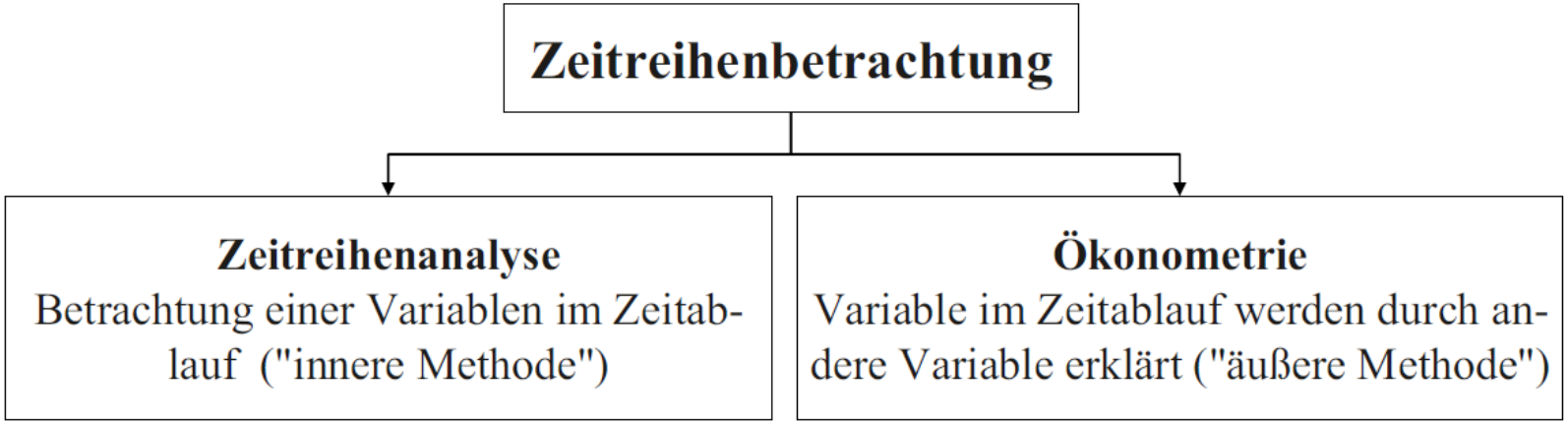
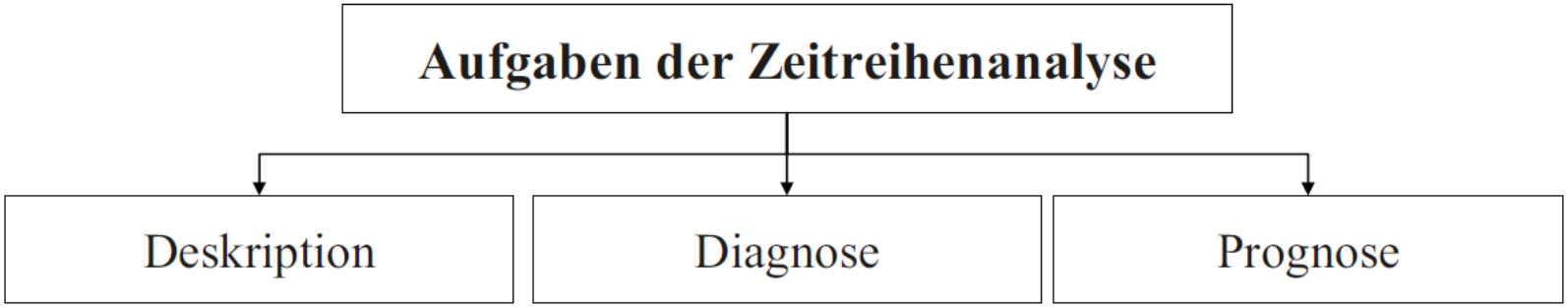
10. Zeitreihenanalyse

10.1 Aufgaben und Ziele

- In der Zeitreihenanalyse wird eine zeitliche Folge von Beobachtungen (Zeitreihe) statistisch untersucht:
  - Bestandsgrößen (z.B. Anzahl Kraftfahrzeuge, Erwerbstätige, Arbeitslose)
  - Stromgrößen (z.B. Umsatz eines Unternehmens, Zahl der Urlauber, Konsumausgaben)
- Es sollen Gesetzmäßigkeiten aufgedeckt werden, denen die Zeitreihe in Abhängigkeit von der Zeit unterliegt:

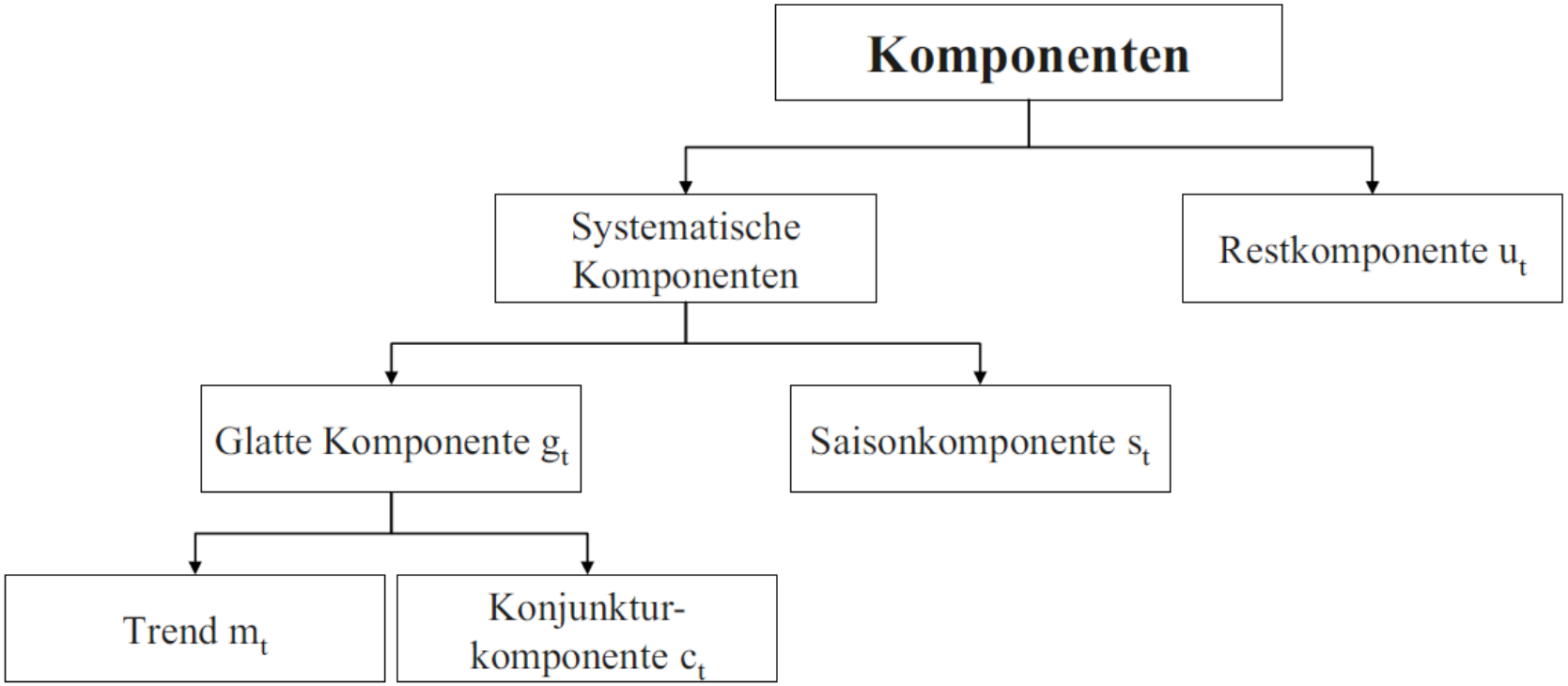




- **Deskription**: z.B. Vergleich der langfristigen Entwicklung von Preisen oder Investitionen in verschiedenen Regionen.
- **Diagnose**: z.B. Einschätzung, ob ein aktuelles Absinken einer Zeitreihe als eine konjunkturelle Auf- oder Abwärtsbewegung interpretiert werden kann oder aber auf saisonale Faktoren zurückzuführen ist.
- **Prognose**: z.B. Umsatzprognose zur sachlich fundierten Planung von Investitionen.

10.2 Zeitreihenzerlegung

- Ökonomische Zeitreihen lassen sich als Resultat eines Zusammenwirkens verschiedener Bewegungskomponenten auffassen:

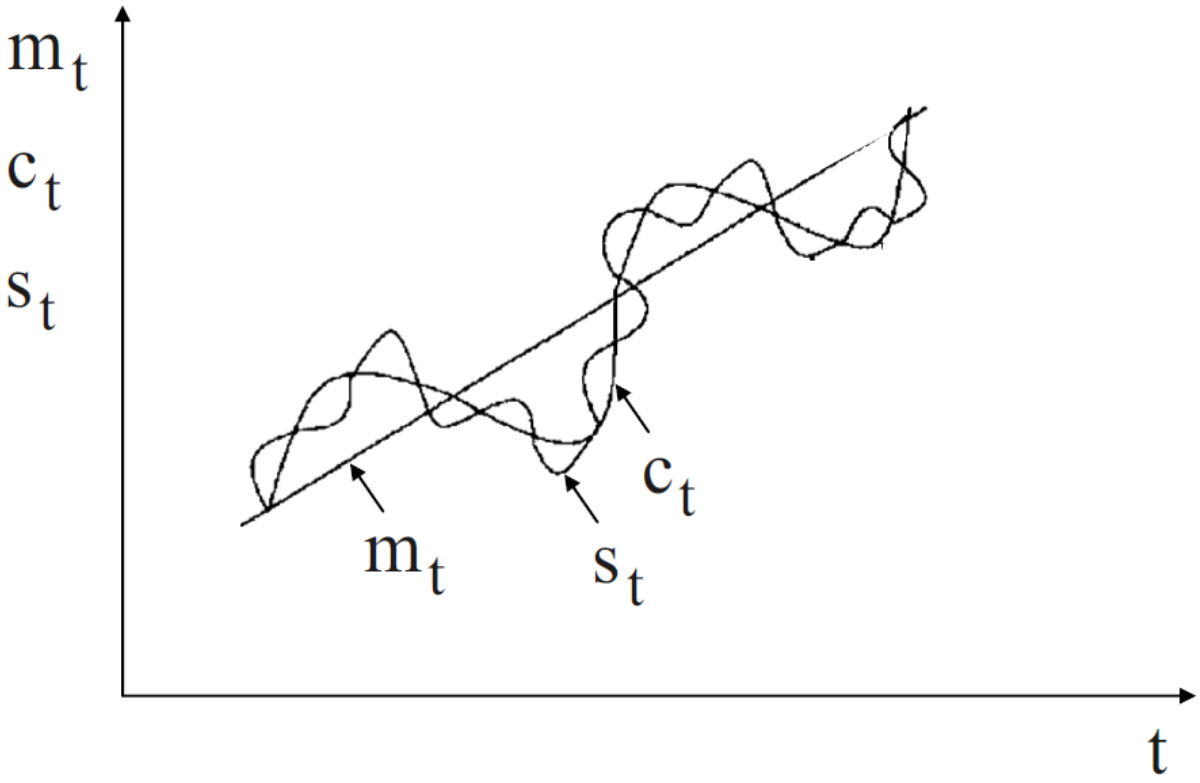


Komponenten ökonomischer Zeitreihen (I)

- **Trend**: Langfristige Entwicklungstendenz
- **Konjunkturkomponente**: Zyklische Schwankungen mit einer Periode von etwa vier bis sechs Jahren
  - Isolierung der Konjunkturkomponente vom Trend setzt lange Zeitreihen voraus und ist aufgrund der unregelmäßigen Zykluslängen sehr schwierig
  - In vielen Fällen fasst man daher den Trend und die Konjunkturkomponente zu einer **glatten Komponente** zusammen.
- **Saisonkomponente**: Erfasst die jahreszeitlichen Einflüsse und überlagert damit den Trend und die Konjunkturkomponente.
  - Umfasst zum einen natürliche Ursachen wie Jahreszeiten und Witterungseinflüsse, aber auch institutionelle Ursachen wie Feiertage, Ferien und Zahlungstermine.
  - Hat eine konstante Periodenlänge.
- **Restkomponente**: Irreguläre Bewegungen (Zufallsschwankungen, einmalige oder unregelmäßige Einflussgrößen wie Streiks oder Frosteinbruch)

Komponenten ökonomischer Zeitreihen (II)

- Zeitreihendiagramm der systematischen Komponenten in einem t,y-Koordinatensystem (Zeitreihenpolygon):



Zeitreihenzerlegung

- Unter einer Zeitreihenzerlegung wird die Separierung der Komponenten einer Zeitreihe (y_t) verstanden.
- **Additives Grundmodell**:
  - Annahme: Zyklische Schwankungen (Konjunktur- und Saisonkomponente) mit konstanter Amplitude:

$$y_t = m_t + c_t + s_t + u_t$$

- **Multiplikatives Grundmodell**:
  - Annahme: Zyklische Schwankungen (Konjunktur- und Saisonkomponente) mit proportional wachsender Amplitude:

$$y_t = m_t \cdot c_t \cdot s_t \cdot u_t$$

  - Nach Logarithmieren in additiver Form (mit $g_t = m_t \cdot c_t$):

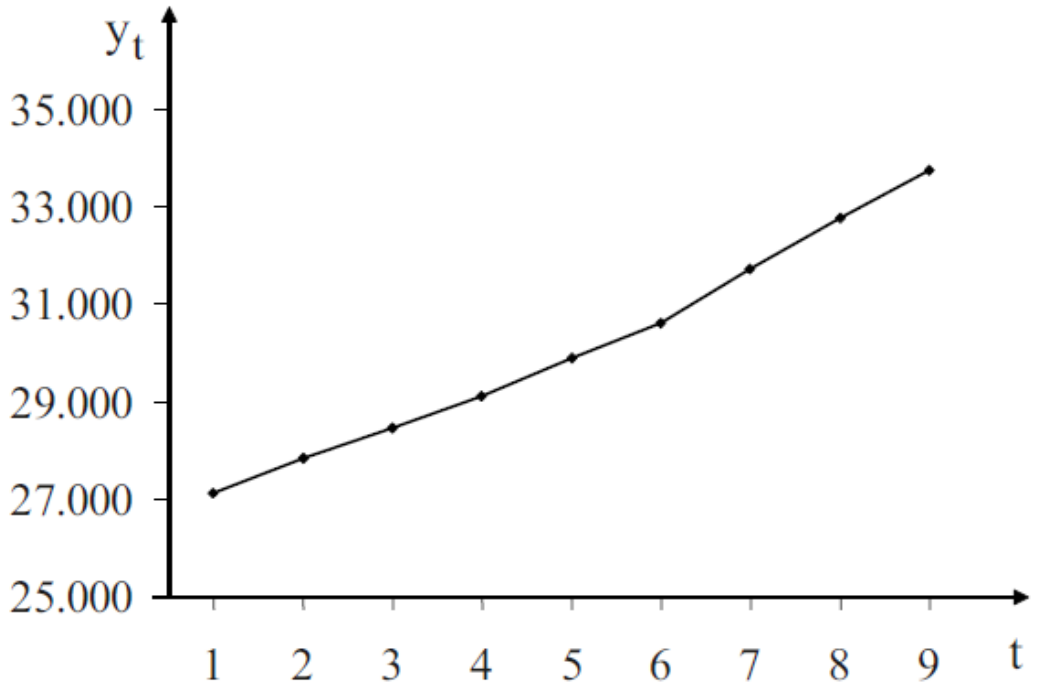
$$\log y_t = \log g_t + \log s_t + \log u_t$$

Beispiel: Linearer Trend bei Jahresdaten

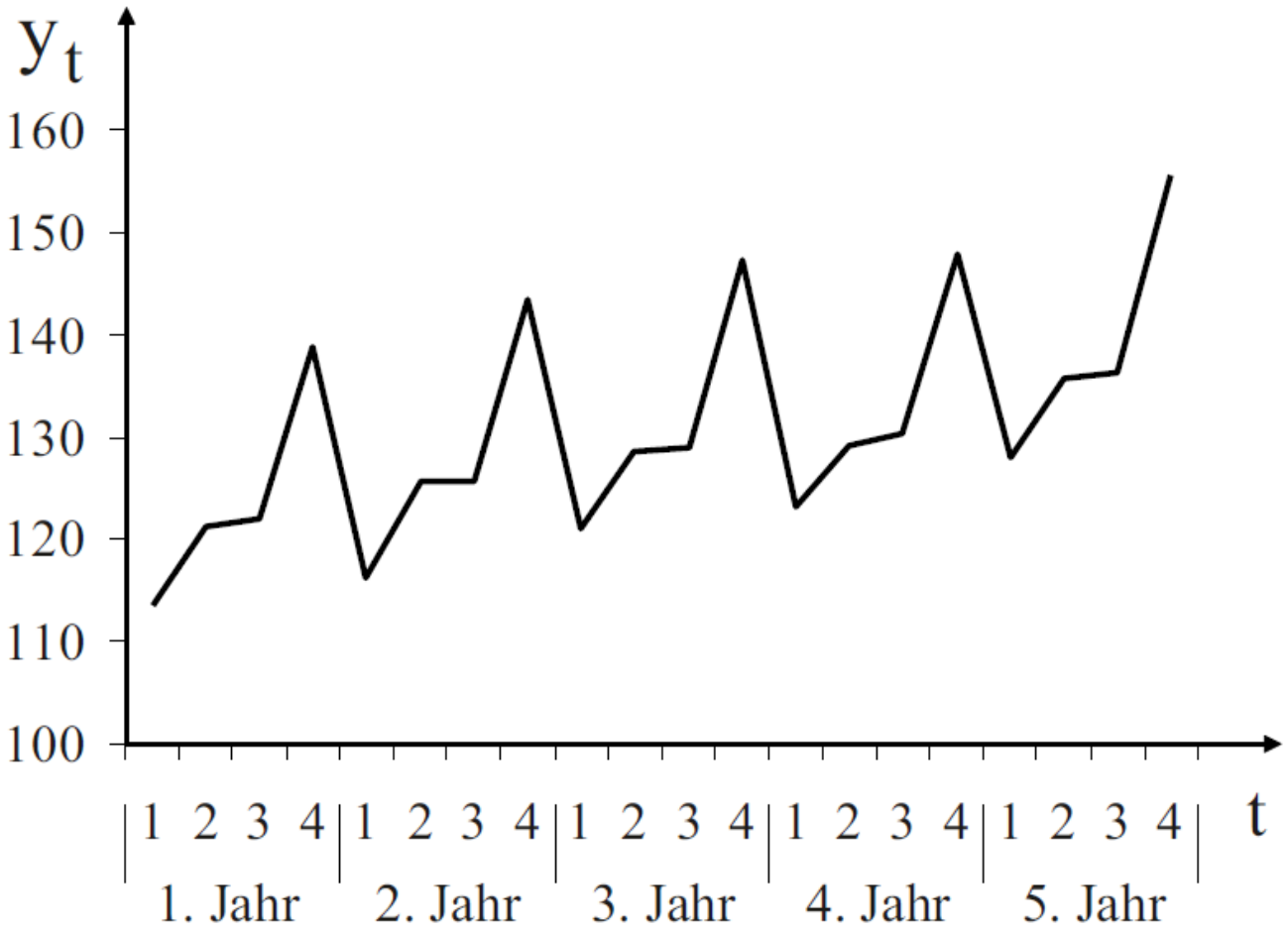
Daten

Jahr t	Bestand an Kraftfahrzeugen (in 1000)
1	27.116
2	27.858
3	28.452
4	29.122
5	29.905
6	30.618
7	31.748
8	32.762
9	33.764

Grafische Darstellung



Beispiel: Saisonale Schwankungen bei unterjährigen Daten



Trendfunktion

- Wenn eine Zeitreihe in einem Zeitintervall keinen Strukturbruch ausweist, ist es oft möglich, ihre Entwicklungstendenz durch eine Funktion der Zeit t zu modellieren.
- Eine solche Funktion, die mittels der Methode der kleinsten Quadrate geschätzt werden kann, heißt **Trendfunktion**:

$$\hat{y}_t = f(t)$$

- Es handelt sich dabei um eine Regressionsfunktion, in der die Zeit t als erklärende Variable auftritt.
- Trendbestimmung beim einfachen Grundmodell:

$$y_t = m_t + u_t$$

- Lineare Trendfunktion (bei annähernd konstanten Zuwächsen):

$$m_t = \hat{y}_t = a + b \cdot t$$

bzw.

$$y_t = a + b \cdot t + u_t$$

Ermittlung der Regressionskoeffizienten (I)

- Kleinst-Quadrate-Kriterium:

$$Q(a,b) = \sum_{t=1}^{n} u_t^2 = \sum_{t=1}^{n} (y_t - \hat{y}_t)^2 = \sum_{t=1}^{n} (y_t - a - b \cdot t)^2 \rightarrow \min_{a,b}$$

- Notwendige Bedingung für ein Minimum:

$$\frac{\partial Q(a,b)}{\partial a} = \sum_{t=1}^{n} 2 \cdot (y_t - a - b \cdot t) \cdot (-1) = 0$$

$$\frac{\partial Q(a,b)}{\partial b} = \sum_{t=1}^{n} 2 \cdot (y_t - a - b \cdot t) \cdot (-t) = 0$$

Ermittlung der Regressionskoeffizienten (II)

- Erste Normalgleichung:

$$\sum_{t=1}^{n} 2 \cdot (y_t - a - b \cdot t) \cdot (-1) = 0$$

$$\Leftrightarrow -\sum_{t=1}^{n} 2 \cdot y_t + \sum_{t=1}^{n} 2 \cdot a + \sum_{t=1}^{n} 2 \cdot b \cdot t = 0$$

$$\Leftrightarrow \sum_{t=1}^{n} 2 \cdot a + \sum_{t=1}^{n} 2 \cdot b \cdot t = \sum_{t=1}^{n} 2 \cdot y_t$$

$$\Leftrightarrow a \cdot n + b \cdot \sum_{t=1}^{n} t = \sum_{t=1}^{n} y_t$$

- Auflösen nach $\hat{a}$ ⇒ Kleinst-Quadrate-Schätzer für das absolute Glied:

$$\hat{a} = \frac{1}{n} \cdot \sum_{t=1}^{n} y_t - \hat{b} \cdot \frac{1}{n} \cdot \sum_{t=1}^{n} t$$

Ermittlung der Regressionskoeffizienten (III)

- Zweite Normalgleichung:

$$\sum_{t=1}^{n} 2 \cdot (y_t - a - b \cdot t) \cdot (-t) = 0$$

$$\Leftrightarrow -\sum_{t=1}^{n} 2 \cdot y_t \cdot t + \sum_{t=1}^{n} 2 \cdot a \cdot t + \sum_{t=1}^{n} 2 \cdot b \cdot t^2 = 0$$

$$\Leftrightarrow \sum_{t=1}^{n} 2 \cdot a \cdot t + \sum_{t=1}^{n} 2 \cdot b \cdot t^2 = \sum_{t=1}^{n} 2 \cdot y_t \cdot t$$

$$\Leftrightarrow a \cdot \sum_{t=1}^{n} t + b \cdot \sum_{t=1}^{n} t^2 = \sum_{t=1}^{n} y_t \cdot t$$

- Einsetzen von $\hat{a}$ und Auflösen nach $\hat{b}$ ⇒ Kleinst-Quadrate-Schätzer für das Steigungsmaß:

$$\hat{b} = \frac{n \cdot \sum_{t=1}^{n} t \cdot y_t - \sum_{t=1}^{n} y_t \cdot \sum_{t=1}^{n} t}{n \cdot \sum_{t=1}^{n} t^2 - (\sum_{t=1}^{n} t)^2}$$

Beispiel: Zuwachs Kraftfahrzeuge (I)

Wie aus der unten stehenden Tabelle hervorgeht, wächst der Bestand an Kraftfahrzeugen (in 1000 Stück) relativ gleichmäßig an, wobei die jährlichen Zuwächse nicht zu stark variieren. Es liegt daher nahe, die Trendkomponente der Zeitreihe durch eine lineare Transformation nachzubilden.

	t	y_t (Kfz-Bestand)	$t \cdot y_t$	t^2
	1	27.116	27.116	1
	2	27.858	55.716	4
	3	28.452	85.356	9
	4	29.122	116.488	16
	5	29.905	149.525	25
	6	30.618	183.708	36
	7	31.748	222.236	49
	8	32.762	262.096	64
	9	33.764	303.876	81
$\sum$	45	271.345	1.406.117	285

Beispiel: Zuwachs Kraftfahrzeuge (II)

- **Steigungsmaß**:

$$\hat{b} = \frac{9 \cdot \sum_{t=1}^{9} t \cdot y_t - \sum_{t=1}^{9} y_t \cdot \sum_{t=1}^{9} t}{9 \cdot \sum_{t=1}^{9} t^2 - (\sum_{t=1}^{9} t)^2} = \frac{9 \cdot 1.406.117 - 271.345 \cdot 45}{9 \cdot 285 - (45)^2} = 823{,}2$$

  - Interpretation: Der Kraftfahrzeugbestand hat sich im Stützzeitraum im Mittel um 823.200 Stück pro Jahr erhöht.

- **Absolutes Glied**:

$$\hat{a} = \frac{1}{9} \cdot \sum_{t=1}^{9} y_t - \hat{b} \cdot \frac{1}{9} \cdot \sum_{t=1}^{9} t = \frac{1}{9} \cdot 271.345 - 823{,}2 \cdot \frac{1}{9} \cdot 45 = 26.033{,}444$$

- **Trendfunktion**:

$$m_t = \hat{y}_t = 26.033{,}444 + 823{,}2 \cdot t$$

Beispiel: Zuwachs Kraftfahrzeuge (III)

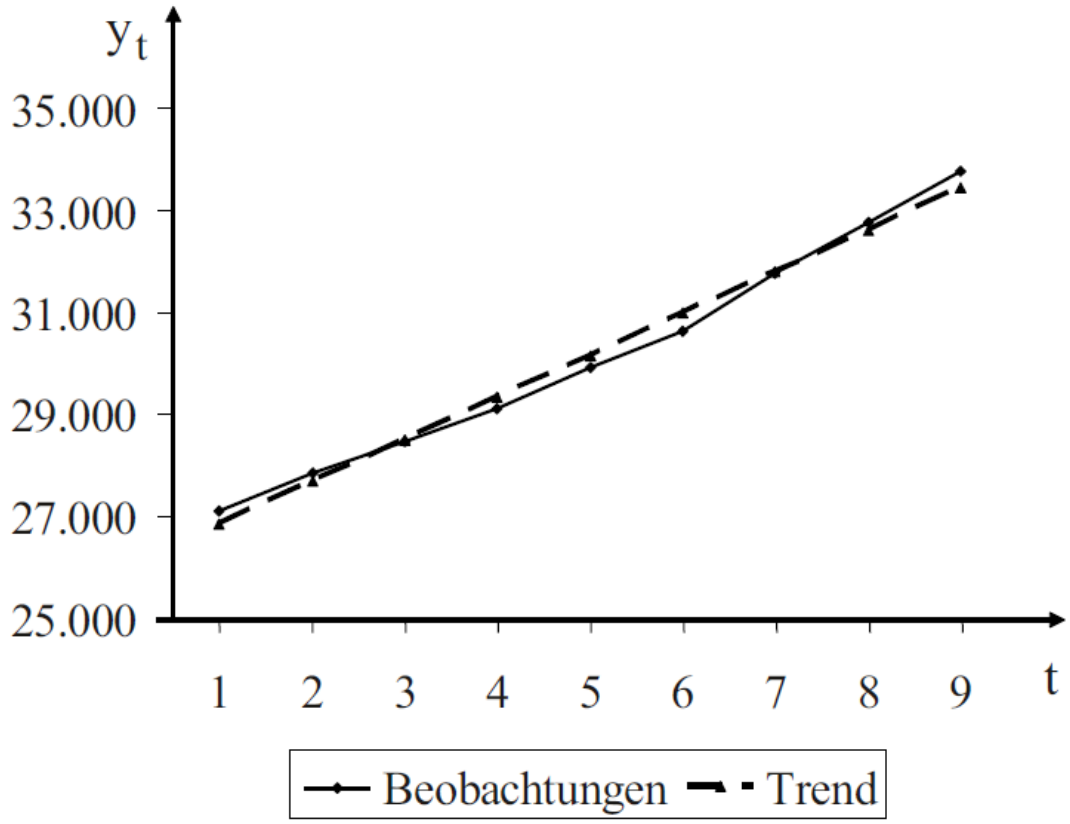
Daten

t	$\hat{y}_t$
1	26.856,644
2	27.679,844
3	28.503,044
4	29.326,244
5	30.149,444
6	30.972,644
7	31.795,844
8	32.619,044
9	33.442,244

z. B.

$$\hat{y}_1 = 26.033{,}444 + 823{,}2 \cdot 1$$

Grafische Darstellung



Beispiel: Zuwachs Kraftfahrzeuge (IV)

- Die Anpassungsgüte der Trendfunktion kann mithilfe des Bestimmtheitsmaßes beurteilt werden. Das Bestimmtheitsmaß gibt den Anteil der Varianz wieder, der auf den Trend zurückgeführt werden kann.
- Zur Berechnung wird das arithmetische Mittel des Kfz-Bestandes benötigt:

$$\bar{y} = \frac{1}{9} \cdot \sum_{t=1}^{9} y_t = \frac{1}{9} \cdot 271.345 = 30.139{,}444$$

t	y_t	$\hat{y}_t = 26.033{,}444 + 823{,}2 \cdot t$	$(y_t - \underbrace{30.149{,}444}_{=\bar{y}})^2$	$(\hat{y}_t - \underbrace{30.149{,}444}_{=\bar{y}})^2$
1	27.116	26.856,644	9.201.782,501	10.842.531,840
2	27.858	27.679,844	5.250.715,605	6.098.924,160
3	28.452	28.503,044	2.881.316,133	2.710.632,960
4	29.122	29.326,244	1.055.641,173	677.658,240
5	29.905	30.149,444	59.752,869	0,000
6	30.618	30.972,644	219.544,725	677.658,240
7	31.748	31.795,844	2.555.381,285	2.710.632,960
8	32.762	32.619,044	6.825.448,853	6.098.924,160
9	33.764	33.442,244	13.065.015,077	10.842.531,840
$\sum$	271.345	271.345,000	41.114.598,221	40.659.494,400

Beispiel: Zuwachs Kraftfahrzeuge (V)

- Der Anteil der **erklärten Varianz** beträgt:

$$R^2 = \frac{\sum_{t=1}^{n}(\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} = \frac{40.659.494{,}400}{41.114.598{,}221} = 0{,}989$$

- **Interpretation**: Durch den Regressionsansatz werden 98,9% der Varianz des Kfz-Bestandes erklärt.

- **Exploration** des Trends für $t = 10$:

$$\hat{y}_{10} = 26.033{,}444 + 823{,}2 \cdot 10 = 34.265{,}444$$

- Unter der Annahme eines linearen Trends lässt sich die Trendexploration als **Prognose** für den Kfz-Bestand verwenden.

Exponentialtrend (I)

- Anwendung bei Wachstumsvorgängen, bei denen eine Variable in einem Zeitraum überproportional steigt.
- **Exponentielle Trendfunktion**:

$$m_t = a \cdot b^t$$

- Beim Exponentialtrend verändern sich die Trendwerte von Periode zu Periode um eine **konstante Wachstumsrate** $b - 1$.
- Der konstante Faktor a gibt den Trendwert einer Zeitreihe für die Periode vor Beginn des Stützzeitraumes wieder.
- Linearisierung der exponentiellen Trendfunktion:

$$\ln m_t = \ln a + t \cdot \ln b$$

- Kleinst-Quadrate-Kriterium:

$$Q(a,b) = \sum_{t=1}^{n} (\ln y_t - \ln a - t \cdot \ln b)^2 \longrightarrow \min_{a,b}$$

Exponentialtrend (II)

- Kleinst-Quadrate-Schätzer:

$$\widehat{\ln a} = \frac{1}{n} \cdot \sum_{t=1}^{n} \ln y_t - \widehat{\ln b} \cdot \frac{1}{n} \cdot \sum_{t=1}^{n} t$$

$$\widehat{\ln b} = \frac{n \cdot \sum_{t=1}^{n} t \cdot \ln y_t - \sum_{t=1}^{n} \ln y_t \cdot \sum_{t=1}^{n} t}{n \cdot \sum_{t=1}^{n} t^2 - (\sum_{t=1}^{n} t)^2}$$

- Trendkoeffizienten:

$$\hat{a} = e^{\widehat{\ln a}}$$

$$\hat{b} = e^{\widehat{\ln b}}$$

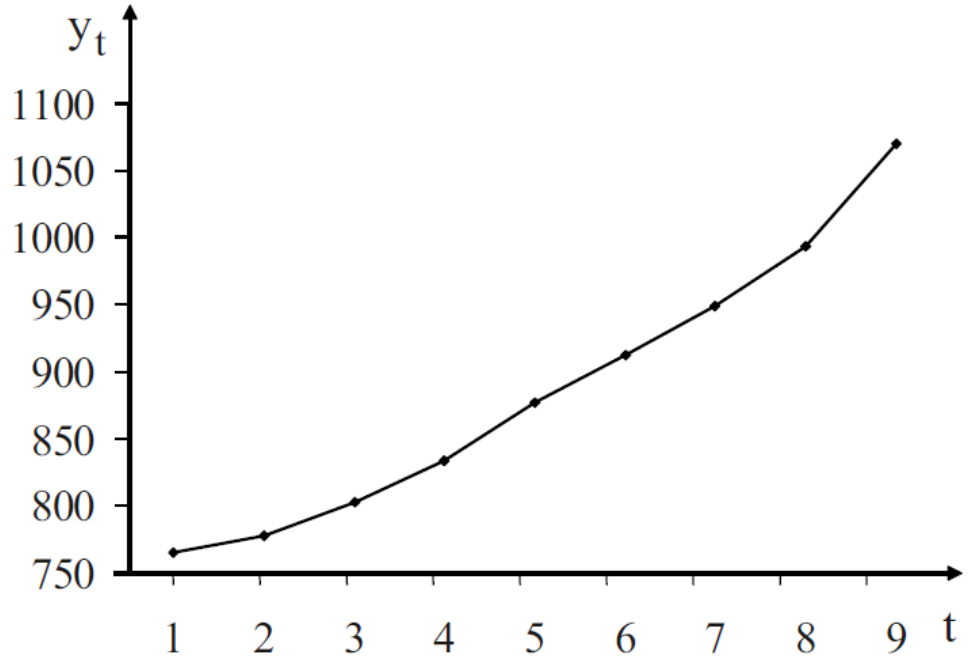
Beispiel: Bruttolohn- und Gehaltssumme (I)

Die Bruttolohn- und Gehaltssumme aus unselbständiger Arbeit ohne Arbeitgeberbeiträge zur Sozialversicherung ist in dem Zeitraum von neun Jahren überproportional angestiegen. Aus diesem Grund lässt sich die zeitliche Entwicklung nicht durch eine lineare Trendfunktion beschreiben. Vielmehr kann der Trend hier unter Verwendung einer konstanten Wachstumsrate modelliert werden (exponentieller Trend).

Daten

t	y_t
1	764,44
2	777,42
3	802,93
4	833,78
5	876,63
6	912,81
7	948,85
8	993,19
9	1070,10

Grafische Darstellung



Beispiel: Bruttolohn- und Gehaltssumme (II)

- Arbeitstabelle zur Berechnung der Trendkoeffizienten (wir arbeiten hier ausnahmsweise mit vier Nachkommastellen, da ansonsten Rundungsabweichungen zu einem Bestimmtheitsmaß > 1 führen können):

	t	t^2	y_t	$\ln y_t$	$t \cdot \ln y_t$
	1	1	764,44	6,6391	6,6391
	2	4	777,42	6,6560	13,3120
	3	9	802,93	6,6883	20,0648
	4	16	833,78	6,7260	26,9039
	5	25	876,63	6,7761	33,8804
	6	36	912,81	6,8165	40,8992
	7	49	948,85	6,8553	47,9868
	8	64	993,19	6,9009	55,2074
	9	81	1.070,10	6,9755	62,7796
Σ	45	285		61,0337	307,6732

Beispiel: Bruttolohn- und Gehaltssumme (III)

- **Steigungsmaß**:

$$\widehat{\ln b} = \frac{9 \cdot \sum_{t=1}^{9} t \cdot \ln y_t - \sum_{t=1}^{9} \ln y_t \cdot \sum_{t=1}^{9} t}{9 \cdot \sum_{t=1}^{9} t^2 - (\sum_{t=1}^{9} t)^2} = \frac{9 \cdot 307{,}6732 - 61{,}0337 \cdot 45}{9 \cdot 285 - 45^2} = 0{,}0417$$

$$\hat{b} = e^{\widehat{\ln b}} = e^{0{,}0417} = 1{,}0426$$

- Trendgemäß wächst die Bruttolohn- und Gehaltssumme jährlich mit einer Wachstumsrate von knapp 4.3%.

- **Absolutes Glied**:

$$\widehat{\ln a} = \frac{1}{9} \cdot \sum_{t=1}^{9} \ln y_t - \widehat{\ln b} \cdot \frac{1}{9} \cdot \sum_{t=1}^{9} t = \frac{1}{9} \cdot 61{,}0337 - 0{,}0417 \cdot \frac{1}{9} \cdot 45 = 6{,}5730$$

$$\hat{a} = e^{\widehat{\ln a}} = e^{6{,}5730} = 715{,}5132$$

Beispiel: Bruttolohn- und Gehaltssumme (IV)

- Trendfunktion:

$$\hat{y}_t = 715{,}5132 \cdot 1{,}0426^t$$

- Das Bestimmtheitsmaß lässt sich aus folgender Arbeitstabelle bestimmen:

	t	$\ln y_t$	$\ln \hat{y}_t$	$(\ln y_t)^2$	$(\ln \hat{y}_t)^2$
	1	6,6391	6,6147	44,0776	43,7453
	2	6,6560	6,6564	44,3023	44,3077
	3	6,6883	6,6982	44,7334	44,8659
	4	6,7260	6,7399	45,2391	45,4263
	5	6,7761	6,7816	45,9155	45,9901
	6	6,8165	6,8233	46,4647	46,5574
	7	6,8553	6,8650	46,9951	47,1282
	8	6,9009	6,9067	47,6224	47,7025
	9	6,9755	6,9485	48,6576	48,2817
Σ	45	61,0337	61,0343	414,0077	414,0141

Beispiel: Bruttolohn- und Gehaltssumme (V)

- Der Anteil der **erklärten Varianz** beträgt:

$$R^2 = \frac{n \cdot \sum_{t=1}^{n} (\ln \hat{y}_t)^2 - (\sum_{t=1}^{n} \ln \hat{y}_t)^2}{n \cdot \sum_{t=1}^{n} (\ln y_t)^2 - (\sum_{t=1}^{n} \ln y_t)^2} = \frac{9 \cdot 414{,}134 - 61{,}043^2}{9 \cdot 414{,}011 - 61{,}034^2} = 0{,}9837$$

- **Interpretation**: 98,37% der Varianz der Bruttolohn- und Gehaltssumme ist durch einen exponentiellen Trend determiniert.

- **Prognose** für $t = 10$:

$$\hat{y}_{10} = 715{,}5132 \cdot 1{,}0426^{10} = 1.085{,}913$$

Methode der gleitenden Durchschnitte (I)

- Bei vielen Zeitreihen lässt sich keine angemessene Trendfunktion mit wenigen Parametern bestimmen.
- Bei saisonalen Zeitreihen ist die Restkomponente des Trendmodells nicht frei von systematischen Einflüssen $\Longrightarrow$ Voraussetzungen für Anwendung der Methode der kleinsten Quadrate nicht erfüllt.
- In diesen Fällen empfiehlt es sich, die glatte Komponente mit einer flexibleren Methode zu ermitteln, die weniger strengen Annahmen unterliegt.
- Mit der **Methode der gleitenden Durchschnitte** wird eine Zeitreihe geglättet, indem man sukzessive Mittelwerte aus einer feststehenden Anzahl benachbarter Zeitreihenwerte ermittelt (gleiche Länge der Stützbereiche), die jeweils der Mitte des betreffenden Zeitintervalls zugeordnet werden.
- Die Folge der Durchschnitte wird als gleitend bezeichnet, weil jeweils der älteste Zeitreihenwert durch den Zeitreihenwert ersetzt wird, der unmittelbar am rechten Rand außerhalb des Stützbereichs liegt.
- Auf diese Art und Weise „gleiten“ die Durchschnitte quasi entlang einer Zeitachse.
- Der Glättungseffekt ergibt sich dadurch, dass extreme Zeitreihenwerte abgewichtet werden (Gewicht < 1).

Methode der gleitenden Durchschnitte (II)

- Die Ordnung eines gleitenden Durchschnitts ist durch die Anzahl p der darin eingehenden Zeitreihenwerte gegeben.
- Man spricht von einem gleitenden Durchschnitt der Ordnung p oder einem p-gliedrigen Durchschnitt $\bar{y}_t^p$.

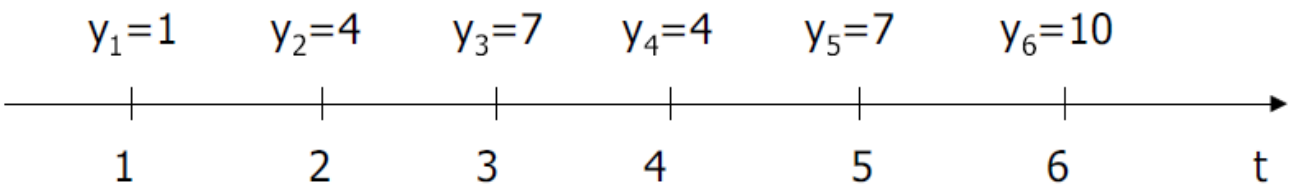
Ungerade Ordnungszahl $p \Longrightarrow$ einfache p-gliedrige gleitende Durchschnitte

- Ein p-gliedriger gleitender Durchschnitt setzt sich aus $p = 2q + 1$ Beobachtungswerten zusammen, wobei jeweils q Werte vor und nach seiner zeitlichen Zentrierung liegen:

$$\bar{y}_t^p = \frac{1}{p} \cdot \left(y_{t-q} + \cdots + y_{t-1} + y_t + y_{t+1} + \cdots + y_{t+q}\right)$$

- In den Rändern lassen sich jeweils q gleitende Durchschnitt nicht berechnen.

Beispiel: Gleitendender Durchschnitt bei ungerader Ordnung ($p = 3$)



- Drei-gliedrige gleitende Durchschnitte: $\bar{y}_t^3 = \frac{1}{3} \cdot (y_{t-1} + y_t + y_{t+1})$
  - $t = 2$: $\bar{y}_2^3 = \frac{1}{3} \cdot (y_{2-1} + y_2 + y_{2+1}) = \frac{1}{3} \cdot (1 + 4 + 7) = 4$
  - $t = 3$: $\bar{y}_3^3 = \frac{1}{3} \cdot (y_2 + y_3 + y_4) = \frac{1}{3} \cdot (4 + 7 + 4) = 5$
  - $t = 4$: $\bar{y}_4^3 = \frac{1}{3} \cdot (y_3 + y_4 + y_5) = \frac{1}{3} \cdot (7 + 4 + 7) = 6$
  - $t = 5$: $\bar{y}_5^3 = \frac{1}{3} \cdot (y_4 + y_5 + y_6) = \frac{1}{3} \cdot (4 + 7 + 10) = 7$
- Die Reihe der gleitenden Durchschnitte ist an den Rändern gegenüber der originären Zeitreihe um jeweils einen Wert vermindert ($3 = 2q + 1 \Longrightarrow q = 1$).

Beispiel: Auftragseingänge (I)

- Das Niveau der Auftragseingänge im Verarbeitenden Gewerbe wird vom Statistischen Bundesamt kalendermonatlich über einen Index gemessen.
- Quartalsmäßig hat sich der Index in einem Zeitraum von drei Jahren wie folgt entwickelt:

Jahr	I. Quartal	II. Quartal	III. Quartal	IV. Quartal
1		106,6	108,6	115,9
2	122,1	123,8	117,8	125,4
3	130,7	124,9	128,5	133,7
4	137,7			

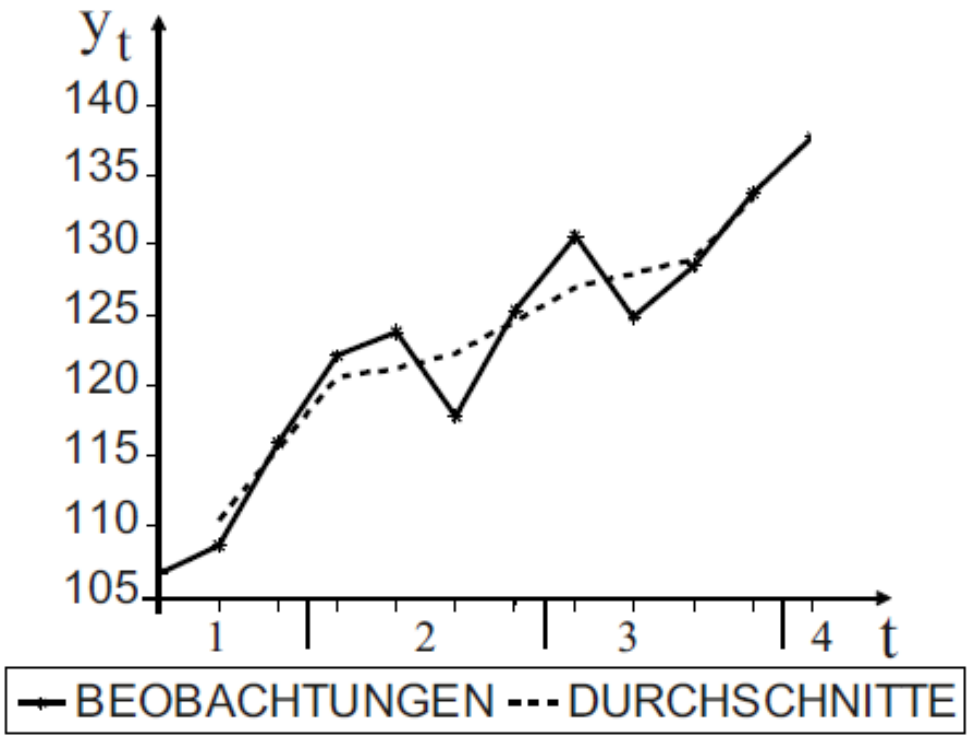
- Bei einer Glättung der Zeitreihe unter Verwendung eines drei-gliedrigen gleitenden Durchschnitts bleiben die erste und letzte Periode des Beobachtungszeitraums unbesetzt.
- Die ersten beiden gleitenden Durchschnitte lauten:
  - $\bar{y}^3_{1/III} = \frac{1}{3} \cdot \left(y_{1/II} + y_{1/III} + y_{1/IV}\right) = \frac{1}{3} \cdot (106{,}6 + 108{,}6 + 115{,}9) = 110{,}4$
  - $\bar{y}^3_{1/IV} = \frac{1}{3} \cdot \left(y_{1/III} + y_{1/IV} + y_{2/I}\right) = \frac{1}{3} \cdot (108{,}6 + 115{,}9 + 122{,}1) = 115{,}5$

Beispiel: Auftragseingänge (II)

Daten

t	y_t	$\bar{y}_t^3$
1/II	106,6	
1/III	108,6	110,4
1/IV	115,9	115,5
2/I	122,1	120,6
2/II	123,8	121,2
2/III	117,8	122,3
2/IV	125,4	124,6
3/I	130,7	127,0
3/II	124,9	128,0
3/III	128,5	129,0
3/IV	133,7	133,3
4/I	137,7	

Grafische Darstellung



Zentrierte p-gliedrige gleitende Durchschnitte (p gerade) (I)

- Wenn eine Zeitreihe saisonale Schwankungen aufweist, wird der Glättungseffekt im Allgemeinen durch einfache gleitende Durchschnitte ungerader Ordnung verzerrt.
- Der Grund liegt darin, dass ein bestimmter Jahresabschnitt (z.B. Monat, Quartal) hierbei in einem gleitenden Durchschnittwert unberücksichtigt bleibt.
- Um saisonale Schwankungen auszuschalten, müssen die gleitenden Durchschnitte so konzipiert werden, dass ihre Ordnung mit der Zykluslänge übereinstimmt.
- Bei **Quartalsdaten** mit sich regelmäßig wiederholenden saisonalen Schwankungen kommen für die glatte Komponente daher vorzugsweise **vier-gliedrige** gleitende Durchschnitte, bei **Monatsdaten** mit diesen Eigenschaften **zwölf-gliedrige** gleitende Durchschnitte in Betracht.

→ **Gleitende Durchschnitte mit gerader Ordnung**!

Zentrierte p-gliedrige gleitende Durchschnitte (p gerade) (II)

- Ein gleitender Durchschnittswert gerader Ordnung lässt sich jedoch keiner Zeiteinheit eindeutig zuordnen, da er auf der Zeitachse genau zwischen den beiden mittleren Perioden oder Zeitpunkten liegt.
- Um dies zu vermeiden, zieht man $p + 1$ Zeitreihenwerte zur Berechnung eines gleitenden Durchschnitts gerader Ordnung heran und gewichtet die beiden äußeren Zeitreihenwerte mit dem Faktor $^1/_2$:

$$\bar{y}_t^p = \frac{1}{p} \cdot \left(\frac{1}{2} \cdot y_{t-q} + \cdots + y_{t-1} + y_t + y_{t+1} + \cdots + \frac{1}{2} \cdot y_{t+q} \right)$$

- An den Rändern lassen sich jeweils $q = {}^p/_2$ Durchschnittswerte nicht berechnen.
- Auf diese Weise werden die gleitenden Durchschnitte zentriert.

→ Bei Zeitreihen mit saisonalen Schwankungen erfolgt die Glättung daher mittels **zentrierter gleitender Durchschnitte**.

Beispiel: Löhne und Gehälter (I)

- Die Löhne und Gehälter je Beschäftigten weisen ein klares Saisonmuster auf. Im I. Quartal eines Jahres liegt der Tiefstand und nach den etwa gleichwertigen beiden mittleren Quartalen wird im IV. Quartal das saisonale Hoch erreicht.
- Die langfristig steigende Tendenz dieser Zeitreihe kann daher am besten durch vier-gliedrige gleitende Durchschnitte beschrieben werden:

Jahr	I. Quartal	II. Quartal	III. Quartal	IV. Quartal
1	113,6	121,3	122,0	138,8
2	116,3	125,7	125,7	143,5
3	121,1	128,6	129,0	147,3
4	123,2	129,2	130,3	147,9
5	128,0	135,7	136,2	155,5

- Berechnung vier-gliedriger gleitender Durchschnitte für das III. Quartal im ersten Jahr:

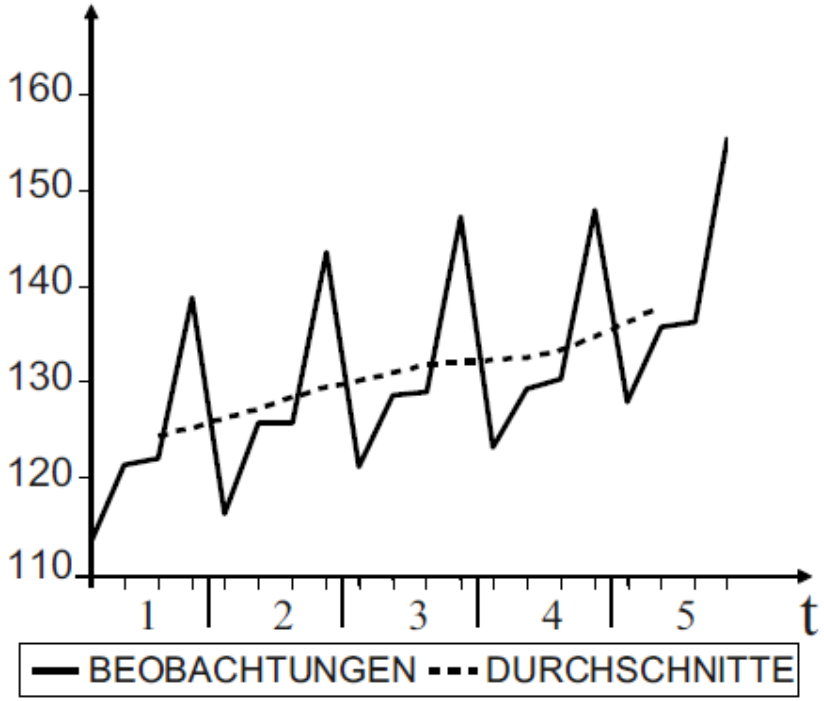
$$\bar{y}^4_{1/III} = \frac{1}{4} \cdot \left(\frac{1}{2} \cdot y_{1/I} + y_{1/II} + y_{1/III} + y_{1/IV} + \frac{1}{2} \cdot y_{2/I} \right)$$
$$= \frac{1}{4} \cdot \left(\frac{1}{2} \cdot 113{,}6 + 121{,}3 + 122{,}0 + 138{,}8 + \frac{1}{2} \cdot 116{,}3 \right) = 124{,}3$$

Beispiel: Löhne und Gehälter (II)

Daten

t	y_t	$\bar{y}_t^4$
1/I	113,6	
1/II	121,3	
1/III	122,0	124,3
1/IV	138,8	125,2
2/I	116,3	126,2
2/II	125,7	127,2
2/III	125,7	128,4
2/IV	143,5	129,4
3/I	121,1	130,1
3/II	128,6	131,0
3/III	129,0	131,8
3/IV	147,3	132,1
4/I	123,2	132,3
4/II	129,2	132,6
4/III	130,3	133,3
4/IV	147,9	134,7
5/I	128,0	136,2
5/II	135,7	137,9
5/III	136,2	
5/IV	155,5	

Grafische Darstellung



Saisonkomponente und Saisonbereinigung (I)

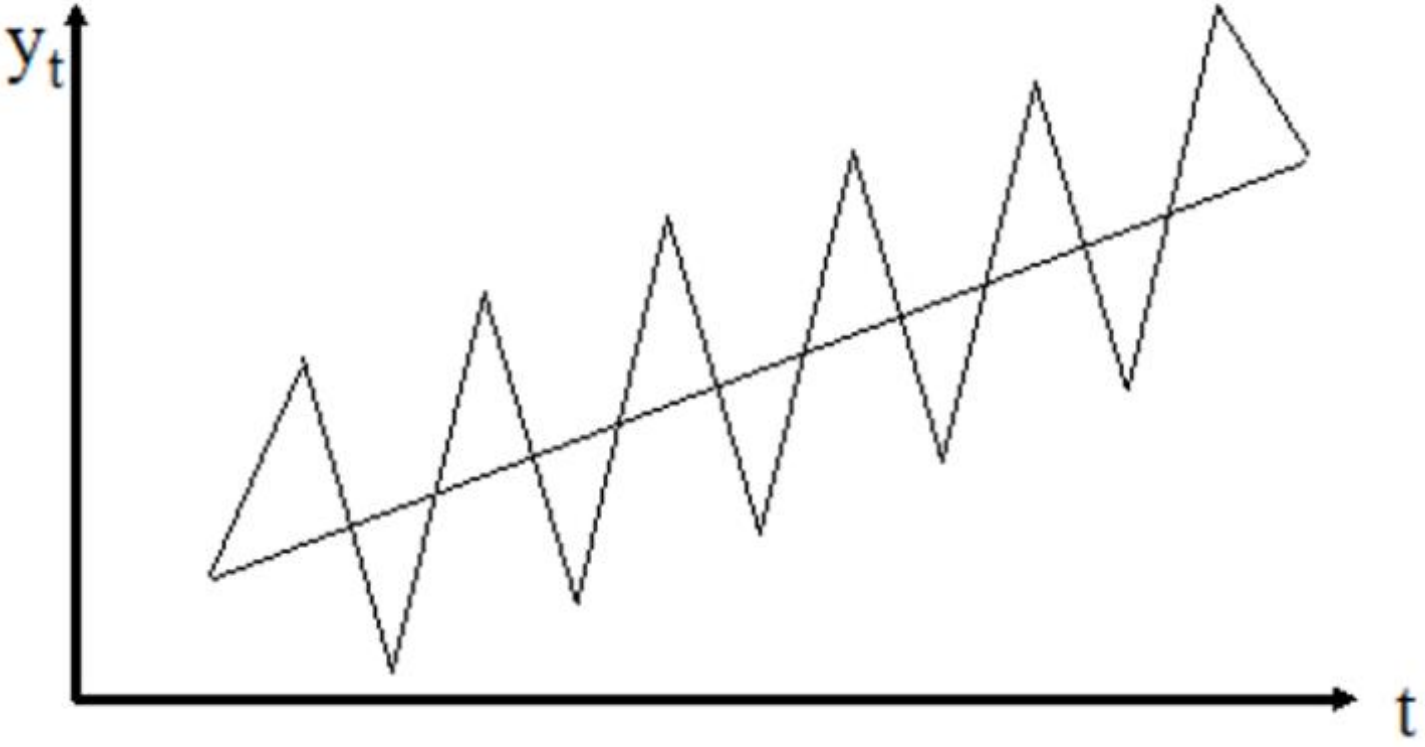
- Viele ökonomische Zeitreihen enthalten saisonale Schwankungen, die eine Einschätzung der Arbeitsmarktsituation und konjunkturellen Lage kaum möglich werden lassen (z.B. starke Witterungsabhängigkeit des Baugewerbes).
- Aus diesem Grund werden **saisonbereinigte Zeitreihen** verwendet, die sich aus den originären Daten nach Ausschaltung der Saisonkomponente s_t ergeben.
- Zuweilen ist auch die **Saisonkomponente** einer Zeitreihe selbst von Interesse.
- Zur Bestimmung der Saisonkomponente einer unterjährigen Zeitreihe (y_t) schaltet man die glatte Komponente vorab aus.
- Beim additiven Modell liegen dann die trendbereinigten (und konjunkturbereinigten) Zeitreihenwerte vor, die nur noch die Saison- und Restkomponente enthalten (mit $g_t = m_t + c_t$):

$$d_t = y_t - g_t = s_t + u_t$$

- Voraussetzung hierfür ist, dass die saisonalen Ausschläge unabhängig vom Trend der Zeitreihe sind $\Longrightarrow$ Saisonausschläge mit konstanter Amplitude

Saisonkomponente und Saisonbereinigung (II)

- Saisonausschläge mit konstanter Amplitude



Bestimmung der Saisonkomponente (I)

- **Trendbereinigte Zeitreihe** mit Jahr i und Phase j (z.B. Monat, Quartal):

$$d_{ij} = y_{ij} - g_{ij} = s_{ij} + u_{ij}, \qquad i = 1, 2, \dots, k_j; j = 1, 2, \dots, p$$

k_j: Anzahl der Jahre für die Phase j

p: Anzahl der Phasen (bei Quartalsdaten $p = 4$, bei Monatsdaten $p = 12$)

Unnormierte Saisonkomponente

- Nach Ausschaltung der glatten Komponente wird zunächst eine unnormierte Saisonkomponente s_j^* gebildet:

$$s_j^* = \frac{1}{k_j} \cdot \sum_{i=1}^{k_j} d_{ij}$$

- Summe der trendbereinigten Werte eines Jahresteils j wird durch die Anzahl k_j der verfügbaren Werte geteilt.

Bestimmung der Saisonkomponente (II)

Normierte Saisonkomponente

- Die unnormierten Saisonkomponenten summieren sich allgemein nicht zu null.
- Diese Tatsache kann die Interpretation einer Phase als saisonal unter- oder überdurchschnittlich erschweren.
- Für die Saisonkomponente s_j fordert man daher:

$$\sum_{j=1}^{p} s_j = 0 \qquad \text{(Normierung)}$$

- Normierte Saisonkomponente:

$$s_j = s_j^* - \bar{d} \quad \text{mit} \quad \bar{d} = \frac{1}{p} \cdot \sum_{j=1}^{p} s_j^*$$

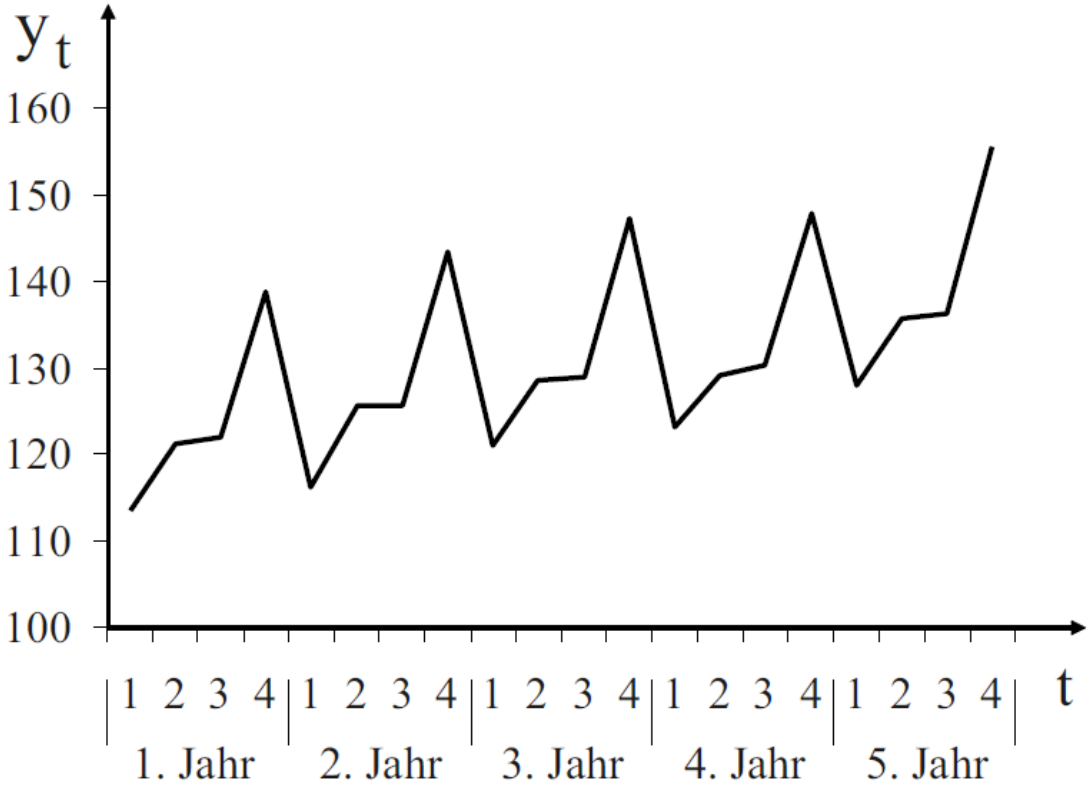
$\bar{d}$: Durchs. jährliche Abweichung der unnormierten Saisonkomponente von null

- **Saisonbereinigte Zeitreihe**:

$$y_{ij}^* = y_{ij} - s_j$$

Beispiel (Fortsetzung): Löhne und Gehälter (I)

Im Zeitreihendiagramm der Lohn- und Gehaltseinkommen je Beschäftigten ist ein klares Saisonmuster erkennbar. Der Index weist jahreszeitlich bedingt jeweils im ersten Quartal eines Jahres einen Tiefstand und im vierten Quartal ein Hoch aus. Außerdem sind keinerlei Anhaltspunkte dafür erkennbar, dass die saisonalen Ausschläge mit wachsendem Trend zunehmen. Daher lässt sich eine Zeitreihenzerlegung adäquat auf der Grundlage des additiven Modells vornehmen.



Beispiel (Fortsetzung): Löhne und Gehälter (II)

Jahr i	Quartal (j)	y_{ij}	$g_{ij} = \bar{y}^4_{ij}$	$d_{ij} = y_{ij} - g_{ij}$
1	I. (1)	113,6		
1	II. (2)	121,3		
1	III. (3)	122,0	124,3	-2,3
1	IV. (4)	138,8	125,2	13,6
2	I. (1)	116,3	126,2	-9,9
2	II. (2)	125,7	127,2	-1,5
2	III. (3)	125,7	128,4	-2,7
2	IV. (4)	143,5	129,4	14,1
3	I. (1)	121,1	130,1	-9,0
3	II. (2)	128,6	131,0	-2,4
3	III. (3)	129,0	131,8	-2,8
3	IV. (4)	147,3	132,1	15,2
4	I. (1)	123,2	132,3	-9,1
4	II. (2)	129,2	132,6	-3,4
4	III. (3)	130,3	133,3	-3,0
4	IV. (4)	147,9	134,7	13,2
5	I. (1)	128,0	136,2	-8,2
5	II. (2)	135,7	137,9	-2,2
5	III. (3)	136,2		
5	IV. (4)	155,5		

Beispiel (Fortsetzung): Löhne und Gehälter (III)

- Aus den in der Arbeitstabelle berechneten trendbereinigten Werten d_{ij} erhält man die unnormierte Saisonkomponente:

$$s_1^* = \frac{1}{4} \cdot \sum_{i=2}^{5} d_{i1} = \frac{1}{4}(-9{,}9 - 9{,}0 - 9{,}1 - 8{,}2) = -9{,}1$$

$$s_2^* = \frac{1}{4} \cdot \sum_{i=2}^{5} d_{i2} = \frac{1}{4}(-1{,}5 - 2{,}4 - 3{,}4 - 2{,}2) = -2{,}4$$

$$s_3^* = \frac{1}{4} \cdot \sum_{i=2}^{5} d_{i3} = \frac{1}{4}(-2{,}3 - 2{,}7 - 2{,}8 - 3{,}0) = -2{,}7$$

$$s_4^* = \frac{1}{4} \cdot \sum_{i=2}^{5} d_{i4} = \frac{1}{4}(13{,}6 + 14{,}1 + 15{,}2 + 13{,}2) = 14{,}0$$

Beispiel (Fortsetzung): Löhne und Gehälter (IV)

- Durchschnitt der unnormierten Saisonziffern:

$$\bar{d} = \frac{1}{p} \cdot \sum_{j=1}^{p} s_j^* = \frac{1}{4} \cdot (-9{,}1 - 2{,}4 - 2{,}7 + 14{,}0) = \frac{1}{4} \cdot (-0{,}2) = -0{,}05 \approx 0$$

- Da $\bar{d}$ annähernd gleich null ist, kann auf eine Korrektur verzichtet werden.
- Die normierte Saisonkomponente s_j stimmt daher mit der unnormierten Saisonkomponente s_j^* überein:

$$s_1 = s_1^* = -9{,}1$$

$$s_2 = s_2^* = -2{,}4$$

$$s_3 = s_3^* = -2{,}7$$

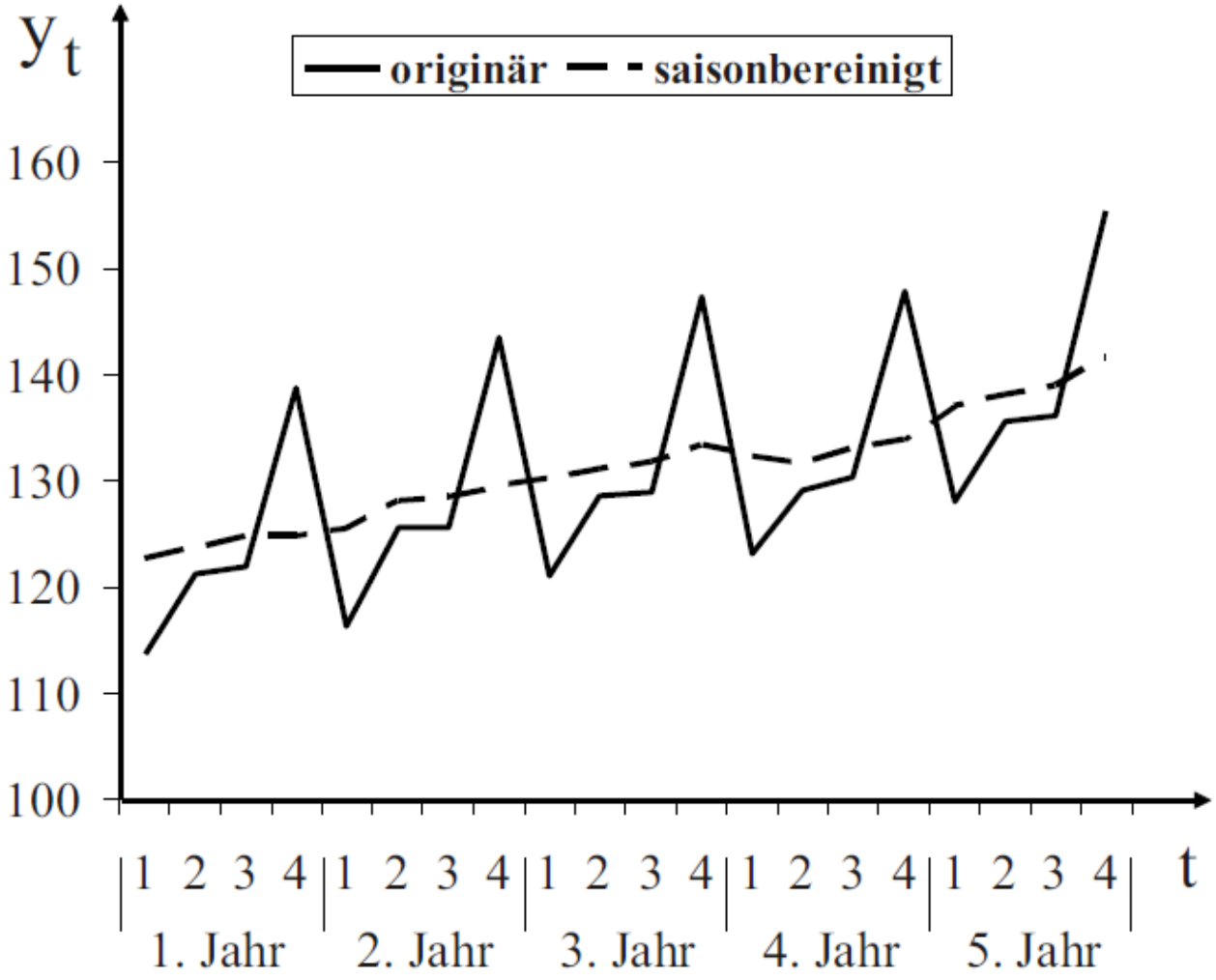
$$s_4 = s_4^* = 14{,}0$$

Beispiel (Fortsetzung): Löhne und Gehälter (V)

Jahr i	Quartal (j)	y_{ij}	s_j	$y^*_{ij} = y_{ij} - s_j$
1	I. (1)	113,6	-9,1	122,7
1	II. (2)	121,3	-2,4	123,7
1	III. (3)	122,0	-2,7	124,7
1	IV. (4)	138,8	14,0	124,8
2	I. (1)	116,3	-9,1	125,4
2	II. (2)	125,7	-2,4	128,1
2	III. (3)	125,7	-2,7	128,4
2	IV. (4)	143,5	14,0	129,5
3	I. (1)	121,1	-9,1	130,2
3	II. (2)	128,6	-2,4	131,0
3	III. (3)	129,0	-2,7	131,7
3	IV. (4)	147,3	14,0	133,3
4	I. (1)	123,2	-9,1	132,3
4	II. (2)	129,2	-2,4	131,6
4	III. (3)	130,3	-2,7	133,0
4	IV. (4)	147,9	14,0	133,9
5	I. (1)	128,0	-9,1	137,1
5	II. (2)	135,7	-2,4	138,1
5	III. (3)	136,2	-2,7	138,9
5	IV. (4)	155,5	14,0	141,5

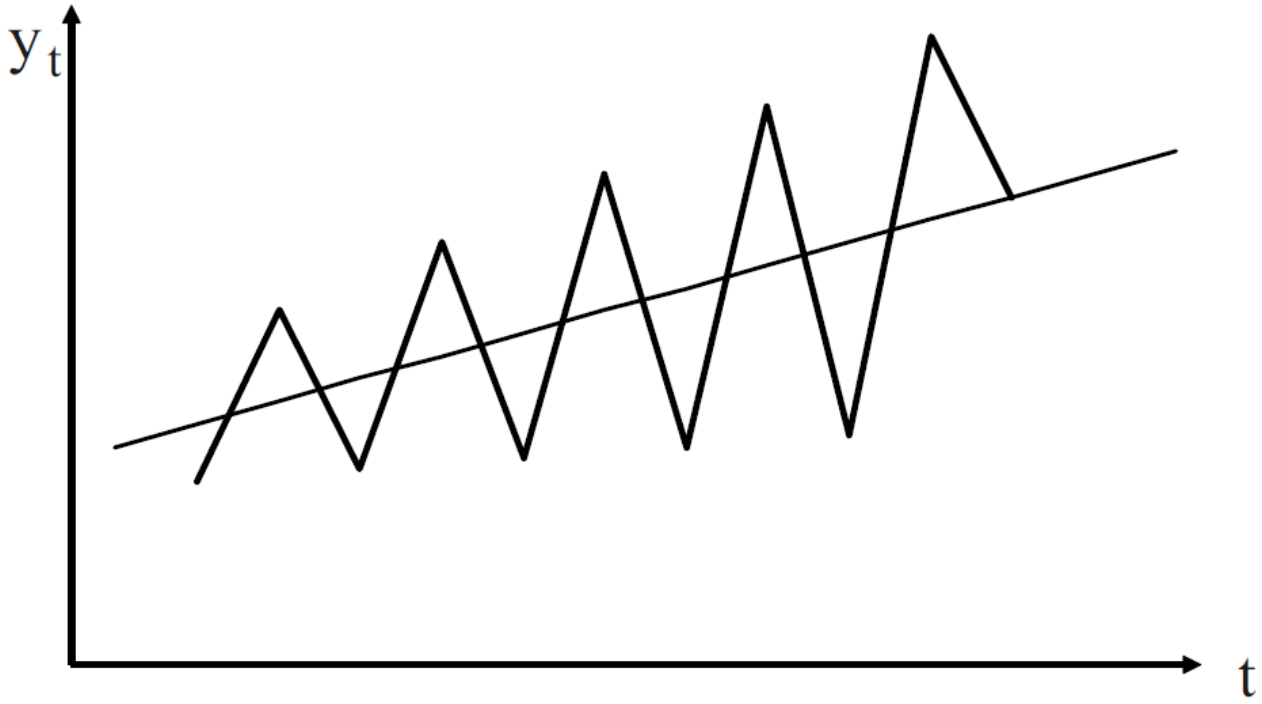
Beispiel (Fortsetzung): Löhne und Gehälter (VI)

- Löhne und Gehälter je Beschäftigten mit saisonbereinigter Zeitreihe:



Saisonkomponente und Saisonbereinigung – multiplikatives Modell (I)

- Das **multiplikative Modell** ist für ökonomische Zeitreihen adäquat, deren Saisonausschläge mit steigendem Trend im Mittel proportional zunehmen:



- Da die Komponenten in den Logarithmen additiv sind, kann die Saisonkomponente im Prinzip wie im additiven Modell bestimmt werden, wenn man die originäre Zeitreihe (y_t) durch die logarithmierte Zeitreihe ($\ln y_t$) ersetzt.

Saisonkomponente und Saisonbereinigung – multiplikatives Modell (II)

- Aus der originären Zeitreihe (y_t) lassen sich die Saisonkomponenten und die saisonbereinigte Zeitreihe wie folgt bestimmen:
  - Trendbereinigte Werte:

$$d_{ij} = \frac{y_{ij}}{g_{ij}} = s_{ij} \cdot u_{ij}, \qquad i = 1, 2, \dots, k_j; j = 1, 2, \dots, p$$

  - Unnormierte Saisonkomponente (gewöhnlich wie im additiven Modell):

$$s_j^* = \frac{1}{k_j} \cdot \sum_{i=1}^{k_j} d_{ij}$$

  - Für die normierte Saisonkomponente gilt approximativ die Eigenschaft:

$$s_1 \cdot s_2 \cdot \dots \cdot s_p = 1 \quad \text{(Normierung)}$$

  - Normierte Saisonkomponente:

$$s_j = \frac{s_j^*}{\bar{d}}$$

  - Saisonbereinigte Zeitreihe:

$$y_{ij}^* = \frac{y_{ij}}{s_j}$$