0}^{\infty} (\sum_{l=1}^{\infty} |\psi_{l+i}|)^2 \leq M_1.$$

Therefore,

$$E(e^{(B(k))}) \leq e^{c_1 ||a||^2}.$$

$$E(e^{A(k)}) \leq e^{c_0 \sum_{i=0}^{k-1} (\sum_{j=0}^{i} a_{k-j} \psi_{i-j})^2}.$$

$$\sum_{i=0}^{k-1} (\sum_{j=0}^{i} a_{k-j} \psi_{i-j})^2 \leq \sum_{i=0}^{k-1} \psi_i^2 \sum_{j=i+1}^{k} a_j^2 + 2 |\sum_{i<j} \psi_i \psi_j \sum_{l=i+1}^{k-(j-i)} a_l a_{l+(j-i)}|.$$

Now,

$$\sum_{l=i+1}^{k-(j-i)} |a_l||a_{l+(j-i)}| \leq \sum_{l=1}^{k-1} |a_l||a_{l+1}| \leq (\sum_{l=1}^{k-1} a_l^2)^{1/2} (\sum_{j=2}^{k} a_j^2)^{1/2} \leq ||a||^2.$$

So,

$$\sum_{i=0}^{k-1} (\sum_{j=0}^{i} a_{k-j} \psi_{i-j})^2 \leq ||a||^2 (\sum_{i=0}^{k-1} \psi_i^2) + ||a||^2 (2 \sum_{i<j} |\psi_i||\psi_j|)$$

$$= ||a||^2 (\sum_0^{k-1} |\psi_j|)^2$$

Therefore,

$$E(e^{A(k)}) \le e^{c_2||a||^2}.$$

Since A(k) and B(k) are independent we get that

$$P\{S_k \ge x\} \le e^{-t_0 x} e^{ct_0^2||a||^2}.$$

Finally, to conclude the proof,we note that

$$P\{S_k \le -x\} = P\{-S_k \ge x\}$$

Lemma 2 *Consider the segmented regression model with the design matrix X_n satisfying either Condition 1 or Condition 2. Assume that the iid errors $\{z_t\}$ are locally exponentially bounded and are independent of X_n. Then*

$$P\{\sup_{\alpha<\beta} T_n(\alpha,\beta) \ge \frac{9p_0^3}{T_0^2}\ln^2 n\} \to 0, \qquad as\ n \to 0, \tag{3.3}$$

where p_0 is the true order of the model and T_0 is the constant associated with the locally exponential boundedness of $\{z_t\}$.

Proof Conditioning on X_n, we have that

$$P\{\sup_{\alpha<\beta} T_n(\alpha,\beta) \ge \frac{9p_0^3}{T_0^2}\ln^2 n|X_n\} = P\{\max_{x_{sd}<x_{td}} \tilde{\epsilon}_n' H_n(x_{sd},x_{td})\tilde{\epsilon}_n \ge \frac{9p_0^3}{T_0^2}\ln^2 n|X_n\}$$

$$\le \sum_{x_{sd}<x_{td}} P\{\tilde{\epsilon}_n' H_n(x_{sd},x_{td})\tilde{\epsilon}_n \ge \frac{9p_0^2}{T_0^2}\ln^2 n|X_n\}.$$

Since $H_n(x_{sd},x_{td})$ is idempotent, it can be decomposed as $H_n(x_{sd},x_{td}) = W'\Lambda W$, where W is orthogonal and $\Lambda = diag(1,\cdots,1,0,\cdots,0)$. Noting that $rank(A) \ge rank(AB)$ we get the following results:

$$\begin{aligned} rank(H_n(x_{sd},x_{td})) &= rank(X_n(x_{sd},x_{td})(X_n(x_{sd},x_{td})'X_n(x_{sd},x_{td}))^- X_n(x_{sd},x_{td})') \\ &\le rank(X_n(x_{sd},x_{td})) \le p_0 \end{aligned}$$

$$\begin{aligned} rank(H_n(x_{sd},x_{td})) &= rank(W'\Lambda W) \\ &\ge rank(W'\Lambda WW') \\ &= rank(W'\Lambda) \ge rank(WW'\Lambda) = rank(\Lambda) \end{aligned}$$

$$rank(H_n(x_{sd}, x_{td})) = rank(W'\Lambda W) \le rank(\Lambda).$$

Thus,

$$\begin{aligned} p_0 \ge p := rank(H_n(x_{sd}, x_{td})) &= rank(\Lambda) \\ &= trace(\Lambda) \\ &= trace(\Lambda W W') \\ &= trace(W'\Lambda W) = trace(H_n(x_{sd}, x_{td})) \end{aligned}$$

Set $Q = (I_p, \mathbf{0})W$. Then Q has full row rank p. Denote $Q' = (\mathbf{q}_1, \cdots, \mathbf{q}_p)$ and $u_l = \mathbf{q}_l'\tilde{\epsilon}_n$, $l = 1, \ldots, p$. Then

$$\tilde{\epsilon}_n' H_n(x_{sd}, x_{td})\tilde{\epsilon}_n = \tilde{\epsilon}_n' Q'Q\tilde{\epsilon}_n = \sum_{l=1}^{p} u_l^2.$$

Since $p \le p_0$ and p_0 is finite, it suffices to show for any l,

$$\sum_{x_{sd} < x_{td}} P\{u_l^2 \ge \frac{9p_0^2}{T_0^2} \ln^2 n | X_n\} \to 0, \qquad as\ n \to 0.$$

Noting that $p = trace(H_n(x_{sd}, x_{td})) = \sum_{l=1}^{p} \| \mathbf{q}_l \|^2$, we have $\| \mathbf{q}_l \|^2 = \mathbf{q}_l'\mathbf{q}_l \le p \le p_0$, $l = 1, \ldots, p$. By Lemma 1, with $t_0 = T_0/p_0$ we have

$$\begin{aligned} \sum_{x_{sd} < x_{td}} P\{|u_l| \ge 3p_0 \ln(n)/T_0 | X_n\} &\le \sum_{x_{sd} < x_{td}} 2\exp(-\frac{T_0}{p_0} \cdot \frac{3p_0}{T_0} ln(n)) \exp(c(T_0/p_0)^2 p) \\ &\le n(n+1)/n^3 \exp(cT_0^2/p_0) \to 0, \end{aligned}$$

as $n \to \infty$, where c is the constant in Lemma 1. Finally, by appealing to the dominated convergence theorem we obtain the desired result.

Lemma 3 *Consider the segmented regression model with the design matrix X_n satisfying either Condition 1 or Condition 2. Assume that the iid errors $\{z_t\}$ are locally exponentially bounded and are independent of X_n. Let τ_r be a threshold. Then for any $\delta \in (0, \min_{1 \le j \le l}(\tau_{j+1} - \tau_j)/4)$,*

$$[S_n(\tau_r - \delta, \tau_r + \delta) - S_n(\tau_r - \delta, \tau_r) - S_n(\tau_r, \tau_r + \delta)]/n \xrightarrow{p} C_r$$

for some $C_r > 0$ as $n \to \infty$.

Proof It suffices to show the case when $l = 1$. Since the proof under either *Condition 1* or Condition 2 is essentially the same, we shall proceed by verifying the lemma under *Condition 1*. Denote $X_1^* = X_n(\tau_1 - \delta, \tau_1)$, $\tilde{\epsilon}_1^* = I_n(\tau_1 - \delta, \tau_1)\tilde{\epsilon}_n$, $X_2^* = X_n(\tau_1, \tau_1 + \delta)$, $\tilde{\epsilon}_2^* = I_n(\tau_1, \tau_1 + \delta)\tilde{\epsilon}_n$,

$X^* = X_n(\tau_1 - \delta, \tau_1 + \delta)$, $\tilde{\epsilon}^* = I_n(\tau_1 - \delta, \tau_1 + \delta)\tilde{\epsilon}_n$ and $\hat{\tilde{\beta}} = (X^{*\prime}X^*)^- X^{*\prime}\mathbf{Y}_n$. Similar to the ordinary regression, we have that

$$S_n(\tau_1 - \delta, \tau_1 + \delta) = ||(I_n(\tau_1 - \delta, \tau_1 + \delta) - H_n(\tau_1 - \delta, \tau_1 + \delta))\mathbf{Y_n}||^2.$$

$$\begin{aligned}
&X_1^*\tilde{\beta}_1 + X_2^*\tilde{\beta}_2 + \tilde{\epsilon}^* \\
&= I_n(\tau_1 - \delta, \tau_1)X_n\tilde{\beta}_1 + I_n(\tau_1, \tau_1 + \delta)X_n\tilde{\beta}_2 + I_n(\tau_1 - \delta, \tau_1)\tilde{\epsilon}_n \\
&= I_n(\tau_1 - \delta, \tau_1 + \delta)I_n(-\infty, \tau_1)X_n\tilde{\beta}_1 + I_n(\tau_1 - \delta, \tau_1 + \delta)I_n(\tau_1, \infty)X_n\tilde{\beta}_2 + I_n(\tau_1 - \delta, \tau_1 + \delta)\tilde{\epsilon}_n \\
&= I_n(\tau_1 - \delta, \tau_1 + \delta)(X_n(-\infty, \tau_1) + X_n(\tau_1, \infty)\tilde{\beta}_2 + \tilde{\epsilon}_n) \\
&= I_n(\tau_1 - \delta, \tau_1 + \delta)\mathbf{Y_n}.
\end{aligned}$$

$$\begin{aligned}
X^*\hat{\tilde{\beta}} &= X_n(\tau_1 - \delta, \tau_1 + \delta)(X_n(\tau_1 - \delta, \tau_1 + \delta)'X_n((\tau_1 - \delta, \tau_1 + \delta))^- X_n(\tau_1 - \delta, \tau_1 + \delta)'\mathbf{Y_n} \\
&= H_n(\tau_1 - \delta, \tau_1 + \delta))\mathbf{Y_n}.
\end{aligned}$$

So,

$$S_n(\tau_1 - \delta, \tau_1 + \delta) = ||X_1^*\tilde{\beta}_1 + X_2^*\tilde{\beta}_2 + \tilde{\epsilon}^* - X^*\hat{\tilde{\beta}}||^2$$

$$X_1^* + X_2^* = (I_n(\tau_1 - \delta, \tau_1) + I_n(\tau_1, \tau_1 + \delta))X_n = I_n(\tau_1 - \delta, \tau_1 + \delta)X_n = X^*.$$

Thus,

$$\begin{aligned}
&S_n(\tau_1 - \delta, \tau_1 + \delta) \\
=&||X_1^*\tilde{\beta}_1 + X_2^*\tilde{\beta}_2 + \tilde{\beta}^* - X^*\hat{\beta}||^2 \\
=&||X_1^*(\tilde{\beta}_1 - \hat{\tilde{\beta}}) + X_2^*(\tilde{\beta}_2 - \hat{\tilde{\beta}}) + \tilde{\epsilon}^*||^2 \\
=&||X_1^*(\tilde{\beta}_1 - \hat{\tilde{\beta}})||^2 + ||X_2^*(\tilde{\beta}_2 - \hat{\tilde{\beta}})||^2 + ||\tilde{\epsilon}^*||^2 + 2\tilde{\epsilon}^{*\prime}X_1^*(\tilde{\beta}_1 - \hat{\tilde{\beta}}) + 2\tilde{\epsilon}^{*\prime}X_2^*(\tilde{\beta}_2 - \hat{\tilde{\beta}}).
\end{aligned}$$

It then follows from the Law of Large Numbers for stationary ergodic stochastic processes that as $n \to \infty$,

$$\frac{1}{n}X^{*\prime}X^* = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i' 1_{\{x_{id}\in(\tau_1-\delta,\tau_1+\delta)\}} \overset{a.s.}{\to} E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1+\delta)\}}\} > 0,$$

$$\frac{1}{n}X_j{*}'X_j^* \overset{a.s.}{\to} \begin{cases} E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}\} > 0, & \text{if j=1,} \\ E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1,\tau_1+\delta)\}}\} > 0, & \text{if j=2,} \end{cases}$$

and

$$\frac{1}{n}X^{*\prime}\mathbf{Y_n} \overset{a.s.}{\to} E\{\mathbf{Y_1}\mathbf{x_1} 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1+\delta)\}}\}.$$

Therefore,

$$\tilde{\hat{\beta}} \overset{a.s.}{\rightarrow} E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1+\delta)\}}\}^{-1} E\{\mathbf{Y_1}\mathbf{x_1} 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1+\delta)\}}\} := \tilde{\beta}^*.$$

Similarly, it can be shown that

$$\frac{1}{n}||X_j^*(\tilde{\beta}_j - \tilde{\hat{\beta}})||^2 \overset{a.s.}{\rightarrow} \begin{cases} (\tilde{\beta}_1 - \tilde{\beta}^*)' E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}\}(\tilde{\beta}_1 - \tilde{\beta}^*), & \text{if j=1,} \\ (\tilde{\beta}_2 - \tilde{\beta}^*)' E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1,\tau_1+\delta)\}}\}(\tilde{\beta}_2 - \tilde{\beta}^*), & \text{if j=2,} \end{cases}$$

$$\frac{1}{n}\tilde{\epsilon}^{*'} X_j^*(\tilde{\beta}_j - \tilde{\hat{\beta}}) \overset{a.s.}{\rightarrow} 0, \quad for \quad j = 1, 2,$$

and

$$\frac{1}{n}||\tilde{\epsilon}^*||^2 \overset{a.s.}{\rightarrow} \sigma^2 P\{x_{1d} \in (\tau_1 - \delta, \tau_1 + \delta)\}.$$

Thus,

$$\lim_{n\to\infty} \frac{1}{n} S_n(\tau_1 - \delta, \tau_1 + \delta)$$
$$= (\tilde{\beta}_1 - \tilde{\beta}^*)' E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}\}(\tilde{\beta}_1 - \tilde{\beta}^*) + (\tilde{\beta}_2 - \tilde{\beta}^*)' E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1,\tau_1+\delta)\}}\}(\tilde{\beta}_2 - \tilde{\beta}^*)$$
$$\sigma_0^2 P\{x_{1d} \in (\tau_1 - \delta, \tau_1 + \delta)\}.$$

It remains to show that $\frac{1}{n}S_n(\tau_1 - \delta, \tau_1)$ and $\frac{1}{n}S_n(\tau_1, \tau_1 + \delta)$ converge to $\sigma_0^2 P\{x_{1d} \in (\tau_1 - \delta, \tau_1)\}$ and $\sigma_0^2 P\{x_{1d} \in (\tau_1, \tau_1 + \delta)\}$ respectively, and either $(\tilde{\beta}_1 - \tilde{\beta}^*)' E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}\}(\tilde{\beta}_1 - \tilde{\beta}^*) > 0$ or $(\tilde{\beta}_2 - \tilde{\beta}^*)' E\{\mathbf{x_1}\mathbf{x_1}' 1_{\{x_{1d}\in(\tau_1,\tau_1+\delta)\}}\}(\tilde{\beta}_2 - \tilde{\beta}^*) > 0$. The latter is a direct consequence of the assumed conditions whie the former can be shown again by the Law of Large Numbers.

$$S_n(\tau_1 - \delta, \tau_1) = ||(I_n(\tau_1 - \delta, \tau_1) - H_n(\tau_1 - \delta, \tau_1))\mathbf{Y_n}||^2$$

Let $\tilde{\tilde{\epsilon}}^* = I_n(\tau_1 - \delta, \tau_1)\tilde{\epsilon}_n$. Then

$$\begin{aligned} X_1^*\tilde{\beta}_1 + \tilde{\tilde{\epsilon}}^* &= I_n(\tau_1 - \delta, \tau_1)X_n\tilde{\beta}_1 + I_n(\tau_1 - \delta, \tau_1)I_n(\tau_1, \infty)X_n\tilde{\beta}_2 + I_n(\tau_1 - \delta, \tau_1)\tilde{\epsilon}_n \\ &= I_n(\tau_1 - \delta, \tau_1)(I_n(-\infty, \tau_1)X_n\tilde{\beta}_1 + I_n(, \tau_1, \infty)X_n\tilde{\beta}_2 + \tilde{\epsilon}_n) \\ &= I_n(\tau_1 - \delta, \tau_1)\mathbf{Y_n} \end{aligned}$$

Let $\tilde{\tilde{\hat{\beta}}} = (X_1^{*'}X_1^*)^- X_1^{*'}\mathbf{Y_n}$. Then $X_1^{*'}\tilde{\tilde{\hat{\beta}}} = H_n(\tau_1 - \delta, \tau_1)\mathbf{Y_n}$. So

$$\begin{aligned} S_n(\tau_1 - \delta, \tau_1) &= ||X_1^*\tilde{\beta}_1 + \tilde{\tilde{\epsilon}}^* - \tilde{\tilde{\hat{\beta}}}||^2 \\ &= ||X_1^*(\tilde{\beta}_1 - \tilde{\tilde{\hat{\beta}}}) + \tilde{\tilde{\epsilon}}^*||^2 \\ &= ||X_1^*(\tilde{\beta}_1 - \tilde{\tilde{\hat{\beta}}}||^2 + ||\tilde{\tilde{\epsilon}}^*||^2 + 2\tilde{\tilde{\epsilon}}^{*'}X_1^*(\tilde{\beta}_1 - \tilde{\tilde{\hat{\beta}}}). \end{aligned}$$

Thus proceeding as before, using the law of large numbers we get that

$$\frac{1}{n}S_n(\tau_1-\delta,\tau_1)\overset{a.s.}{\to}(\tilde{\beta}_1-\tilde{\tilde{\beta}}^*)'E[\mathbf{x}_1\mathbf{x}_1'1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}](\tilde{\beta}_1-\tilde{\tilde{\beta}}^*)+\sigma_0^2P\{x_{1d}\in(\tau_1-\delta,\tau_1)\},$$

where $\tilde{\tilde{\beta}}^*=E[\mathbf{x}_1\mathbf{x}_1'1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}]^{-1}E[Y_1\mathbf{x}_1 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}]$. Similarly

$$\frac{1}{n}S_n(\tau_1,\tau_1+\delta)\overset{a.s.}{\to}(\tilde{\beta}_2-\tilde{\tilde{\tilde{\beta}}}^*)'E[\mathbf{x}_1\mathbf{x}_1'1_{\{x_{1d}\in(\tau_1,\tau_1+\delta)\}}](\tilde{\beta}_2-\tilde{\tilde{\tilde{\beta}}}^*)+\sigma_0^2P\{x_{1d}\in(\tau_1,\tau_1+\delta)\},$$

where $\tilde{\tilde{\tilde{\beta}}}^*=E[\mathbf{x}_1\mathbf{x}_1'1_{\{x_{1d}\in(\tau_1,\tau_1+\delta)\}}]^{-1}E[Y_1\mathbf{x}_1 1_{\{x_{1d}\in(\tau_1,\tau_1+\delta)\}}]$. But

$$\begin{aligned}E[Y_1\mathbf{x}_1 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}] &= E[\mathbf{x}_1[\mathbf{x}_1'\tilde{\beta}_1 1_{\{x_{1d}\in(-\infty,\tau_1)\}}+\mathbf{x}_1'\tilde{\beta}_2 1_{\{x_{1d}\in(\tau_1,\infty)\}}+\epsilon_1]1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}]\\ &= E[\mathbf{x}_1\mathbf{x}_1'1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}]\tilde{\beta}_1+E[\epsilon_1]E[\mathbf{x}_1 1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}]\\ &= E[\mathbf{x}_1\mathbf{x}_1'1_{\{x_{1d}\in(\tau_1-\delta,\tau_1)\}}]\tilde{\beta}_1.\end{aligned}$$

So, $\tilde{\tilde{\beta}}=\tilde{\beta}_1$, and similarly $\tilde{\tilde{\tilde{\beta}}}=\tilde{\beta}_2$. Thus

$$\begin{aligned}\frac{1}{n}S_n(\tau_1-\delta,\tau_1)&\overset{a.s.}{\to}\sigma_0^2P\{x_{1d}\in(\tau_1-\delta,\tau_1)\},\\ \frac{1}{n}S_n(\tau_1,\tau_1+\delta)&\overset{a.s.}{\to}\sigma_0^2P\{x_{1d}\in(\tau_1,\tau_1+\delta)\},\end{aligned}$$

which gives the result.

Lemma 3.4 *Consider the segmented regression model (2.1) with the design matrix X_n satisfying either Condition 1 or Condition 2. Assume that the iid errors $\{z_t\}$ are locally exponentially bounded and are independent of X_n. Let l^0 denote the true l.Let $(\tau_1^0,\ldots,\tau_l^0)$ denote the true $\tau_i's$. Then*

(i) $\forall\ l<l^0$, $P\{\hat{\sigma}_l^2>\sigma^2+C_r\}\to 1,\ n\to 0$ for some $C_r>0$, and (ii) $\forall\ l$ such that $l^0\le l\le L$, where L is an upper bound of l^0,

$$0\le 1/n\sum_{t=1}^{n}\epsilon_t^2-\hat{\sigma}_l^2=O_p(ln^2(n)/n),$$

where $\hat{\sigma}_l^2=\frac{1}{n}S_n(\hat{\tau}_1,\ldots,\hat{\tau}_l)$

Proof (i) Fix $(\tau_1,\ldots,\tau_l)\in\Re^l$, and $\delta\in(0,\min_{1\le j\le l^0}(\tau_{j+1}^0-\tau_j^0)/4)$. For all i let r_i be such that $1\le r\le l^0$ and $|\tau_r^0-\tau_i|$ is a minimum. Since $l<l^0$ there exists $1\le r\le l^0$ such that $r\ne r_i\forall 1\le i\le l$. Thus for all $1\le i\le l$,

$$\begin{aligned}\delta<|\tau_r^0-\tau_{r_i}^0|&=\frac{1}{4}|\tau_r^0-\tau_i+\tau_i-\tau_{r_i}^0|\\ &\le\frac{1}{4}\{|\tau_r^0-\tau_i|+|\tau_i-\tau_{r_i}^0|\}\\ &\le\frac{1}{2}|\tau_r^0-\tau_i|\end{aligned}$$

So there exists $1 \le r \le l^0$ such that $(\tau_1, \ldots, \tau_l) \in A_r := \{(\tau_1, \ldots, \tau_l) : |\tau_s - \tau_r^0| > \delta, \forall\ s = 1, \ldots, l\}$. Hence it suffices to show that for each r, $1 \le r \le l^0$, with probability approaching to 1,

$$\min_{(\tau_1, \cdots, \tau_l) \in A_r} S_n(\tau_1, \cdots, \tau_l)/n > \sigma_0^2 + C_r,$$

for some $C_r > 0$. For any $(\tau_1, \cdots, \tau_l) \in A_r$, let $\xi_1 \le \cdots \le \xi_{l+l^0+1}$ be the ordered

$$\{\tau_1, \ldots, \tau_l, \tau_1^0, \ldots, \tau_{r-1}^0, \tau_r^0 - \delta, \tau_r^0 + \delta, \tau_{r+1}^0, \ldots, \tau_{l^0}^0\}$$

and let $\xi_0 = -\infty$, $\xi_{l+l^0+2} = \infty$. Then it follows from the previous lemmas and the law of large numbers that uniformly in A_r,

$$\begin{aligned}
&\frac{1}{n} S_n(\tau_1, \cdots, \tau_l) \\
\ge& \frac{1}{n} S_n(\xi_1, \cdots, \xi_{l+l^0+1}) \\
=& \frac{1}{n} \sum_{j=1}^{l+l^0+2} S_n(\xi_{j-1}, \xi_j) \\
=& \frac{1}{n} [\sum_{j:\ \xi_j \ne \tau_r^0 + \delta} S_n(\xi_{j-1}, \xi_j) + S_n(\tau_r^0 - \delta, \tau_r^0) + S_n(\tau_r^0, \tau_r^0 + \delta) \\
&+ S_n(\tau_r^0 - \delta, \tau_r^0 + \delta) - S_n(\tau_r^0 - \delta, \tau_r^0) - S_n(\tau_r^0, \tau_r^0 + \delta)] \\
=& \frac{1}{n} \tilde{\epsilon}_n' \tilde{\epsilon}_n - \frac{1}{n} \sum_{j:\ \xi_j \ne \tau_r^0 + \delta} T_n(\xi_{j-1}, \xi_j) - \frac{1}{n} T_n(\tau_r^0 - \delta, \tau_r^0) - \frac{1}{n} T_n(\tau_r^0, \tau_r^0 + \delta) \\
&+ \frac{1}{n} (S_n(\tau_r^0 - \delta, \tau_r^0 + \delta) - S_n(\tau_r^0 - \delta, \tau_r^0) - S_n(\tau_r^0, \tau_r^0 + \delta)) \\
=& \frac{1}{n} \tilde{\epsilon}_n' \tilde{\epsilon}_n + O_p(ln^2(n)/n) + \frac{1}{n} (S_n(\tau_r^0 - \delta, \tau_r^0 + \delta) - S_n(\tau_r^0 - \delta, \tau_r^0) - S_n(\tau_r^0, \tau_r^0 + \delta)) \\
=& \sigma_0^2 + C_r + o_p(1),
\end{aligned}$$

where C_r is as defined in lemma 4.

(ii) Let $\xi_1 \le \cdots \le \xi_{l+l^0}$ be the ordered

$$\{\hat{\tau}_1, \cdots, \hat{\tau}_l, \tau_1^0, \cdots, \tau_{l^0}^0\}$$

, $\xi_0 = \tau_0^0 = -\infty$ and $\xi_{l+l^0+1} = \tau_{l^0+1}^0 = \infty$. Using a similar argument as in (i), we get that

$$\begin{aligned}
n\hat{\sigma}_l^2 = S_n(\hat{\tau}_1, \cdots, \hat{\tau}_l) &\le S_n(\tau_1^0, \cdots, \tau_{l^0}^0) \\
&= \tilde{\epsilon}_n' \tilde{\epsilon}_n - \sum_{i=1}^{l^0+1} \tilde{\epsilon}_n' H_n(\alpha, \beta) \tilde{\epsilon}_n \\
&\le \tilde{\epsilon}_n' \tilde{\epsilon}_n,
\end{aligned}$$

and,

$$
\begin{aligned}
n\hat{\sigma}_l^2 = S_n(\hat{\tau}_1, \cdots, \hat{\tau}_l) &\geq S_n(\hat{\tau}_1, \cdots, \hat{\tau}_l, \tau_1^0, \cdots, \tau_{l^0}^0) \\
&= \tilde{\epsilon}_n' \tilde{\epsilon}_n - \sum_{j=1}^{l^0+1} \sum_{\tau_{j-1}^0 \leq \xi_{k-1} < \xi_k \leq \tau_j^0} T_n(\xi_{k-1}, \xi_k) \\
&= \tilde{\epsilon}_n' \tilde{\epsilon}_n + O_p(ln^2(n)),
\end{aligned}
$$

which proves (ii).

Proof of Theorem 1 It follows from Lemma 4 (i) that $P\{\hat{l} \geq l^0\} \to 1$ as $n \to \infty$. By Lemma 4 (ii), for $l^0 < l \leq L$, $0 \geq \hat{\sigma}_l^2 - \hat{\sigma}_{l^0}^2 = O_p(\ln^2(n)/n)$, and $\hat{\sigma}_{l^0}^2 = \tilde{\epsilon}_n' \tilde{\epsilon}_n/n + O_p(\ln^2(n)/n) = \sigma_0^2 + o_p(1)$. Hence $(\hat{\sigma}_l^2 - \hat{\sigma}_{l^0}^2)/\hat{\sigma}_{l^0}^2 = O_p(ln^2(n)/n)$. Also for small $x > 0$, $\ln(1-x) > -2x$. Therefore,

$$
\begin{aligned}
MIC(l) - MIC(l^0) &= \ln(\hat{\sigma}_l^2) - \ln(\hat{\sigma}_{l^0}^2) + (l - l^0)(\ln(n))^{2+\delta}/n \\
&= \ln(1 - (\hat{\sigma}_l^2 - \hat{\sigma}_{l^0}^2)/\hat{\sigma}_{l^0}^2) + (l - l^0)(\ln(n))^{2+\delta}/n \\
&\geq -2O_p(\ln^2(n)/n) + (l - l^0)(\ln(n))^{2+\delta}/n \\
&> 0
\end{aligned}
$$

for sufficiently large n. Whence $\hat{l} \xrightarrow{p} l^0$ as $n \to \infty$.

In concluding the theoretical section of this paper some mention should be made as to the need for further research into the statistical properties of these models. For instance it would be useful to have likelihood based tests for testing the existence of thresholds. Also of interest would be the asymptotic distributions of the estimated parameter values under mild stationarity assumptions.

Table 1

Forward Rate Models

	Random Walk Model	Forward Rate Model
Unadjusted		
sse	1.67	1.66
per cent reduction		1.04
five year mfse	0.45	0.42
per cent reduction		6.67
Adjusted for Inflation		
sse	1.29	1.31
five year mfse	0.44	0.44
With Intercept Unadjusted		
sse	1.67	1.65
per cent reduction		1.14
five year mfse	0.45	0.43
per cent reduction		4.44
Adjusted for Inflation		
sse	1.29	1.30
five year mfse	0.44	0.47

Table 2

Segmented Models–Standard Principle Components

model description	Percent reduction in root mse from random walk model
Interest Differential	
-Principle Component 1(Lag 1)	
Two Segments	16.69
Three Segments	23.61
Two Segments (random walk coeff.=1)	6.75
– Principle Component 1(Lag 2)	
Two Segments	15.20
- Principle Component 1(Lag 3)	
Two Segments	9.85
Trade Balance Differential	
-Principle Component 1(Lag 1)	
Two Segments	6.28
Two Segments (random walk coeff.=1)	2.46
–Principle Component 1(Lag 2)	
Two Segments	4.92
–Principle Component 1(Lag 3)	
Two Segments	6.90
Money Supply Differential	
-Principle Component 1(Lag 1)	
Two Segments	13.46
Two Segments (random walk coeff.=1)	5.98
–Principle Component 1(Lag 2)	
Two Segments	15.00

model description	Percent reduction in root mse from random walk model
Interest Diferential -Principal Component 1 (lag1) and Trade Balance Differential -Principal Component 1 (lag1)	
I.D. segmentation variable	
–Two Segments	16.38
–Three Segments	22.77
T.B. segmentation variable	
–Two Segments	11.45
Interest Diferential -Principal Component 1 (lag1) and Money Supply Diferential -Principal Component 1 (lag1)	
I.D. segmentation variable	
–Two Segments	20.76
–Three Segments	26.03
T.B. segmentation variable	
–Two Segments	13.37
Money Supply Diferential -Principal Component 1 (lag1) and Trade Balance Diferential -Principal Component 1 (lag1)	
T.B. segmentation variable	
–Two Segments	7.03
M.S. segmentation variable	
–Two Segments	12.66

model description	Percent reduction in root mse from random walk model
Interest Diferential	
-Principal Component 1 (lag1) and	
Money Supply Diferential	
-Principal Component 1 (lag1) and	
Trade Balance Diferential	
-Principal Component 1 (lag1)	
I.D. segmentation variable	
–Two Segments	21.28
–Three Segments	27.56
T.B. segmentation variable	
–Two Segments	11.67
–Three Segments	12.55

Table 3

Segmented Models–Standard Principle Components Forecasting Results

Model Description	Sum of Squared Forecasting Error (SSFE)	SSFE (Random Walk)	Percent Reduction in SSFE
Interest Differential			
Two Segments	0.299	0.437	31.56
Three Segments	0.368	0.437	15.63
Interest Differential and Trade Balance Differential			
Two Segments	0.424	0.437	2.92
Three Segments	0.490	0.437	
Interest Differential and Money Supply Differential			
Two Segments	0.636	0.437	
Three Segments	0.252	0.437	42.33
Interest Differential and Money Supply Differential and Trade Balance Differential			
Two Segments	0.630	0.437	
Three Segments	0.507	0.437	

Table 4

Monetary Models

Model	root MSE	percent reduction root MSE	SSFE	percent reduction SSFE	(without ln transf.)
Random Walk	0.063		0.0908		
Frenkel-Bilson	0.2219		0.7780		
Dornbusch-Frankel	0.192		0.8594		
Hooper-Morton	0.194		2.6440		
Vector Autoregressions					
Frenkel-Bilson -2 lags	0.0598	5.06	0.1003		0.86
Frenkel-Bilson -3 lags	0.061	3.15	0.1244		
Dornbusch-Frankel -2 lags	0.0583	7.44	0.0887	2.26	12.36
Dornbusch-Frankel -3 lags	0.0604	4.11	0.1154		
Hooper-Morton -2 lags	0.0545	13.48	0.1025		

Table 5

Segmented Models–Tsay Principle Components

model description	Percent reduction in root mse from random walk model
Interest Differential -Principle Component 1(Lag 1)	
Two Segments	15.83
Three Segments	15.53
Trade Balance Differential -Principle Component 1(Lag 1)	
Two Segments	5.46
Three Segments	13.23
Money Supply Differential -Principle Component 1(Lag 1)	
Two Segments	12.90
Three Segments	15.33
Interest Diferential -Principal Component 1 (lag1) and Trade Balance Differential -Principal Component 1 (lag1)	
I.D. segmentation variable	
–Two Segments	15.03
–Three Segments	15.78
T.B. segmentation variable	
–Two Segments	9.14
–Three Segments	15.20

model description	Percent reduction in root mse from random walk model
Interest Diferential -Principal Component 1 (lag1) and Money Supply Diferential -Principal Component 1 (lag1)	
I.D. segmentation variable	
–Two Segments	16.39
–Three Segments	16.43
M.S. segmentation variable	
–Two Segments	14.72
–Three Segments	15.88
Money Supply Diferential -Principal Component 1 (lag1) and Trade Balance Diferential -Principal Component 1 (lag1)	
T.B. segmentation variable	
–Two Segments	8.71
–Three Segments	13.14
M.S. segmentation variable	
–Two Segments	14.34
–Three Segments	17.72
Interest Diferential -Principal Component 1 (lag1) and Money Supply Diferential -Principal Component 1 (lag1) and Trade Balance Diferential -Principal Component 1 (lag1)	
I.D. segmentation variable	
–Two Segments	15.68
–Three Segments	16.55

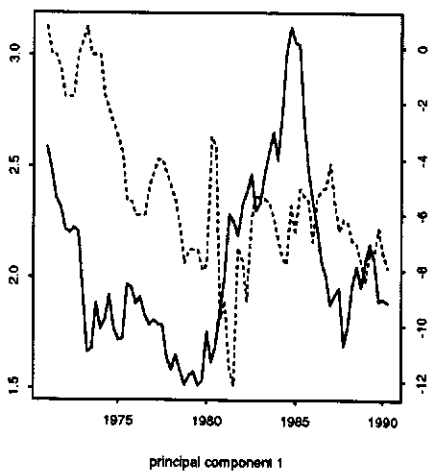
Table 6

Segmented Models–Tsay Principle Components Forecasting Results

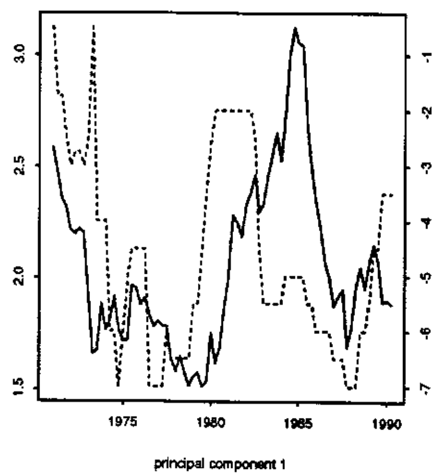
Model Description	Sum of Squared Forecasting Error (SSFE)	SSFE (Random Walk)	Percent Reduction in SSFE
Interest Differential			
Two Segments	0.329	0.437	24.65
Three Segments	0.360	0.437	17.65
Interest Differential and Money Supply Differential			
Two Segments	0.331	0.437	24.15
Three Segments	0.378	0.437	13.38
Interest Differential and Trade Balance Differential			
Two Segments	0.378	0.437	13.52
Three Segments	0.416	0.437	4.69
Interest Differential and Money Supply Differential and Trade Balance Differential			
Two Segments	0.403	0.437	7.65
Three Segments	0.443	0.437	

Time Series Plots - Interest Differentials

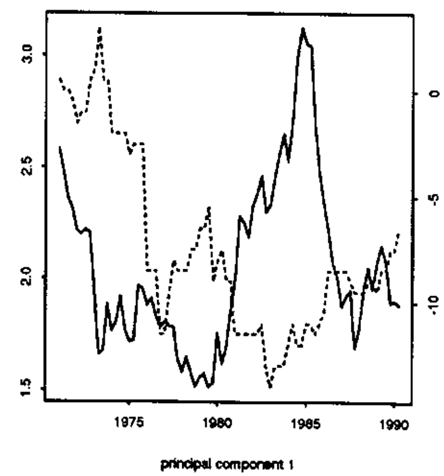
(Germany-Canada)



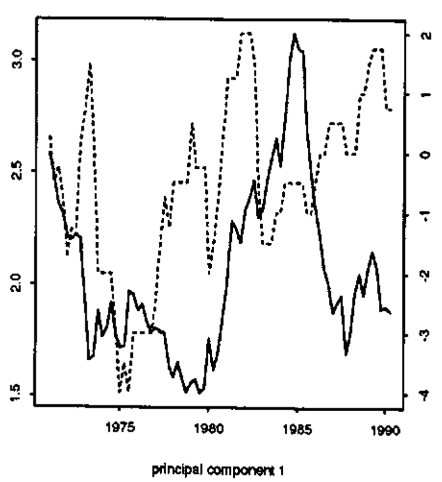
(Germany-France)



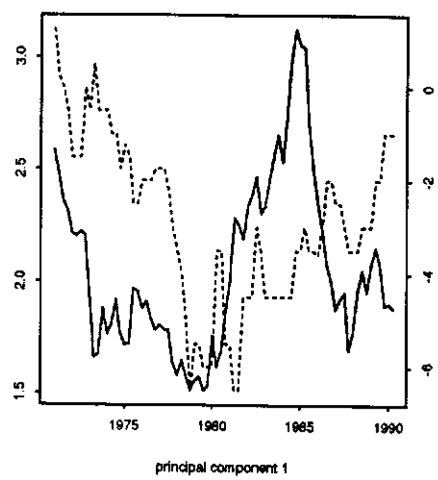
(Germany-Italy)



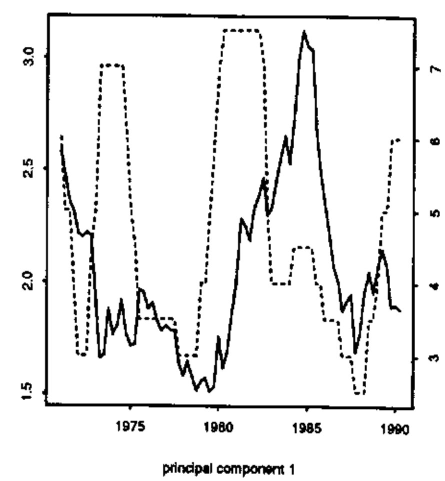
(Germany-Japan)



(Germany-U.S.)

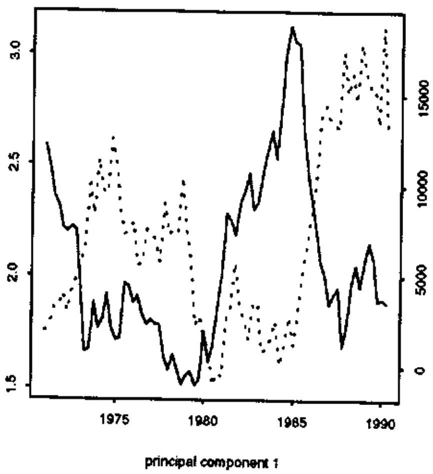


German Interest Rates

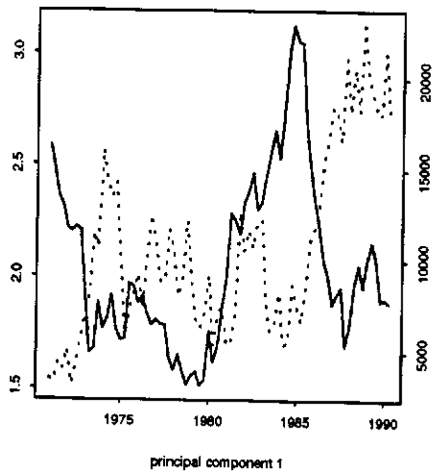


Time Series Plots - Trade Balance Differences Mill U.S.$ (1985)

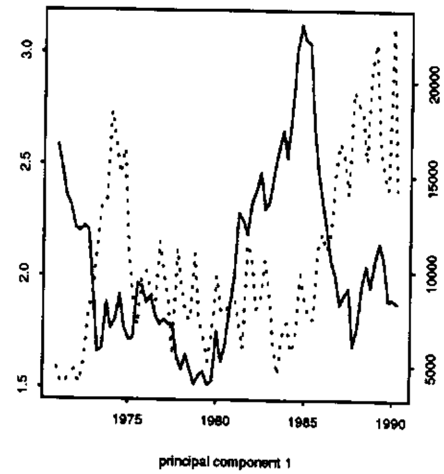
(Germany-Canada)



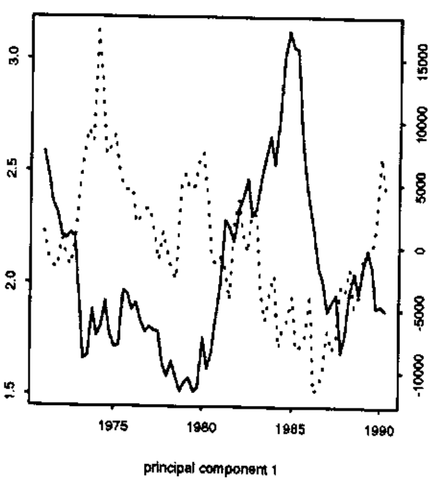
(Germany-France)



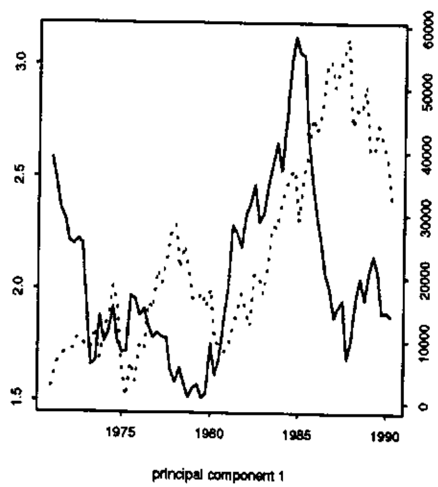
(Germany-Italy)



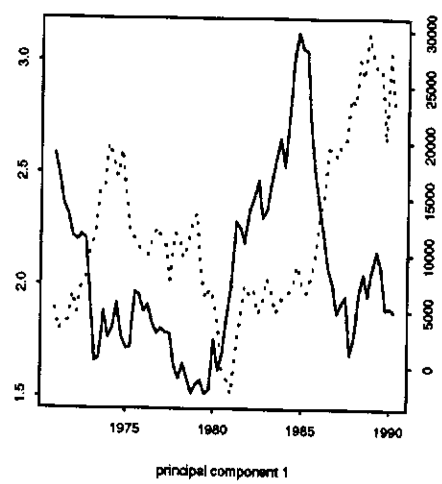
(Germany-Japan)



(Germany-U.S.)

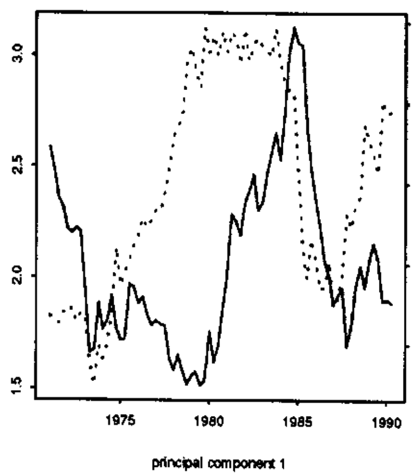


(Germany-U.K.)

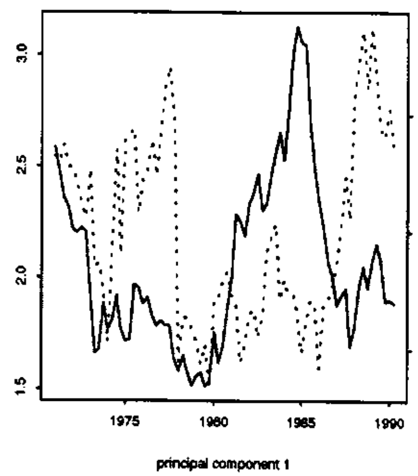


Time Series Plots - Money Supply Diff. Standardized

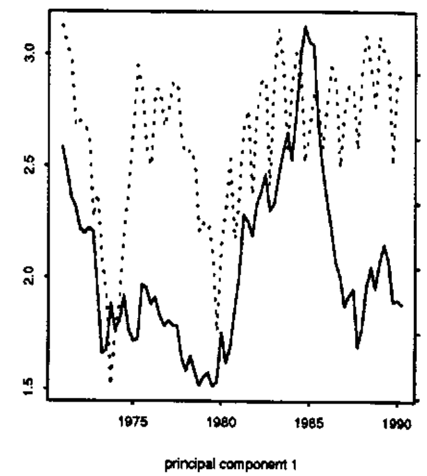
(Germany-Canada)



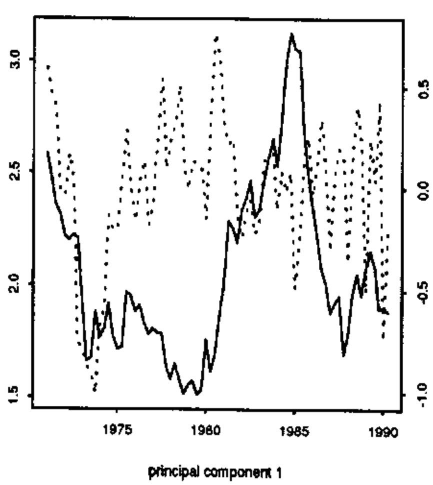
(Germany-France)



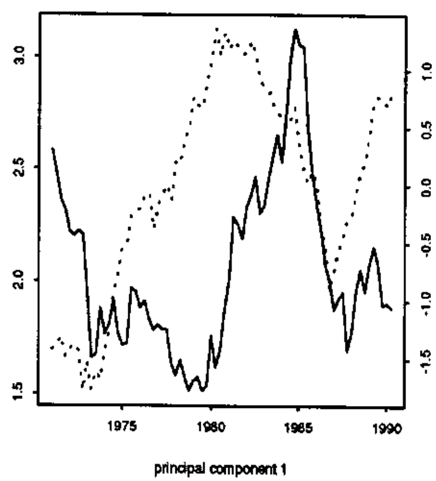
(Germany-Italy)



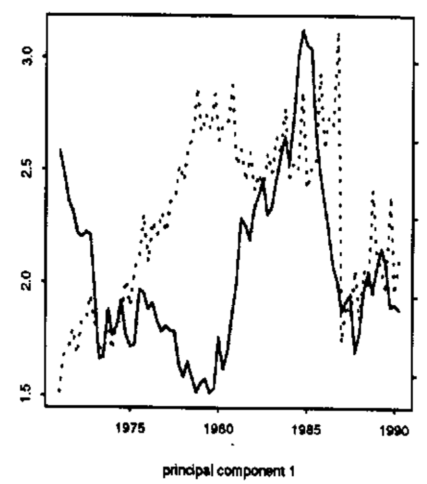
(Germany-Japan)



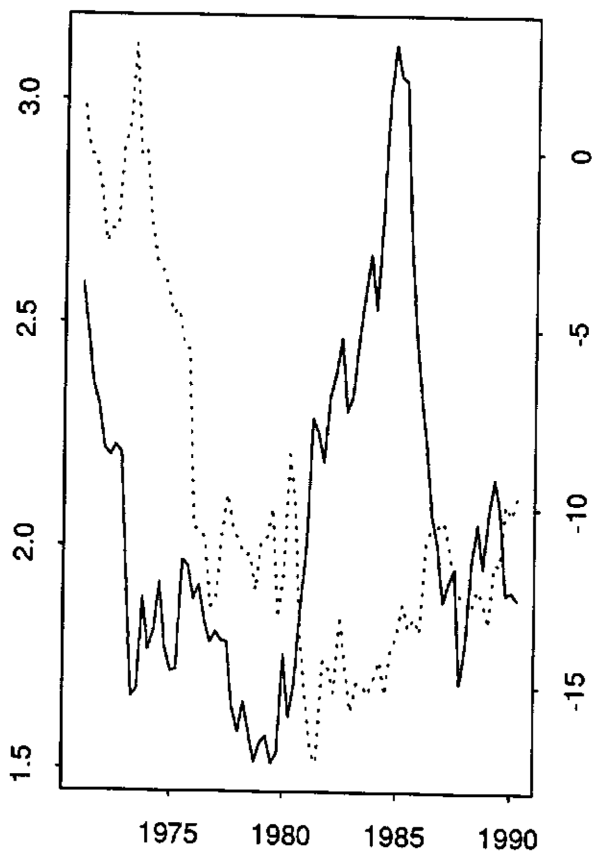
(Germany-U.S.)



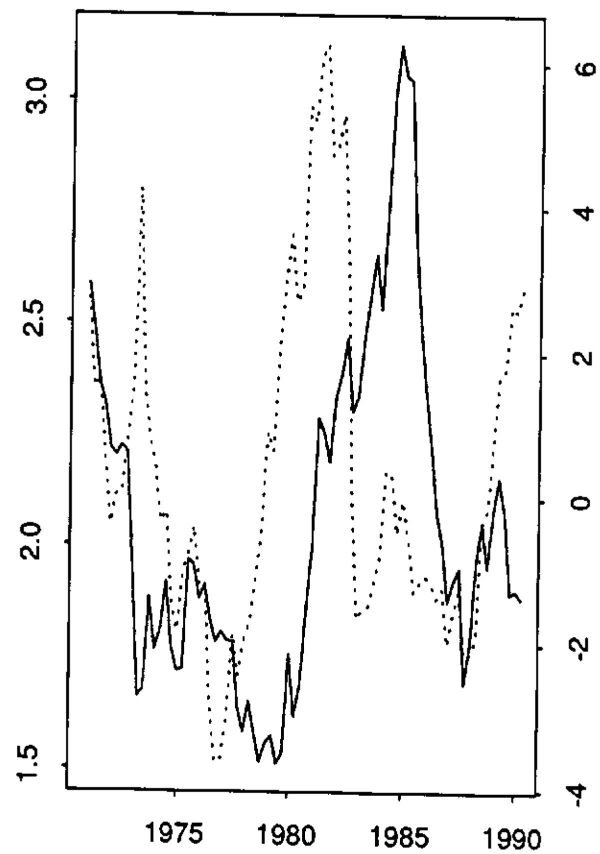
(Germany-U.K.)



Time Series Plots - Interest Diff. Prin Comp.

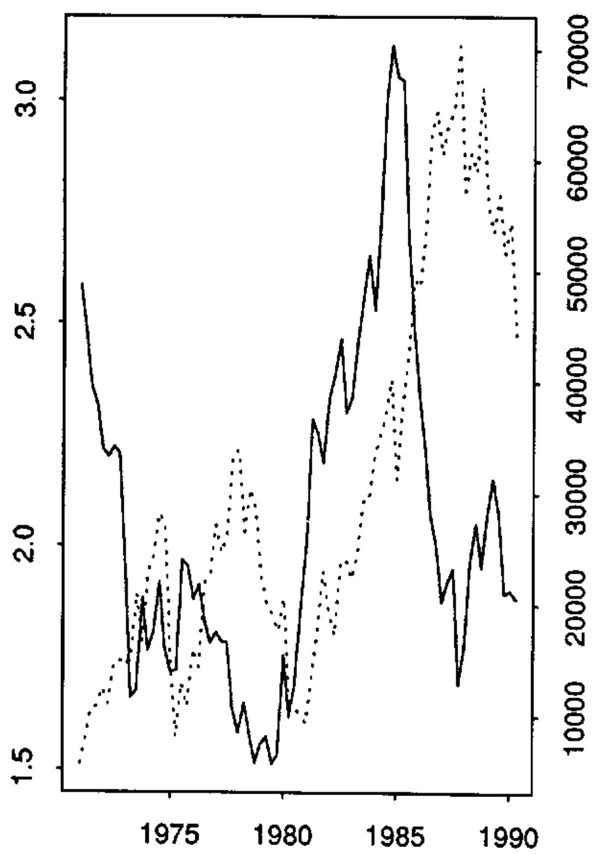


principal component 1

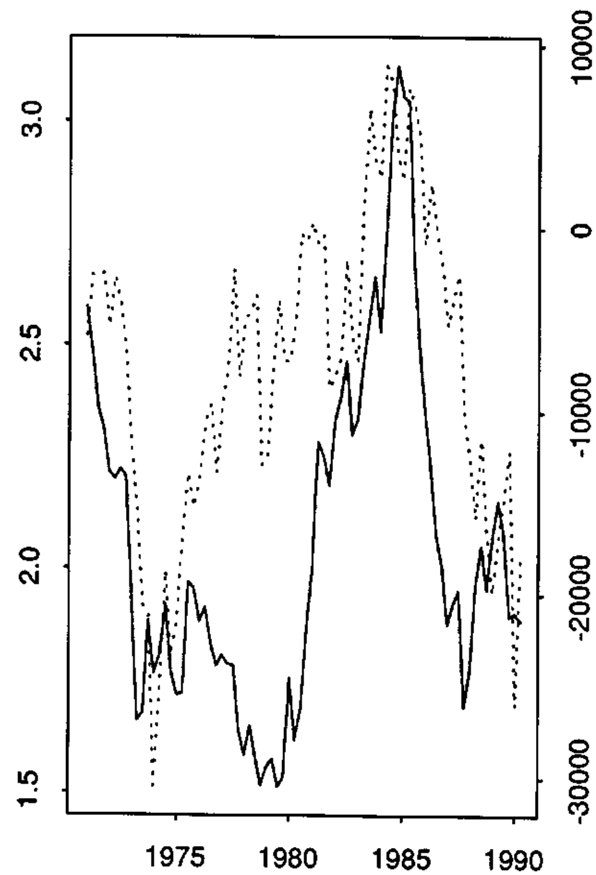


principal component 2

Time Series Plots - Trade Balance Principal Component

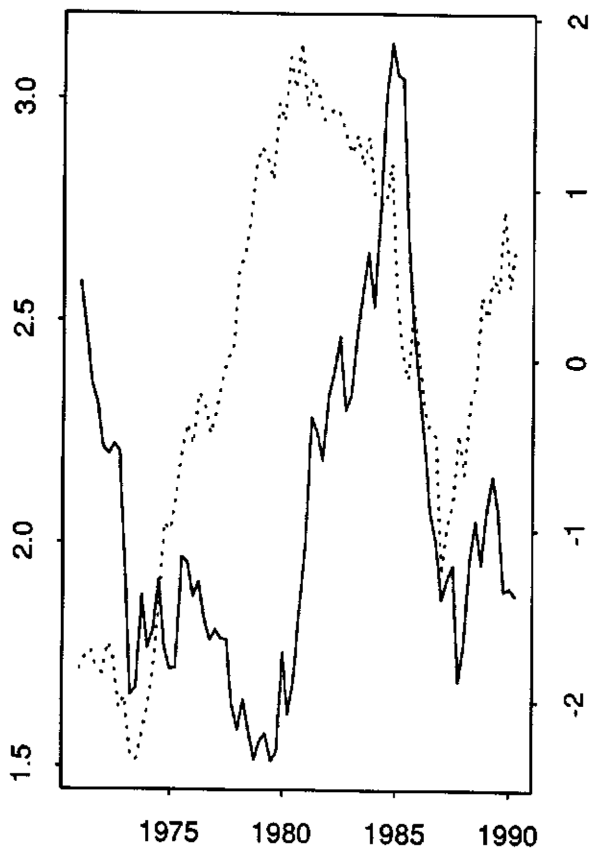


principal component 1

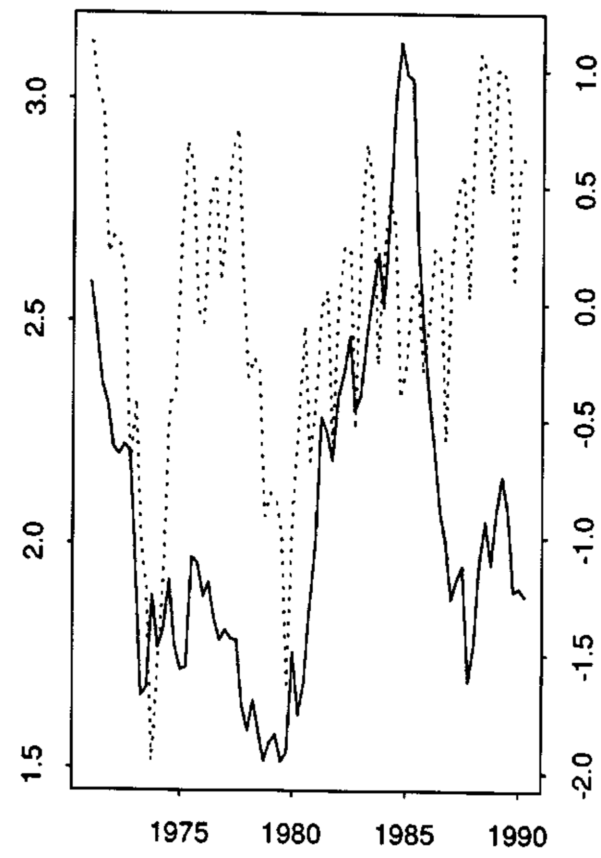


principal component 2

Time Series Plots - Money Supply Diff. Prin Comp.

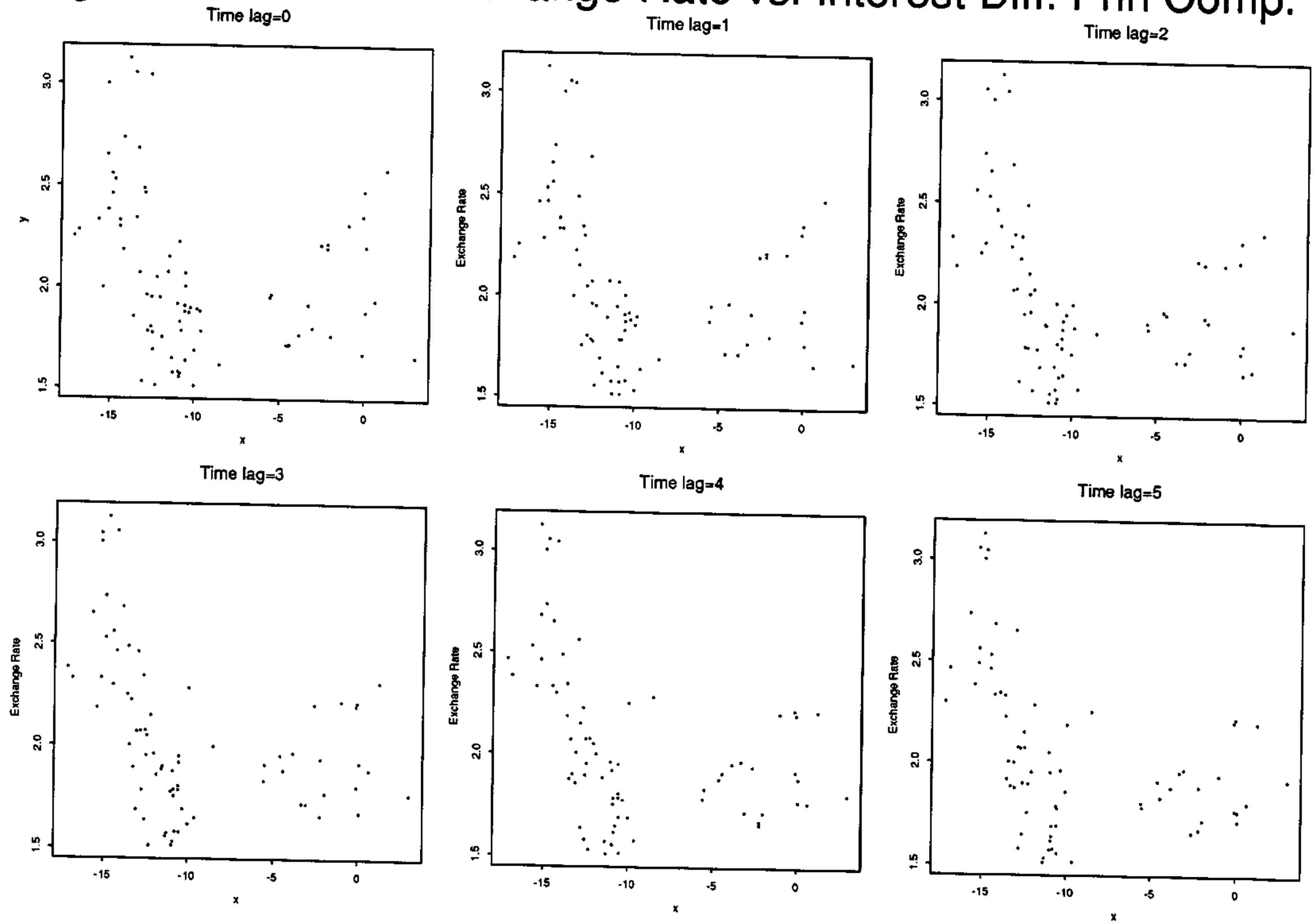


principal component 1

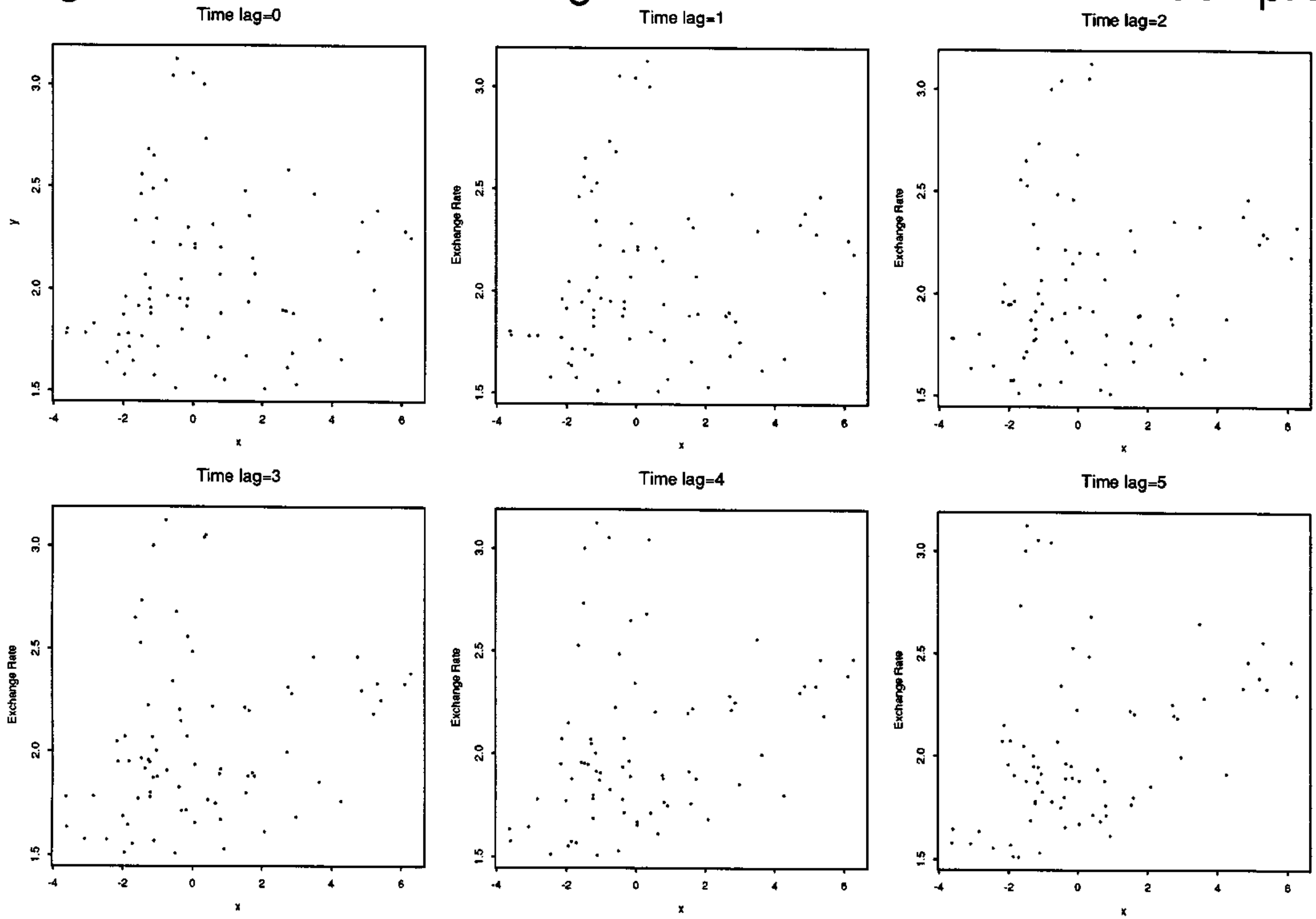


principal component 2

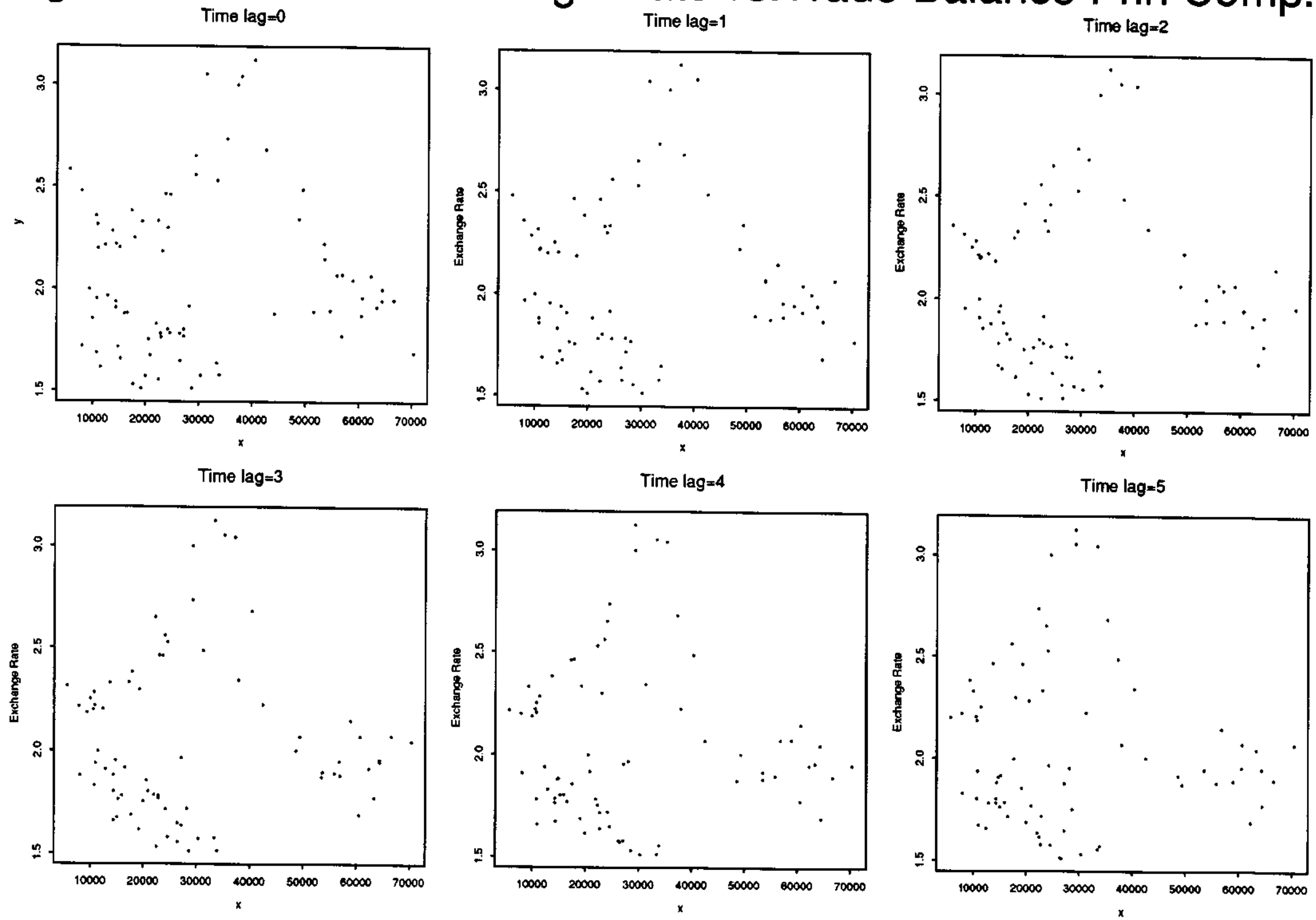
Lag Plots - German Exchange Rate vs. Interest Diff. Prin Comp. 1



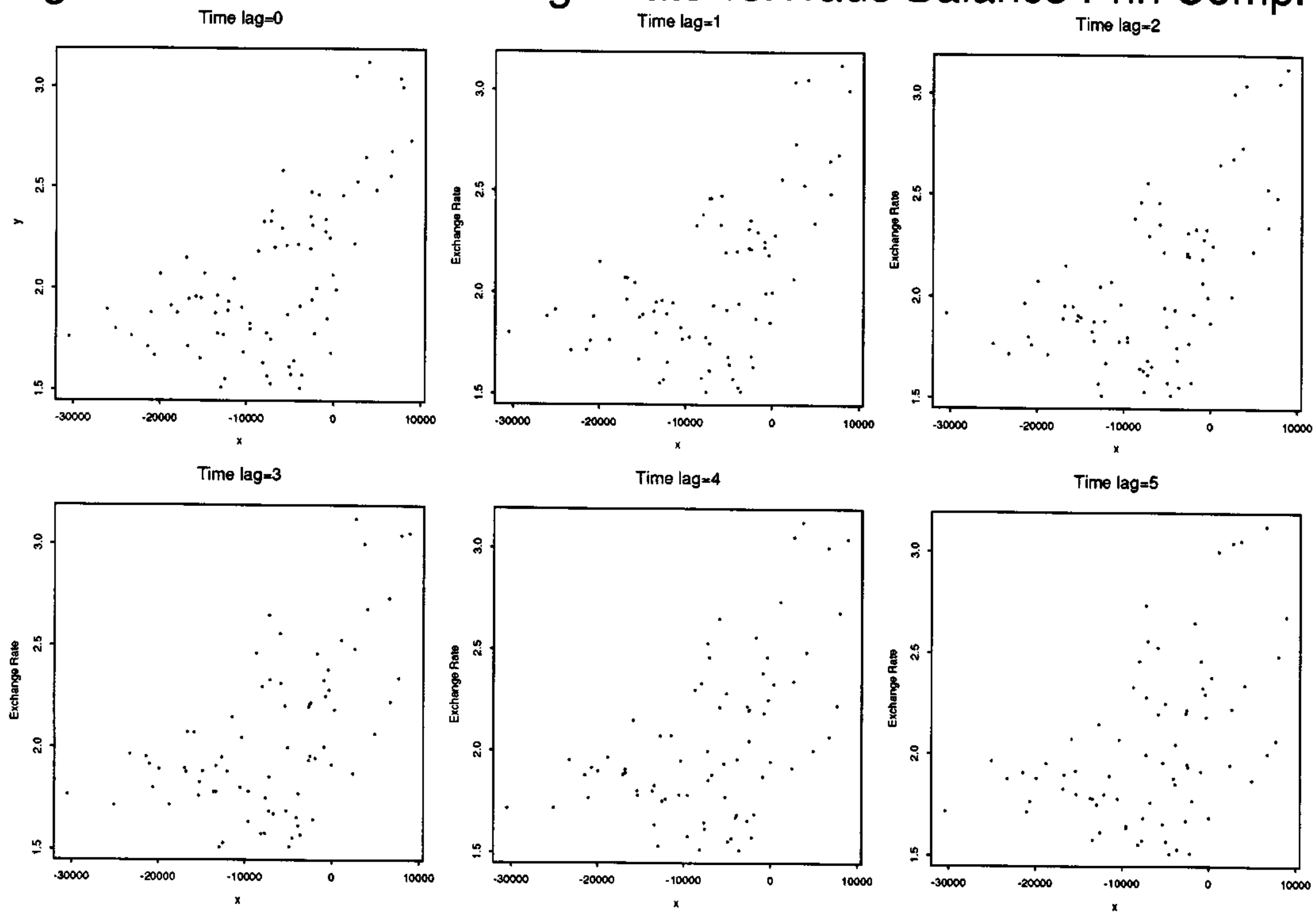
Lag Plots - German Exchange Rate vs. Interest Diff. Prin Comp. 2



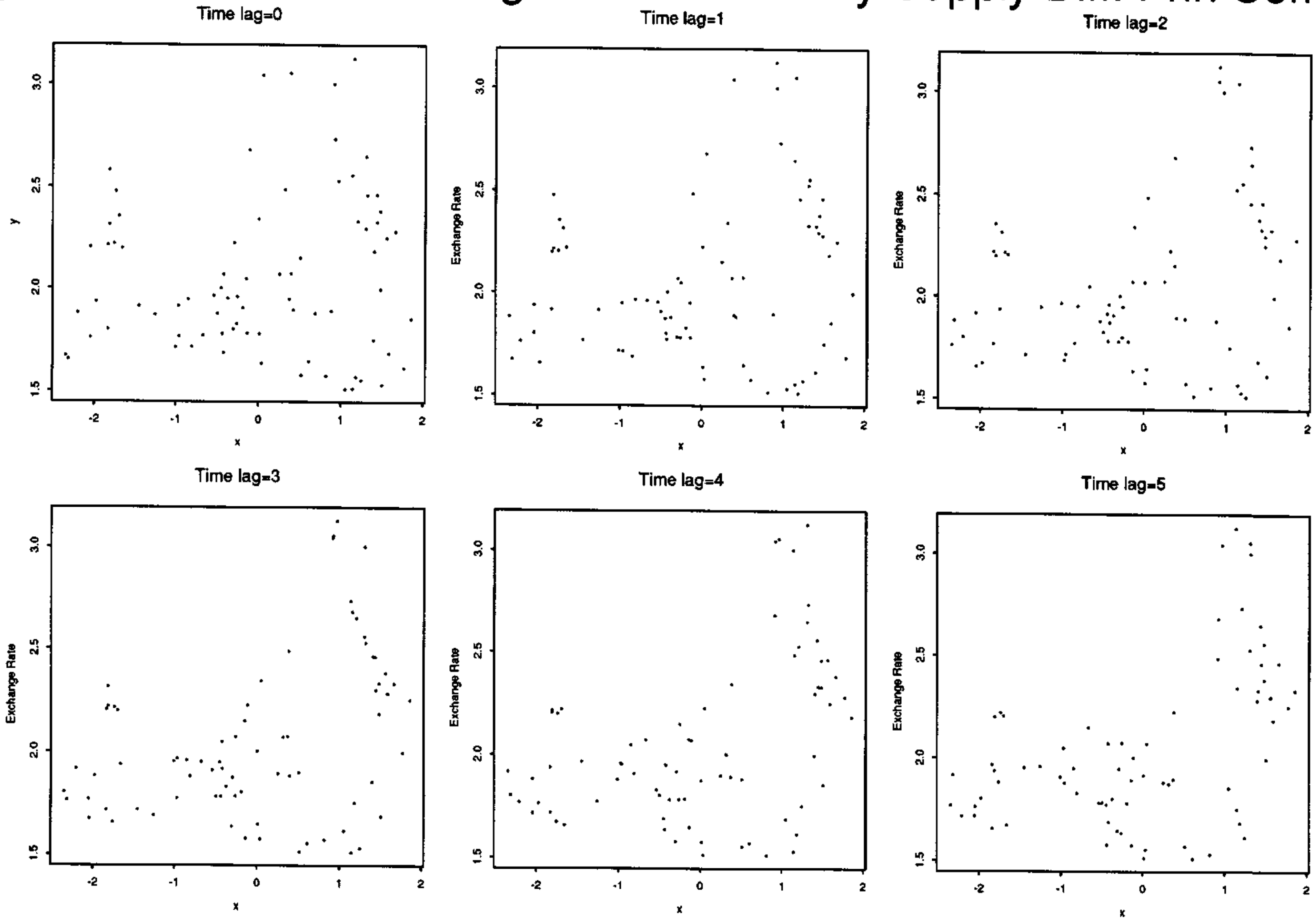
Lag Plots - German Exchange Rate vs.Trade Balance Prin Comp. 1



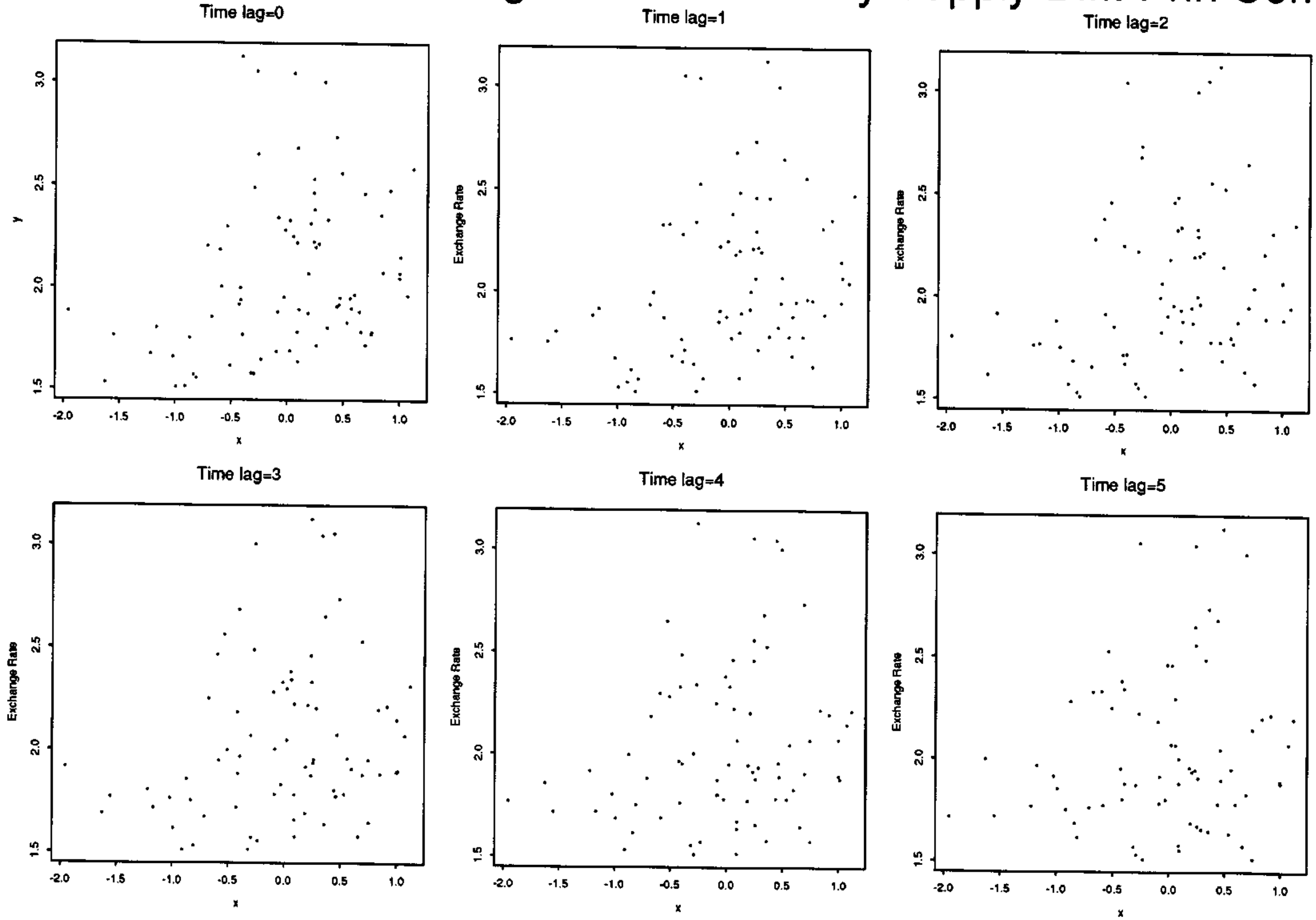
Lag Plots - German Exchange Rate vs.Trade Balance Prin Comp. 2



Lag Plots - German Exchange Rate vs. Money Supply Diff. Prin Comp. 1

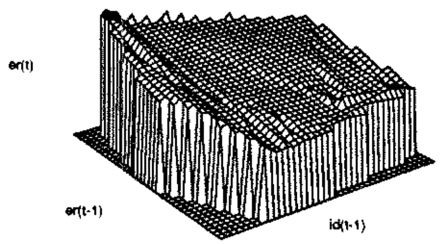


Lag Plots - German Exchange Rate vs. Money Supply Diff. Prin Comp. 2

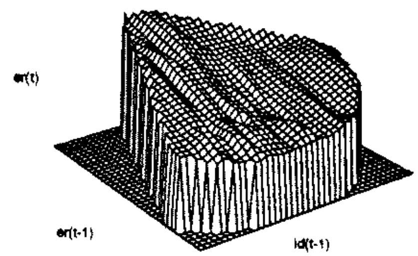


Three Dimensional Plots - Principle Components

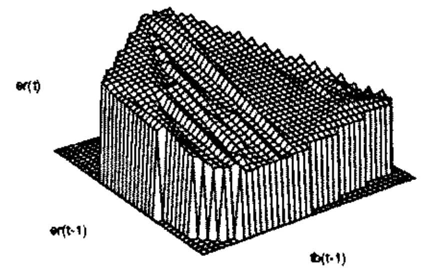
Int Diff. Prin Comp. 1



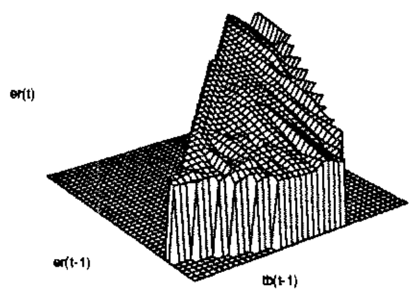
Int Diff. Prin Comp. 2



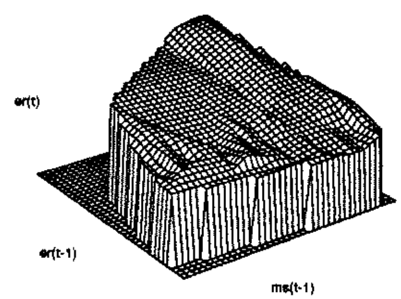
Trade Bal. Prin. Comp. 1



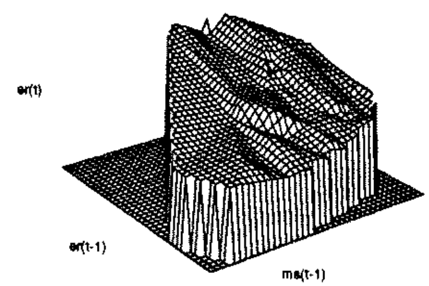
Trade Bal. Prin. Comp. 2



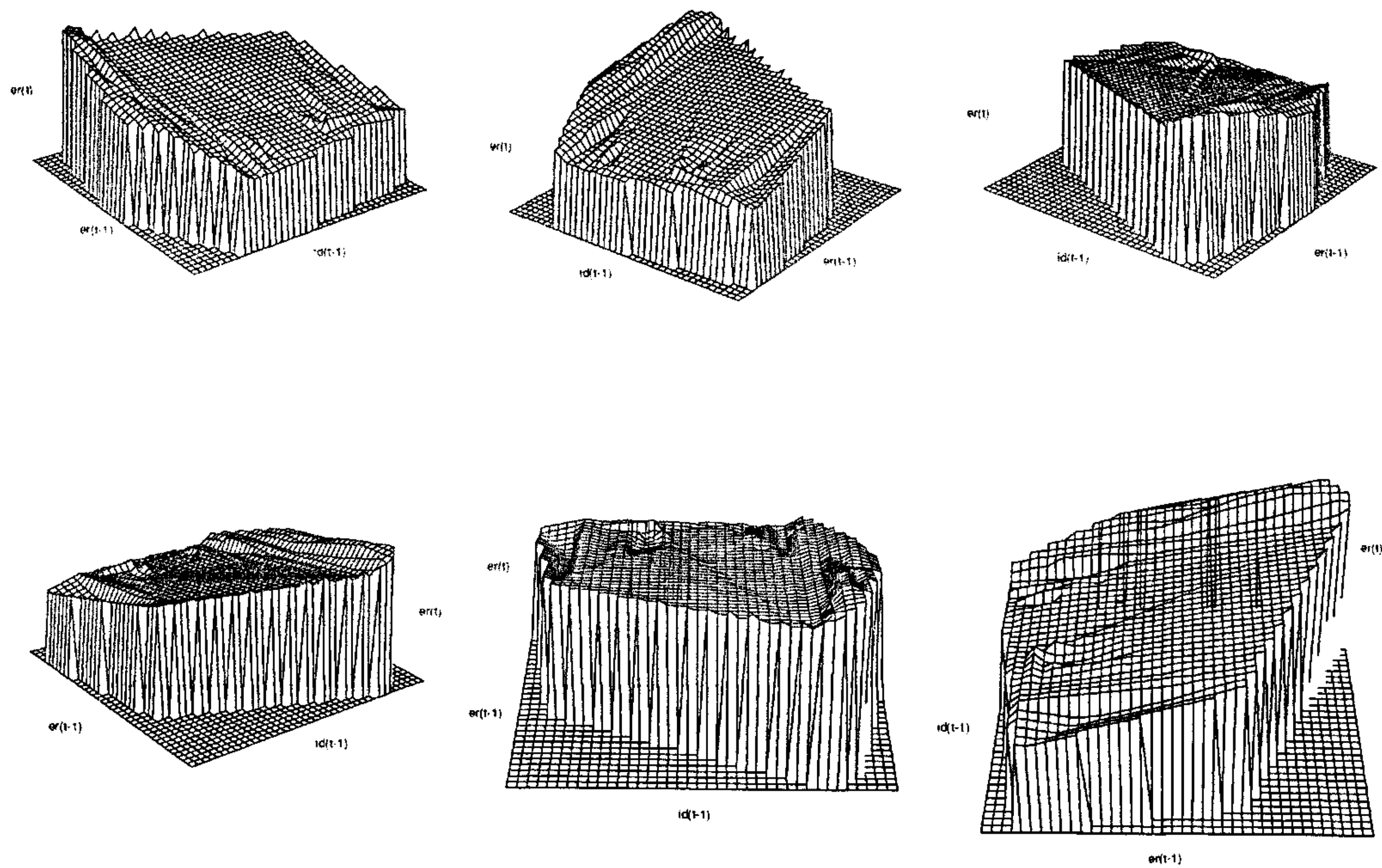
Money Supply Prin. Comp. 1



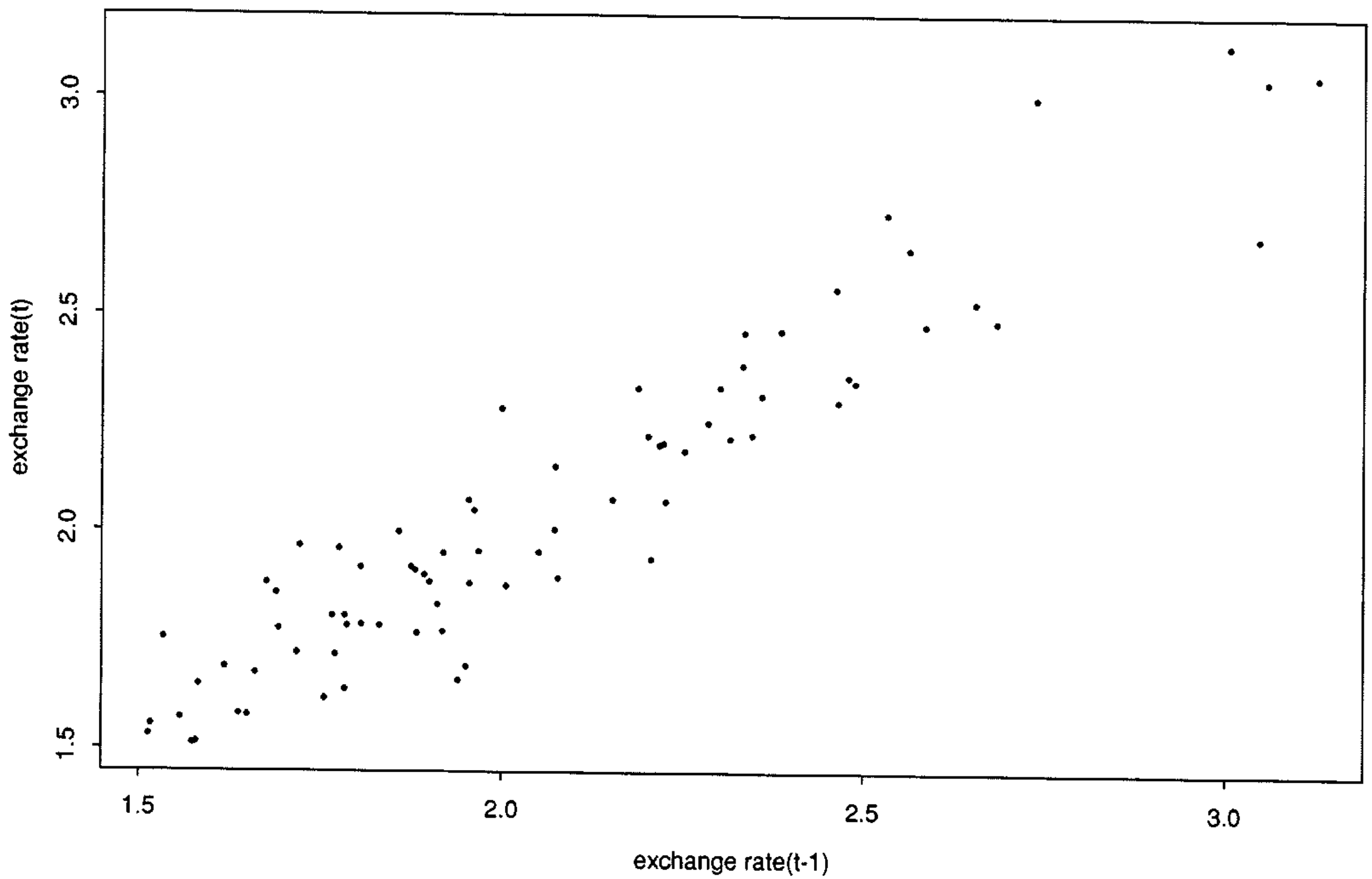
Money Supply Prin. Comp. 2



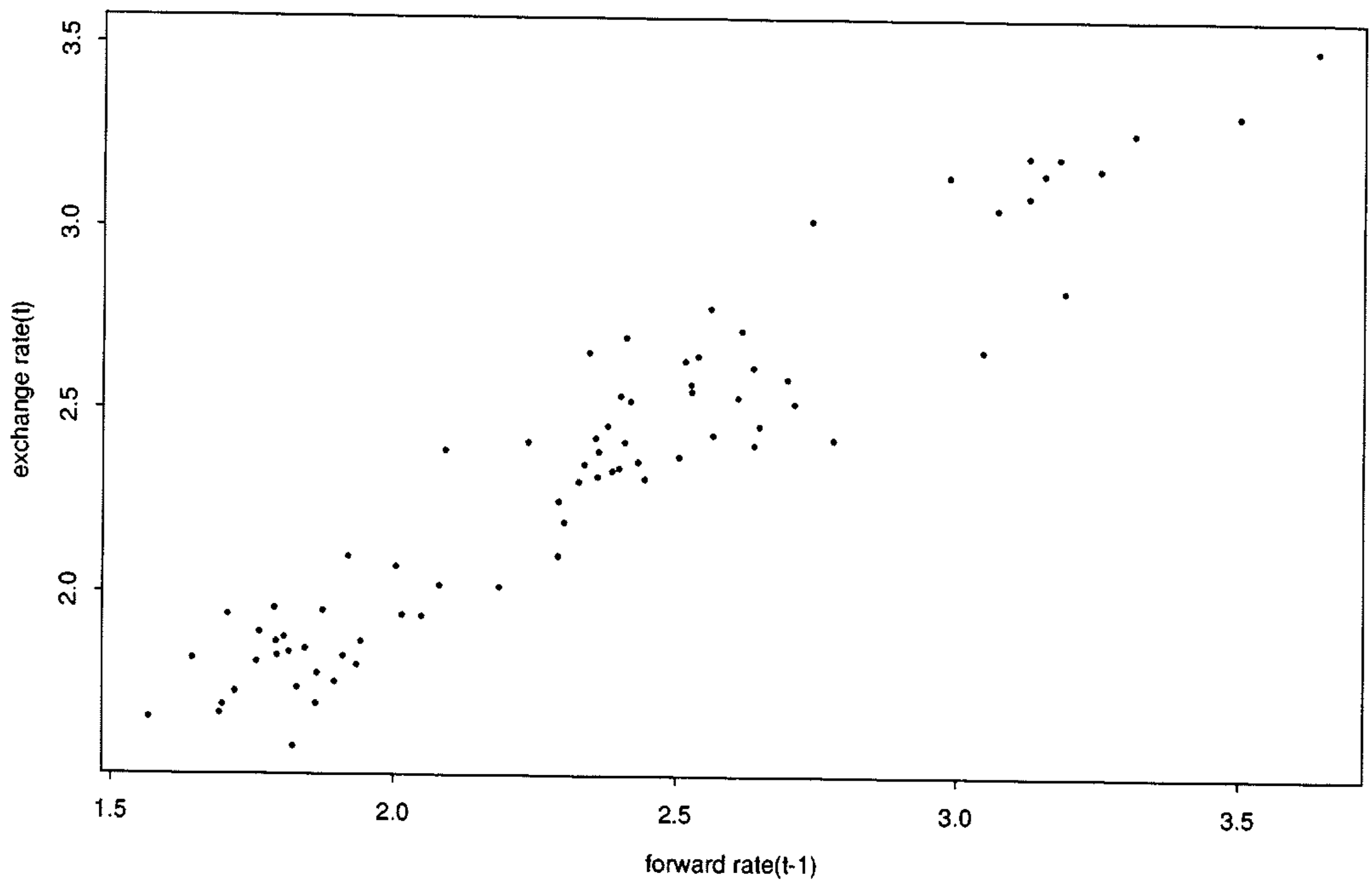
Three Dimensional Plots- Interest Differential Principle Component 1(Lag 1)



Lag 1 Plot German Exchange Rate

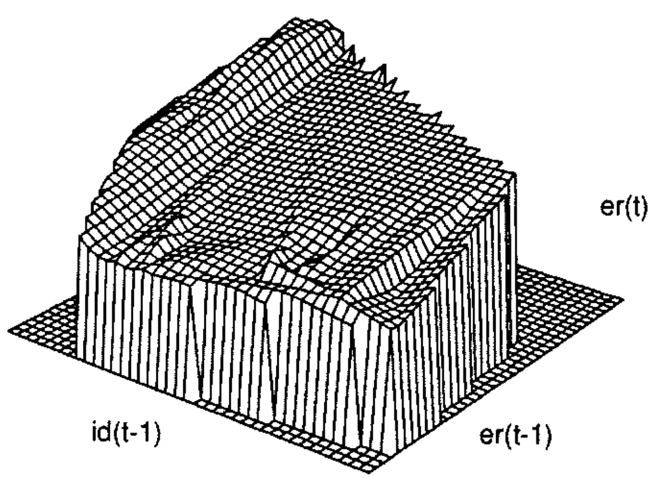


Scatter Plot - Forward Rates

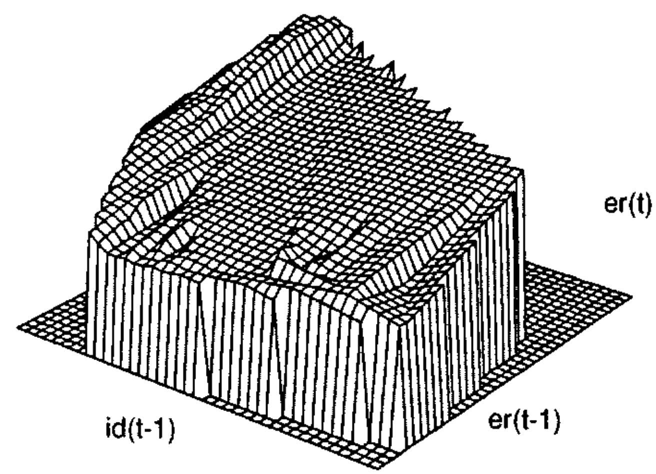


Three Dimensional Plots--Interest Rate Differential

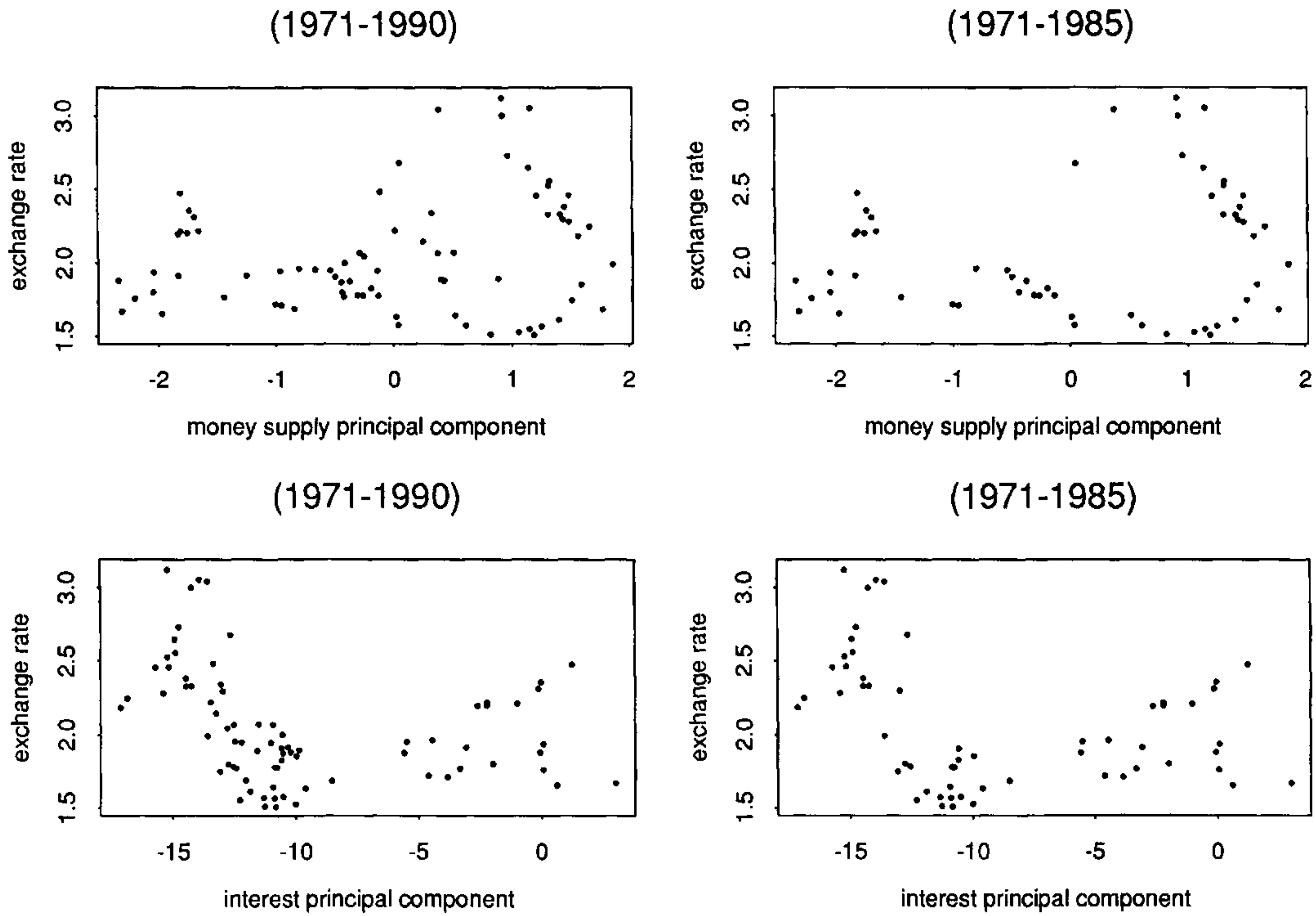
1971-1985



1971-1990



Two Dimensional Plots



Bibliography

Baillie, R. and McMahon, P. (1989), The foreign exchange market. Theory and econometric evidence (Cambridge University Press).

Bilson, J.F.O. (1978), Rational expectations and exchange rates, in J.A. Frenkel and H.G. Johnson, The Economics of Exchange Rates: Selected Studies, 75-76.

Bilson, J.F.O. (1979), The deutsche mark/dollar rate - A monetary analysis, in: Karl Brunner and Allan H. Meltzer, eds., Policies for employment, prices and exchange rates, Carnegie - Rochester Conference 11 (North Holland Publishing Company, Amsterdam).

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

Dornbusch, R. (1976), Exchange rate expectations and monetary policy, Journal of International Economics, **6**, 231-244.

Feder, P.I. (1975a) On asymptotic distribution theory in segmented regression problems-identified case, Annals of Statistics, **3**, 49-83.

Feder, P.I. (1975b) The log likelihood ratio in segmented regression, Annals of Statistics, **3**, 84-97.

Frankel, J.A. (1979), On the mark: theory of floating exchange rates based on real interest rate differentials, American Economic Review, **69**, 610-622.

Frenkel J.A. (1976), A monetary approach to the exchange rate: doctrinal aspects and empirical evidence., Scandinavian Journal of Economics, **78**, 200-224.

Frenkel, J.A. (1983), An introduction to exchange rates and international macroeconomics, in: J.A. Frenkel ed., Exchange rates and international macroeconomics, 1-18, University of Chicago Press.

Hooper, P. and Morton, J.E. (1982), Fluctuations in the dollar: A model of nominal and real exchange rate determination, Journal of International Money and Finance, **1**, 39-56.

Issard, P. (1983), An accounting framework and some issues for modeling how exchange rates respond to the news, in: J.A. Frenkel ed., Exchange rates and international macroeconomics, 19-56, University of Chicago Press.

Kouri, P.J.K (1976), The exchange rate and balance of payments in the short run and in The long run, a monetary approach, Scandinavian Journal of Economics, **78**, 280-304.

Liu, J., Wu, S., Zidek, J.V. (1991), On Segmented Multivariate Regressions, Dept. of Statistics Technical Report No. 109, University of British Columbia.

Meese, R.A. and Rogoff, K. (1983), Empirical exchange rate models of the seventies: Do they fit out of sample? Journal of International Economics, **14**, 3-24.

Serfling, R.J. (1981), Approximation theorems of mathematical statistics, Wiley, New York.

Tsay, R.S. (1990), Correlation transformation and components of stock prices, unpublished manuscript.

Yeh, M.P., Gardner, R.M., Adams, T.D., Yanowitz, F.G., and Crapo, R.O. (1983), "anaerobic threshold':Problems of determination and validation. J Appl. Physiol. Respirit. Envioron. Excercise Physiol., **55**, 1178-1186.