

Zerlegung ökonomischer Zeitreihen Ein deterministischer und stochastischer Ansatz

by

**RALF PAULY
und
EKKEHART SCHLICHT**

Allgemeines Statistisches Archiv

1984

volume 68, number 2

pages 161– 175



published by

www.semverteilung.vwl.uni-muenchen.de

A. Abhandlungen

Zerlegung ökonomischer Zeitreihen Ein deterministischer und stochastischer Ansatz

Von RALF PAULY, Osnabrück und EKKEHART SCHLICHT, Darmstadt

Zusammenfassung: Gegenstand der Arbeit ist die Analyse einer Saisonhypothese in einem gewichteten Regressionsansatz zur Zerlegung einer beobachteten Zeitreihe in einen Trend, einen saisonalen Teil und einen Störterm. Es wird gezeigt, daß wie im deskriptiven Verfahren von Schlicht (1981) auch für die modifizierte Regression eine eindeutige Zerlegung ohne Vorgabe der Anfangswerte für den Trend und die Saison existiert. Es stellt sich heraus, daß sowohl die Zerlegungen mit als auch ohne Vorgabe der Anfangswerte als bedingte Erwartungswerte interpretierbar sind, wenn die Störungen unabhängig normalverteilte Zufallsvariablen sind. Die Zerlegung und auch die Prognose nach dem Regressionsansatz sind anhand von Beispielen illustriert.

Summary: The paper discusses a new seasonality hypothesis which is one part of a weighted regression approach for the decomposition of a time series into a trend, a seasonal component and an irregular component. It is shown that there exists a regression formulation leading, as in the descriptive approach in Schlicht (1981), to a unique decomposition without having recourse to initial values. It turns out that both solutions of the descriptive regression are conditional expected values in the stochastic specification. The decomposition as well as the prediction are illustrated by examples.

I. Problemstellung

Die Zerlegung einer beobachteten Zeitreihe x_t in der Periode t in einen Trend y_t , eine Saisonkomponente z_t und einen irregulären Störterm u_t erweist sich für ökonomische Analysen als zweckmäßig. Die im folgenden betrachtete additive Zerlegung

$$x_t = y_t + z_t + u_t, \quad t = 1, 2, \dots, T, \quad (1.1)$$

ist wenig restriktiv, wenn die Zerlegungshypothese für transformierte Zeitreihen x_t aufgestellt ist. Die Trendspezifikation

$$y_t = 2y_{t-1} - y_{t-2} + v_t, \quad t = 1, 2, \dots, T, \quad (1.2)$$

mit den Anfangswerten y_0 und y_{-1} basiert auf der Vorstellung, daß der Trend eine glatte Entwicklung hat, deren Tendenz sich nur allmählich mit der Zeit ändert. Dabei führen die Änderungen, welche die Störungen v_t in (1.2) bewirken, nicht zu sich regelmäßig wiederholenden Entwicklungsmustern. Unter der Annahme einer regelmäßigen Saisonfigur ergibt die Anwendung der verallgemeinerten Methode der kleinsten Quadrate auf (1.1) und (1.2) in Leser (1963) eine Zerlegung für kurze Zeitreihen. Die Zerlegungsverfahren in Akaike (1980) und in Schlicht (1981) können ebenfalls als gewichtete Regression dargestellt werden. Die Regressionsform enthält den Wiederholungs- und den Summationsaspekt der Periodizität in den beiden getrennten Saisonhypothesen

$$\left. \begin{array}{l} \text{a) } z_t = z_{t-s} + w_{1t} \\ \text{und} \\ \text{b) } z_t = \sum_{\tau=1}^{s-1} (-z_{t-\tau}) + w_{2t}, \quad t = 1, 2, \dots, T, \end{array} \right\} (1.3)$$

wobei s die Zahl der saisonalen Ausschläge in einem Jahr angibt. Die Anfangswerte sind hier $z_0, z_{-1}, \dots, z_{-s+1}$. Nach (1.3) ist die Saisonfigur eines Jahres ein fast regelmäßiges Muster, das sich nur langsam ändert. Die separate Spezifikation der beiden Grundvorstellungen der Periodizität führt jedoch zu Schwierigkeiten sowohl für die Stochastisierung als auch für die Prognose. Zudem beinhaltet sie unerwünschte Effekte der Störungen w_{1t} und w_{2t} auf die Musteränderung der Saison.

Gegenstand dieser Arbeit ist die Modifikation der Saisonhypothese in der gewichteten Regression. Es wird gezeigt, daß wie im deskriptiven Verfahren von Schlicht (1981) auch für die modifizierte Regression eine eindeutige Zerlegung ohne Vorgabe der Anfangswerte für den Trend und die Saison existiert. Es stellt sich heraus, daß sowohl die Zerlegung mit als auch ohne Vorgabe der Anfangswerte als bedingte Erwartungswerte interpretierbar sind, wenn die Störungen unabhängig normalverteilte Zufallsvariablen sind. Unter dieser Bedingung sind die einzelnen Komponenten nichtstationäre stochastische Prozesse. Das sich darin stellende Problem der Anfangswertbestimmung wird auf eine andere Weise als in Akaikes stochastischem Ansatz (1980) gelöst.

Im Abschnitt II werden die Eigenschaften der modifizierten Saisonhypothese analysiert. Die Zerlegungen im deskriptiven und stochastischen Ansatz werden im Abschnitt III. vorgeführt. Abschnitt IV. gibt ausgewählte Zerlegungen und Prognosen an.

II. Eigenschaften der Saisonspezifikation

Beide Grundvorstellungen der Periodizität aus (1.3) sind in der modifizierten Saisonspezifikation

$$z_t = z_{t-s} + (1 - g) \left[\sum_{\tau=1}^{s-1} (-z_{t-\tau}) - z_{t-s} \right] + w_t, \quad 0 \leq g \leq 1. \quad (2.1)$$

vereint. Nach (2.1) kann sich die fast zyklische Bewegung mit der Zeit ändern, Wie in (1.3a) wiederholt sich der Saisonausschlag z_{t-s} fast in gleicher Höhe nach einem Jahr. Ähnlich wie in (1.3) bewirkt ein negativer Impuls w_t eine Reduktion der saisonalen Komponenten z_t . Die Reduktion wird verstärkt durch die Übertragung einer Schrumpfungstendenz der saisonalen Bewegung in dem Zeitraum $[t - s, t - 1]$ auf

die Periode t . Die Differenz $\left[\sum_{\tau=1}^{s-1} (-z_{t-\tau}) - z_{t-s} \right]$ gibt das Ausmaß der Schrumpfungstendenz an, und $1 - g$ ist der Anteil, der in der Periode t erhalten bleibt. Demnach wird auch der Impuls w_t auf die folgenden Perioden übertragen.

Um die Wirkung einer Störung w_t auf die Musteränderung der Saison herauszustellen, zeichnen wir den Effekt eines einzelnen Impulses $w_{t_0} = 1$ in der Periode t_0 , $t_0 \in \{1, 2, \dots, T\}$, auf die Änderung der Saisonfigur in den darauffolgenden Perioden nach. Die übrigen Störungen seien Null, $w_t = 0$ für $t \neq t_0$. Die Referenz ist die ungestörte Saisonentwicklung z_t^* mit $w_t = 0$ für $t = 1, 2, \dots, T$. Die Differenz zwischen $z_{t_0+\theta}^*$ und der in der Periode t_0 gestörten Entwicklung $z_{t_0+\theta}$ gibt die Auswirkung der einmaligen Störung $w_{t_0} = 1$ auf die Änderung der Saisonfigur in der Periode $\theta = s(n - 1) + i$ an, wobei $i = 1, 2, \dots, s$ und $n = 1, 2, \dots, N$. Für $0 \leq g < 1$ sind die s saisonalen Unterschiede im Jahr n :

$$z_{t_0+s(n-1)+i} - z_{t_0+s(n-1)+i}^* = \begin{cases} (g - 1) g^{i-1} \left(\frac{g^{s \cdot n} - 1}{g^s - 1} \right) \\ \text{für } i = 1, 2, \dots, s - 1, \\ 1 - (1 - g) g^{i-1} \left(\frac{g^{s \cdot n} - 1}{g^s - 1} \right) \text{ für } i = s. \end{cases} \quad (2.2)$$

Nach (2.2) sind die langfristigen Unterschiede der saisonalen Ausschläge

$$\lim_{n \rightarrow \infty} (z_{t_0+s(n-1)+i} - z_{t_0+s(n-1)+i}^*) = \begin{cases} (g - 1) g^{i-1} \frac{1}{1 - g^s} \\ \text{für } i = 1, 2, \dots, s - 1, \\ \frac{1 - g^{i-1}}{1 - g^s} \text{ für } i = s. \end{cases} \quad (2.3)$$

Aus (2.3) folgt, daß die Summe der Abweichungen in einem Jahr Null ist:

$$\lim_{n \rightarrow \infty} \sum_{i=1}^s (z_{t_0+s(n-1)+i} - z_{t_0+s(n-1)+i}^*) = \lim_{n \rightarrow \infty} g^{s \cdot n} = 0. \quad (2.4)$$

Demnach resultiert aus dem einmaligen Impuls w_{t_0} ein neues Saisonmuster mit der wünschenswerten Eigenschaft, daß sich die negativen und positiven Saisonausschläge in einem Jahr ausgleichen. Nach (2.3) erhöht der Impuls $w_{t_0} = 1$ den Ausschlag in der s -ten Saisonperiode permanent um $(1 - g^{s-1})/(1 - g^s)$. Diese Erhöhung wird in den darauffolgenden $s - 1$ Saisonperioden in Form einer geometrischen Reduktion der Saisonausschläge vollständig kompensiert, so daß die Summe der Saisonausschläge in einem Jahr Null ist. Wie ein Vergleich von (2.3) mit (2.2) zeigt, sind die Abweichungen der saisonalen Bewegungen nach der Störung von dem neuen Muster sehr gering und die Konvergenz sehr schnell.

Die Abbildung 1 kennzeichnet die Musteränderung für $w_{t_0} = 1$ und $t_0 = 0$ im Modell (2.1) anhand von zwei Kurvenzügen. Die eine Kurve gibt die Änderung der Saisonausschläge $z_{s(n-1)+i} - z_{s(n-1)+i}^*$ nach (2.2) und die andere Kurve die Summe der Abweichungen $\sum_{i=1}^s (z_{s(n-1)+i} + z_{s(n-1)+i}^*)$ in einem Jahr an. Die beiden Entwicklungen sind für Monatsdaten, $s = 12$, und für ausgewählte Gewichte g angegeben.

Mit $g = 1$ und $g = 0$ enthält (2.1) die Hypothesen in (1.3) als Spezialfälle. In beiden Spezifikationen ist das Muster unmittelbar im ersten Jahr nach der Störung erreicht. Die Art der Musteränderung ist aber unbefriedigend. Die Summationsbedingung gemäß (2.4) ist in (1.3a) verletzt; dies wird deutlich in der letzten Graphik mit $g = 1$ in der Abbildung 1. Nach (1.3b) wird ein einmaliger Anstieg der saisonalen Bewegung in der anschließenden Periode sofort vollkommen ausgeglichen; dies veranschaulicht die erste Graphik mit $g = 0$ in der Abbildung 1. Zudem ist in den Ansätzen in Akaike (1980) und Schlicht (1981) undurchsichtig, wie die gewichtete Regression die beiden unerwünschten Effekte der separaten Spezifikation (1.3) kombiniert¹⁾.

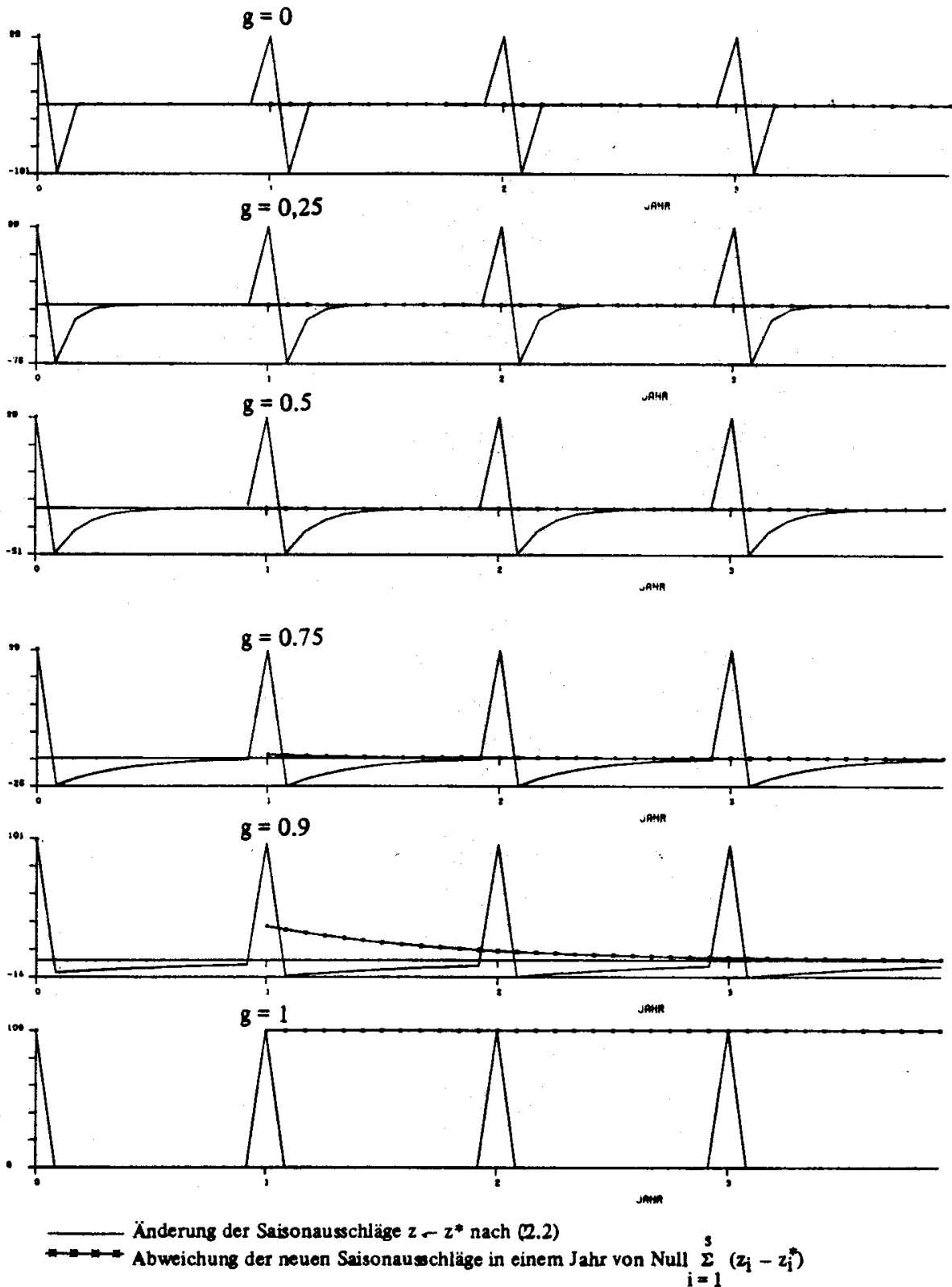
Stellen die Impulse w_t in (2.1) reine Zufallsstörungen dar, so wird man den Effekt der Zufallsstörung w_{t_0} auf die Wandlung der Saisonfigur mittels der Folge der Korrelationskoeffizienten von w_{t_0} und $z_{t_0+\theta}$ messen. Sind analog zur deskriptiven Analyse die übrigen Störungen w_t , $t \neq t_0$, vorgegeben, dann sind bei festen Anfangswerten in (2.1) die Korrelationskoeffizienten $\rho(z_{t_0+\theta}, w_{t_0})$ zwischen $z_{t_0+\theta}$ und w_{t_0} gleich den Differenzen aus (2.2) in der Periode θ , die durch den einmaligen Impuls w_{t_0} verursacht sind,

$$\rho(z_{t_0+\theta}, w_{t_0}) = (z_{t_0+\theta} - z_{t_0+\theta}^*) \quad (2.5)$$

Werden die übrigen Zufallsstörungen im Zeitraum $[1, T]$ nicht als gegeben betrachtet, dann sind die Kovarianzen $\text{cov}(w_{t_0}, z_{t_0+\theta})$ zwischen w_{t_0} und $z_{t_0+\theta}$ proportional zu den entsprechenden Abweichungen von $z_{t_0+\theta}$ und $z_{t_0+\theta}^*$ in (2.2). Die Proportionali-

¹⁾ Diese Kritik führt in Schlicht/Pauly (1983) zu einem weiteren Zerlegungsverfahren mit einer neuen Saisonhypothese.

Abbildung 1: Änderung des Saisonmusters im Modell (2.1) für $s = 12$ und ausgewählte Werte von g



tät sichert, daß die Wandlung des Saisonmusters in der stochastischen Interpretation von (2.1) sich in ähnlicher Weise vollzieht wie in der deterministischen Betrachtung. Nach Stochastisierung ist (2.1) ein nichtstationäres Zeitreihenmodell. Die Varianz der darin erklärten Saisonvariablen $z_{t_0+\theta}$ steigt mit θ . Der Anstieg hat zur Konsequenz, daß der Änderungseffekt der Zufallsstörung w_{t_0} , der anhand des Korrelogramms von w_{t_0} und $z_{t_0+\theta}$ gemessen wird, mit der Entfernung von t_0 abnimmt. Eine ähnliche Eigenschaft hat das stationäre Zeitreihenmodell, das in Hannan (1964) zur Erklärung der saisonalen Musteränderung diskutiert wird. Im Gegensatz zu stationären Prozessen hängt hier aber der Störeffekt nicht nur von der Entfernung von t_0 ab, sondern auch von der Periode t_0 , in der die Musteränderung durch die Zufallsstörung w_{t_0} beginnt.

III. Die gewichtete Regression als Zerlegungsverfahren

Sowohl der deterministische als auch der stochastische Ansatz bestehen aus den Spezifikationen (1.1), (1.2) und (2.1). Der additive Aufbau der Beobachtung x_t in (1.1) kann matriziell geschrieben werden in Form von

$$x = (I \ D) \begin{pmatrix} y \\ z \end{pmatrix} + u, \quad (3.1)$$

wobei $x' = (x_T, x_{T-1}, \dots, x_t, \dots, x_1)$. Entsprechend sind die anderen Vektoren y , z und u aufgebaut. I ist die $T \times T$ Einheitsmatrix. In Matrixnotation werden die Gleichungen der Trend- und Saisonspezifikation (1.2) und (2.1), in denen Anfangswerte auftreten, von den übrigen Gleichungen getrennt ausgewiesen. Die Spezifikation (1.2) ist in Matrixform

$$\begin{bmatrix} 0 \\ y^0 \end{bmatrix} = \begin{bmatrix} P^* \\ P^0 \end{bmatrix} y + \begin{bmatrix} v^* \\ v^0 \end{bmatrix}, \quad (3.2)$$

wobei der Trendvektor mit Anfangswerten

$$y^0 = (y_T - 2y_0 - y_{-1}) \text{ ist,}$$

$$(v^* \vdots v^0) = (v_T, v_{T-1}, \dots, v_3 \vdots v_2, v_1) = v'$$

und die $T \times T$ Matrix aus (3.2) definiert ist gemäß

$$\left[\frac{P^*}{P^0} \right] = \left[\begin{array}{cccccc} -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & & & & & 0 \\ 0 & & & 0 & -1 & 2 & -1 \\ \hline 0 & & & & & & \\ 0 & & & & & & \end{array} \right] = P \quad (3.3)$$

Die Saisonhypothese ist

$$\left[\frac{0}{z^0} \right] = \left[\frac{Q^*}{Q^0} \right] z + \left[\frac{w^*}{w^0} \right] \quad (3.4)$$

mit dem $s \times 1$ Anfangswertfaktor

$$z^0' = (-g \cdot z_0, -g \cdot z_{-1} - (g-1) z_0, \dots, -g z_{-s+1} - (g-1) \sum_{\tau=1}^{s-1} z_{1-\tau}),$$

$$(w^* \vdots w^0) = (w_T, w_{T-1}, \dots, w_{s+1} \vdots w_s, w_{s-1}, \dots, w_1) = w'$$

und der $T \times T$ Matrix

$$\left[\frac{Q^*}{Q^0} \right] = \left[\begin{array}{cccccc} \overbrace{-1 \quad (g-1) \quad (g-1) \quad \dots \quad (g-1)}^{s-1} & g & 0 & \dots & 0 \\ 0 & & & & 0 \\ 0 & & & 0 & -1 & (g-1) \quad (g-1) \quad \dots \quad (g-1) & g \\ \hline 0 & & & & & & \\ 0 & & & & & & \\ & & & & & & \vdots \\ & & & & & & g-1 \\ 0 & & & & & & -1 \end{array} \right] = Q \quad (3.5)$$

Aus (3.1), (3.2) und (3.4) ergibt sich der Regressionsansatz

$$\begin{bmatrix} x \\ -\frac{0}{y^0} \\ -\frac{0}{z^0} \end{bmatrix} = \begin{bmatrix} I & I \\ P & 0 \\ 0 & Q \end{bmatrix} \begin{pmatrix} y \\ z \end{pmatrix} + \begin{bmatrix} u \\ v \\ w \end{bmatrix}. \quad (3.6)$$

Die Glätte des Trends und die Stabilität des Saisonmusters sind die beiden elementaren Vorstellungen, die zur Spezifikation von (3.6) geführt haben. Die Variation der Störungen v und w bestimmen den Grad der Glätte und der Stabilität. Das Ausmaß der Streuung ist in der Regression dadurch steuerbar, daß man die Quadratsummen der Störungen $u'u$, $v'v$ und $w'w$ bei Anwendung der Methode der kleinsten Quadrate unterschiedlich gewichtet.

Aus der Minimierung von

$$L(y, z; y^0, z^0) = u'u + \alpha(v'^*v^* + v^0'y^0) + \beta(w'^*w^* + w^0'z^0), \quad \alpha, \beta > 0, \quad (3.7)$$

bezüglich y und z resultieren unter Beachtung von (3.6) die Normalgleichungen der gewichteten Regression

$$\begin{bmatrix} I + \alpha(P'^*P^* + P^0'P^0) & I \\ I & I + \beta(Q'^*Q^* + Q^0'Q^0) \end{bmatrix} \begin{pmatrix} \hat{y} \\ \hat{z} \end{pmatrix} = \begin{bmatrix} x + \alpha P^0'y^0 \\ x + \beta Q^0'z^0 \end{bmatrix}. \quad (3.8)$$

Die Lösung von (3.8) für $(\hat{y}'\hat{z}')$ setzt die Vorgabe der Anfangswerte in y^0 und in z^0 voraus. Das Problem der Anfangswertsetzung kann vermieden werden, indem man in (3.2) und (3.4) die Gleichungen eliminiert, welche Anfangswerte enthalten. Das Minimierungsergebnis ist dann

$$\begin{bmatrix} \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} I + \alpha P'^*P^* & I \\ I & I + \beta Q'^*Q^* \end{bmatrix}^{-1} \begin{bmatrix} x \\ x \end{bmatrix}, \quad 0 \leq g < 1. \quad (3.9)$$

Wie für (3.8) existiert auch für (3.9) eine eindeutige Lösung.

Theorem: Für eine beliebige Zeitreihe x ist das Minimum von

$$L^*(y, z) = u'u + \alpha v'^*v^* + \beta w'^*w^* \quad (3.10)$$

bezüglich y und z eindeutig durch (3.9) festgelegt.

Beweis: Es gilt

$$\frac{1}{2} \frac{\partial L^*}{\partial \begin{pmatrix} y \\ z \end{pmatrix}} = H^* \begin{pmatrix} \hat{y} \\ \hat{z} \end{pmatrix} - \begin{pmatrix} x \\ x \end{pmatrix}$$

und

$$\frac{1}{2} \frac{\partial^2 L^*}{\partial \begin{pmatrix} y \\ z \end{pmatrix}^2} = H^* = \begin{bmatrix} I + \alpha P^* P^* & I \\ I & I + \beta Q^* Q^* \end{bmatrix}.$$

Da

$$H^* = G'G \tag{3.11}$$

mit

$$G' = \begin{bmatrix} \sqrt{\alpha} P^* & \sqrt{1/2} I & 0 & \sqrt{1/2} I \\ 0 & \sqrt{1/2} I & \sqrt{\beta} Q^* & \sqrt{1/2} I \end{bmatrix}, \tag{3.12}$$

ist die Matrix H nichtnegativ definit. Zum Beweis der Existenz und der Eindeutigkeit der Lösung von (3.9) ist der Beweis der Nichtsingularität von H* ausreichend. Aus (3.11) und (3.12) folgt, daß

$$r(H^*) = r(G') \geq T + r \begin{pmatrix} Q^* \\ -P^* \end{pmatrix} \geq r(M) + T,$$

wobei M die folgende T x T Teilmatrix von $\begin{pmatrix} Q^* \\ -P^* \end{pmatrix}$ ist:

$$\begin{bmatrix} -1 & (g-1) & (g-1) & \dots & (g-1) & g & 0 & \dots & 0 \\ 0 & -1 & (g-1) & & & (g-1) & g & 0 & \dots & 0 \\ \hline & & & & & -P^* & & & & \end{bmatrix}$$

Analog zu Schlicht (1981) kann man nun zeigen, daß $\det M = ((1-g) \cdot s)^2$. Demnach hat die Matrix für $g \neq 1$ den Rang T, und die Matrix H* ist nichtsingulär.

Im stochastischen Ansatz sind die Störungen als reine Zufallsfehler betrachtet. Unter der Annahme, daß der Zufallsvektor (u^*, v^*, w^*) normalverteilt ist gemäß $N(0, \Sigma)$, wobei

$$\Sigma = \begin{bmatrix} \sigma_u^2 I & 0 & 0 \\ 0 & \sigma_v^2 I & 0 \\ 0 & 0 & \sigma_w^2 I \end{bmatrix}, \tag{3.13}$$

ist für gegebene Anfangswerte nach (3.6) die Lösung der gewichteten Regression aus (3.8) gleich dem bedingten Erwartungswert von den Trend- und Saisonkomponenten,

$$E \begin{pmatrix} y \\ z \end{pmatrix} | x, y^0, z^0 = \begin{pmatrix} \hat{y} \\ \hat{z} \end{pmatrix}, \quad (3.14)$$

wobei

$$a = \sigma_u^2 / \sigma_v^2 \quad \text{und} \quad \beta = \sigma_u^2 / \sigma_w^2.$$

Ersetzt man (2.1) durch die separate Saisonspezifikation (1.3), so ist die neue gewichtete Regressionslösung im stochastischen Modell ebenfalls der bedingte Erwartungswert der Komponenten. Vorausgesetzt ist dabei, daß die Störungen w_{1t} und w_{2t} stochastisch unabhängig sind²⁾. Diese Voraussetzung ist jedoch nicht für feste Anfangswerte z^0 gegeben, wie dies direkt aus (1.3) für $t = 1$ ersichtlich ist.

Die Trend- und Saisonkomponenten aus (3.8) bzw. aus (3.14) sind erst nach Vorgabe der Anfangswerte in y^0 und z^0 berechenbar. Das Einsetzen von y aus (3.2) und von z aus (3.4) in (3.1) ergibt ein verallgemeinertes Regressionsmodell mit den Regressionskoeffizienten y^0 und z^0 . Bei gegebenen Gewichten g , a , und β ist die Regressormatrix fest vorgegeben, so daß der GLS-Schätzer die besten Schätzungen für y^0 und z^0 ergibt³⁾. Eine damit übereinstimmende Schätzung von y^0 und z^0 resultiert aus (3.9), worin die Anfangswerte impliziert bestimmt sind⁴⁾. Die „Rückwärtsprognosen“ nach den Formeln (3.2) und (3.4) führen zu folgenden Schätzungen:

$$\hat{y}^0 = P^0 \hat{y} \quad \text{und} \quad \hat{z}^0 = Q^0 \hat{z}. \quad (3.15)$$

Die Zerlegung (3.9) und die Schätzer (3.15) erhalten ihre Bedeutung im stochastischen Rahmen durch das folgende Theorem und Korollar.

Theorem: Für eine beliebige Zeitreihe x fallen die Zerlegungen nach (3.8) und (3.9) dann und nur dann zusammen, wenn die unbekanntenen Anfangswerte in (3.8) gemäß (3.15) festgelegt sind.

Korollar: Die Schätzer in (3.15) sind unverzerrt, $E_x(\hat{y}^0 | y^0, z^0) = y^0$ und $E_x(\hat{z}^0 | y^0, z^0) = z^0$.

²⁾ Vergleiche die Lösung in Akaïkes Ansatz.

³⁾ Siehe hierzu Pauly (1982). Die Gewichte g , a und β können nach dem Maximum-Likelihood-Prinzip bestimmt werden. Dazu wendet man das Prinzip auf die Dichtefunktion von x an, die aus (3.1) nach Substitution von y aus (3.2) und von z aus (3.4) resultiert.

⁴⁾ Wegen dieses bemerkenswerten Ergebnisses siehe Pauly (1984).

Beweis: Ersetzt man (x', x') in (3.8) durch die entsprechenden Vektoren aus (3.9) und y^0 sowie z^0 in (3.8) durch ihre Schätzungen aus (3.15), dann sehen wir, daß $\hat{y} = \hat{y}$ und $\hat{z} = \hat{z}$. Demnach folgt aus (3.15)

$$\hat{y}^0 = P^0 \hat{y} \text{ und } \hat{z}^0 = Q^0 \hat{z}.$$

Da

$$\begin{aligned} E_x \{ \hat{y} | z^0, y^0 \} &= E_x \{ E(y | x, y^0, z^0) \} = P^{-1} \begin{pmatrix} 0 \\ y^0 \end{pmatrix} - E_x \{ P^{-1} E(v | x, y^0, z^0) \} \\ &= P^{-1} \begin{pmatrix} 0 \\ y^0 \end{pmatrix} \end{aligned}$$

und da

$$P^0 P^{-1} \begin{pmatrix} 0 \\ y^0 \end{pmatrix} = y^0,$$

ist der Schätzer \hat{y}^0 unverzerrt. Entsprechende Gleichungen gelten für \hat{z} . Demnach ist auch \hat{z}^0 unverzerrt. Es fehlt noch der Beweis, daß aus $\hat{y} = \hat{y}$ und aus $\hat{z} = \hat{z}$ die Gleichungen (3.15) resultieren. Diese Implikation folgt direkt aus (3.8). Für $\hat{y}^0 = P^0 \hat{y}$ ist zu beachten, daß

$$P^{0'} y^0 = P' \begin{pmatrix} 0 \\ y^0 \end{pmatrix} \text{ und } P'^{-1} P^{0'} P^0 \hat{y} = \begin{pmatrix} 0 \\ P^0 \hat{y} \end{pmatrix}.$$

Entsprechend gilt für \hat{z}^0 in (3.15).

Bemerkenswert ist zudem die folgende Beziehung zwischen den Summen der gewichteten Quadrate L in (3.7) und L^* in (3.10). Aus (3.6), (3.7) und (3.10) ist ersichtlich, daß für beliebige Anfangswerte (y^0, z^0)

$$L(\hat{y}, \hat{z}; y^0, z^0) \geq L^*(\hat{y}, \hat{z}).$$

Die Wahl der in (3.9) implizierten Anfangswerte führt zu

$$L(\hat{y}, \hat{z}; \hat{y}^0, \hat{z}^0) = L^*(\hat{y}, \hat{z}),$$

was äquivalent ist mit $\hat{y} = \hat{y}$ und $\hat{z} = \hat{z}$; die Anfangswerte aus (3.15) minimieren die Quadratsumme L .

IV. Numerische Resultate

Monatliche Arbeitslosenzahlen aus der Bundesrepublik Deutschland sind nach dem Verfahren (3.9) zerlegt worden. Verschiedene Gewichte sind gewählt worden. Die Abbildungen 2 und 3 sowie die Tabelle „Ausgewählte Saisonzerlegungen und -prognosen“ enthalten einige Ergebnisse. Die Gewichte α und β steuern die Glätte des Trends und die Regularität des Saisonmusters. Eine Heraufsetzung der Gewichte α und β erhöht die Glätte und die Stabilität. Dies geht allerdings zu Lasten eines Varianzanstiegs der irregulären Komponenten. Die beiden Abbildungen 2 und 3 belegen den Stabilitätseffekt, den eine Variation von β verursacht. Die Saisonentwicklung ist bei $\beta = 1000$ wesentlich regelmäßiger als bei $\beta = 10$. Dabei ist die Grundstruktur der jährlichen Bewegung in den beiden Zerlegungen ähnlich. Die Herabsetzung des Gewichts von 1000 auf 10 läßt die Schrumpfungstendenz in den sechziger Jahren und die leichte Wandlungstendenz des Saisonmusters in den siebziger Jahren prägnanter hervortreten.

Tabelle: Ausgewählte Saisonzerlegungen und -prognosen ($\alpha = 1000$ und $\beta = 10$)

Monatliche Arbeitslosenzahlen in der Bundesrepublik (in 1000)

g = 0,80						g = 0,95					
Z			P			Z			P		
	1979	80	81	82	83		1979	80	81	82	83
10	-76	-79	-81	-81	-80	10	-77	-79	-80	-80	-79
11	-27	-27	-27	-27	-26	11	-27	-28	-26	-26	-25
12	78	84	90	90	91	12	77	83	90	91	91
	1980	81	82	83	84		1980	81	82	83	84
1	238	247	257	257	258	1	236	245	256	256	257
2	205	211	218	219	220	2	204	210	218	219	219
3	71	75	79	79	81	3	70	74	79	79	80
4	-14	-17	-18	-18	-16	4	-15	-18	-17	-17	-17
5	-100	-104	-107	-107	-103	5	-100	-104	-106	-106	-105
6	-120	-126	-130	-130	-126	6	-120	-126	-130	-129	-129
7	-70	-72	-76	-75	-70	7	-71	-74	-76	-76	-75
8	-73	-76	-79	-79	-72	8	-75	-77	-79	-79	-78
9	-123	-126	-126	-126	-118	9	-123	-126	-126	-126	-125
	-11	-10	0	+2	+39	Σ	-21	-20	+3	+6	+14

Z = Zerlegung, P = Prognose

Abbildung 2: Zerlegung I ($\alpha = 1000; \beta = 1000; g = 0,5$)
 Arbeitslose in der Bundesrepublik Deutschland 1950/1-1982/9

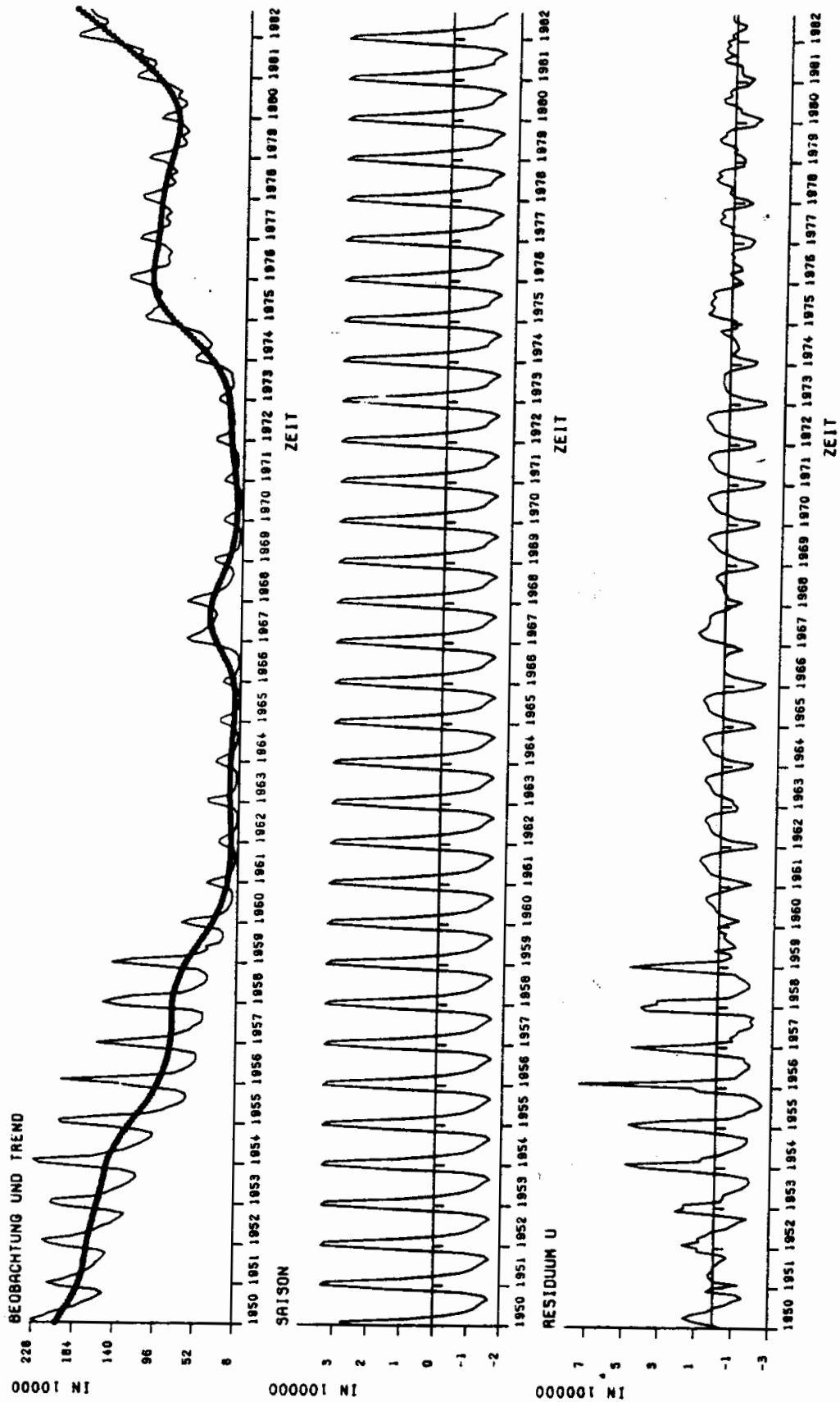
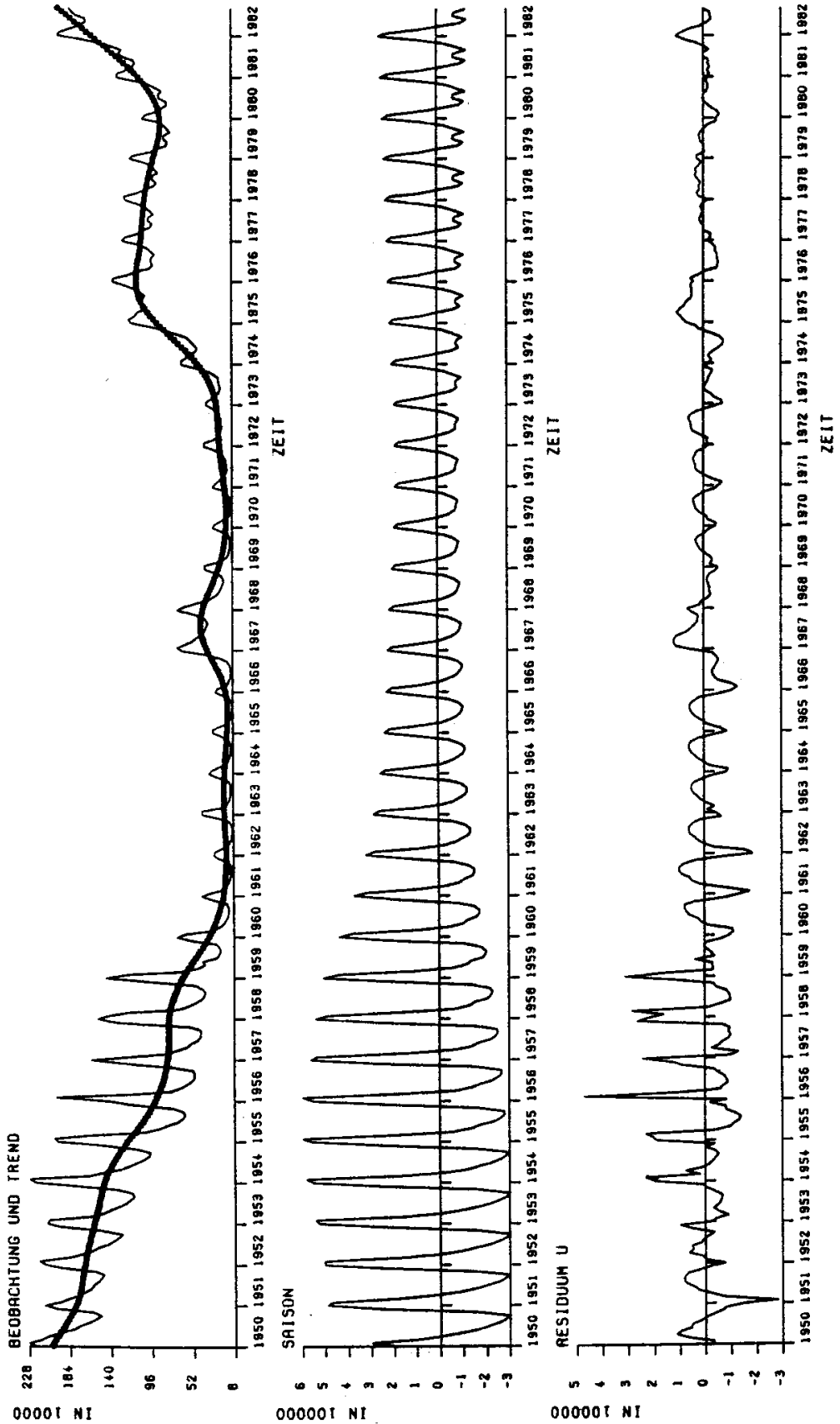


Abbildung 3: Zerlegung II ($\alpha = 1000; \beta = 10; g = 0,5$)

Arbeitslose in der Bundesrepublik Deutschland 1950/1-1982/9



Die Änderung der Saisonfigur hängt nach (2.1) von der Gewichtung g ab. Wie wir aus (2.2) und (2.3) wissen, verlangsamt eine Erhöhung von g den Anpassungsprozeß nach einer Störung an das neue Saisonmuster. In den Untersuchungen zeigt sich, daß die Wahl von g wenig Einfluß auf die Zerlegung hat. Insbesondere am oberen Rand einer Zeitreihe, der für ökonomische Analysen von zentraler Bedeutung ist, ist das Zerlegungsverfahren wenig sensitiv gegenüber Parametervariationen von g . Dagegen hängt die Regularität der Saisonprognose ganz erheblich von g ab. Die Stabilität der Prognose erhöht sich mit wachsendem g und β . Für hohe Werte von g bleibt die prognostizierte Saisonfigur über mehrere Jahre stabil. Dies geht aus der saisonalen Bewegung über fünf Jahre in der Tabelle „Ausgewählte Saisonzerlegungen und -prognosen“ hervor. Die Differenz zwischen den beiden prognostizierten Saisonmustern und den Zerlegungen in den drei vorangegangenen Jahren sind für die beiden hohen Werte $g = 0,8$ und $g = 0,95$ nur geringfügig. Die Prognosewerte sind nach der Formel (2.1) berechnet, wobei die Störungen w Null gesetzt sind.

Analog zu (2.1) lassen sich über die Formel (1.3) Prognosen ermitteln. Die separate Spezifikation der Saison in (1.3) erbringt jedoch, wie wir wissen, zwei voneinander abweichende Prognosen, die sich bereits nach der ersten Prognoseperiode Jahr für Jahr wiederholen. Demnach ist auch die Musteränderung weniger flexibel als in (2.1). Anders als in (2.1) existiert aber für (1.3) kein Stabilitätsproblem.

Literaturverzeichnis

- Akaike, H. (1980): Seasonal Adjustment by a Bayesian Modeling; *Journal of Time Series Analysis*, Vol. 1 no. 1, pp. 1–13.
- Giebel, F. (1982): Documentation of a Software Package which Decomposes, Analyses and Forecasts Economic Time Series (I), *Beiträge des Fachbereichs Wirtschaftswissenschaften der Universität Osnabrück*, Available on request (mimeo).
- Hannan, E. J. (1964): The Estimation of a Changing Seasonal Pattern; *Journal of the American Statistical Association*, Vol. 59, pp. 1063–1077.
- Leser, C. E. V. (1963): Estimation of Quasi-Linear Trend and Seasonal Variation; *Journal of the American Statistical Association*, Vol. 58, pp. 1033–1043.
- Pauly, R. (1982): Zerlegung und Analyse ökonomischer Zeitreihen; *Statistische Hefte*, pp. 291–303.
- Pauly, R. (1984) Regressionsansätze zur Zeitreihenzerlegung, in: J. Frohn (Hrsg.), *Zur Spezifizierung und Analyse ökonometrischer Modelle*, Sonderhefte zum Allgemeinen Statistischen Archiv, Göttingen.
- Schlicht, E. (1981): A seasonal Adjustment Principle and a Seasonal Adjustment Method Derived from this Principle; *Journal of the American Statistical Association*, Vol. 76, no. 374, pp. 374–378.
- Schlicht, E. and R. Pauly (1983) Descriptive Seasonal Adjustment by Minimizing Perturbations, *Empirica*, pp. 15–28.