Forecasting Changing Seasonal Components in German and US Unemployment using Periodic Correlations
Tinbergen Institute Discussion Paper 95-77
ISSN 0929-0834

Philip Hans Franses   Marius Ooms

March 1995

Econometric Institute
Erasmus University Rotterdam
P.O. Box 1738
NL- 3000 DR Rotterdam
The Netherlands
email: ooms@ect.few.eur.nl

Abstract

An important underlying idea in seasonal adjustment procedures is the presumed orthogonality of seasonal and nonseasonal components. In this paper we examine this orthogonality assumption for official quarterly unemployment figures in Germany and the USA. Although *nonperiodic* correlations do not seem to reject the orthogonality assumption, a *periodic* analysis based on correlation functions that vary with the seasons indicates the violation of orthogonality. For example, we find the correlation between the seasonal and nonseasonal components to be significantly positive in the first quarter and negative in other quarters. This implies that seasonally adjusted figures of the number of unemployed are likely to be "too optimistic" in the first quarter in times of rising unemployment. We find significant predictability of seasonal components when we base such forecasts on periodic autoregressions with unit roots. Similar results are obtained when we construct forecasts for seasonal components of so-called structural time series models [Harvey(1989)]. In addition to these empirical findings, we discuss the usefulness of periodic models with unit roots to assess the relative adequacy of additive and multiplicative models of (long-run) seasonality.

Keywords

seasonal adjustment, periodicity, unit roots, predictability.

Acknowledgements

Philip Hans Franses gratefully acknowledges financial support from the Royal Netherlands Academy of Arts and Sciences. Earlier versions of this paper were presented at the $(EC)^2$-meeting in Oxford (1993) and at econometrics seminars at the University of Montreal and the Southern Methodist University in Dallas, the EUROSTAT/INSEE workshop on quarterly national accounts in Paris (1994), and the North American Winter meeting of the Econometric Society in Washington (1995). We thank two anonymous referees and participants at the seminars for helpful comments. We thank Siem Jan Koopman for his permission to use recent versions of the computer package STAMP.

1 Introduction and Summary

An important underlying concept in most seasonal adjustment procedures is the presumed orthogonality of seasonal and nonseasonal components. Therefore, the empirical correlation between contemporaneous (changes) in estimates of the components of the series is typically small. In this paper we suggest new methods to examine this empirical orthogonality for estimated seasonal components.. For this purpose, we use periodic models, which allow for seasonally varying parameters. Additionally, we evaluate several seasonal adjustment methods.

One of the basic ideas in this paper is to investigate the empirical correlation between seasonal and nonseasonal components by computing the correlations between yearly changes in the components for each quarter separately. In this paper we show for two specific examples that these so-called periodic correlations can be far from zero and that they can differ significantly from quarter to quarter. We find convincing evidence of this phenomenon in German and US series of the total number of unemployed. For example, for the German data the correlation is large and significantly positive in the first quarter and negative in the other quarters. In an unobserved components interpretation this entails that estimated changes in the seasonal pattern are biased upwards with rising unemployment in the first quarter, while they are biased downwards for the other quarters. In a sense, this implies that seasonally adjusted figures in times of rising unemployment can be "too optimistic" in the first quarter and "too pessimistic" in the other quarters. We find this bias in the first quarter to be about 10 percent of the yearly change in unemployment. Since these findings are based on official OECD data, where variants of Census X-11 methods are used for seasonal adjustment, we investigate the robustness of our findings using data-dependent Structural Time Series Models [STM] advocated by Harvey (1989).

In order to study the periodic variation in the correlations between seasonal and nonseasonal components we use linear AR models with periodically changing coefficients to describe the unadjusted series. The behavior of the seasonal pattern and its relation with the trend-cycle can be evaluated using standard tests for integration and cointegration. We find that the number of independent sources of nonstationarity (i.e. the number of common (stochastic) trends) is often smaller than in traditional nonperiodic seasonal ARIMA models. This indicates that a concept of an independent seasonal component as it exists in the STM framework is then no longer applicable. Furthermore, we provide an alternative interpretation of cointegration of the seasonal component with the trend component with seasonally varying cointegrating vectors. This interpretation suggests an alternative method of modeling stochastic seasonality of a multiplicative nature, which, in contrast to the log transformations, can be applied to series with both positive and negative values.

The structure of this paper is as follows. In section 2 we present a predominantly graphical analysis of the periodic correlations. In section 3 we discuss estimation and interpretation of a convenient class of so-called Periodically Integrated Auto-Regressive [PIAR] Models, which can be useful to describe the periodic correlations between the seasonal and nonseasonal components. In section 4 we discuss the estimation results for the German and US unemployment data. The method works especially well for the German unemployment data without a prior log transformation. In an appendix we treat the effects of seasonal adjustment using Harveys STMs in more detail. We apply STMs to obtain alternative estimates of seasonal and nonseasonal components.

In section 5 we show that changes in the seasonal components from different seasonal adjustment methods are significantly predictable. The forecast function for the German data is strongly periodic. In section 6 we conclude this paper with some remarks and suggestions for practical application.

2 Some Stylized facts of Unemployment in Germany and the USA

In this section we first analyze the untransformed German and US unemployment data and their official seasonally adjusted (SA) and unadjusted (NSA) time series. The quarterly time series are taken from the OECD Main Economic Indicators and they span the period 1960.1-1987.4. Next, we analyze the same data when transformed using the natural logarithmic transformation. Finally we present the results of a sensitivity analysis with respect to the seasonal adjustment method and the effect of influential observations.

First we introduce some notation. Let y_t denote the (not seasonally adjusted) observations for n subsequent quarters, $t = 1, 2, \ldots, n$. Suppose we have N years of subsequent observations with quarterly data. Let $Y_{s,T}$ denote these observations, $s = 1, 2, 3, 4$, $T = 1, 2, \ldots, N$, so that s denotes the number of the quarter and T denotes the number of the year. Suppose we have complete years with observations starting in quarter 1, so that $t = s + 4 \cdot (T - 1)$. Note that we analyze only quarterly data in this paper, but that most of the material can easily be applied to, e.g., monthly time series.

Insert Figures 1 *and* 2

In Figures 1 and 2 the unemployment series $Y_{s,T}$ are compared with previous years' levels, $Y_{s,T-1}$, separately for each quarter. It is clear that the levels are not stationary in mean, i.e. it seems that there is positive mean growth or persistence of upward shocks in the German series. The distance between the level series of the different quarters does not seem to be stationary either, which suggests persistent changes in the seasonal pattern. These changes may be due to a partly multiplicative nature of the seasonal movements. It is for this reason that we repeat our analysis for the log levels later in this section.

Insert Figures 3 *and* 4

In order to get a clearer picture of the nonstationarities in seasonal pattern we plot the *quarterly changes*, $Y_{1,T} - Y_{4,T-1}$, and $Y_{s,T} - Y_{s-1,T}$, $s = 2, 3, 4$, in Figures 3 and 4. It is clear that in downturns of the business cycle, e.g. in Germany in 1967 and in the US in 1975, the growth of unemployment in the first quarter increases relative to the growth rates in the other quarters in Germany and the US whereas the growth rate in the second quarter shows a corresponding relative drop. Apart from these seemingly cyclical changes in the seasonal pattern, we observe trend-like behavior in the first and second quarter growth rates of German unemployment. This corresponds to the movement of the level of the first quarter away from the other quarters observed in Figure 1. The differences between the other quarters decrease while this gap increases, which suggests that the seasonal pattern may not be purely multiplicative, since in that case one would expect proportional growth of the overall level and the size of the seasonal pattern.

Insert Figures_5_and_6

Next we depict the "official" seasonal components in Figures 5 and 6, as published by the OECD. Let $Y_{Ss,T}$ denote the (unobserved) seasonal component in quarter s of year T. Let $\widehat{Y_{Ss,T}}$ denote its estimate. Let $Y_{Ns,T}$ denote the nonseasonal component, and let $\widehat{Y_{Ns,T}}$ denote its estimate. These components obey the theoretical decomposition $Y_{s,T} = Y_{Ss,T} + Y_{Ns,T}$. In most cases the empirical decomposition satisfies the same restriction: $Y_{s,T} = \widehat{Y_{Ss,T}} + \widehat{Y_{Ns,T}}$. In this paper we study the empirical decomposition of $Y_{s,T}$ (or possibly $\ln Y_{s,T}$). The stylized facts about the seasonal components $\widehat{Y_{Ss,T}}$, seen in the previous figures show up clearly in Figures 5 and 6. The adjustment procedures allow for substantial yearly changes in the seasonal components with a maximum of about 79,000 people in Germany (1967.1) and 336,000 people in the USA (1975.1). Obviously, these estimates reduce the increase in the adjusted unemployment figures substantially.

Correlation Seasonal and Nonseasonal components

Next we examine the connection between the yearly changes in the estimates of the seasonal components $\Delta\widehat{Y_{Ss,T}} = \widehat{Y_{Ss,T}} - \widehat{Y_{Ss,T-1}}$, and the nonseasonal components $\widehat{Y_{Ns,T}}$ more closely. First we present the ordinary nonperiodic (i.e. pooled) regression results in Figures 7 and 8.

Insert Figures 7 *and* 8

At first sight there seems no significant correlation. Hence, the orthogonality assumption seems to reflect itself rather well in the component estimates. Notice that this seems to contrast with the strong relationships between the components visible in Figures 5 an 6.

To investigate the robustness of the orthogonality assumption across the four seasons, we depict $\Delta\widehat{Y_{Ss,T}}$ and $\Delta\widehat{Y_{Ns,T}}$ for each of the seasons in Figures 9 and 10.

Insert Figures 9 *and* 10

This idea of periodic regression has been used in earlier studies to investigate the relationship between trend and seasonal component. For example, Shiskin and Eisenpress (1957, p. 441) applied it to the *levels* of the (provisional) seasonal and trend-cycle component. One expects a significant linear relationship in case of a purely multiplicative seasonality and no significant linear relationship in case of purely additive seasonality. The relative advantage of our approach is that the application of the yearly difference operator makes both components approximately stationary in many situations. This facilitates the statistical interpretation of the observed correlations. Below we discuss transformation to stationarity in more detail.

One observes from Figures 9 and 10 that it is not a coincidence that large changes in the seasonal component correspond with rapid rises in the yearly level of unemployment. Note that the periodic regressions are based on *annual* data. There appears to be a strong positive correlation between yearly changes in the seasonal component and yearly changes in the nonseasonal component in the first quarter. The regression "explains" about 30,000 of the increase in the seasonal component of German unemployment in 1967.1 by a change in the nonseasonal component. This indicates that the seasonal and nonseasonal components are not independent[1]. Note that total unemployment at that time was only about 500,000 people! A similar underestimation of the nonseasonal component occurred in in Germany in 1975.1, 1982.1 and 1983.1. In the USA, a systematic underestimate of about 300,000 people can be noted in 1975.1. The opposite effect occurs when there is a sharp drop in unemployment. An overestimate of more than 200,000 people can be seen for 1984.1.

Next we crudely estimate the statistical significance of the periodic correlations. Since we observe from Figures 9 and 10 that the correlations vary periodically, and hence that inference based on a pooled regression is not reliable, we modify the regression in two ways. First, most seasonal adjustment procedures introduce (near unit root) MA type correlation at yearly lags[2] and therefore we take the unknown autocorrelation structure of the error term of the

[1]A significant correlation between changes in the *estimated* seasonal and nonseasonal component is not a weakness of a seasonal adjustment method per se. Well known methods like the additive versions of the structural time series decomposition proposed by Harvey (1989) and the canonical decomposition advocated by Maravall and Gomez (1992) have this feature even if they are applied to the (nonperiodic) seasonal ARIMA processes for which they are designed. However, a *periodically changing* correlation does not appear as a result of their methods.

[2]In appendix B we elaborate on the autocorrelation function of the seasonal component and it cross-correlation function with the non-seasonal component, corresponding to explicit seasonal adjustment methods.

periodic regressions into account by allowing for a second order MA structure in the yearly observations. A second modification of the regression is the reduction of the effect of influential observations by the introduction of dummy variables. This makes the normal approximation more appropriate. In short we run the following regressions:

$$\Delta\widehat{Y_{Ss,T}} = \alpha_s + \beta_s \Delta\widehat{Y_{Ns,T}} + \delta_{1s}\widetilde{O}_{s1T} + \delta_{1s}\widetilde{O}_{s2T} + E_{s,T} + \theta_{1s}E_{s,T-1} + \theta_{2s}E_{s,T-2}, \qquad (1)$$

where α_s, β_s, δ_{1s}, δ_{2s}, θ_{1s}, and θ_{2s} denote scalar regression parameters, s denotes the quarter, $\widetilde{O}_{sjT}$ denotes a dummy variable which equals 1 in year j and zero elsewhere, and $E_{s,T}$ denotes the error term at time T. For Germany we set j equal to 1967 and 1974, while for the USA j corresponds to 1975 and 1982.

Insert Table 1

Table 1 presents the results of these modified regressions, which give evidence against the null of no correlation in most seasons at very low significance levels. It is clear from this table that in case one wants to describe these unemployment series, one should allow for periodically varying parameters. Table 1 also indicates that there is a very notable difference from quarter to quarter in the correlation coefficient and the residual variance, even though we reduce the influence of the two most striking observations by the introduction of dummy variables. These regressions clearly indicate the relevance of a periodic analysis of the raw data.

Results for data in logs

It may of course be that the periodic correlation structure found in Figures 9 and 10 and in Table 1 can be attributed to the presence of ordinary multiplicative seasonality. Therefore, we replicate the above analysis for the unemployment data after applying the natural logarithmic transformation. In case of multiplicative seasonality one would expect smaller absolute correlations.

Insert Figures 11 *and* 12

In Figures 11 and 12 we present the *pooled* nonperiodic scatter plots for the relation between yearly changes in *log* levels of the seasonal and the nonseasonal component. Similar to figures 7 and 8, the presumed nonperiodic orthogonality of these components seems well established.

The periodic analysis in Figures 13 and 14 (to be compared with Figures 9 and 10) is again more interesting. From figure 13 it is clear that the seasonal component in the German data is not purely multiplicative. This is most obvious for the first quarter. The picture is now even more dominated by the recession and recovery figures (around 1967) where the most positive and most negative relative changes in the number of unemployed occurred, which here correspond to relative changes in the seasonal component *with opposite sign*. There is also severe time (and unemployment level) dependent heteroskedasticity in the changes in the seasonal component, with high variance steadily declining from the 60s to the 80s: The changes belonging to the late 70s and the 80s depicted in Figure 13 are all relatively close to zero. Apparently, a standard multiplicative seasonal adjustment procedure does not work satisfactorily. Other (non-linear and phase-of-the business-cycle dependent) relationships between trend-cycle and season have been suggested for this kind of data, see e.g., Van der Hoeven and Hundepool (1986) . However, the scatter plots in Figure 13 do not suggest a clear alternative here.

For the US data in Figure 14 we observe that a multiplicative model for the seasonal component may be more useful. The exceptional correlation in the first quarter is far less pronounced than in Figure 10. It is still not negligible though. Note that for the USA data there is no changing sign phenomenon as for the German data. Another striking phenomenon which comes out more clearly in the logarithmic data is the permanent change in the seasonal component (quarters 2, 3 and 4) in 1967, since the observation for this year lies far away form the regression lines in the scatter plots in Figure 14. Notice that, using a nonperiodic analysis, this phenomenon would probably go unnoticed.

Insert Table_2

In Table 2 we report on repeated estimations for the auxiliary regression in (1) where data are transformed by taking logs. The patterns suggested by the graphs in Figures 11 and 12 carry over to the significance of the parameter estimates $\widehat{\beta_s}$ for Germany with opposite sign as in Table 1.

Sensitivity Analysis

In the last part of this section we investigate whether the significance of the periodically changing correlations is specific to the seasonal adjustment method of the OECD and the treatment of influential observations. We also investigate the robustness with respect to the treatment of residual autocorrelation by varying the order of the MA process. More precise, we examine 3 MA orders, 2 degrees of outlier correction, and 3 methods of seasonal adjustment. The analysis results in 18 different estimates of the statistical significance of the periodic correlations. We apply the same procedure to the data in logs. The first adjustment method is the one applied by the OECD. The same adjustment method is used for the levels and the log levels. The second adjustment method uses a standard simple structural time series model from Harvey (1989), which is explained in more detail in appendix B. Here the seasonal component in the model for the levels is additive, the seasonal component in the model for the log levels is multiplicative for the original series. The third method is also based on an STM and it accounts for the most influential observations using extra interventions in the irregular component of the structural time series model. The model-based decomposition of Harvey (1989) is well defined in a statistical context, and it is easily performed using the computer package STAMP, Koopman et al. (1995).

Insert Tables 3 *and* 4

The outcomes are presented in Tables 3 and 4. For the German data in levels the correlation of the first quarter remains significantly positive for all methods. For the log levels this correlation is significantly negative for most methods. For the second quarter the results are comparable with a reverse sign. This corroborates the idea that neither pure multiplicative nor pure additive modeling of seasonality leads to satisfactory results. The results in, e.g., the fourth-quarter rows in Table 3 indicate that the accommodation of influential observations can reduce the significance of the correlation, or it can increase it. The correlation for the OECD data in logs seems to be dominated by two observations. The correlation for the Structural Time Series adjustment seems to be hidden by the effect of two observations.

The correlations for the US data are smaller than for the German data, cf. Table 4. The application of the STM leads to smaller correlations compared with the OECD adjusted data. The smoothness restrictions on the seasonal component estimates of the applied version of the STM are rather effective, resulting in comparatively small year to year changes. Furthermore, the additive adjustment leads to an exceptionally large correlation in the first quarter, see also Table 1. The multiplicative adjustment fares better, but is leaves a large correlation in the fourth quarter.

Summarizing the empirical results in this section, we conclude that the significant periodically varying correlations between changes in the seasonal and nonseasonal component are not likely to be due to the specific adjustment method or the influence of two exceptional observations only. To investigate the properties of the unemployment data, we proceed in the next section with periodic time series models.

3 Periodic models for unemployment

In this section we discuss two partly equivalent models for the periodic correlations of interest: a univariate model with changing coefficients for y_t and a multivariate model with constant coefficients for $Y_T = (Y_{1,T}\ Y_{2,T}\ Y_{3,T}\ Y_{4,T})$. First we present notation. Then we discuss unit root restrictions and integration properties in these models. The case with one unit root restriction turns out to be especially relevant. We treat this case where three long run equilibrium relations between the variables $Y_{s,T}$ appear in more detail. We show how this one-unit-root model can lead to long-run multiplicative seasonality without the usual log transformation. Next we discuss the treatment of deterministic growth components and innovation outliers. All these issues are discussed in order to motivate the possible usefulness of periodic models with unit roots for modeling the two unemployment variables at hand. Finally we present the estimation and testing procedure which we apply in the next section.

Notation

The point of departure is a periodic AR model of order p, i.e.

$$y_t = \phi_{1s}y_{t-1} + \phi_{2s}y_{t-2} + \cdots + \phi_{ps}y_{t-p} + \epsilon_t, \quad t = 1, 2, \ldots, n, \tag{2}$$

with ϵ_t a white noise process, probably with seasonally varying mean μ_s (the well known seasonal dummies) and variance σ_s^2. The AR parameters ϕ_{is} vary with the seasons, $i = 1, 2, \ldots, p$, $s = 1, 2, 3, 4$. We condition on the starting values $y_0, y_{-1}, \ldots, y_{-p}$. For the statistical analysis of this model it is useful to put it in a *companion form* which explains the values of the different quarters $Y_{s,T}, s = 1, 2, 3, 4,\ T = 1, \ldots, N,\ T = n/4$. See Tiao and Grupe (1980) and Lütkepohl (1991) for detailed descriptions of this procedure. Recall that the time series plots in Figures 1 and 2 represent these different quarters. In vector notation the model in (2) can be written as

$$A_0 Y_T = A_1 Y_{T-1} + \cdots + A_P Y_{T-P} + \mu + E_T, \tag{3}$$

with $Y_T' = (Y_{1,T}, Y_{2,T}, Y_{3,T}, Y_{4,T})$, $A_{0ii} = 1$, $A_{0ij} = 0$, $j > i$, $A_{0ij} = -\phi_{i-j,j}$, $j < i$, $A_{kij} = \phi_{i+4k-j,j}$, $i = 1, 2, 3, 4$, $j = 1, 2, 3, 4$, $k = 1, 2, \ldots, P$, $P = 1 + [(p-1)/4]$ if $p > 0$, and μ is a 4×1 vector of constants. $E_T' = (E_{1T}, E_{2T}, E_{3T}, E_{4T})$ is the notation for the zero mean innovations corresponding to Y_T'. Note that periodic AR models up to order $p = 4$ can be captured in a VAR model of order $P = 1$. We present an example with $p = 2$ and $P = 1$ below.

Unit roots and integration properties

The vector system (3) has constant parameters. Assuming appropriate starting conditions the process in (3) is stationary if E_T is stationary and all roots of $D(z) = \det(A_0 - \sum_{i=1}^{P} A_i z^i) = 0$ are outside the unit circle. If so-called "unit roots" $z = 1$ apply, these can easily be incorporated in the estimation and testing procedures for the vector system. The unit root restrictions on the univariate periodic model (2) are nonlinear functions of the parameters ϕ_{is}, see Franses and Paap (1994) for some examples. For $p = 1$ one gets $D(z) = 1 - \phi_{11}\phi_{12}\phi_{13}\phi_{14}z = 0$, so that the restriction

$$\phi_{11}\phi_{12}\phi_{13}\phi_{14} = 1$$

incorporates the unit root. In combination with the inequality $\phi_{1s} \neq \phi_{1s'}$, $s \neq s'$, for at least one s and s' Osborn and Smith (1989) called this property *periodic integration* to distinguish it from *seasonal integration*, the concept known from the Box-Jenkins ARIMA class of models. Franses (1994) proposed a method for the comparison of these two model types.

If all the variables $Y_{s,T}$, are integrated of at most order 1, i.e. I(1) in the sense of Engle and Granger (1987), some roots of $D(z) = 0$ are equal to one and all other roots are outside the unit circle, one can apply the Granger Representation Theorem to rewrite the system in the so-called Vector Error Correction Model (VECM), see e.g., Johansen (1991), as

$$\Delta Y_T = \gamma(\alpha' Y_{T-1}) + B_1 \Delta Y_{T-1} + \cdots + B_{P-1} \Delta Y_{T-P+1} + A_0^{-1} (\mu + E_T), \tag{4}$$

where Δ is the difference operator such that $\Delta Y_T = Y_T - Y_{T-1}$, where α is the $4 \times r$ matrix of *cointegrating vectors*, which makes $\alpha' Y_T$ a stationary r dimensional vector process, and γ is a corresponding $4 \times r$ matrix of "*error correction*" *coefficients*, so that $\gamma\alpha' = A_0^{-1} \left(\sum_{i=1}^{P} A_i - I \right)$. The lagged variables ΔY_{T-i} do not appear for models with $p \leq 4$, since then $P \leq 1$.

For the *seasonally integrated* model, $y_t - y_{t-4} = \epsilon_t$ with ϵ_t homoskedastic white noise, one has $\phi_{41} = \phi_{42} = \phi_{43} = \phi_{44} = 1$ and all other parameters ϕ_{js} in (3) equal to zero. That is the simplest model to satisfy the *nonperiodicity restrictions*: $\phi_{js} = \phi_{js'}$, $\sigma_s^2 = \sigma_{s'}^2$, $j = 1, 2, \ldots p$, $s = 1, 2, 3, 4$, $s' = 1, 2, 3, 4$, *and* 4 unit root restrictions: $D(z) = (1 - z)^4$. A more general AR model satisfying these restrictions, i.e. a nonperiodic AR model solely in terms of $y_t - y_{t-4}$, corresponds in the multivariate notation with $r = 0$, or no cointegrating vectors. In our empirical analysis below, we firmly reject these unit root restrictions of the seasonally integrated model. In fact, the case with *cointegrating rank* $r = 3$ turns out to be empirically more relevant.

Parameterization of three independent cointegrating vectors

If there are three cointegrating vectors it is convenient to normalize α in the following way:

$$\alpha' = \begin{pmatrix} -\alpha_2 & 1 & 0 & 0 \\ 0 & -\alpha_3 & 1 & 0 \\ 0 & 0 & -\alpha_4 & 1 \end{pmatrix}, \text{ so that } \alpha' Y_T = \begin{pmatrix} Y_{2,T} - \alpha_2 Y_{1,T} \\ Y_{3,T} - \alpha_3 Y_{2,T} \\ Y_{4,T} - \alpha_4 Y_{3,T} \end{pmatrix}, \tag{5}$$

which implies stationarity for $Y_{1,T} - 1/(\alpha_2\alpha_3\alpha_4))Y_{4,T-1}$ as well. Let α_1 denote $(1/\alpha_2\alpha_3\alpha_4)$. This gives a first simple, well understood transformation of the data to stationarity. In the simple case where the *ordinary differences* of the series are stationary one has $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$, which implies stationarity for $Y_{1,T} - Y_{4,T}$ as well as for $Y_{1,T} - Y_{4,T-1}$. As discussed in the previous section, the Figures 3 and 4 for the unemployment data indicate that the ordinary differences of the time series do not seem to be stationary. Apparently we need a more flexible model with different cointegrating vectors for the different seasons. For the German data presented in Figures 1 and 3 it is clear that α_1 has to be greater than one and α_2 smaller than one in order to make $Y_{1,T} - \alpha_1 Y_{4,T-1}$ and $Y_{2,T} - \alpha_2 Y_{1,T}$ look stationary. In other words, for this series we wish to investigate the case of one unit root leading to *periodic integration*, i.e. $\alpha_1\alpha_2\alpha_3\alpha_4 = 1$, where not all the parameters α_s are equal to one.

Modeling three independent cointegrating vectors

In order to understand the impact of periodic integration we rewrite model (4) with the cointegrating vectors in (5) as dependent variables. As an illustration we first consider the following periodic AR(2) model with periodic integration, i.e.,

$$(y_t - \alpha_s y_{t-1}) = \beta_s (y_{t-1} - \alpha_{s-1} y_{t-2}) + \epsilon_t. \tag{6}$$

The periodic integration restriction can be written as $\alpha_1\alpha_2\alpha_3\alpha_4 = 1$, where it is understood the subscripts s obey modulo-4 arithmetic, i.e. that $\alpha_s = \alpha_{s+4k}$, $\beta_s = \beta_{s+4k}$, $k = \cdots, -2, -1, 0, 1, 2, \cdots$. Note that deviations from equilibrium are propagated to the next quarter if $\beta_s \neq 0$. This model allows for periodically changing error correction coefficients next to the periodically changing cointegrating vectors. One rewrites (6) as the VAR model (3) using the following matrices for A_0 and A_1:

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\alpha_2 - \beta_2 & 1 & 0 & 0 \\ \beta_3\alpha_2 & -\alpha_3 - \beta_3 & 1 & 0 \\ 0 & \beta_4\alpha_3 & -\alpha_4 - \beta_4 & 1 \end{pmatrix} \text{ and } A_1 = \begin{pmatrix} 0 & 0 & -\beta_1\alpha_4 & \alpha_1 + \beta_1 \\ 0 & 0 & 0 & -\beta_2\alpha_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Note that the parameters β_s appear in the s-th *rows* of the parameter matrices, and that the parameters α_s appear in the $(s-1)$-th *columns* of the parameter matrices corresponding to the effect *of* the deviation from the equilibrium involving $Y_{s-1,T}$. The characteristic roots of the system are the solutions of $\det(A_0 - A_1 z) = (1 - \alpha_1\alpha_2\alpha_3\alpha_4 z)\cdot\ (1 - \beta_1\beta_2\beta_3\beta_4 z) = 0$. The periodic integration restriction $\alpha_1\alpha_2\alpha_3\alpha_4 = 1$ is seen to lead to one unit root. We assume that the other root is stable: $\mid \beta_1\beta_2\beta_3\beta_4 \mid < 1$. The annual series are I(2) processes if $\beta_1\beta_2\beta_3\beta_4 = 1$ as well, which is most easily seen for the case where all parameters α_s and β_s equal one.

A more general model for three cointegrating vectors can be derived from (4) using (5) . It reads:

$$\begin{aligned} A_0 G^{-1} \begin{pmatrix} Y_{1T} - \alpha_1 Y_{4,T-1} \\ \alpha' Y_T \end{pmatrix} &= F_1 \begin{pmatrix} Y_{1,T-1} - \alpha_1 Y_{4,T-2} \\ \alpha' Y_{T-1} \end{pmatrix} + \cdots + \\ & F_P \begin{pmatrix} Y_{1,T-P} - \alpha_1 Y_{4,T-P-1} \\ \alpha' Y_{T-P} \end{pmatrix} + A_0 G A_0^{-1} (\mu + E_T) \quad (7) \end{aligned}$$

with

$$G = \begin{pmatrix} e_1' \\ \alpha' \end{pmatrix}, \text{ and } e_1' \text{ a } 4 \times 1 \text{ row selection vector,}$$

and where $F_1, \ldots, F_P$ are 4×4 matrices describing short run dynamics. Note that $Y_{1,T} - \alpha_1 Y_{4,T-1} = e_1' \Delta Y_T + l'\alpha' Y_{T-1}$, with $l = (-1/\alpha_2,\ -1/\alpha_2\alpha_3,\ -\alpha_1)$, so that the derivation is trivial once one has rewritten the VECM in (4) to a model for $\begin{pmatrix} e_1'\Delta Y_T & \alpha' Y_T \end{pmatrix}$. The roots of $\det(A_0 G^{-1} - \sum_{i=1}^{P} F_i z^i) = 0$ for (7) are assumed to be outside the unit circle, to ensure that the vectors $y_t - \alpha_s y_{t-1}$ are stationary processes. As an example, for model (6) one has:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\alpha_2 & 1 & 0 & 0 \\ 0 & -\alpha_3 & 1 & 0 \\ 0 & 0 & -\alpha_4 & 1 \end{pmatrix},\ A_0 G^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\beta_2 & 1 & 0 & 0 \\ 0 & -\beta_3 & 1 & 0 \\ 0 & 0 & -\beta_4 & 1 \end{pmatrix},\ F_1 = \begin{pmatrix} 0 & 0 & 0 & \beta_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

In general both A_0 and G from (7) are lower triangular with unit diagonal elements. $A_0 G^{-1}$ and F_k have the same structure as A_0 and A_k in (3). Whereas $A_{0ij} = 0$ for $i < j - p$, $(A_0 G^{-1})_{ij} = 0$ for $i < j - p + 1$: the highest lag for y_t is p, the highest lag for the periodically varying difference $y_t - \alpha_s y_{t-1}$ is $p - 1$. The *periodic adjustment coefficients* in the off-diagonal elements of $A_0 G^{-1}$ and F_k measure the impact of deviations from previous quarters' equilibria on the changes in the current quarter. Formulated otherwise: the parameters in $A_0 G^{-1}$ and F can be used to model short run variation in the levels of Y_T and its seasonal pattern. Note that model (6) does not allow for clear cyclical movements at frequencies lower than one per year, since the second root of the model is real. Just as in nonperiodic AR models one needs at least one extra lag in the model to get a pair of complex roots next to the unit root.

Multiplicative seasonality without taking logs

An interesting feature of the equivalent models (4) and (7), which is important for our present purpose of describing German and US unemployment, is the possibility of a strong interrelationship between the seasonal variation and the trend(-cycle) variation. In structural time series models, the conditional long-run forecast of the seasonal pattern, which equals the conditional expectation, i.e.

$$\lim_{N\to\infty} E(Y_{T+N} - \overline{Y}_{T+N} \mid Y_T,\ Y_{T-1},\ldots)\,,$$

with $\overline{Y}_{T+N} = (1/4)\iota\iota' Y_{T+N}$, and ι a $4{\times}1$ vector of ones, is sometimes used as an estimate of the seasonal component $Y_{Ss,T}$ at time T. In general this limit does not exist for models with periodic integration. Therefore, a corresponding definition of the seasonal component does not apply.

However, in the case of three cointegrating vectors one can compute *long-run seasonal factors*, denoted by $1+\lambda_s$, which give the average position of the series in a specific quarter relative to the "nonseasonal" component of the series, where the nonseasonal component is estimated using conventional seasonal adjustment filters to the *untransformed* series[3]. The long run predictions of the periodic model with one unit root will satisfy the (cointegrating) restrictions reparameterized using $1+\lambda_s$, since

$$\lim_{N\to\infty} E(Y_{s,T+N} - (1+\lambda_s)Y_{Ns,T+N} \mid Y_T,\ Y_{T-1},\ldots) = c_s,\ s=1,2,3,4 \tag{8}$$

where $Y_{Ns,T+N}$ is an estimate of the nonseasonal component in quarter s of year $T+N$, and c_s is a constant depending on the matrices A_l and the vector μ in (3), see *e.g.* Johansen (1994). The expression in (8) implies that the long run prediction of the seasonal pattern is proportional to the long run prediction of the nonseasonal component.

In order to investigate the properties of a periodically integrated time series with respect to (8) we first define $Y_{N1T} = (1/4)\iota' Y_T$ as the nonseasonal component. For simplicity and illustrative purposes we take the nonseasonal component equal across seasons in one year: $Y_{N1T} = Y_{N2T} = Y_{N3T} = Y_{N4T}$. We define $Y_{Ss,T} = Y_{s,T} - Y_{N1T}$ as the seasonal component $s=1,2,3,4$. In general, Y_{N1T} and $Y_{Ss,T}$ are not stationary in the periodic model with one unit root. However, the linear combinations $Y_{Ss,T} - \lambda_s Y_{N1T}$ are stationary for particular choices of λ_s. The *long-run seasonal factors* are then equal to $1+\lambda_s$. For the parameterization of α in (5) we find the following solutions for the parameters λ_s :

$$\lambda_s = \frac{3-\alpha_{s+1}(1+\alpha_{s+2}(1+\alpha_{s+3}))}{1+\alpha_{s+1}(1+\alpha_{s+2}(1+\alpha_{s+3}))},\ s=1,2,3,4. \tag{9}$$

where the expression again obeys modulo-4 arithmetic: $\alpha_s = \alpha_{s+4k}$, $k=\cdots,-2,-1,0,1,2,\cdots$, see appendix A for details. The reverse relations read:

$$\alpha_s = \frac{1+\lambda_s}{1+\lambda_{s-1}},\ s=1,2,3,4. \tag{10}$$

[3]In the widely used Census X-11 method and related procedures, the nonseasonal component is based on centered moving averages, *both* in the *additive and* in the *multiplicative* version of the method, see Shiskin et al. (1967). Neither version uses a centered moving average of the log transformed data. The application of (data-dependent, model based) moving averages to log transformed data, is followed in structural time series modeling of time series with multiplicative seasonality. In the model based methods one carefully distinguishes the underlying seasonal component and its estimate. In the Census procedure (a model for) the "true" underlying seasonal component is not introduced.

which show that the "*periodic differencing*" *parameters* α_s are easily interpreted as ratios of subsequent long-run seasonal factors. It should be understood that the observed seasonal factors, i.e. the ratios $Y_{s,T}/\widehat{Y_{Ns,T}}$, can be far from their long-run equilibrium values $1+\lambda_s$ if the return to equilibrium is slow. The speed of return to equilibrium depends on the other parameters in the model.

The foregoing analysis suggests that λ_s can be estimated using a *cointegrating regression* of $\widehat{Y_{Ss,T}}$ on $\widehat{Y_{Ns,T}}$. So far, we have not defined the appropriate class of estimators for the nonseasonal component $Y_{Ns,T}$. In appendix A we derive that this regression is valid for all *nonperiodic seasonal adjustment filters* $\phi(L)$, $\phi(L)y_t = \widehat{y_{N_t}}$, with $\widetilde{\phi}(z) = 0$ for $z = 1, -1$, $\pm i$, $i^2 = -1$, where $\widetilde{\phi}(z) = \phi(z) - \left(1 + z^2 + z^3 + z^4\right)/4$. This holds e.g. for the linear part of the Census X-11 filter as defined in Laroque (1977), which we use in simulations below. Application of seasonal adjustment filters of this type to series generated by periodic AR models with three different cointegrating vectors is expected to lead to a clear linear long-run relationship between the estimated seasonal and nonseasonal component in each quarter. The regressions of $\widehat{Y_{Ss,T}}$ on $\widehat{Y_{Ns,T}}$ are *valid cointegrating regressions* in these cases. The parameters λ_s can be estimated (super-) consistently, yet inefficiently, using OLS for the levels of the seasonal and the nonseasonal component, see Stock and Watson (1993) for more efficient regression estimators. Finally, note that all λ_s are 0 and that all seasonal factors equal 1 in the case of "ordinary differences", i.e. when $\alpha_s = 1$, $s = 1, 2, 3, 4$. In our empirical application of VAR models like (3) in section 4 we use the Gaussian ML estimators from a VAR representation to compute the parameters α_s and λ_s.

Periodic deterministic growth

Until now we did not explicitly focus on the role of the constant term μ in (3) in cases with unit roots. Under the restriction of at least one unit root one can model nonzero yearly mean growth in Y_T and a nonzero mean for the cointegrating vectors by this vector μ in (3). The vector μ leads to yearly mean growth if $A_0^{-1}\mu$ is not collinear with the columns of γ in (4), see Johansen (1994). Consider again the case with one unit root, and suppose that $Y_{Ns,T} = \iota' Y_T/4$ for simplicity. In this case, the process satisfies the mean stationarity restrictions

$$E(Y_{s,T} - (1+\lambda_s)\iota' Y_T) = c_s, \; s = 1, 2, 3, 4, \tag{11}$$

since each $e_s - (1+\lambda_s)\iota'$ is a cointegrating vector of the system, with e_s the s-th unit vector of dimension 4, and c_s is the constant mean of the cointegrating combination. The means, c_s, can easily be calculated from (7). If $\iota' Y_T$ contains a deterministic trend, it is clear from (11) that a *deterministic trend in the seasonal pattern* follows if there is one $\lambda_s \neq 0$. [4]

Multiplicative deterministic seasonality reflects itself in the property that mean yearly growth rates differ from season to season. In empirical examples this phenomenon is best observed as an ever lasting growing apart of the first differences of the different seasons. From Figure 3 we notice that this property may be relevant for the German data. Seasonally varying yearly growth shows clearly in the time series plots of the seasonal component in Figure 5 as well. The figure also shows how deterministic parts of the periodic trends can help to identify the long-run seasonal factors.

Innovation outliers

As already mentioned when considering the estimation of the auxiliary regression in (1), both the German and USA unemployment data show a number of outlying observations in the unadjusted data which have to be taken into account in order to make the assumption of

[4] Bowerman, Koehler and Pack (1990) discussed the modeling of trends in the seasonal pattern using nonperiodic models.

Gaussianity workable. In the ML procedure below, we model these surprising observations as innovation outliers by using extra dummies in the regression equations of the VAR model (12).

Suppose we add a dummy explanatory variable o_t which equals one at observation t and zero elsewhere, to the right hand side of (2). This extra variable appears as $A_0^{-1}O_{s,T}$ in (4), with $O_{s,T}$ the sth unit vector in year T and a zero vector elsewhere. Given this transformation of $O_{s,T}$, and since only the lower triangular structure of A_0^{-1} is known, we model the innovation outliers by introducing the same dummy variable in the each equation for each $Y_{s,T}$, i.e. we include in model (4) the variables $\widetilde{O}_{1,T}$ and $\widetilde{O}_{2,T}$, which take a value of one in year T and zero in other years; cf. Table 1. An additional advantage of this model is that conditional ML estimation of the VAR in (3) and the VECM in (4) can be done using OLS.

Estimation

In the interesting case of one unit root and three cointegrating vectors, there are several ways to estimate α. The first way is to estimate (2) under the (nonlinear) restriction $D(1) = 0$, see Boswijk and Franses (1995). If σ_s^2 does not vary from quarter to quarter, this can be done using a standard nonlinear least squares algorithm. An alternative less efficient, but (seasonal heteroskedasticity) robust way is to estimate the reduced form of VAR model (3),

$$Y_T = C_1 Y_{T-1} + \cdots + C_P Y_{T-P} + A_0^{-1} E_T, \tag{12}$$

with $C_k = A_0^{-1}A_k,\ k = 1, 2, \ldots, P$, prior to imposing a single unit root restriction using a VECM parameterization (4) as in Johansen (1991). Note that the elements of the parameter matrix for the highest lag, C_P, satisfy $C_{Pij} = 0$ for $j \leq 4P - p$: The possibility of a low order univariate parameterization still reflects itself in leading zero columns in the parameter matrix belonging to the highest lag order, cf. the matrix A_1, corresponding to (6) above. If one incorporates these last restrictions one estimates $4^2(P) - 4(4P - p) = 4p$ dynamic parameters in C_1 to C_P and 10 parameters in the variance-covariance matrix of the disturbances, whereas one would estimate $4p + 1$ parameters in (2) under seasonal homoskedasticity and $4p + 4$ parameters without that restriction. The only information one does not use in the VAR is the weak exogeneity of $Y_{s,T}$ in the equations for $Y_{s'T}$ for all $s' < s$, which corresponds to 6 extra restrictions. The efficiency loss should not be so large for large p and a reasonable number of observations. On the other hand, it is clearly more important if one wants to use the vector approach for monthly data.

Franses (1994) used the vector approach in (12) in a testing strategy for a range of models for quarterly seasonality. This can be done in four ways using the following permutations of Y_T' :

$$\begin{aligned}
(Y_T^1)' &= Y_T' = (Y_{1,T}, Y_{2,T}, Y_{3,T}, Y_{4,T}), \\
(Y_T^2)' &= (Y_{2,T}, Y_{3,T}, Y_{4,T}, Y_{1,T-1}), \\
(Y_T^3)' &= (Y_{3,T}, Y_{4,T}, Y_{1,T-1}, Y_{2,T-1}), \text{ and} \\
(Y_T^4)' &= (Y_{4,T}, Y_{1,T-1}, Y_{2,T-1}, Y_{3,T-1}).
\end{aligned}$$

Each model associated with a permutation $Y_T^s,\ s = 1, 2, 3, 4$ corresponds to a *one-step ahead forecasting model for* $Y_{s,T}$ and to two to four step ahead forecasting models for the other quarters. In our case where the changes in the first quarter are of prime importance as we discussed in the previous section, we prefer to use Y_T^1. For the estimation of the "long-run" equilibrium relations the choice of Y_T^s should not matter too much if the model applies adequately.

To summarize our discussion on the model representation that we consider in the next section for the relevant unemployment data, we mention that we use

$$Y_T = C_1 Y_{T-1} + \cdots + C_P Y_{T-P} + \widetilde{\mu} + D_1 \tilde{O}_{1,T} + D_2 \tilde{O}_{2,T} + W_T, \tag{13}$$

with $\widetilde{\mu}$ a 4×1 vector of constants and with D_1 and D_2 as (4×1) coefficient vectors of the dummy variables $\tilde{O}_{1,T}$ and $\tilde{O}_{2,T}$ which equal one in particular years T_1 and T_2, respectively. After checking its validity, we impose one unit root restriction on (13) and use conditional maximum likelihood as in Johansen (1991) to estimate the parameters. Note that $\widetilde{\mu} = A_0^{-1}\mu$, $W_T = A_0^{-1} E_T$ and $I - \sum_{i=1}^{P} C_i = -\gamma\alpha'$, where A_0 is defined under (3) and E_T has a diagonal variance/covariance matrix and zero mean.. The cointegrating vector matrix α is defined and subsequently normalized under (4). We estimate its free elements and compute seasonal factors from (9).

Testing

Several *tests* can be applied to the model. First one can test for the *number of unit roots.* Given that the variables for the separate quarters are I(1) the finding of *four* unit roots, i.e. no cointegrating vectors, can be interpreted as an indication of (nonperiodic) *seasonal integration* (at frequencies 1/(2 quarters) and (1/(4 quarters)), cf. (4). The model with *one* unit root has an intuitive interpretation described above. If the parameters α_s in that parameterization all *equal one* (and the data are *not* in logs), there is *no* sign of multiplicative seasonality. The differenced time series, $y_t - y_{t-1}$, is then mean stationary, so that "*ordinary differencing*" is sufficient. Given a nonzero vector μ one can then still observe *deterministic seasonality,* which is of course uncorrelated with the trend-cycle variation in the series. The restrictions of deterministic seasonality and the restriction of (nonperiodic) seasonal integration can be tested in (2) and (3). For this purpose we use the Johansen (1991, 1994) likelihood ratio tests for unit roots and restrictions on the cointegrating space, which are based on a Gaussian model for the disturbances in (3).

4 Estimation results

In this section we estimate models as (13) for the quarterly unadjusted German and USA unemployment data. The goal of this empirical analysis is to show that correlations captured in the VAR estimates of the model can replicate the most striking aspects of the data discussed in section 2. We estimate models for levels and log-levels. We discuss the results for the levels first. A VAR model of order 1 proves sufficient for the German data, whereas a we need a second order VAR for the US data. Note that this corresponds to AR orders 4 and 8 for the periodic model (2) for the first quarter. The US model can be reduced to order 6 using conventional significance levels in likelihood ratio tests (as suggested under (12)), but this is not pursued in order to stay within the standard unrestricted VAR framework for testing unit roots.

Insert Tables 5 *and* 6.

Levels

We present the main results in Tables 5 and 6. For the German and US data in levels the hypothesis of more than one unit root in the system can be rejected at the 5% level. The incorporation of three cointegrating vectors seems thus supported, and hence we can consider the parameterization in the previous section. The second largest root in the German VECM model is .453+.387i ($i^2 = -1$), the second largest one for the US data is .847. We thus do not find clear indications of extra unit roots. The test for $\alpha_s = 1\ \forall s$, which corresponds to a model with additive deterministic seasonality only, rejects at extremely low significance levels. For the German data this is mainly due to the values for α_1 and α_2. In the US data these long-run coefficients are not as clearly identified. Yearly growth in the USA unemployment is smaller than one regression standard error for the first quarter, whereas the mean growth in the German data is about twice the regression standard error, compare $\widehat{\sigma}_1$ and $\widehat{m}_s$ in Tables 5 and 6.

From the values $\widehat{m}_s$ it can be observed that for German unemployment the first quarter values clearly "outgrow" the values of the other quarters by about 6,500 people per year, i.e. about 175,000 people in 27 years of the sample, see also Figure 5. In the USA, unemployment in the first quarter and the third quarter seems to grow (marginally) faster than in the other quarters, but as Figure 6 shows this effect is not obvious for the entire sample.

Both the German and the US data show evidence of at least two surprising innovation outliers. The largest German one (about 200,000 people) occurs in 1967.1. In the US data the largest one appears in 1975.1 (1500,000 people). These surprising innovations are considerably smaller in the seasonally adjusted data, as pointed out in section 2.

Log levels

In Tables 5 and 6, we also report the results of estimating and evaluating VAR models as (13) for log-transformed data. Let us turn to the results for the log-levels in Tables 5 and 6. As far as the analysis of the cointegrating vectors is concerned, the hypothesis of multiplicative seasonality in the German data seems sustainable: i.e. the first differences of the *log*-levels seem stationary. Although the estimated α_s still differ widely from 1, the p-value of the test for the restriction increases from $3.4{\cdot}10^{-6}$ to 0.044. Moreover, note the small growth rate for the first quarter, which is only 81% of the overall growth rate. Other diagnostics lead us to reject the specification in log-levels, however. The hypothesis of an additive homoskedastic error term is not warranted. The residuals for the model in log-levels are far from normal and downward trending heteroskedasticity (not reported in Table 5) is pronounced. This trending variance makes inference based on this model unreliable. It appears that the observations in the first part of the data set are very influential. Forecasts from the model for the log-levels are not sensible due to the vastly different trends. Note however that the mean growth rates are not significantly different from zero in either the level or the log-level model for the German data.

For the US data in Table 6 the model in log-levels makes more sense. The timing of the innovation outliers was adjusted for the logged US data, where the influence of the 1967 observations is very pronounced in the logged data, see also Figure 12. The residuals for the logs are closer to normality. The estimated growth rates are never significantly different from zero at reasonable significance levels. The implicit seasonal factors are not too far from one.[5] The indicated number of cointegrating vectors is smaller for the logged data. Multiplicative seasonality can still be an option. The possibility of seasonal unit roots (more than one unit root in the VAR) indicates that trends in seasonal movements probably are not "caused" by the trend in the overall level of the series only.

Altogether, these empirical results seem to favor a model in levels for the German data and a model in log-levels for the US data. In the subsequent analysis we use the VAR(1) for the levels allowing for nonzero mean growth as a model for the German data. For the US data we use a VAR(2) for the log-levels restricting mean growth to zero. In both models we impose a single unit root.

Cointegrating regressions for the seasonal component in German unemployment

[5]Sensitivity analysis with respect to the the lag length of the VAR and timing of the innovation outliers showed these growth rates to be poorly identified, with estimates ranging from -.12 to .02. Estimates of the seasonal factors are more stable.

The empirical results for the German data suggest a valid interpretation of the cointegrating vectors in terms of long-run seasonal factors. In Table 7 we present the results of direct regression methods to estimate these seasonal factors using different estimates of seasonal components. We compare the results for the official adjustment and adjustment using a standard additive STM, discussed in more detail in appendix B. These cointegrating regressions provide a simple way to check the interpretation of long-run multiplicative seasonality. One has to be careful with the statistical interpretation, as always when dealing with nonstationary regressors and residual serial correlation.

Insert Table 7

For the levels we get comparable results across the methods. The results for the logged data are however less easy to interpret, because of the downward trending heteroskedasticity and the large influence of the early outlying observations. Remember that the analysis in the previous section is based on moving average estimation of the nonseasonal component on the levels. The effect of a linear MA filter applied to the levels does not lead to a similar effect on the log levels, unless relative changes are very small. The regression approach need not work for both the levels and the log-levels, if one uses only one (given) estimate of the seasonal component.

The results in Table 7 also show the importance of the adjustment for *level-shifts* on the estimates of the long-run seasonal factors. Such level-shifts decrease, at least for the log-levels, the (unreasonably large) absolute values of the long-run seasonal factors λ_1 and λ_3 to a large extent. This holds for all three methods presented here. The adjustment for level-shifts breaks the apparent decreasing trend in the seasonal component into three parts. Table 7 also shows that the estimated VARs for the log-levels pick up the relative downward trend for the first quarter, which corresponds to a long-run seasonal factor $(1 + \lambda_1)$ smaller than one, although the first quarter values are the highest within sample, see Figure 1. These simple regressions help to clarify why the periodic VAR models for the log levels produce incredible long run forecasts in this case, cf. (8) above.

We do not claim our VAR models to be excellent representations of data generating processes. We only show that cointegrating relationships between the levels of the seasonal and nonseasonal components can be modeled by the VAR. We have not yet shown how the VAR estimates translate to the correlations between their yearly *changes*. We do this next by simulating long samples from the estimated VAR models.

Simulation

We use 264 year samples generated by estimated VAR models using starting values of 1986 and 1987. In order to mimic the major effects of the "official" seasonal adjustment we use the linear part of the Census X-11 filter for quarterly data as presented in Laroque (1977). This is a 57-term symmetric moving average filter with 6 seasonal unit roots, see Ooms (1994). We do not use the end-of-sample corrections of the Census procedure and just generate the extra post-sample observations in order to apply the same filter everywhere. We retain 250 years, i.e. 1000 quarterly observations.

Insert Table 8

If we compare the results on the simulated data in Table 8 with those in Table 1 we see that the VAR for the German data is able to replicate significant periodically varying correlations between the changes in the seasonal and the nonseasonal component. The correlations in the simulated data have the correct sign. They are substantially smaller, however. The variances of the changes in the estimated seasonal component are also much smaller. This suggests that the linear part of the Census-filter produces a smooth seasonal component, compared with the official adjustment.

Comparing the USA results in Table 8 with those in Table 2 one observes the replication of the signs of the correlations of interest for the US data as well. Again the correlations in the simulated data set are much smaller. The differences across the variances of the changes in the seasonal component for the different quarters are also smaller in the simulated data set. If one uses multiplicative adjustment by application of the filter to the logs and subsequently recomputing the levels, the correlations (not reported in Table 8) are larger and become similar to the clear correlations obtained in section 2.

Again, we perform a sensitivity analysis as in section 2, although we do not provide detailed results to save space. We re-do the regressions with different MA-orders, applying STMs to estimate seasonal components for the levels and the log-levels, as in Tables 3 and 4. Increasing the MA order does not make a big difference in most cases. The STM approach applied to the simulated level data for Germany results in larger coefficients with the same sign as the linear Census procedure. In the other cases results are mixed, but at least one of the four correlations turns out to be statistically significant from zero with signs corresponding to those in Tables 3 and 4. The most outstanding regression coefficients for the STM adjusted data appear for the German levels in quarter 1, (0.059), for the log-levels in quarter 2, (0.06), and for the USA log-levels in quarter 4, (0.019), which are all significant at very low significance levels. Note that the last coefficient has the opposite sign of its empirical counterpart.

5 Forecasting changes in the seasonal component

Until now, we focused on contemporaneous within-sample correlations. It is interesting to know how well one can forecast the changes in the seasonal component, in order to assess the dangers of using seasonally adjusted data in (tests of) models for rational expectations, see Sims (1993). Note that adjusted data are the results of "historical filtering" or "smoothing" so that adjustment for time T uses previous as well as later observations, see Maravall (1985) or Harvey and Koopman (1992). If the possible predictability applies to "on-line" seasonal adjustment remains an open question.

The correlations in section 2 are well identified, especially for the first quarters of the German data. The periodic models in section 4 slow large degrees of predictability for yearly changes in unadjusted employment. We combine both models to forecast changes in the seasonal component. We examine whether there are useful correlations between changes in the seasonal component and forecasts for the yearly changes in the unadjusted series. We re-estimate the equations of Table 1 where we use the fitted values for $Y_{s,T} - Y_{s,T-1}$ instead of the seasonally adjusted data $Y_{Ns,T} - Y_{Ns,T-1}$. We use the same dummies as in the construction of the forecasting model. The results are presented in Table 9. They cover the same set of seasonally adjusted data as used for Tables 3 and 4

Insert Table 9

Due to the good fit of the periodic models for the unadjusted data and the relatively small difference between yearly changes in the unadjusted and the adjusted data the results are similar to those reported in Tables 1, 3 and 4. The estimates presented for the seasonal components identified by the STM are also close to those in the regressions using $Y_{Ns,T} - Y_{Ns,T-1}$. The coefficients for the German data have the same sign as in Table 1 for the OECD adjusted data, but they are considerably smaller in absolute value. For the log-levels of the USA, the pattern for the STM-adjusted data differs from the pattern for the OECD data. Where the OECD data show a significant negative correlation in the third quarter, the STM adjusted data show it in the fourth. The OECD adjustment leads to clear periodicity in the forecast function, while the STM adjustment does not.

6 Conclusion

In this paper we suggest a set of tools to analyze the relation between the estimates of the seasonal and nonseasonal components in quarterly time series. We apply these tools in detail to quarterly unemployment series for Germany and the USA. One tool is the *periodic* regression of yearly changes in the seasonal component on yearly changes in the nonseasonal component, which is appropriate if one considers the nonseasonal component to be I(1), integrated of order one. Another tool is the periodic regression of the seasonal component on the nonseasonal component, which applies if one considers these components to be cointegrated. One can also regress yearly changes in the seasonal component on yearly changes in quarterly changes of the nonseasonal component if one considers the nonseasonal component to be I(2), integrated of order two, an assumption often made in Structural Time series Models. Finally, one can regress yearly changes in the seasonal component on forecasts for the overall yearly changes to get a first idea of (periodic) predictability of the seasonal component. If the nature of the adjustment is not known, application to both levels and log-levels may convey information on which method has been used.

Our analysis of the unemployment data reveals that periodic time series models yield useful descriptions. One of our main results is that yearly changes in the seasonal component of the number of unemployed seem to be linearly predictable by past changes in the nonseasonal component if one allows for a periodically changing forecast function. This may have the implication that adjusted unemployment figures in the first quarter are "too optimistic" in times of rising unemployment. A second result is that one can replicate this property to some extent in long data series generated by periodic AR models, estimated for the unadjusted data. The unadjusted data show evidence of so-called *periodic integration*. The periodic predictability is found for different decompositions into seasonal and nonseasonal components.

Linear predictability in either the level changes or the log-level changes may be due to the wrong decision with respect to additive versus multiplicative adjustment. We indicate how linear periodic AR models with unit roots allow for multiplicative seasonality without a log transformation of the data. A periodic cointegration analysis of the seasonal and nonseasonal component applied to the levels of the series may indicate long-run multiplicative seasonality.

Likelihood ratio tests applied to periodic AR models indicate *periodic cointegration* between the seasonal and nonseasonal component of German unemployment. This interpretation is confirmed by periodic level regressions of the seasonal on the nonseasonal component for different seasonal adjustment methods. Therefore a linear time series model for the log-levels of German unemployment is seriously misspecified. Periodicity in the prediction function for changes in the seasonal component of the log-levels of US unemployment is not so strong. Furthermore, it is different in nature across different seasonal adjustment methods. One finds strong periodicity in the correlations between seasonal and nonseasonal components in the levels only if one applies a linear seasonal adjustment filter to the log-levels.

The combined analysis of the periodic regressions and periodic AR models for the unadjusted data provides useful insight in the adequacy of different models for seasonal time series and corresponding seasonal adjustment methods.

We have several suggestions for further research. First, the sampling properties of the regression tests considered in this paper can be investigated using Monte Carlo methods. This requires an explicit modeling strategy for the seasonal component, and is therefore not included in this paper. Second, in the parametric context of structural time series models, it may be possible to do more efficient inference on alternatives allowing for periodicity of certain coefficients. One can then use the time series properties of the estimated models while taking the implied smoothing process into account. In that way one can obtain information how these models can be extended to allow for the periodic correlations observed in for example the German unemployment data. Such an extension implicitly allows for periodic MA coefficients in a "reduced form" representation of the process, which could complement the analysis based on periodic AR models. Estimation of unrestricted periodic ARMA models is not an attractive option, since it has all the drawbacks of unrestricted Vector ARMA modeling, in particular time-consuming procedures of identification and estimation.

Appendices

A Long-Run Seasonal Factors

In this appendix we first derive the formula (9) for long-run seasonal factors for the model with periodic integration and periodic differencing parameters given in (5). Next we derive conditions for nonperiodic linear seasonal adjustment filters that are sufficient to identify long-run seasonal factors in cointegrating regression.

Derivation of long-run seasonal factors from cointegrating vectors

One has to choose λ_s such that the re-parameterized cointegrating vectors, to be denoted as $e_s - (1+\lambda_s)\iota/4$, lie in the space spanned by the columns of α in (5), so that $Y_{s,T} - (1+\lambda_s)Y_{N1T}$ is stationary, where e_s is the s-th unit vector of dimension 4, and where the nonseasonal component, Y_{N1T}, equals $\iota' Y_T/4$ with ι a 4×1 vector of ones. One derives $\lambda = (\lambda_1\ \lambda_2\ \lambda_3\ \lambda_4)'$ explicitly from:

$$e_s - \iota(1+\lambda_s)/4 = \alpha v_s$$

with ι a vector of ones, α the 4×3 matrix with cointegrating vectors and v_s a 3×1 vector of unknown coefficients. Rearrangement gives

$$e_s - \iota/4 = (\alpha : \iota/4) \begin{pmatrix} v_s \\ \lambda_s \end{pmatrix}.$$

Combining the equations for $s = 1, 2, 3, 4$ gives

$$I_4 - \iota\iota'/4 = (\alpha : \iota/4) \begin{pmatrix} \nu \\ \lambda' \end{pmatrix},$$

where I_4 is an identity matrix of dimension 4, and ν is a 3×4 matrix of unknown coefficients. Premultiplication of both sides with $(\alpha : \iota/4)^{-1} e_4'$ and rearrangement gives

$$\lambda' = (\alpha : \iota/4)^{-1} (e_4' - \iota'/4). \tag{14}$$

Substituting the parameterization of α in (5) one gets (9). One proves equation (10) by substitution of equation (9) and the equality $\alpha_s = 1/(\alpha_{s+1}\alpha_{s+2}\alpha_{s+3})$.

Equivalent identification of seasonal factors using other adjustment filters

Suppose we use a more sophisticated nonseasonal component estimator, like the first stage, so-called 4 term centered moving average, estimator of the trend-cycle in the Census-X11 procedure, see Shiskin et al. (1967, p. 14, 21).

$$Y^*_{N1T} = (Y'_{T-1}\ Y'_T)((0\ 0\ (1/8)\ (1/4)\ (1/4)\ (1/4)\ (1/8)\ 0)')$$

and define Y^*_{S1T} as $Y_{1,T} - Y^*_{N1T}$. Note that Y^*_{N1T} is cointegrated with Y_{N1T} (defined above) with cointegrating vector $\begin{pmatrix} 1 & -1 \end{pmatrix}$ since $Y^*_{N1T} - Y_{N1T} = (1/8)\Delta_1 Y_{3,T} + (1/4)\Delta_1 Y_{4,T}$, which is stationary. Consequently Y_{S1T} is cointegrated with Y^*_{S1T} with cointegrating vector $\begin{pmatrix} 1 & -1 \end{pmatrix}$ as well. Therefore Y^*_{N1T} and Y^*_{S1T} are cointegrated with the same cointegrating vector as Y_{N1T} and Y_{S1T}.

In practice one often uses long nonperiodic moving average filters to estimate the nonseasonal component. Let $\phi(L)$ denote a nonperiodic seasonal adjustment filter for the quarterly series y_t, so that $\widehat{y_{Nt}} = \phi(L)y_t$. The result in (14) carries over to these cases as well because of the following arguments.

The series $\phi(L)y_t$ is cointegrated with $(1/4)\left(1+L+L^2+L^3\right)y_t$ with cointegrating vector $(1\ \ -1)$ if $\widetilde{\phi}(L) = (\phi(L) - (1/4)(1+L+L^2+L^3))$ contains a factor $(1-L^4)$, i.e. if $\widetilde{\phi}(z) = 0$, for $z = 1,\ -1,$ and $\pm i$, since $(1-L^4)y_t$ is (periodically) stationary by definition in all quarters s, $t = 4\,(T-1)+s$, $s = 1,2,3,4$.

B Seasonal adjustment using "structural time series models"

In this appendix we discuss the STM method and show that it can result in correlations between the seasonal and nonseasonal components. We compute these correlations for an STM for the German unemployment data.

By way of comparison we analyse the seasonal components using time series models proposed by Harvey (1989), implemented in the computer package STAMP by Koopman et al. (1995). The (nonperiodic) structural time series model reads:

$$\begin{aligned} y_t &= \mu_t + \gamma_t + \epsilon_t \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t \\ \beta_t &= \beta_{t-1} + \zeta_t \end{aligned} \tag{15}$$

with y_t the unadjusted series, μ_t the level component, γ_t the seasonal component, ϵ_t the irregular component and β_t the slope component. For quarterly data there are three options for γ_t. The models read in increasing order of flexibility:

$$\gamma_t = \omega_t/(1+L+L^2+L^3) \quad \text{, or} \tag{16}$$

$$\gamma_t = (\omega_{1t} + \omega^*_{1,t-1})/(1+L^2) + \omega_{2t}/(1+L) \quad \text{, or} \tag{17}$$

$$\gamma_t = \gamma_{1t} + \gamma_{2t}, \tag{18}$$

$$\begin{aligned} \gamma_{1t} &= \frac{(1-\rho\cos\lambda L)\omega_{1t} + \rho\sin\lambda L\omega^*_{1t}}{1+2\rho\cos\lambda L + \rho^2L^2} \\ \gamma_{2t} &= \omega_{2t}/(1+\rho L), \end{aligned}$$

They are labeled "dummy variable", "trigonometric" and "two cycle" seasonality, respectively. The parameter ρ denotes the "damping factor" of the seasonal cycle. In most applications it is close to 1. The parameter λ denotes the angular frequency of the longest seasonal cycle. It is close to $\pi/2$ if applied to quarterly data.. The Gaussian zero mean white noise innovations ϵ_t, η_t, ζ_t and the seasonal innovations ω_{it} are mutually uncorrelated. We define the "*seasonally adjusted" series* as $\widehat{y_{Nt}} = y_t - \widehat{\gamma}_t$, and the *estimated seasonal component* $\widehat{y_{St}}$ as $\widehat{\gamma}_t$.

The parameters, viz. the variances of ϵ_t, η_t, ζ_t and ω_{it}, $i = 1,2$, σ^2_ϵ, σ^2_η, σ^2_ζ, $\sigma^2_{\omega_i}$, are estimated by maximum likelihood. Relatively very small variances are put to zero in the estimation process. The variances of ω_{1t} and ω^*_{1t} are set equal. In (17) the variances of ω_{1t} and ω_{2t} are set equal as well.

Estimation results

For our data series the "dummy variable" model (16) produced unsatisfactory seasonally adjusted data with high spectral power at the yearly frequency $\pi/2$. The other two models (17) and (18)produce reasonable, and very similar estimates of the seasonal component. For the German data the variance of the level component, σ^2_η drops to zero in the "two cycle" model. Estimates in (18) of λ are close to $\pi/2$ and estimates of ρ are close to 1. In most cases the variance of the irregular component σ_ϵ^2 is negligible and it is set to zero during the ML estimation. The irregular component only comes into play in the outlier analysis where it accommodates the outlying observations. If one estimates a structural model for the corresponding first differences, the former level component naturally becomes the irregular and the former slope component becomes the level component. Here we present only results for the estimation in levels. The estimation in differences leads to larger correlations between changes in the seasonal and the nonseasonal component, both in the nonperiodic and the periodic analysis. Furthermore, the standard error of the seasonal component doubles, while the standard errors of the other components stay approximately the same.

For the German data we apply model (15)-(17) to the levels. The most informative diagnostic (provided in STAMP) in favor of levels over log-levels is the CUSUM-of-Squares plot, which indicates far too steep a decline in residual variance for the log-level model[6]. For the US data we apply model (15)-(17) to the log-levels, which has the more stable residual variance compared to the level model. Since we consider quarterly observations, we introduce 8 interventions to the irregular component in 1967 and 1974 for the German data and in 1967 and 1975 for the USA data. They do not seem to be necessary in the latter case. We present estimation results in Table 10.

Insert_Table_10

The seasonal variance in the German data is small compared to the variance in the slope component. It is however substantial compared to the variance in the level component, which means that it is clearly important when analyzing first differences. In the log-level data for the USA, the seasonal component is not as influential as in the German data.

Implicit correlations of seasonal and nonseasonal component

We simulate 10,000 observations from these models in order to compute implicit correlations between different typical component estimates of these models. Comparing the theoretical correlations with empirical ones is useful in judging model adequacy of structural time series models, see, e.g,. Harvey and Koopman (1992). In practice these correlations depend on the number of observations and the distance of the observations from the end points of the sample. By taking sample correlations of a very large sample we abstract from the last two problems. The results mainly characterize the central observations of a large sample.

Insert_Table_ 11

One sees in Table 11 that a number of correlations between different lags of the estimated seasonal and nonseasonal component can take substantial values. If one applies the linear part of the Census X-11 filter to the same simulated series, the absolute values of the *cross correlations* at low lags are substantially smaller than the cross correlations for STM components. The Census X-11 cross-correlations are larger at higher lags. On the other hand, one observes that the *autocorrelation* function of the changes in the Census seasonal component tails off at higher lags. The presence of these correlations, especially at yearly lags, is the reason for allowing an MA-error structure in Table 1.

Given these correlations in Table 11 it should come as no surprise that lagged changes in the growth rate of the nonseasonal component can help to forecast yearly changes in the seasonal component in sample, even if one uses the true model to estimate these components in a large sample. It is a general property of estimated components in this type of models.

[6]The smoothing procedure complicates statistical interpretation of standard regression diagnostics. For a survey on diagnostic checking in the context of unobserved component models, see Koopman (1992) and Harvey and Koopman (1992).

However, the correlations should not vary with the seasons as they do in our empirical example concerning German and US unemployment.

Evaluation of the periodic regressions in the context of STMs

It is clear from (18) that $\Delta_4\widehat{y_{Nt}}$ is not stationary if $\sigma_\zeta^2 > 0$. The same holds for its estimate. In order to get an interpretable correlation between the seasonal and the nonseasonal component from this model we regress $\Delta_4\widehat{y}_{St}$ on $\Delta_1\Delta_4\widehat{y}_{Nt}$ instead of using the regression from $\Delta_4\widehat{y}_{St}$ on $\Delta_4\widehat{y}_{Nt}$. These regression coefficients are small but non negligible, see the last column of Table 10. We check the values of these correlations in our data using the periodic version of this regression. We regress the yearly changes in the seasonals on $(Y_{Ns,T} - Y_{Ns,T-1}) - (Y_{Ns-1,T} - Y_{Ns-1,T-1})$. For the log-levels of the US data we obtain nonperiodic results in correspondence with the coefficient in the last column of Table 10. For the German data in levels this regression leads to significant periodically varying correlations, still indicating misspecification. The latter results are not reported to save space, but can be obtained from the authors upon request.

References

[1] Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis.* John Wiley and Sons, New York, USA, second edition.

[2] Boswijk, H.P, and Franses, P.H. (1992). Testing for periodic integration. Technical Report 16, Econometric Institute.

[3] Bowerman, B.L., Koehler, A.B., and Pack, D.J. (1990). Forecasting Time Series with Increasing Seasonal Variation. *Journal of Forecasting*, 9:419–436.

[4] Engle, R.F., and Granger, C.W.J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55:251–276.

[5] Franses, P.H. (1994). A Multivariate Approach to Modeling Univariate Seasonal Time Series. *Journal of Econometrics*, 63:133–151.

[6] Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman Filter.* Cambridge University Press, Cambridge, UK.

[7] Harvey, A.C., and Koopman, S.J. (1992). Diagnostic Checking of Unobserved Components Time Series Models. *Journal of Business and Economic Statistics*, 10:377–389.

[8] Harvey, A.C., Koopman, S.J., Doornik, J.A., and Shephard, N.G. (1995). *STAMP 5.0, Structural Time Series Analyser, Modeller and Predictor.* Chapman and Hall, Andover, England.

[9] Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, 59:1551–1580.

[10] Johansen, S. (1994). The Role of the Constant and Linear Terms in Cointegration Analysis of Nonstationary Variables. *Econometric Reviews*, 13:205–229.

[11] Koopman, S.J. (1992). *Diagnostic checking and Intra-Daily effects in Time series models.* Thesis Publishers, Amsterdam, The Netherlands.

[12] Laroque, G. (1977). Analyse d'une Méthode de Désaisonnalisation: le Programme X 11, du US Bureau of Census, Version Trimestrielle. *Annales de l'INSEE*, 28:105–127.

[13] Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis.* Springer Verlag, Berlin, Germany.

[14] Maravall, A. (1985). On Structural Time Series Models and the Characterization of Components. *Journal of Business and Economic Statistics*, 3:350–355.

[15] Maravall, A., and Gómez, V. (1992). Signal Extraction in ARIMA Time Series Program SEATS. Technical Report ECO 92/65, Economics Department, European University Institute, Florence, I- 50016 San Domenico di Fiesole (FI), Italy.

[16] Mardia, K.V. (1970). Measures of Multivariate Skewness and Kurtosis with Applications. *Biometrika*, 57:519–530.

[17] Ooms, M. (1994). *Empirical Vector Autoregressive Modeling*, volume 407 of *Lecture Notes in Economics and Mathematical Systems.* Springer-Verlag, Berlin, Germany.

[18] Osborn, D.R. (1991). The Implications of Periodically Varying Coefficients for Seasonal Time-series Processes. *Journal of Econometrics*, 48:373–384.

[19] Osborn, D.R., and Smith, J.P. (1989). The Performance of Periodic Autoregressive Models in Forecasting Seasonal UK Consumption. *Journal of Business and Economic Statistics*, 7:117–127.

[20] Shiskin, J. and Eisenpress, H. (1957). Seasonal Adjustment by Electronic Computer Methods. *Journal of the American Statistical Association*, 52:415–449.

[21] Shiskin, J., Young, A.H., and Musgrave, J.C. (1967). The X-11 variant of the Census method II Seasonal Adjustment Program. Technical report, Technical Paper no. 15, Bureau of the Census, U.S. Department of Commerce, U.S. Government Printing Office., Washington D.C., USA.

[22] Sims, C. (1993). Rational Expectations Modeling with Seasonally Adjusted Data. *Journal of Econometrics*, 55:9–19.

[23] Stock, J.H., and Watson, M.W. (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica*, 61:783–820.

[24] Tiao, G.C., and Grupe, M.R. (1980). Hidden Periodic Autoregressive-Moving Average Models in Time Series Data. *Biometrika*, 67:365–373.

[25] Van Der Hoeven, H., and Hundepool, A.J. (1986). A Method for Seasonally Adjusting Time Series With Variation in the Seasonal Amplitude. *Journal of Business and Economic Statistics*, 4:455–471.

Table 1: *Testing the null of no correlation between changes in the seasonal and nonseasonal components of unemployment using the auxiliary regression in (1).*

Germany: 1961-1987				
Quarter s	$\hat{\beta}_s$	t-val	$\hat{\sigma}_{Es}$	$\hat{\sigma}_{\hat{Y}s}$
1	.113	9.87	11.35	32.0
2	-.040	-8.22	4.93	12.6
3	-.070	-6.42	10.51	22.2
4	-.018	-3.10	5.38	6.9
USA: 1961-1987				
Quarter s	$\hat{\beta}_s$	t-val	$\hat{\sigma}_{Es}$	$\hat{\sigma}_{\hat{Y}s}$
1	.0984	16.12	26.7	117
2	-.019	-1.89	45.9	51.7
3	-.039	-4.03	38.9	58.2
4	-.039	-5.41	32.3	50.8

NOTES: The yearly changes in the estimates of the seasonal components $\Delta\widehat{Y_{Ss,T}} = \widehat{Y_{Ss,T}} - \widehat{Y_{Ss,T-1}}$ are regressed on a constant and the contemporary yearly change in the nonseasonal component $\Delta\widehat{Y_{Ns,T}} = \widehat{Y_{Ns,T}} - \widehat{Y_{Ns,T-1}}$ allowing for MA error structure of order 2 and two dummy variables, $s = 1, 2, 3, 4$, $T = 1, \cdots, 27$. Regression parameters change from quarter to quarter. Dates of dummies $\tilde{O}_{1,T}$ and $\tilde{O}_{2,T}$ are 1967 and 1974 for Germany and 1975 and 1982 for the USA. $\hat{\sigma}_{\hat{Y}s}$ denotes standard deviation of dependent variable, $\hat{\sigma}_{Es}$ denotes standard error of the regression.

Table 2: *Testing the null of no correlation between changes in the seasonal and nonseasonal components of* **log** *unemployment using the auxiliary regression in (1).*

Germany: 1961-1987				
Quarter s	$\hat{\beta}_s$	t-val	$\hat{\sigma}_{Es}$	$\hat{\sigma}_{\hat{Y}s}$
1	-.149	-6.33	.036	.093
2	.080	5.17	.021	.037
3	.094	5.64	.022	.051
4	.012	1.00	.014	.017
USA: 1961-1987				
Quarter s	$\hat{\beta}_s$	t-val	$\hat{\sigma}_{Es}$	$\hat{\sigma}_{\hat{Y}s}$
1	.030	4.42	.0049	.0080
2	-.006	-.85	.0053	.0137
3	-.036	-6.29	.0042	.0123
4	.007	1.64	.0034	.0056

NOTE: See Table 1. Dates of dummies $\tilde{O}_{1,T}$ and $\tilde{O}_{2,T}$, 1967 and 1974 for Germany and 1975 and 1982 for the USA.

Table 3: *Sensitivity analysis of periodic correlations between yearly changes seasonal and non-seasonal components for German unemployment using the auxiliary regression in (1).*

number of unemployed											
			OECD			*STM*			*STM-I*		
Dummies	s	MA	0	2	3	0	2	3	0	2	3
no	1		+	+	+	+	+	+	+	+	+
no	2		−	−	−	−	−	−	−	−	−
no	3		−	−	−	−	o	o	−	o	−
no	4		o	−	−	o	o	o	o	−	o
yes	1		+	+	+	+	+	+	+	+	+
yes	2		−	−	−	−	−	o	−	o	o
yes	3		−	−	−	o	o	−	−	−	−
yes	4		o	−	−	−	−	−	o	−	o
log number of unemployed											
			OECD			*STM*			*STM-I*		
Dummies	s	MA	0	2	3	0	2	3	0	2	3
no	1		−	−	−	−	−	−	o	o	−
no	2		+	+	+	+	+	+	+	+	+
no	3		+	+	+	o	+	+	o	o	+
no	4		o	+	+	o	o	o	o	+	o
yes	1		−	−	−	−	−	−	o	o	−
yes	2		+	+	+	+	+	+	+	o	o
yes	3		+	+	+	o	o	o	o	o	o
yes	4		o	o	o	o	o	o	+	+	+

NOTES: Upper panel gives results for level data, lower panel for logged data. *Dummies*: indicates presence of dummies for 1967 and 1974 in periodic regressions as in table 1. s indicates quarter of year. *OECD* indicates seasonal adjustment as published by the *OECD*. *STM* indicates seasonal adjustment by a Structural Time Series model with trigonometric seasonality, see appendix B. *STM-I* indicates seasonal adjustment using an *STM* with intervention dummies in the irregular at 1967 and 1974.
The "+", "−", "o" in the main panels indicates positive, negative or no significant correlation at a 5% level using standard t-values. MA indicates the order of the MA process for the errors.

Table 4: *Sensitivity analysis of periodic correlations between yearly changes seasonal and non-seasonal components for USA unemployment using the auxiliary regression in (1).*

number of unemployed											
				OECD			*STM*			*STM-I*	
Dummies	s	MA	0	2	3	0	2	3	0	2	3
no	1		+	+	+	+	+	+	+	o	+
no	2		−	−	−	o	o	o	o	o	o
no	3		−	−	−	o	o	o	o	o	o
no	4		o	−	−	o	o	o	o	o	o
yes	1		+	+	+	o	o	+	o	o	o
yes	2		o	−	−	o	o	o	o	o	o
yes	3		−	−	−	o	o	o	o	o	o
yes	4		−	−	−	o	o	o	o	o	o
log number of unemployed											
				OECD			*STM*			*STM-I*	
Dummies	s	MA	0	2	3	0	2	3	0	2	3
no	1		+	+	+	o	o	+	o	o	o
no	2		o	o	o	o	o	o	o	o	o
no	3		−	−	−	−	o	o	o	o	o
no	4		o	+	+	+	+	o	+	+	+
yes	1		+	+	+	o	o	o	o	o	o
yes	2		−	o	−	o	o	o	o	o	o
yes	3		−	−	−	−	o	o	−	o	o
yes	4		+	o	o	+	+	+	+	+	+

NOTES: see Table 3. Dates for dummies level data USA: 1975 and 1982. Dates for log level data: 1967 and 1974.

Table 5: *Periodic AR models for German unemployment*

Test results using the VAR model in (13)

	levels				*log-levels*			
r	0	1	2	3	0	1	2	3
LR_r	166.9*	63.2*	26.5*	.001	150.92*	44.17*	7.9	0.20
$LR_{\alpha=1}$				$3.4 \cdot 10^{-6*}$				$4.38 \cdot 10^{-2*}$
$LR_{m=0}$		.37	.36	.19		.64	.48	.25

NOTE: LR_r: Likelihood ratio test statistic, -2(lnL(H_0)-lnL(H_1)), on H_0 of $4-r$ unit roots against no unit roots, where r indicates the number of cointegrating vectors. An asterisk indicates rejection of hypothesis of $4-r$ unit roots at the 5% significance level. $LR_{\alpha=1}$: p-value of LR test for $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$, $LR_{m=0}$: p-value of LR test for zero mean growth given r cointegrating vectors, see Johansen(1994), computed using small sample correction from Anderson (1984), see Ooms (1994, p. 195).

Estimation results and diagnostics when one unit root restriction is imposed

		levels					*log-levels*			
	s	1	2	3	4		1	2	3	4
$\hat{\alpha}_s$		1.095	.914	1.014	.985		.822	1.321	1.075	0.856
$1+\hat{\lambda}_s$		1.065	.974	0.988	.973		.807	1.066	1.146	0.981
$\hat{m}_s$		78.6	71.8	72.8	71.7		.054	.072	.077	.066
$\hat{\sigma}_1$		37.9					.0685			
$Norm_3$				.64					.002*	
$Norm_4$					.77					.026*
$\tilde{O}_{s,T}$	1967	1974				1967	1974			
D_1	222.2					.0214				

NOTE: The $\hat{\alpha}_s$ are the coefficients in the cointegrating relations $Y_{s,T} - \alpha_s Y_{s-1,T}$, $s = 2, 3, 4$, $\hat{\alpha}_1 = 1/\hat{\alpha}_2\hat{\alpha}_3\hat{\alpha}_4$, estimated under restriction of one unit root. The row starting with $1+\hat{\lambda}_s$ present the long-run seasonal factors, see (9) . $\hat{m}_s$: expectation of $Y_{s,T} - Y_{s,T-1}$ under the restriction of one unit root. $\hat{\sigma}_1$: standard error of regression equation for first quarter. $Norm_j$: asymptotic p-value of Mardia's (1970) multivariate normality test applied to residuals on jth moment, see Ooms (1994). $\tilde{O}_{s,T}$: years for which dummies for innovation outliers are included. D_1: size of innovation outlier in first quarter with the biggest impact. Order of VAR: $P = 1$. The vector of constants is included in the regressions.

Table 6: *Periodic AR models for USA unemployment*

Test results using the VAR model in (13)

	levels				log-levels			
r	0	1	2	3	0	1	2	3
LR_r	104.7*	44.0*	16.8*	3.1	57.77*	25.46	8.12	0.536
$LR_{\alpha=1}$				$2.4 \cdot 10^{-4*}$				$8.0 \cdot 10^{-3*}$
$LR_{m=0}$		.93	.85	.58		.73	.57	.45

NOTE: See Table 5.

Estimation results and diagnostics when one unit root is imposed

	levels				log-levels			
s	1	2	3	4	1	2	3	4
$\hat{\alpha}_s$	1.06	.92	1.07	.96	.931	.964	1.102	1.011
$1+\hat{\lambda}_s$	1.037	.955	1.026	.982	.976	.940	1.036	1.048
$\hat{m}_s$	173.9	160.0	172.0	164.6	-.0744	-.0717	-.0790	-.0799
$\hat{\sigma}_1$	203.9				.047			
$Norm_3$			.47				.89	
$Norm_4$				.13				.91
$\tilde{O}_{s,T}$	1975	1982			1967	1975		
D_1	1506.4					0.167		

NOTE: See Table 5. Order of VAR: $P = 2$. The vector of constants is included in the regressions.

Table 7: *Cointegrating regressions for German unemployment 1960-1987, Dependent variable is* $\widehat{Y_{Ss,T}}$, *Explanatory variable is* $\widehat{Y_{Ns,T}}$

Levels					
Dummies	I		II		
Quarter	λ_s OECD	λ_s VAR	λ_s OECD	λ_s VAR	λ_s STM
1	0.069	0.065	0.070	0.075	0.056
2	-0.029	-0.026	-0.017	-0.015	-0.019
3	-0.015	-0.012	-0.009	-0.013	-0.046
4	-0.019	-0.027	-0.018	-0.047	-0.017
Log-levels					
Dummies	I		II		
Quarter	λ_s OECD	λ_s VAR	λ_s OECD	λ_s VAR	λ_s STM
1	-0.182	-0.193	-0.072	-0.067	-0.032
2	0.041	0.066	0.022	0.027	0.050
3	0.168	0.146	0.041	0.055	0.024
4	0.018	-0.019	0.008	-0.016	-0.035

NOTE: $1+\lambda_s$: seasonal factor in quarter s, see (9). Dummy set I: temporary dummies in '67 and '74, as in table 5. Dummy set II: temporary and permanent change dummy variables, i.e. dummies equal to zero before and equal to one on and after the breakpoint. Breakpoints in '67 and '74. OECD: seasonal factors computed using levels regression for seasonal component published by OECD. VAR: seasonal factors computed from VAR(1) with one unit root, see also table 5 lower panel. STM: seasonal factors computed using levels regression for components using Structural Time Series Method, with additive trigonometric seasonality for levels and log-levels. Regression estimation allowing for MA(2) errors.

Table 8: *Simulation results on testing the null of no correlation between yearly changes in seasonal and nonseasonal components of simulated unemployment data*

Periodic model for Germany					Periodic model for the USA				
Quarter	$\hat{\beta}_s$	t-val	$\hat{\sigma}_{\epsilon s}$	$\hat{\sigma}_{\hat{Y}Ss}$	Quarter	$\hat{\beta}_s$	t-val	$\hat{\sigma}_{\epsilon s}$	$\hat{\sigma}_{\hat{Y}Ss}$
1	0.0097	5.96	4.43	7.68	1	.0009	1.39	0.0016	0.0039
2	-0.0087	-14.63	1.51	2.88	2	-0.0052	-6.58	0.0018	0.0046
3	-0.0034	-6.64	1.26	3.22	3	-0.0022	-2.69	0.0018	0.0044
4	-0.0127	-17.10	1.87	3.56	4	0.0055	6.05	0.0021	0.0036

NOTE: These results are based on the auxiliary regression (1), where the data are generated from VAR models like (3) with one unit root imposed and starting values taken equal to the 1987 and 1986 observations. See table 1. No dummies are included in test regression. We allow for MA(4) errors in (1). In the simulation model for Germany we allow for a deterministic trend in the levels, and for the USA model we do not allow for deterministic trends. The first 7 yearly observations are only used in seasonal adjustment procedure, not in the test regressions, which are based on 250 yearly observations.

Table 9: *Testing the null of no correlation between yearly changes in the seasonal components and one step ahead forecasts of yearly changes in the unadjusted series for the number of unemployed, based on the auxiliary regression (1)*

Germany: levels: dependent variable is: $\Delta\widehat{Y_{Ss,T}}$
Explanatory variables are one-step-ahead forecasts for $\Delta Y_{s,T}$ and dummies.
1961-1987, 27 observations

	OECD		STM		STM-I	
Quarter s	$\hat{\beta}_s$	t-val	$\hat{\beta}_s$	t-val	$\hat{\beta}_s$	t-val
1	.096	8.06	.029	7.19	.041	5.69
2	-.040	-5.18	-.008	-2.12	-.014	-1.95
3	-.072	-4.26	-.002	-0.51	-.009	-1.49
4	-.027	-3.91	-.019	-3.63	-.015	-2.16

USA: **log** levels: dependent variable is: $\Delta\widehat{Y_{Ss,T}}$
Explanatory variables are one-step-ahead forecasts for $\Delta Y_{s,T}$ and dummies.
1962-1987, 26 observations

	OECD		STM		STM-I	
Quarter s	$\hat{\beta}_s$	t-val	$\hat{\beta}_s$	t-val	$\hat{\beta}_s$	t-val
1	.032	5.88	.003	0.77	.004	0.75
2	-.008	-1.09	.002	0.56	-.003	-0.56
3	-.040	-5.49	-.011	-1.86	-.004	-0.45
4	.008	1.63	-.021	-4.49	-.019	-3.53

NOTES: OECD: Data adjusted by OECD. STM: Seasonal component estimated using structural time series model. STM-I: STM, while allowing for interventions in irregular component in two years, see also Table 3. Forecasting model for $\Delta Y_{s,T}$ for Germany is a VAR(1) with one unit root, unrestricted constant and dummies in 1967 and 1974, see table 5; for the USA it is a VAR(2) with one unit root, restricted constant and dummies in 1967 an 1975. We allow for MA(2) errors and two dummies in regressions (1), where the forecasts for $\Delta Y_{s,T}$ replace $\Delta\widehat{Y_{Ns,T}}$.

Table 10: *Estimates of Structural Time Series model (15)-(17)*

		$\hat{\sigma}_\epsilon$	$\hat{\sigma}_\eta$	$\hat{\sigma}_\zeta$	$\hat{\sigma}_\omega$	p.s.e.	$\beta(\Delta_4\widehat{y_{St}}, \Delta_4\Delta\widehat{y_{Nt}})$
Germany	levels + I	0	14.58	33.09	4.63	54.64	0.025
USA	log-levels	0	.0283	.0376	.0035	.0636	0.014
USA	log-levels + I	0	.0269	.0296	.0037	.0567	0.021

NOTES: The p.s.e. is the one-step-ahead prediction standard error. The $\beta(\Delta_4\widehat{y_{St}}, \Delta_4\Delta\widehat{y_{Nt}})$ is the regression coefficient β based on simulated series of length 10,000, where $\widehat{y_{S,t}} - \widehat{y_{S,t-4}}$ is regressed on $\widehat{y_{N,t}} - \widehat{y_{N,t-1}}$ - $\widehat{y_{N,t-4}} - \widehat{y_{N,t-5}}$. The OLS standard-error of β is 0.0006. The "+ I" denotes estimates allowing for intervention dummies in irregular component in 1967 and 1974 (Germany) or 1967 and 1975 (USA).

Table 11: *"Empirical" Asymptotic (Nonperiodic) Autocorrelations of the Seasonal Component and Cross-correlations of the Seasonal Component with the Nonseasonal Component*

	Germany				USA			
	Structural		Census		Structural		Census	
no.	2	3	4	5	6	7	8	9
Lag	$\widehat{y_{St}}$	$\widehat{y_{St}}\widehat{y_{Nt}}$	$\widehat{y_{St}}$	$\widehat{y_{St}}\widehat{y_{Nt}}$	$\widehat{y_{St}}$	$\widehat{y_{St}}\widehat{y_{Nt}}$	$\widehat{y_{St}}$	$\widehat{y_{St}}\widehat{y_{Nt}}$
0	1.00	0.20	1.00	-0.02	1.00	0.21	1.00	-0.03
1	-0.63	-0.00	-0.49	0.00	-0.64	0.00	-0.43	0.00
2	0.16	-0.20	-0.02	0.02	0.23	-0.21	-0.11	0.03
3	-0.25	0.41	-0.42	0.03	-0.39	0.38	-0.34	0.00
4	0.40	-0.22	0.87	-0.04	0.59	-0.23	0.80	-0.00
5	-0.18	-0.04	-0.46	0.00	-0.32	0.02	-0.41	-0.03
6	-0.04	-0.06	-0.02	-0.08	0.05	-0.12	-0.07	-0.07
7	-0.05	0.18	-0.30	0.20	-0.18	0.23	-0.23	0.19
8	0.14	-0.06	0.68	-0.14	0.33	-0.10	0.60	-0.10
9	-0.03	-0.05	-0.38	0.02	-0.16	-0.04	-0.34	-0.03
10	-0.06	-0.02	-0.01	-0.10	0.00	-0.04	-0.01	-0.06
11	-0.00	0.08	-0.16	0.22	-0.10	0.12	-0.10	0.19
12	0.05	-0.01	0.41	-0.14	0.21	-0.04	0.30	-0.10
13	-0.00	-0.04	-0.25	0.02	-0.10	-0.04	-0.21	-0.01
14	-0.03	-0.00	0.01	-0.09	-0.01	-0.02	0.06	-0.07
15	-0.00	0.04	-0.06	0.19	-0.06	0.09	-0.04	0.17
16	0.02	0.01	0.19	-0.12	0.12	-0.04	0.09	-0.10
17	-0.01	-0.03	-0.14	0.01	-0.04	-0.02	-0.11	-0.01
18	-0.01	-0.00	0.03	0.00	-0.02	-0.02	0.09	0.03
19	0.01	0.03	0.00	0.03	-0.03	0.06	0.01	-0.02
20	-0.01	-0.01	0.03	-0.01	0.06	-0.03	-0.07	0.01

NOTES: Columns no. 2,4,6,8: auto correlations $\Delta_4\widehat{y_{St}}$. Columns no. 3,5,7,9: cross-correlations $\Delta_4\widehat{y_{St}}$ and $\Delta_4\Delta_1 y\widehat{_{Nt-l}}$. *Lag* denotes lag l. Structural: Decomposition and simulation based on structural time series model of Table 10. Census: Decomposition based on 58-lag MA-filter approximating linear part of Census X-11 filter as in Laroque (1977). For "Census" and "Structural" we use the same simulated series. $\widehat{y_{St}}$: Estimate Seasonal component. $\widehat{y_{Nt}}$: Estimate Nonseasonal Component. Series length: 10,000.

Figure 1: Unadjusted unemployment in thousands of persons in Federal Germany 1960-1987. Numbers at end of plots indicate quarter of the year. Plots connect levels in a specific quarter $Y_{s,T}$, with levels in the same quarter in the previous year, $Y_{s,T-1}$.

Figure 2: Unadjusted unemployment in thousands of persons in the USA 1960-1987. Numbers at end of plots indicate quarter of the year. Plots connect levels in a specific quarter $Y_{s,T}$, with levels in the same quarter in the previous year, $Y_{s,T-1}$.

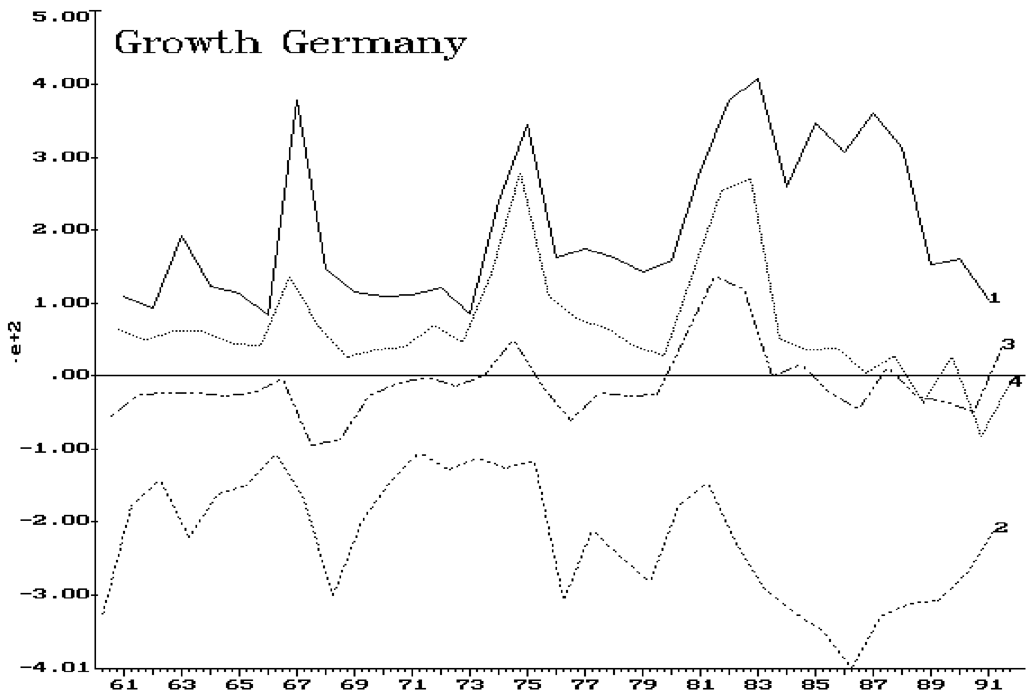

Figure 3: Changes in unemployment in thousands of persons in Federal Germany, 1960-1987. Numbers at end of plots indicate quarter of the year. Plots connect quarterly changes, $Y_{s,T} - Y_{s-1,T}$, in a specific quarter with the change in the same quarter of the previous year, $Y_{s,T-1} - Y_{s-1,T-1}$, where $Y_{0,T} = Y_{4,T-1}$.

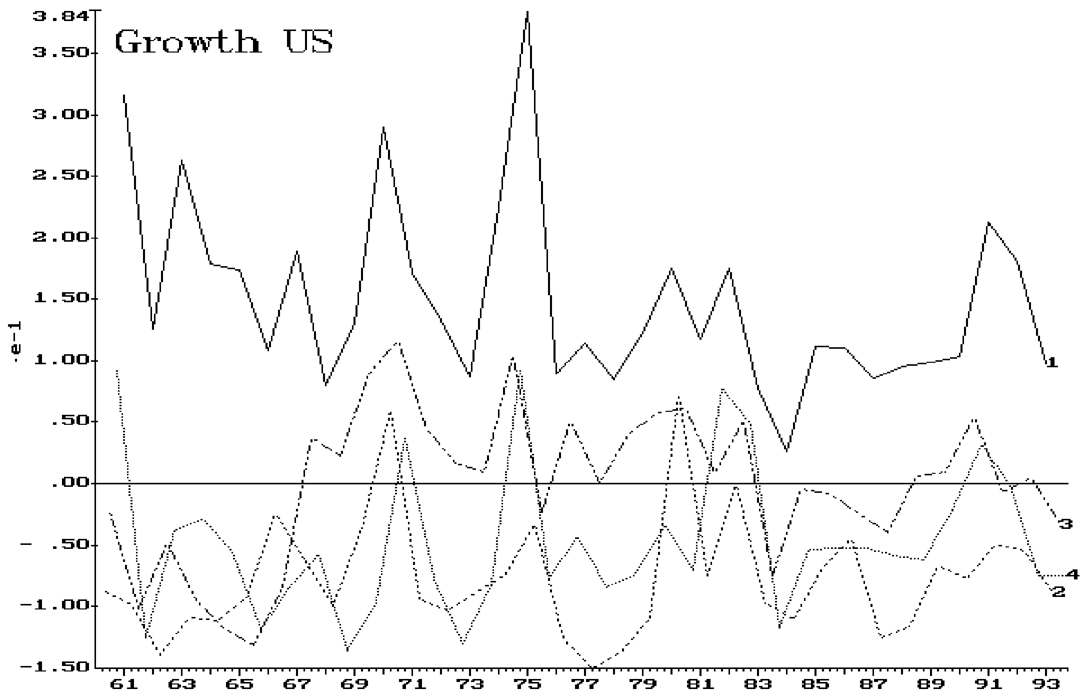

Figure 4: Changes in unemployment in thousands of persons in the USA, 1960-1987. Numbers at end of plots indicate quarter of the year. Plots connect quarterly changes in a specific quarter $Y_{s,T} - Y_{s-1,T}$ with the changes in the same quarter in the previous year, $Y_{s,T-1} - Y_{s-1,T-1}$, where $Y_{0,T} = Y_{4,T-1}$.

Figure 5: The "official" seasonal component of unemployment in thousands of persons in Germany, 1960-1987. Numbers at end of plots indicate quarter of the year. Plots connect the level in a specific quarter, $\widehat{Y_{Ss,T}}$, with the level in the same quarter in the previous year, $\widehat{Y_{Ss,T-1}}$.

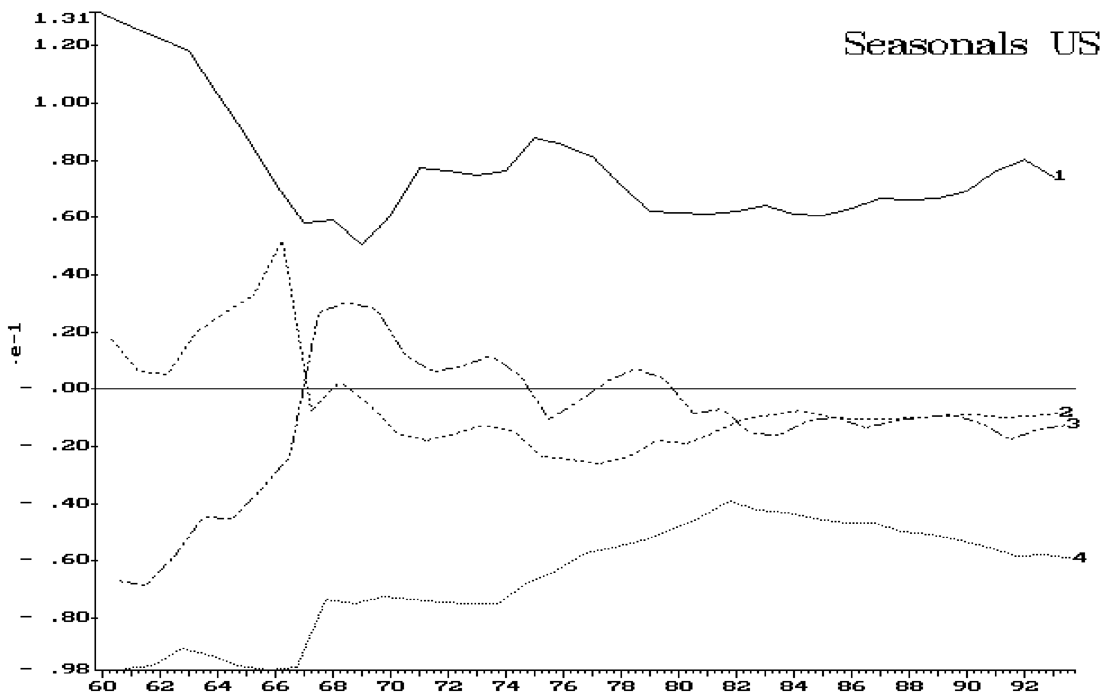

Figure 6: The "official" seasonal component of unemployment in thousands of persons in Germany, 1960-1987. Numbers at end of plots indicate quarter of the year. Plots connect the level in a specific quarter, $\widehat{Y_{Ss,T}}$, with the level in the same quarter in the previous year, $\widehat{Y_{Ss,T-1}}$.

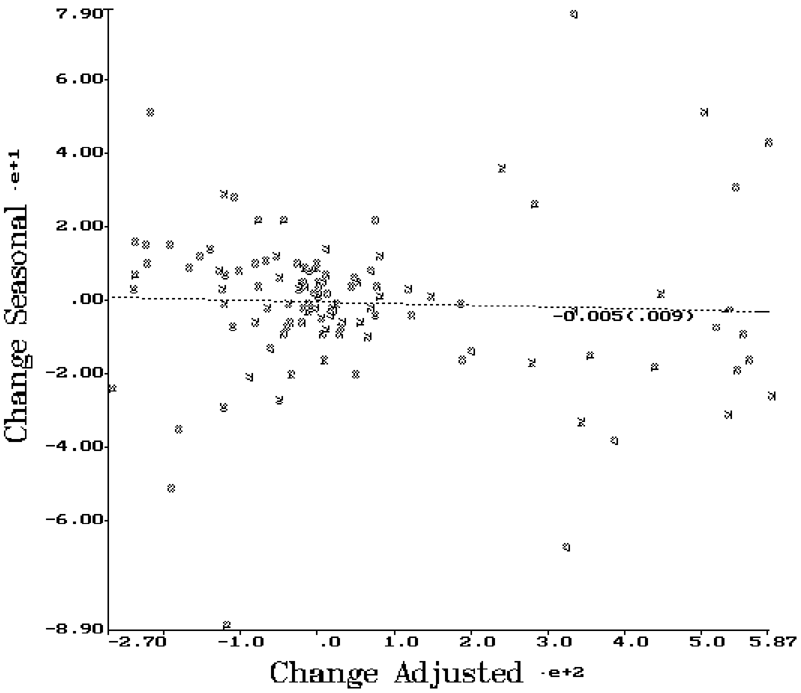

Figure 7: Pooled scatter plot of yearly changes in the adjusted number of unemployed, $\Delta_4 \widehat{y_{Nt}}$, versus yearly changes in the estimated seasonal component, $\Delta_4 \widehat{y_{St}}$ in Germany (1960-1987) with OLS estimate of regression coefficient (with estimated OLS standard error in parentheses). Points indicated by number of year.

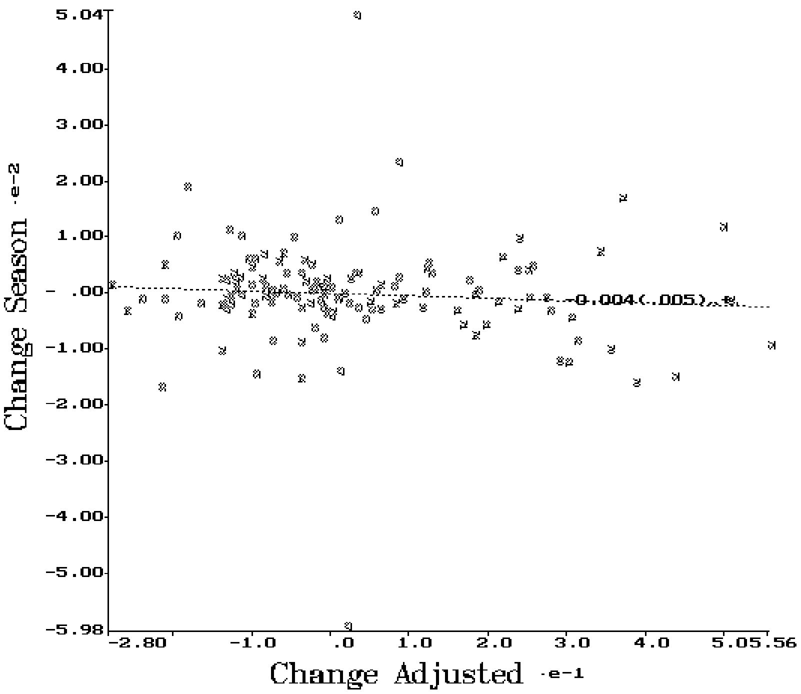

Figure 8: Pooled scatter plot of yearly changes in the adjusted number of unemployed, $\Delta_4 \widehat{y_{Nt}}$, versus yearly changes in the estimated seasonal component, $\Delta_4 \widehat{y_{St}}$ in the USA (1960-1987) with OLS estimate of regression coefficient (with estimated OLS standard error in parentheses). Points indicated by number of year.

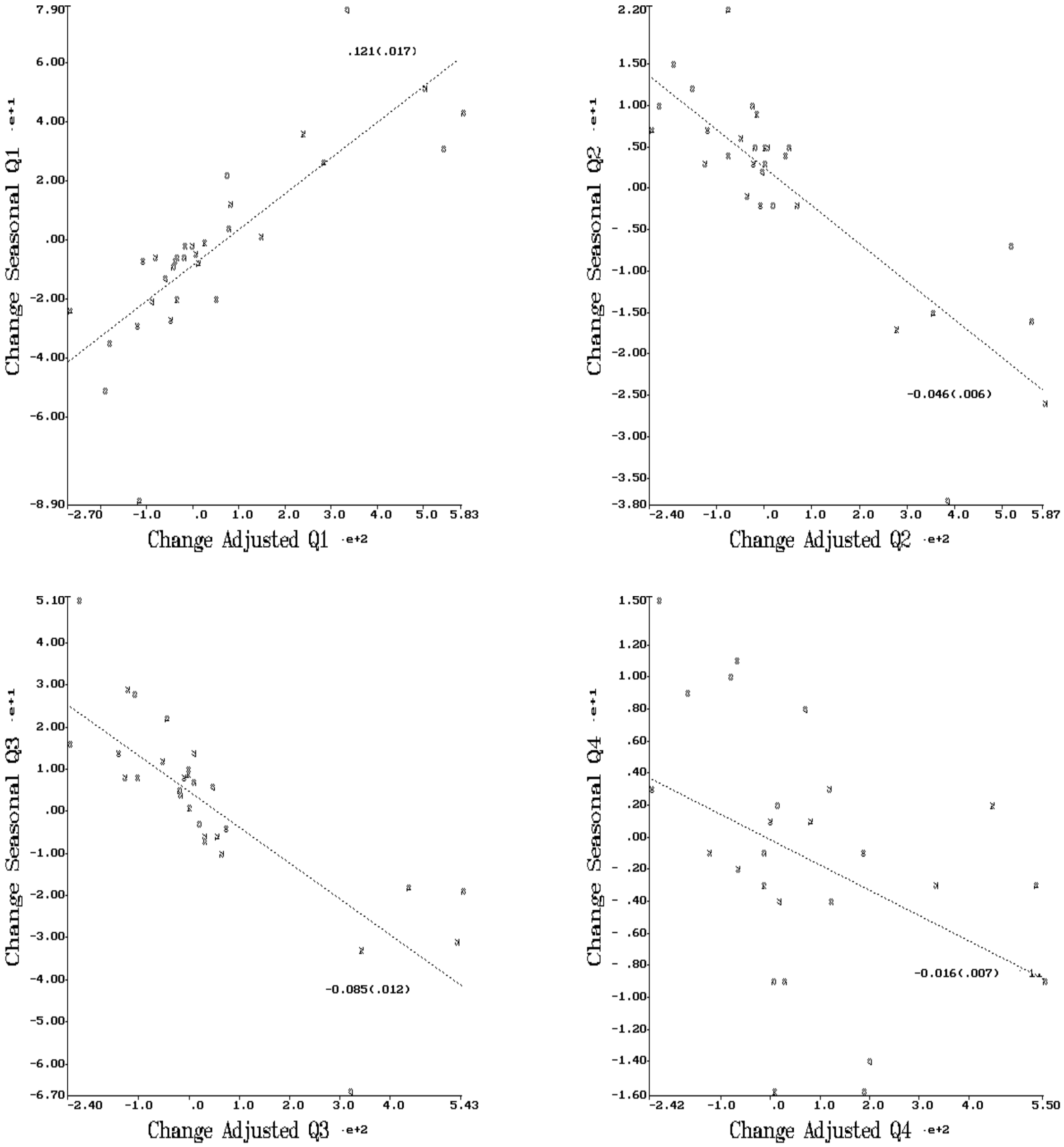

Figure 9: Scatter plots of yearly changes in the adjusted number of unemployed versus yearly changes in the estimated seasonal component for each of the 4 quarters Q_s, $s = 1, 2, 3, 4$ in Germany (1960-1987) with OLS estimate of regression coefficient (with OLS estimate of standard error in parentheses). Points indicated by number of year.

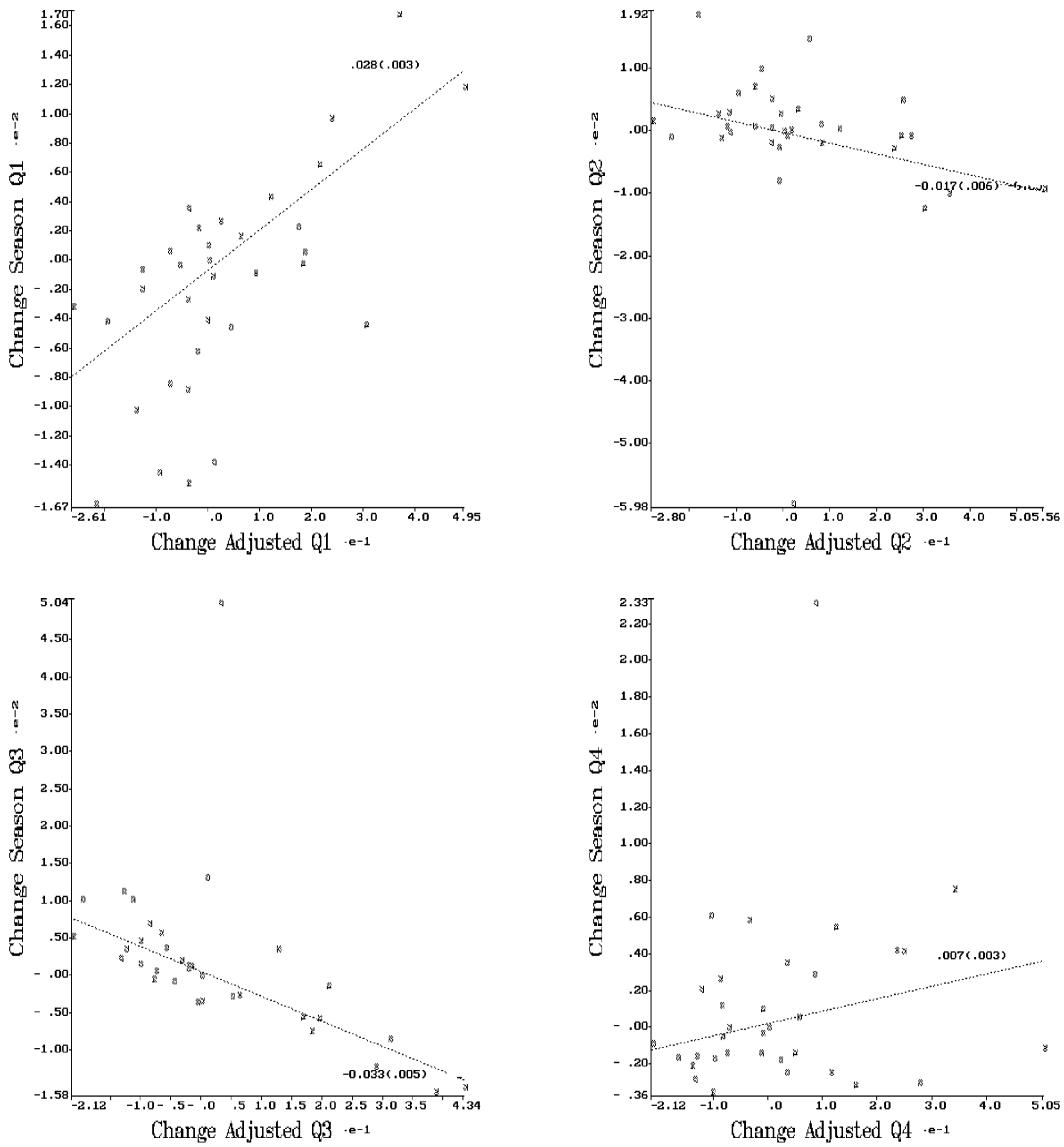

Figure 10: Scatter plots of yearly changes in the adjusted number of unemployed versus yearly changes in the estimated seasonal component for each of the 4 quarters Q_s, $s = 1, 2, 3, 4$ in the USA (1960-1987) with OLS estimate of regression coefficient (with OLS estimate of standard error in parentheses). Points indicated by number of year.

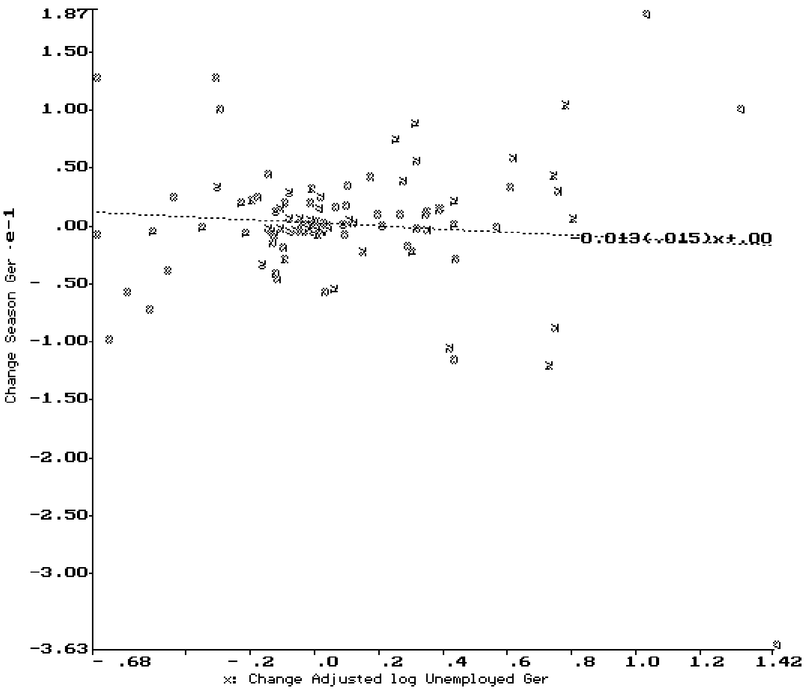

Figure 11: Pooled scatter plot of yearly changes in the **logarithm** of the adjusted number of unemployed versus yearly changes in the estimated seasonal component in Germany 1960-1987 with OLS estimate of regression coefficient (with estimated standard error in parentheses). Points indicated by number of year. Outlying observations for 1967.

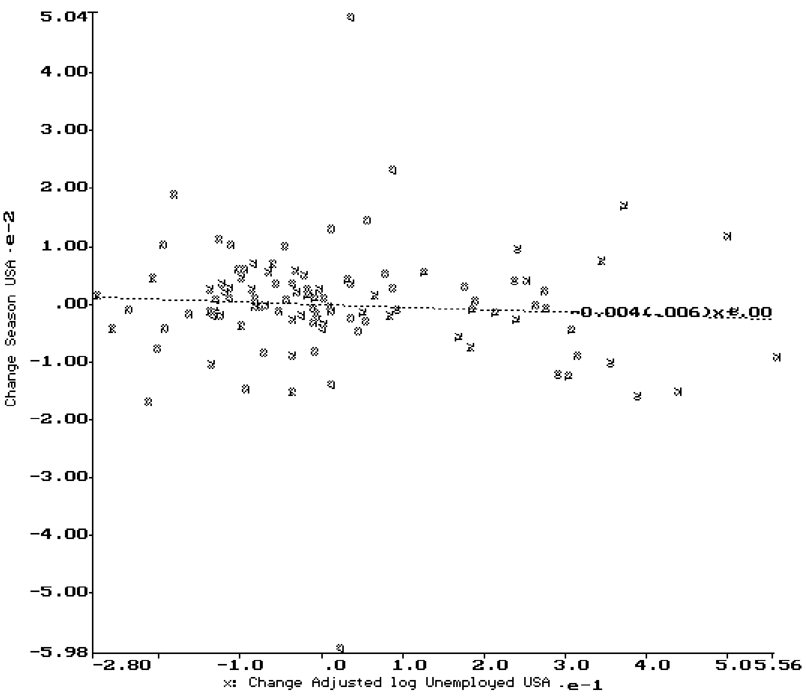

Figure 12: Pooled scatter plot of yearly changes in the **logarithm** of the adjusted number of unemployed versus yearly changes in the estimated seasonal component in the USA (1960-1987) with OLS estimate of regression coefficient (with estimated standard error in parentheses). Points indicated by number of year. Outlying observations for 1967.

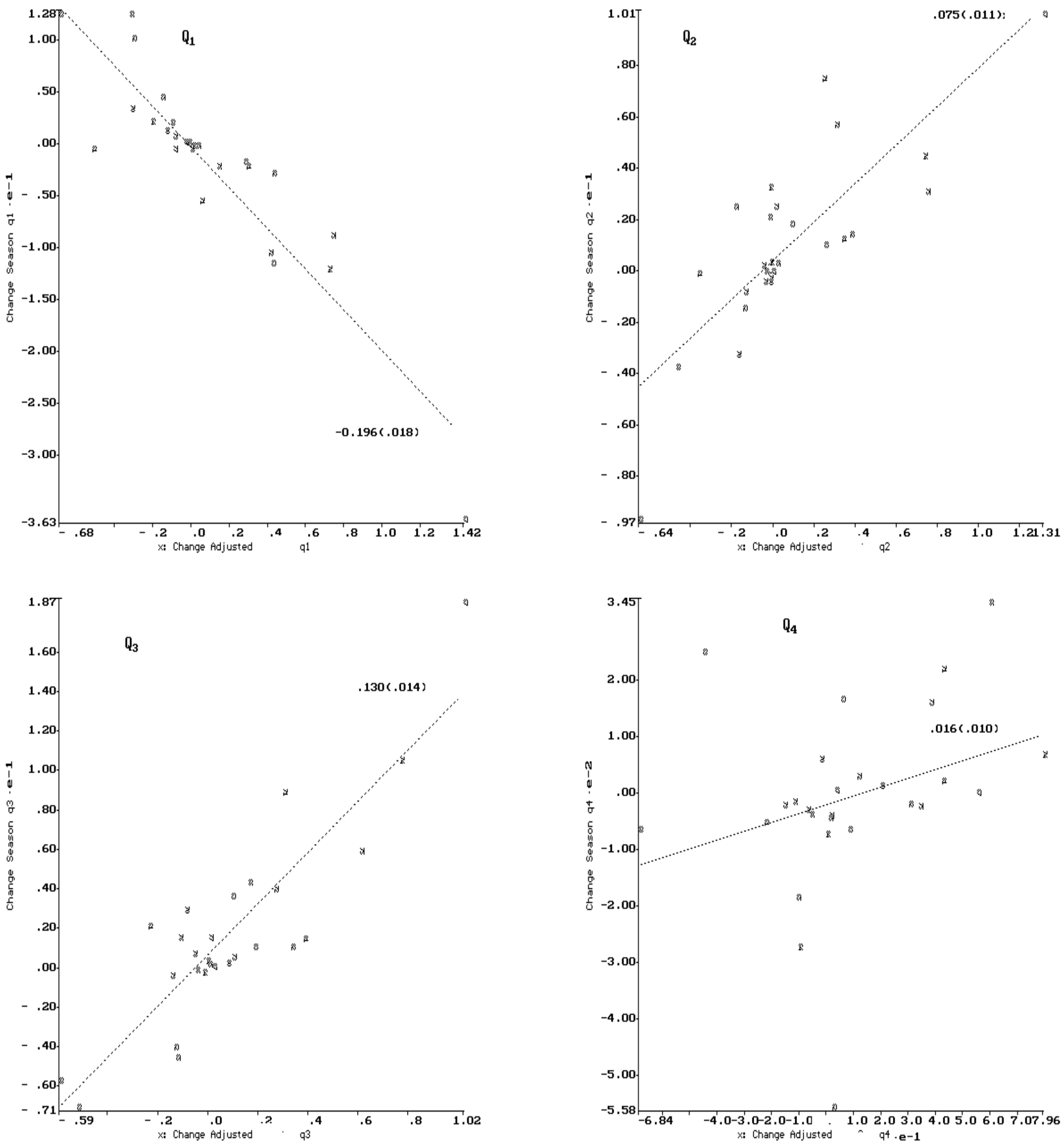

Figure 13: Yearly changes in the logarithm of the adjusted number of unemployed versus yearly changes in the estimated seasonal component for each of the 4 quarters $Q_s,\ s=1,2,3,4$ in Germany (1960-1987) with OLS estimate of regression coefficient (with OLS estimate of standard error in parentheses). Points indicated by number of year.

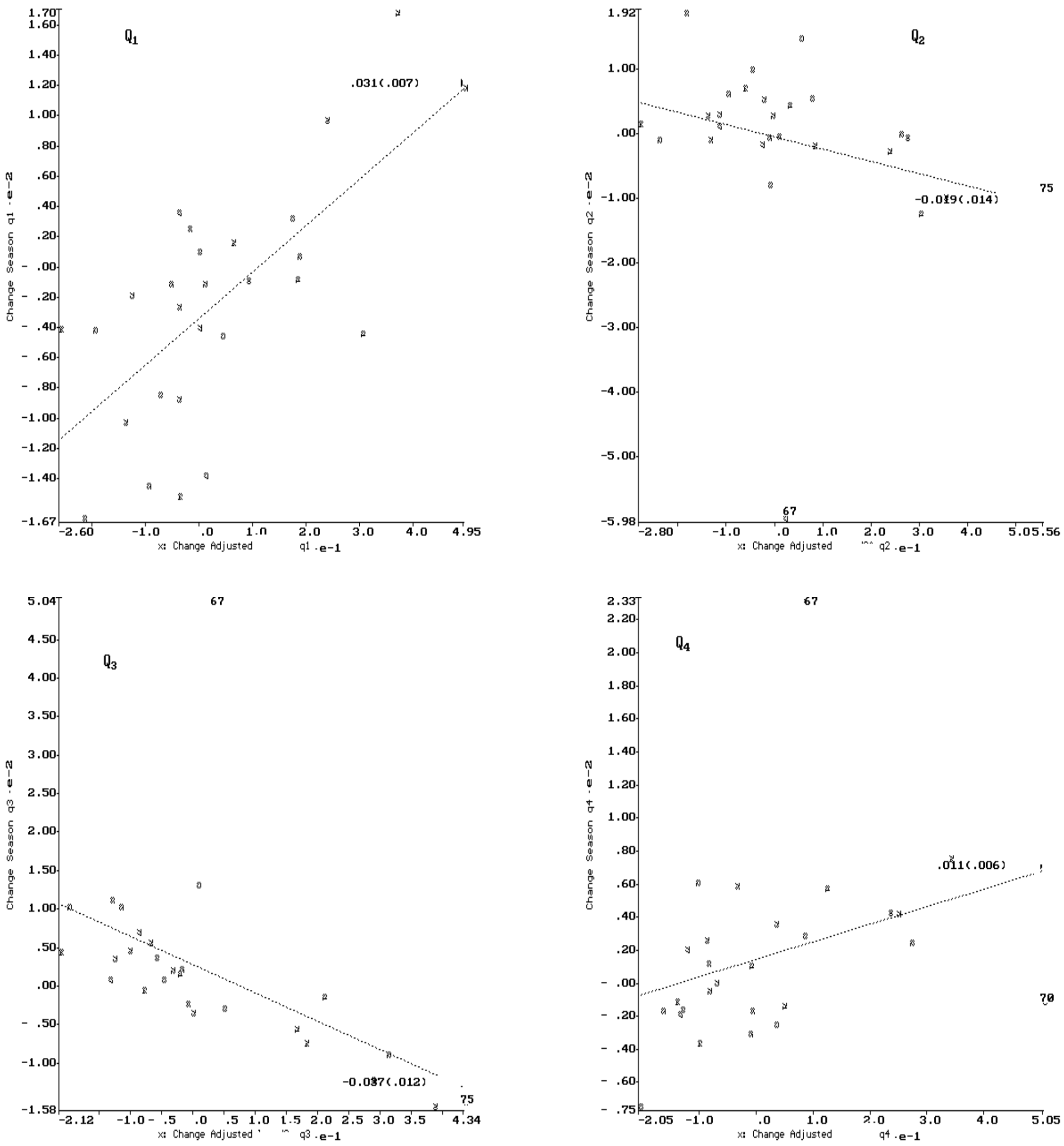

Figure 14: Yearly changes in the logarithm of the adjusted number of unemployed versus yearly changes in the estimated seasonal component for each of the 4 quarters Q_s, $s = 1, 2, 3, 4$ in the USA (1960-1987) with OLS estimate of regression coefficient (with OLS estimate of standard error in parentheses). Points indicated by number of year. Extreme outlying observation for 1967 in quarters 2 (low), 3 (high), and 4 (high).