

### Trends, cycles, and co-integration: some issues in modelling long-term development in time series analysis

Thome, Helmut

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

**Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:**

GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Thome, H. (1996). Trends, cycles, and co-integration: some issues in modelling long-term development in time series analysis. *Historical Social Research*, 21(4), 3-23. <https://doi.org/10.12759/hsr.21.1996.4.3-23>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

## Trends, Cycles, and Co-Integration. Some Issues in Modelling Long-Term Development in Time Series Analysis

*Helmut Thome\**

**Abstract:** In this article, the work of Namenwirth and Weber on the long term cyclical nature of culture change has been taken as a starting point in a discussion of more general issues in identifying trends and cycles and structural relationships between variables in time series analysis. A brief introduction into the statistical concept of cointegrated processes is offered. This concept clarifies conditions under which equilibrium relationships between variables exhibiting stochastic trends can be modeled.

### Introduction

In this paper, I take the work of Namenwirth and Weber (1987) on the long term cyclical nature of culture change as my point of departure to discuss more general issues that are relevant to any situation where trends, cycles and possibly relationships between variables are to be identified and modeled on the basis of time series data that contain deterministic and stochastic components. I should like to emphasize that these issues do not only arise in the field of economic history, they are relevant in other fields of social and cultural history as well. This can be illustrated by a brief reference to Namenwirth and Weber's work on cultural dynamics (section 1). This will carry me to problems of detrending time series data (section 2) and the identification of cycles (section 3). In section 4, I will introduce the notion of »cointegration« which will be useful in modelling long-term equilibrium relationships.

---

\* Address all communications to Helmut Thome, Institut für Soziologie, Martin-Luther-Universität Halle-Wittenberg, D- 06099 Halle (Saale), Tel.: 0345 / 5524260, Fax: 0345 / 5527149, E-Mail: [thome@soziologie.uni-halle.de](mailto:thome@soziologie.uni-halle.de).

Slightly extended version of a paper presented at the European Social Science History Conference, Economics - Session 24, at Noordwijkerhout, Holland, May 11, 1996. I am grateful to Rainer Metz and Thomas Rahlf for their helpful comments.

## 1. Namenwirth and Weber on the Cyclical Nature of Culture Change

In a series of articles published since the late sixties - culminating in a book that appeared in 1987<sup>1</sup> - these authors have advanced a theory of culture change which heavily rests on the seemingly cyclical behavior of certain value indicators. Based on the results of computer aided content analysis of U.S. party platforms 1844 to 1964 and British »Speeches from the Throne« from 1790 to 1972 they claim to have identified »Long Cycles« (L-cycles) of value change with a period of about 150 years overlapping with shorter cycles (K-cycles) of about 50 years which are supposed to correspond to the famous (much disputed) Kondratieff waves of economic growth. How have they done this? First, the words in each document were classified by computer into a smaller set of content categories (like wealth, power, respect, rectitude) defined by the dictionary. This resulted in a frequency count of words in each of the documents for each content category (or subcategory), thereby building up a set of time series data representing the relative weight of each category in the sequence of time ordered documents.

Content categories were assigned to four comprehensive »themes« adopted from Parsons' (and Bales') AGIL-scheme of functional prerequisites: Social systems have to attend to the solution of four types of functional problems labelled as »expressive«, »adaptive«, »instrumental«, and »integrative« problems. »However, in the process of solution, some problems take precedence over others in a determined sequence or phase movement« (NW, p. 74, henceforth NW) which gives rise to a cyclical pattern in the value indicators; they move in different phases depending on the thematic concern they represent. Namenwirth with the American and Weber with the British data both found that the social system takes about 150 years to work its way through the full cycle of functional problems, each of them moving - at different phases - from high to low or low to high priority.

Cycles were identified by iteratively fitting sine curves to the original or the detrended<sup>2</sup> time series data. The data were detrended by linearly regressing the value indicator series on time and calculating the residuals. Figure 1\* presents a

<sup>1</sup> Henceforth, references to Namenwirth & Weber (1987) are given by »NW«.

<sup>2</sup> This seems to be the case only for Weber's analysis of the British data. Namenwirth has apparently fitted sine curves to the original strongly trending data, a procedure on which I comment in section 3.

\* The initial estimates for the parameters of the sine curve were derived from visual inspection of plots of these OLS residuals. The estimates were subsequently adjusted in a stepwise manner until neither a change in wavelength nor a change in peak would further improve the fit of the sine curve to the data. Depending on country and epoch, for one half to two thirds of the value indicators long cycles could be identified in this way. Their individually observed wavelengths varied considerably. For the capitalist period in Great Britain, for example, L-cycles from 80 (»Power Arena«)

panel of eight value indicators: those two of each theme for which the fitted sine curves explain the largest proportion of variance left after detrending the series.

The authors rightly point out that spectral analysis cannot be used to identify cycles which do not repeat itself several times within the period of observation. They also seek to justify their methodological escape route, i. e., their sine curve fitting technique, on substantive grounds presuming »... that the underlying process is deterministic rather than stochastic... Thus we prefer the deterministic to the stochastic model... Textual evidence strongly suggests, and the theoretical arguments presented support the contention, that long-term deterministic cycles in changing thematic concerns exist in these texts. Thus we prefer the deterministic to the stochastic model« (NW, p. 94). This argument is misleading on several points. For example, being able to fit a deterministic sine curve tells nothing about the deterministic or stochastic nature of the underlying process. Sine curves can even be fitted to random walks and white noise (cf. M. Eisner 1990). Generally, it is rather risky to fit deterministic functions to processes which are indeed stochastic.

Actually, Namenwirth and Weber have fitted two kinds of deterministic functions, first a function of time in order to eliminate a presumably linear trend and, second, the sine curves themselves. We thus have to discuss, first, the appropriateness of alternative methods of detrending time series.

## 2. Problems in Detrending Data

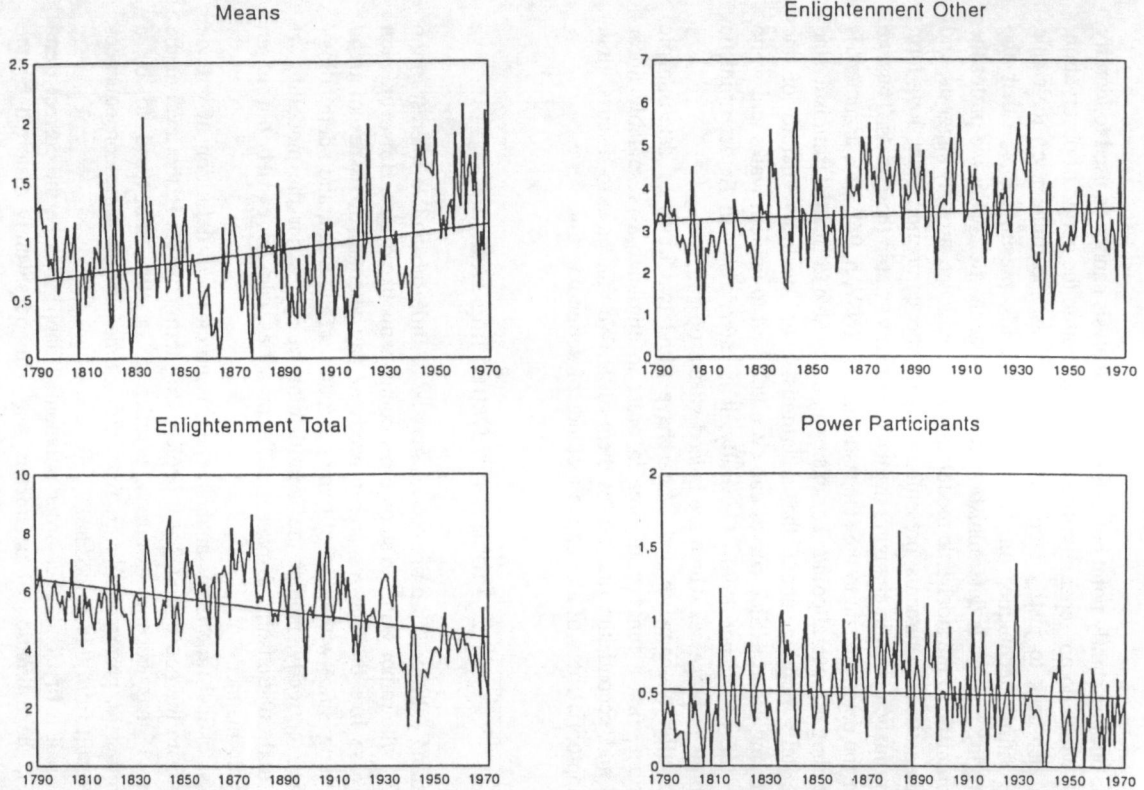
The trend exhibited in a time series may be of little substantive interest to the analyst who rather wants to study other components like cycles. However, even in a case like this the researcher needs to pay attention to issues of trend modelling, since without identifying the nature of the trend, and »detrending« the data correctly, neither the identification of cycles nor the modelling of structural relationships between different series can be carried out in an adequate manner.

In statistical time series analysis there is no clear cut definition of »trend« apart from the conception that it depicts long term changes in the level of the series. It has become common practice to assume the trend to be a deterministic, preferably linear, function of time. Troublesome consequences may result from this procedure.

Even if a first or higher order polynomial happens to be the correct trend model, the estimation of parameters may still be severely distorted by the particularities of the sampled time period: there is hardly any social or cultural

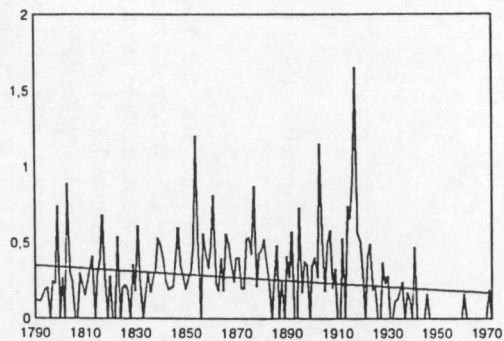
---

to 200 years (»Power Total«) and K-cycles from 27 (»Well-being Total«) to 89 years (»Time Space«) were observed, with median wave lengths of 148 and 52 years respectively.

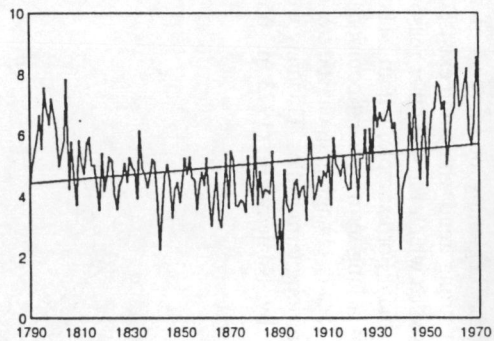


*Fig. 1:* Eight Examples of Weber's Value Indicator Series

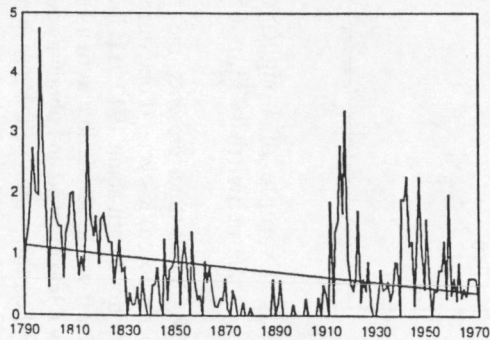
Rectitude Gains



Ends



Selves



Transaction Gains

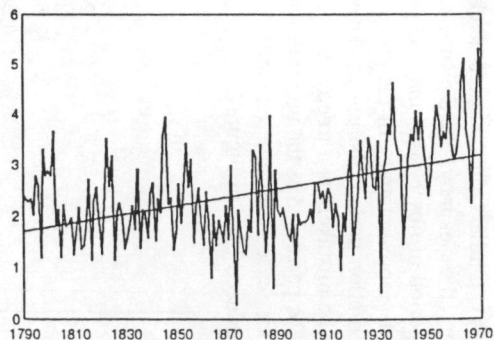


Fig. 1: continued

7

indicator whose long term observation is not affected by irregular components like measurement error and outlier effects (cf. McCleary and Hay, Jr., 1980, p. 33-36). Outliers or a structural change at the beginning or the end of an observed series may give the appearance of a strong trend where non exists. Or the period of observation may cut out a time slice which for still other reasons is not »representative« of the process under study. For example, within a short observation period, the up- or downswing of a long term cyclical component may be mistaken as a trend - or vice versa: locally changing stochastic trends may be perceived as the up- or downswings of long term cycles. This may well be the case with the example presented by Namenwirth (1973 p. 654) reproduced here in Figure 2.

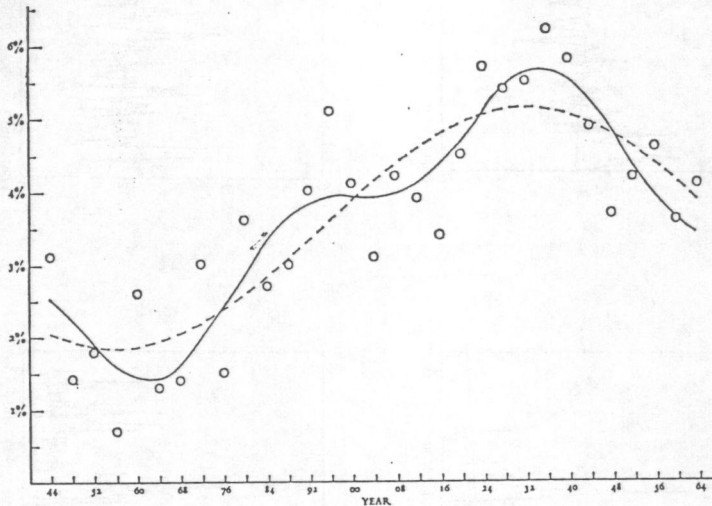


Fig. 2: Example of Long Cycle Fitting by Namenwirth (Source: Namenwirth 1973, p. 654).

Fig. 2: Example of Long Cycle Fitting by Namenwirth (Source: Namenwirth 1973, p. 654).

The dotted line presents the fitted L-cycle, the solid line the fitted (superimposed) K-cycle. If one disregards the first value (1844) there is a fairly linear trend until about 1940. It is indeed possible to fit a sine-curve to the data, just as it would be possible to fit a 1st or 3rd degree polynomial. One could also draw a line by hand and obtain a coefficient of determination of at least  $r^2 = .50$ . Together, this would neither prove nor disprove that there is a cyclical process of 152 years generating the data. It is simply a matter of faith or a priori theory and not, as Namenwirth wants us to believe, »the data« that suggest the presence of a cycle in a case like this.

The problem to be dealt with is not simply one of finding the correct deterministic function, be it a lower or higher order polynomial or some other

function of time. None of these functions is appropriate for defending a series if the series is not deterministic but has been generated by a random process that can be formalized by the following equations - or some variants on it:

$$\begin{array}{ll}
 y_t = y_{t-1} + \varepsilon_t & y_t = y_{t-1} + \mu + \varepsilon_t \\
 \text{1a} & \text{1b} \\
 y_t = y_0 + \sum_{i=1}^t e_i & E(Y_t) = y_0 + t * \mu
 \end{array}$$

The last terms, the e's, in each equation represent the stationary component in a series. If it is »white noise«, the first model is called a »simple random walk«, the second a »random walk with drift«. The characteristic feature of these processes is that each random shock has a permanent impact upon the series. Figures 3 and 4 represent various *RW* realizations.

Please note, that even a simple random walk without drift may, within a limited time period, appear to produce a trending series and or a cyclical movement. But I do not want to discuss the formal properties of such processes in any detail. I only want to mention some of the unpleasant consequences which arise from fitting a deterministic trend to a random process:<sup>3</sup>

1. If the series has been generated by a simple ~~random walk~~ (without drift), an *OLS* regression on the time-index *t* produces a spurious coefficient of determination that averages  $r^2 = .44$ , which does not decrease by increasing sample size. If the random walk contains a »drift«, a coefficient of determination larger than zero makes some sense, but it tends to be overestimated approaching unity as the series moves towards infinity. This holds independently from the size of the drift parameter.

2. Standard tests of significance (based on Student's *t* statistic) are biased. In Monte Carlo experiments with simple random walks and a series length of  $n = 100$  observations, the correct null-hypothesis (stating that the slope coefficient of the time index should be zero) was rejected in 87 % of all the cases with a nominal significance level of 5 %. This »spurious« regression causes correlated error structures. Applying an estimation procedure which takes a first-order autoregressive error structure into account, however, reduced the rate only slightly to 73 % in these experiments.

3. The average size of the autocorrelation within the residual series (resulting from regressing a random walk series on time) depends on the length of the series. The autocorrelation at lag 1 is given by  $1 - (10/n)$ ,  $n$  being the size of the sample. Even more important for our discussion is the pattern of these autocorrelation coefficients. They show cyclical behaviour with a period length

<sup>3</sup> For the following see Nelson & Kang (1981; 1984) and Banerjee et al. (1993).



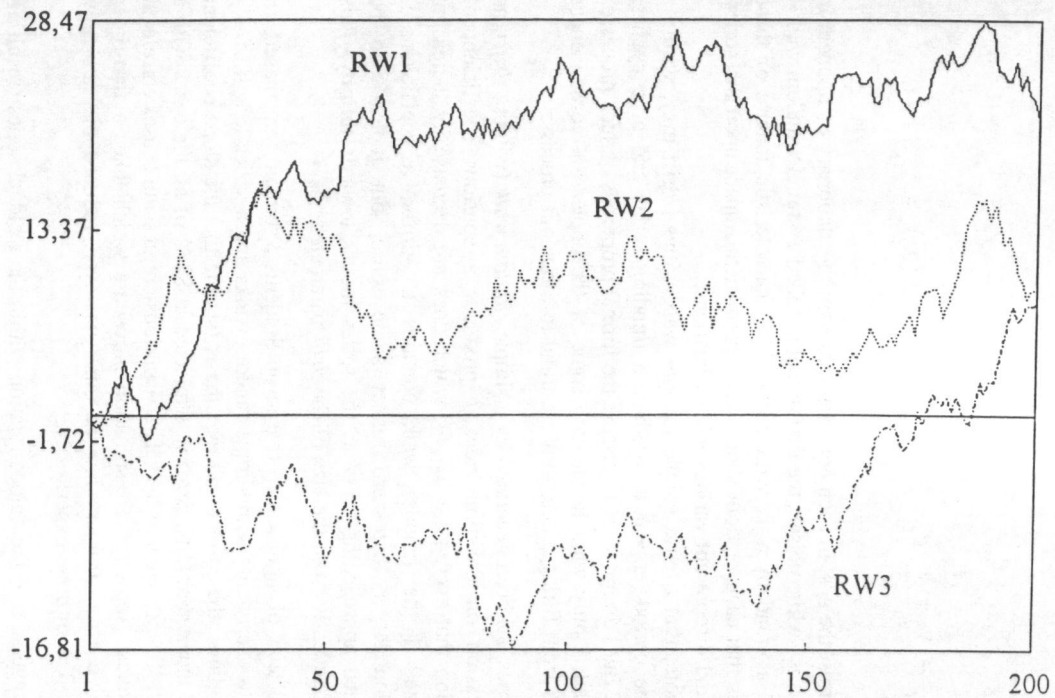


Fig. 3: Three Realizations of a Simple Random Walk Process

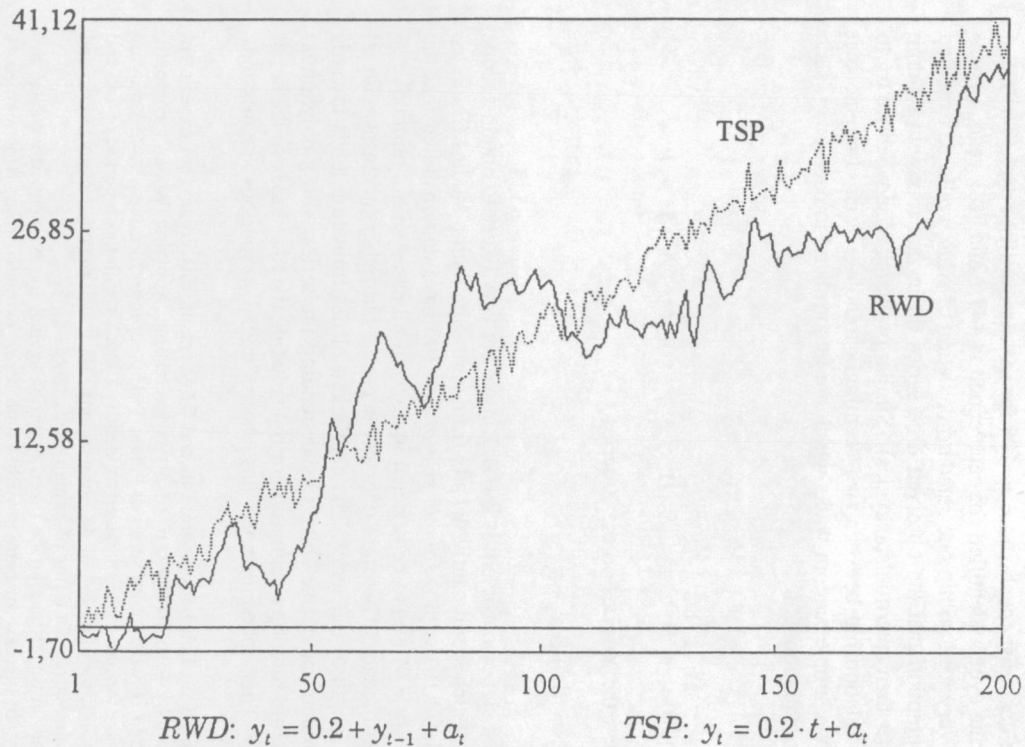


Fig. 4: Realization of a Random Walk with Drift (RWD) in Comparison with a Linear Trend and Added Error Component (TSP)

which has been analytically determined to be 83 % of sample size. Monte Carlo experiments evidenced a mean period corresponding to .65 of sample length with a standard deviation of .21 of sample size.

Now, the size of Weber's sample of British speeches from the Throne is  $n_t = 177$  for the capitalist epoch and  $n_m = 106$  for the mercantilist epoch. The mean period for the cycles identified for each epoch is 147.7 and 70.3 years.<sup>4</sup> These figures lie very well within one standard deviation of what would be expected on the assumption that the observed time series do in fact represent random walks (rather than deterministic cycles).<sup>5</sup> Similar problems arise when we try to model the relationship between two seemingly trending series. Imagine that you have two simple random walks which are uncorrelated to each other, such as stated in equation (2):

$$(2) \quad \begin{aligned} Y_t &= Y_{t-1} + u_t \\ X_t &= X_{t-1} + v_t \\ E(u_t v_t) &= 0 \quad \forall t; \quad E(u_t u_{t-k}) = E(v_t v_{t-k}) = 0 \quad \forall k \neq 0 \end{aligned}$$

If we regress the Y-series on the X-series

$$(3) \quad y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

the expected slope coefficient is  $\beta = 0$ . But we are very likely to obtain a slope coefficient which significantly departs from zero. Simulation studies (Banerjee et al. 1993: 74 ff.) with a sample size of  $n = 100$  have demonstrated that the correct null ( $\gamma = 0$ ) was to be rejected in 75 % of all cases with a nominal 5 % level of significance. This ratio does not decrease but rather increases with an increasing length of the series. The coefficient of determination tends towards 1. This problem cannot be solved by detrending the series with a polynomial function before running a regression or by including the time index in the set of regressors. This phenomenon has been called »spurious regression«

<sup>4</sup> The reported **median periods** are 148 and 72 years. In this context an experiment carried out by Eisner (1990, p. 610f.) is interesting: Applying Weber's computer algorithm to 10 white noise series of length  $T=150$ , after performing a 9-year moving average transformation, he managed to identify L-cycles with mean period of 143 and K-cycles with mean period of 48 observations. They also matched Namenwirth and Weber's sine-curves in terms of explained variance.

<sup>5</sup> This alone, however, cannot be **taken as proof** against the Namenwirth/Weber methodology. In the case of Namenwirth's time series extracted from American party platforms the mean of the identified L-cycles is 165 years, although the sample size is only  $n = 121$  (which is, for other reasons, a dubious result in itself). The reported median is 152 years, the range is from 104 to 232 years! It is very curious that out of 55 series 20 are reported to have a wavelength of exactly 152 and 11 to have a length of 184 years. In fact, the distribution of the observed wavelengths (reported in NW, p. 66) is inconsistent with the assumption of random variation used to justify the standardization of wave lengths and peaks.

(Granger/Newbold 1974). A rule of thumb is: If you have a Durbin/Watson-Statistic lower than the coefficient of determination, you are likely to have spurious regression (Banerjee et al. 1993, p. 81).

What needs to be done, then? First, we have to decide whether or not the series that we are looking at are stationary or not. If they are stationary, then, of course, one should not do any detrending at all. If the series are non-stationary one has to find out whether the trend is deterministic or stochastic. If it is deterministic, applying a polynomial or some other function of time or including the time index (possibly raised to powers larger than one) in the regression equation - these are both appropriate means in taking account of the trend. The processes are called »trend stationary«. These techniques are inappropriate, however, if the trend is stochastic. In such a case, differencing the series once or twice is the appropriate means to render the series stationary and to avoid the pitfalls of spurious regression. This type of processes is called »difference stationary«. One also speaks of »integrated processes« of »order one« - if the series needs to be differenced once to become stationary - or of »order two«, if the series that has been differenced once needs to be differenced once more. (The term »integration« refers to the fact that the level series may be retrieved by cumulatively adding up the differences.) Still another terminology used in this context speaks of »unit root processes«, because, mathematically speaking, differencing as a means of handling trend is appropriate only if the autoregressive lag-polynomial contains one or more unit roots. This is the common ground for various »unit root tests« that have been designed to test for the presence of a unit root against the alternative of level- or trend-stationarity. There is an ongoing, vivid exchange of arguments on the merits and flaws of these tests, which I am not going to discuss here.<sup>6</sup>

In case of doubt, whether the trend is deterministic or stochastic, it has generally been argued that with differencing one should be on the safer side causing little harm even in cases where the trend is deterministic. However, differencing may have rather troublesome consequences in at least two instances: first, if the series is not integrated and contains a cyclical component; second, if two or more series, whose structural relationships one wants to model, are not just integrated, but are »co-integrated«. The common ground for both problems is the fact that differencing eliminates or reduces the weight of low-frequency components, i. e., it not only eliminates the trend, but also cycles that are either deterministic or emanate from stationary second- or higher order autoregressive processes. I will only offer cursory remarks on the problem of identifying cycles and then turn to co-integration.

---

<sup>6</sup> See R. Metz (1995) for a comprehensive review.

### 3. Problems in Identifying Cycles

The standard method of identifying cycles is spectral analysis. This approach, as already mentioned, fails if the series is not stationary and/or if the cycle one wants to study does not repeat itself several times during the observation period. So, at any case, one needs to make a decision on stationarity or non-stationarity first.

As already noted, during a limited period of observation the *random walk* mechanism may generate a series which might give the appearance of a cyclical process. In this case, if the cycle disappears when taking first differences, this is perfectly alright. If the series is integrated, differencing leaves intact any other components such as deterministic or stochastic cycles. If the series is not integrated, differencing does distort or eliminate cycles. So, again, one needs to know in advance whether or not the series is stationary, trend-stationary or difference-stationary. Experts argue that unit root tests generally have too little power to discriminate, within finite series, between the different mechanisms that may have produced observations that give the appearance of some cyclical movement (cf. Metz 1995) The problem becomes aggravated by the fact that it is nearly impossible to model long-term stochastic cycles by estimating second order autoregressive models. Although there is a theoretical relationship between the coefficients of such a model and the wave-length of the cycle, this relationship is non-linear, and in the low frequency domain very small changes in parameter estimates may make for verly large differences in estimated wave length.<sup>7</sup>

Linear filter techniques have often been interpreted as a theory- and model-independent instrument for determining the cyclical properties of a time series (cf. Brandner and Neusser 1992; Smeets 1992). In the light of several papers published recently (see Nelson & Plosser 1982; Harvey & Jäger 1993; King & Rebelo 1993; Jäger 1994) this claim is unwarranted. It has been shown, for example, that, depending on a certain filter parameter and the variance of the noise component, the filter produces artificial cycles when applied to simple random walk processes. The filter technique is optimal only in cases where the time series follows an additive components model, in which the oscillations pertaining to different components do not overlap. The noise component introduces further problems. Obviously, any filter which extracts a certain range of frequencies from a white noise process, will produce a cycle corresponding to those frequencies. Also, if two overlapping cycles produce large sidelopes in the spectrum of a finite series, a filter which extracts a frequency band that includes lefthand and righthand portions of these sidelopes may create an artificial cycle out of the spectral mass of the sidelopes. Generally speaking, the filter design approach works more reliably in

---

<sup>7</sup> See the example given in Thome (1994) and the detailed discussion in Rahlf (1996).

disproving than in proving the presence of a hypothesized cycle. On a descriptive level, a well designed filter reproduces and translates into the time domain any prespecified frequency components, which are present in a time series, with only minor errors. If the filter fails to reproduce a certain cyclical component assumed to be present in the series, sufficient evidence is provided that such a component is indeed missing in the series. In a recent paper, Thome & Rahlf (*forthcoming*) this property of the filter was used to examine the cycles which Weber assumes to be present in the British value indicator series. Our results call into doubt the claims which he has put forward. But I do not want to get into this here in more detail; instead, I want to turn to the concept of cointegration.

#### 4. The Concept of Cointegration

Two or more time series are said to be co-integrated if two conditions prevail: 1st, the series must be integrated to the same order, 2nd, there must exist at least one linear combination among the series which is stationary. If there is only one such linear combination, we can easily obtain it by way of regressing one series upon the other(s):

$$(4) \quad y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

The residuals are a linear combination of the other two series. Generally, linear combinations of first order integrated series are again integrated to the first order. But under specific conditions they are stationary. Without getting lost in formal derivations, we can develop the following line of reasoning:

Imagine two time series, unit root processes, each dominated by stochastic trend components. If they are integrated to different orders they cannot be structurally related in their long-term development. If they are integrated to the same order, these stochastic movements may or may not be related to each other. If we find a close correspondence between the up and down movements of different series, either positively or negatively, we are quite confident that there is indeed a structural, causal relationship between the series, precisely because of their stochastic nature. Without a maintained causal relationship, stochastic movements cannot be expected to be coordinated in such a way. So, if in a static regression of two equally integrated series we produce stationary residuals, this supports the assumption that there must exist a long-term equilibrium relationship between them. Externally induced departures from the equilibrium trigger more or less rapid readjustments. Differencing the series would eliminate this long term co-movement, and one might possibly find no indication for a structural relationship at all, thus falling victim not to spurious causality but to spurious non-causality.

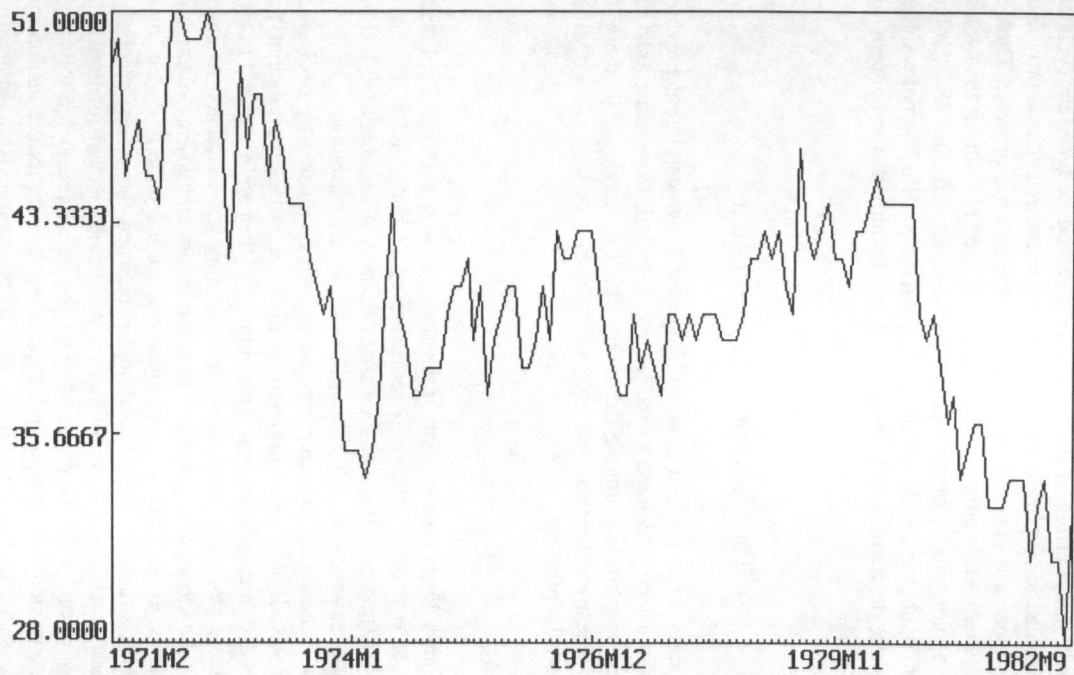


Fig. 5: Percentage Shares of SPD Supporters, Febr. 1971 - Sept. 1982

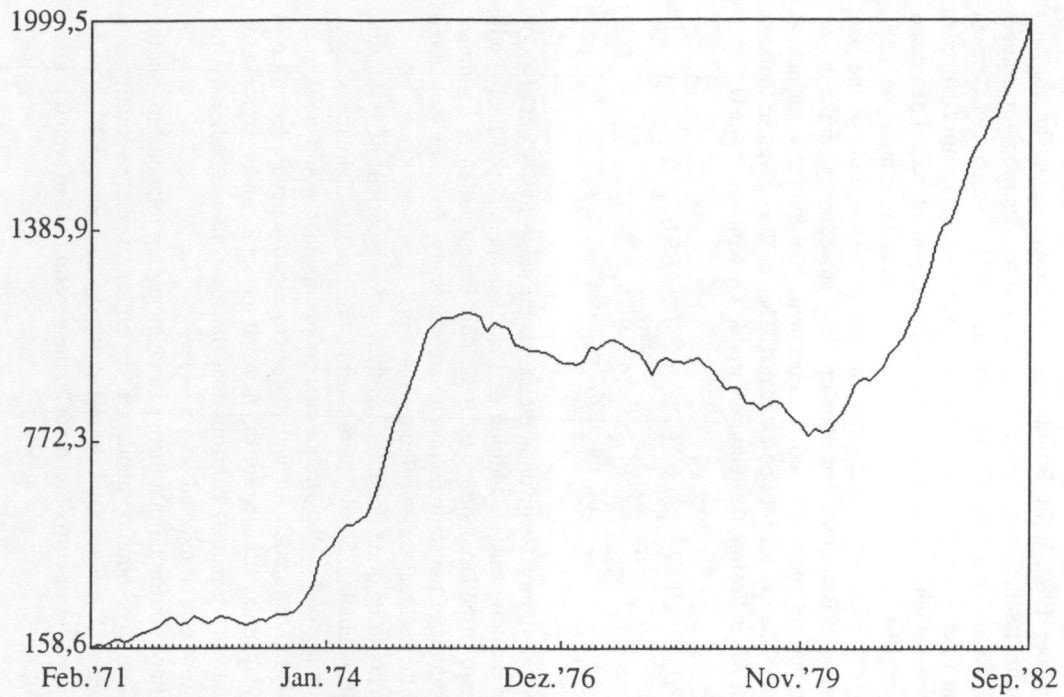


Fig. 6: Desaisonalized Unemployment Figures (in Thousands)



The concept of co-integration can be illustrated by way of giving an example. The two series to look at are presented in Figures 5 and 6. The first of these figures represents the percentage share of supporters of the social democratic party within the German electorate polled monthly from February 1971 to September 1982. Willy Brandt was the German chancellor till May 1974 when Helmut Schmidt took over till the end of the coalition government of social and liberal democrats. Figure 6 presents the deseasonalized German unemployment figures. Unit root tests confirm that both series are first order integrated processes without a deterministic trend component. Next I examined the assumption that the two series are cointegrated by running a static regression in which I included a dummy variable to take account of the last, agonizing months of Brandt's chancellorship. This procedure is admittedly *ad hoc*, but without this sort of outlier-adjustment, the long-run impact of unemployment could not reliably be estimated. The result is given in the next equation (with *PK* indicating the dummy variable for political context):

$$(5) \quad SPD_t = 48.76 - 9.076PK_t - .009ARBLO_t - \varepsilon_t$$

$$\quad \quad \quad (.459) \quad (.952) \quad (.00047)$$

$$\quad \quad \quad \bar{R}^2 = 0.74 \quad DW = .716$$

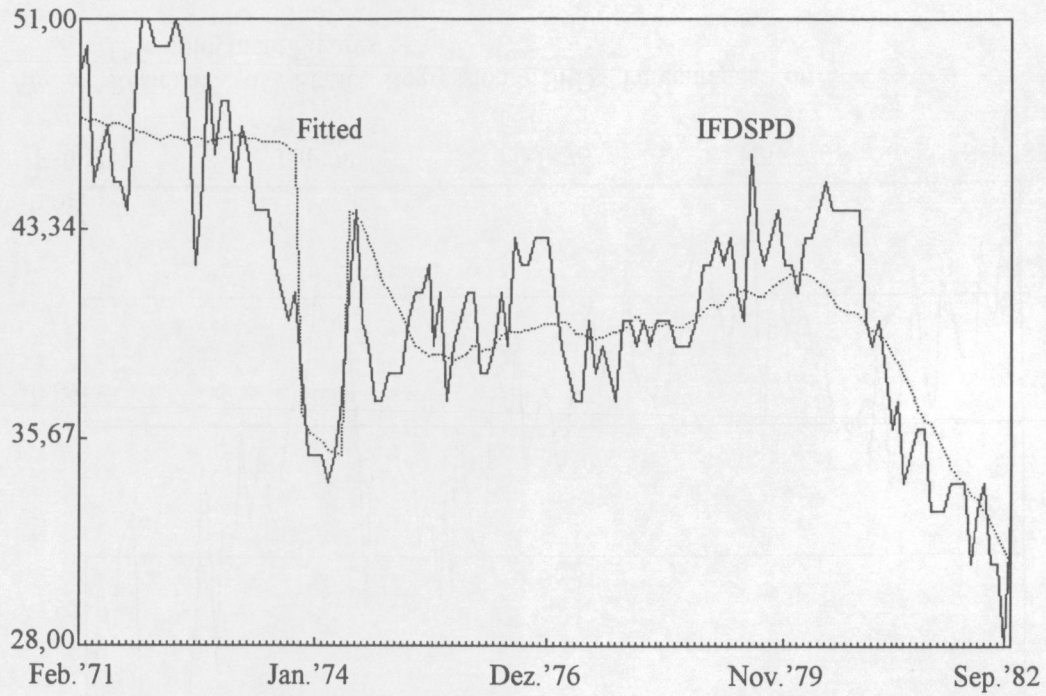
Since employment figures are given in thousands the slope coefficient of  $-.009$  means that, in the long run, an addition of 1 million unemployed people will reduce the share of SPD supporters by 9 %. This interpretation is legitimate only if the residuals of the static regression are stationary. So let us look at the model fit and the residual series (Figures 5 and 6).

The fit appears to be quite good, though there are sub-periods of systematic over- and underestimation. Surely, there must be other political situations, apart from the agonizing end of the Brandt chancellorship, that would need to be incorporated within the model. Anyway, the augmented Dickey/Fuller unit root test (with critical values adjusted for the fact that the series under examination is not observed but estimated) confirms that the residuals are stationary, the null of a unit root can be rejected on a 5 % level.

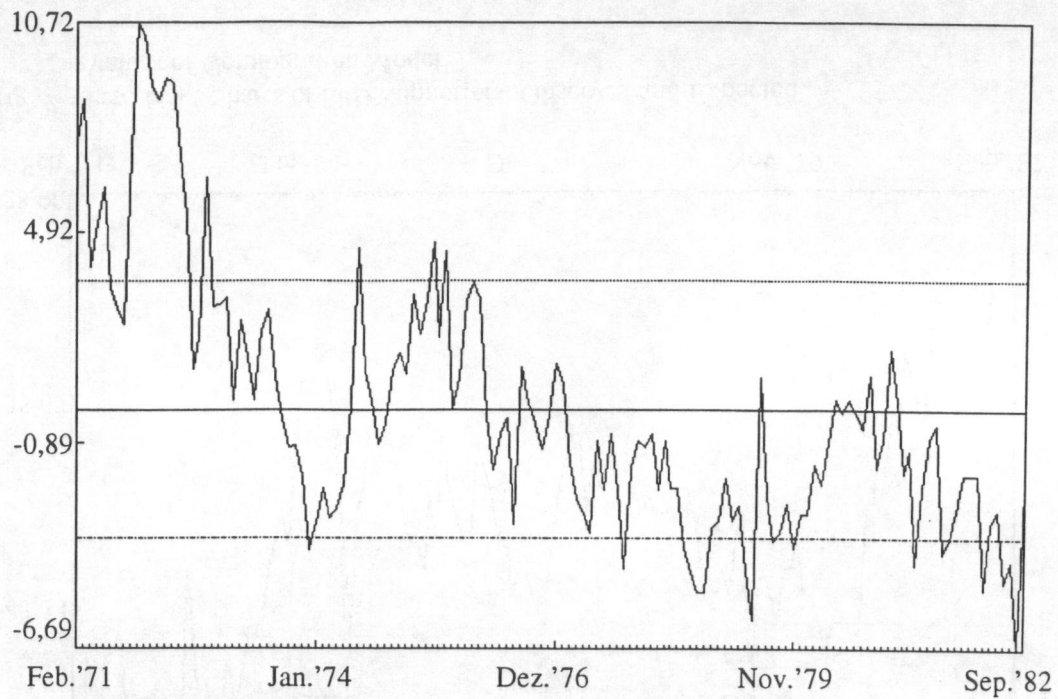
So far, we have confirmed that there is a cointegrating relationship, but the static regression says nothing about the dynamics of the re-equilibration processes. Engle and Granger have suggested that the dynamic specification be given in a so-called »error correction model«, the general form of which is given in equation (6):

$$(6) \quad \Delta Y_t = \alpha + \sum_{j=1}^k \delta_j (\Delta Y_{t-j}) + \sum_{j=0}^l \beta_j (\Delta X_{t-j}) + \gamma Z_{t-1} + \varepsilon_t$$

It contains the departure from the equilibrium, i. e. the residuals of the co-integrating equation as one of the regressors. Note that the dependent



*Fig. 7:* Percentage Shares of SPD Supporters: Observed and Expected Values of Cointegration Model



*Fig. 8:* Residuals of Static Regression: SPD Percentages on Unemployment Figures

variable and the regressor variables are all stationary. The (negative) coefficient  $\gamma$  gives the rate of readjustment towards the equilibrium. A coefficient  $\gamma = 0.60$ , for example, means that a dis-equilibrium  $z_t$  will be reduced by a rate of 60 % in the following interval; the remaining dis-equilibrium will be reduced again by 60 % in the next interval, and so on. In our example, the following results were obtained:

$$(7) \quad \Delta SPD_t^* = .07 - .0148 (\Delta ARBLO_t) - .384 (SPD - 48.76 + 9.076 PK + .009 ARBLO)_{t-1} + \varepsilon_t$$

(.176) (.006)  
(.078)

The rate of re-equilibration is about 38 %. There is an instantaneous reaction to a change of the unemployment figures which exceeds the long-term impact: each 100,000 increase in unemployment instantaneously reduces SPD support by 1.4 %, though the long-term decrease in support is, as we have already noted, «only» 9 %, not 14 %.

This two stage modelling procedure: first the static co-integration regression and then the error correction model each estimated by OLS, has some drawbacks. The OLS coefficients of the static regression are consistent, but not efficient, and may be quite biased in small samples - depending, among other things, upon the coefficient of determination related errors (Hamilton 1994, p. 587). The estimates, however, are not efficient and biased in small samples; the usual t-tests are not reliable (Muscatelli & Hum 1995, p. 173). Alternative estimations procedures based on the maximum likelihood principle have been developed by Johansen (for references see Muscatelli & Hum 1995 or Hamilton 1994).<sup>8</sup> A number of alternatives have been suggested, e. g., Engel and Yoo have proposed a full information maximum likelihood estimation to be realized by adding a third estimation stage. Another proposal is to estimate the dynamic specification and the long-term equilibrium simultaneously in a single equation:

$$(8) \quad \Delta SPD_t^* = \alpha + \beta (\Delta ARBLO_t) + \gamma SPD_{t-1} - \theta ARBLO_{t-1} + \varepsilon_t$$

In this equation the cointegration coefficient is given by  $\omega/\gamma$  which can be estimated by least-squares methods. In our example we obtain  $\tau = -.003158$  and  $\gamma = -.38887$ , thus  $\omega = -.00827$  - an estimate which is only slightly below the one we obtained in the static regression. The instantaneous effect rises

<sup>8</sup> The consistency of the estimates rests upon the assumption (a) that all variables of the cointegrating regression are integrated to the same order, (b) that there exists only one (no more) cointegrating function among the variables (Muscatelli & Hum 1995). Consistency also applies in the case of autocorrelated errors (Hamilton 1994, p. 587). The estimates, however, are not efficient and biased in small samples; the usual t-tests are not reliable (Muscatelli & Hum 1995, p. 173). Alternative estimations procedures based on the maximum likelihood principle have been developed by Johansen (for references see Muscatelli & Hum 1995 or Hamilton 1994).

slightly from  $\beta = -.0148$  to  $\beta = -.016$ . The re-equilibration coefficient remains nearly the same with  $\gamma = -.3888$ .

Equation (8), too, is valid only if both series are indeed integrated to the same order. Thus, the problematics of unit-root testing carries over to co-integration modelling. Confidence in our results is increased if they are obtained by applying alternative testing and estimation procedures. The possibility to estimate simultaneously long-term equilibrium effects and short-run dynamics is a considerable advancement. The concept of cointegration helps to resolve the long-standing debate about the virtues and flaws of level vs. difference regression, and it helps to avoid the dilemma of either falling victim to spurious causality or to spurious non-causality.

## Bibliography

- Banerjee, A./ J. J. Dolado/ J. W. Galbraith/ D. F. Hendry (1993): Co-integration, error correction, and the econometric analysis of non-stationary data. Oxford: Oxford University Press.
- Brandner, P./ K. Neusser (1992): Business Cycles in Open Economies. Stylized Facts for Austria and Germany. In: *Weltwirtschaftliches Archiv* 128, 67-87.
- Eisner, M. (1990): Long-term dynamics of political values in international perspective: Comparing the results of content analysis of political documents in the USA, GB, FRG, and Switzerland. In: *European Journal of Political Research* 18, 605-621.
- Granger, C. W. J./ P. Newbold (1974): Spurious regressions in econometrics. In: *Journal of Econometrics*, 2, 111-120.
- Hamilton, J. D. (1994): *Time series analysis*. Princeton, N. J.: Princeton University Press.
- Harvey, A. C./ A. Jäger (1993): Detrending, Stylized Facts and the Business Cycle. In: *Journal of Applied Econometrics* 8, 231-247.
- Jäger, A. (1994): Mechanical Detrending by Hodrick-Prescott Filtering: A Note. In: *Empirical Economics* 19, 493-500.
- King, R. G./ S. T. Rebelo (1993): Low frequency filtering and real business cycles. In: *Journal of Economic Dynamics and Control*, 17, 207-231.
- McCleary, R./ J. r. Hay (1980): *Applied Time Series Analysis for the Social Sciences* (with E. E. Meidinger and D. McDowall). Beverly Hills and London: Sage.
- Metz, R. (1995): *Stochastische Trends und langfristige Wachstumsschwankungen. Neue Forschungsansätze und ihre theoretische und empirische Relevanz für die Wirtschaftsgeschichte* (Habilitationsschrift). Universität St. Gallen.
- Muscattelli, V. A./ Hurn, S. (1995): Econometric modelling using cointegrated time series. In: L. Oxley/D. A. R. George/C. J. Roberts/S. Sayer (Hrsg.), *Surveys in Econometrics*. Oxford: Blackwell, 171-214.

- Namenwirth, J. Z. (1973): *Wheels of Time and the Interdependence of Value Change in America*. In: *Journal of Interdisciplinary History* 3, 649-683.
- Namenwirth, J. Z./ R. P. h. Weber (1987): *Dynamics of Culture*. Boston: Allen & Unwin.
- Nelson, Ch. R./Kang, H. (1981): *Spurious periodicity in inappropriately detrended time series*. *Econometrica* 49: 741-751.
- Nelson, Ch. R/Kang, H. (1984): *Pitfalls in the use of time as an explanatory variable in regression*. *Journal of Business & Economic Statistics* 2: 73-82.
- Nelson, Ch. *RJ* Ch. I. Plosser (1982): *Trends and random walks in macro-economic time series. Some evidence and implications*. In: *Journal of Monetary Economics* 10, 139-162.
- Rahlf, Th. (1996): *Zur Modellierung stochastischer Zyklen als AR(2)-Prozeß. Der Hallesche Graureiher 96-3. Forschungsberichte des Instituts für Soziologie*.
- Smeets, H.-D. (1992): *»Stylized Facts« zum Konjunkturverlauf in der Bundesrepublik Deutschland*. In: *Jahrbücher für Nationalökonomie und Statistik* 210, 512-532.
- Thome, H. (1994): *Univariate Box/Jenkins-Modelle in der Zeitreihenanalyse. Historical Social Research / Historische Sozialforschung* 19, Heft 3: 5-77.
- Thome, H. (with Thomas Rahlf, in press): *Dubious cycles: A methodological critique of the Namenwirth-Weber thesis on cultural change with an introduction into filter design methods (appears in: Quality & Quantity)*.