**Dr. Erich Oltmanns**

hat Volkswirtschaftslehre an der Philipps-Universität Marburg studiert und promovierte dort zu einem Thema aus den Volkswirtschaftlichen Gesamtrechnungen. Seit 2002 ist er in verschiedenen Bereichen im Statistischen Bundesamt tätig, derzeit im Referat „Erwerbstätigenrechnung“.

# HOFFERSCHE FORMEL

## Ein Verfahren der temporalen Disaggregation von Zeitreihen

Dr. Erich Oltmanns

**Schlüsselwörter:** Volkswirtschaftliche Gesamtrechnungen – Woffersche Formel – ökonomische Zeitreihen – temporale Disaggregation

## ZUSAMMENFASSUNG

Bei der Erstellung der Volkswirtschaftlichen Gesamtrechnungen kommt es vor, dass Vierteljahresdaten aus Jahresdaten generiert werden müssen, ohne dass Zusatzinformationen zur unterjährigen Entwicklung vorliegen. Zu diesem Zweck wurde im Statistischen Bundesamt ein als Woffersche Formel bezeichnetes Verfahren entwickelt. Dieser Aufsatz erläutert die Woffersche Formel und ihre Genauigkeit im Vergleich zu anderen Verfahren der temporalen Disaggregation.

***Keywords:*** *national accounts – Hoffer's formula – economic time series – temporal disaggregation*

## *ABSTRACT*

*The compilation of national accounts aggregates sometimes requires the disaggregation of yearly data into quarterly data without any additional information on the infra-annual development. For this purpose, a method has been developed at the Federal Statistical Office that is referred to as Hoffer's formula. This article discusses Hoffer's formula and its accuracy compared with other methods of temporal disaggregation.*

# 1

## Einleitung

Den Berechnungen der Volkswirtschaftlichen Gesamtrechnungen (VGR) liegen grundsätzlich Zeitreihen mit Jahresdaten und mit Vierteljahresdaten zugrunde. Allerdings gibt es Fälle, in denen für ein Merkmal zwar Jahresdaten vorliegen, jedoch Vierteljahresdaten erforderlich sind. Um dem Nutzerbedarf entsprechen zu können, ist es nicht möglich, hier eine Datenlücke zuzulassen. Vielmehr ist diese Lücke mit einem geeigneten Verfahren zu schließen. Im günstigsten Fall liegen hierfür unterjährige Informationen in Form von Indikatorreihen vor, mit denen die sogenannte temporale Disaggregation, also die Zerlegung von Jahresdaten in Quartalsdaten, vorgenommen werden kann. Ist dies nicht der Fall, verbleiben rein mathematische Verfahren als Lösung.

Ein mathematisches Verfahren der temporalen Disaggregation, welches in gewissen Bereichen der Volkswirtschaftlichen Gesamtrechnungen des Statistischen Bundesamtes seit längerem angewandt wird, ist die sogenannte Hoffersche Formel.[1]

Dieser Aufsatz soll die Hoffersche Formel näher erläutern: Zunächst wird ein kurzer Überblick zu den Verfahren der temporalen Disaggregation gegeben. Anschließend werden der Ansatz von Lisman und Sandee (1964), ein mathematisches Verfahren der Disaggregation, und die Hoffersche Formel selbst, die einen Spezialfall des Ansatzes von Lisman und Sandee darstellt, näher erklärt.

Im darauffolgenden Kapitel wird die Leistungsfähigkeit der Hofferschen Formel untersucht. Hier werden die verschiedenen Verfahren der temporalen Disaggregation auf eine Zeitreihe angewandt, bei denen sowohl Jahres- als auch Quartalsangaben zur Verfügung stehen und auch veröffentlicht werden: das saison-, kalender- und preisbereinigte Bruttoinlandsprodukt für Deutschland. Anhand ausgewählter Datensituationen werden die verschiedenen Verfahren mithilfe von gängigen Fehlermaßen dargestellt. Der Aufsatz schließt mit Bemerkungen zur Leistungsfähigkeit der Hofferschen Formel im Vergleich zu anderen Verfahren.

1 Die Hoffersche Formel ist benannt nach Dr. Helmut Hoffer, ehemals Leiter der Gruppe „Monatsstatistiken im Bergbau und Verarbeitenden Gewerbe, Indizes". Dr. Hoffer war von 1960 bis 1988 im Statistischen Bundesamt tätig.

# 2

## Verfahren der temporalen Disaggregation

### 2.1 Überblick

Das Problem der temporalen Disaggregation ist bereits älterer Natur und die vorgeschlagenen Lösungsansätze sind entsprechend zahlreich (Pavía-Miralles, 2010). Grundsätzlich können zwei Arten von Methoden der temporalen Disaggregation unterschieden werden: Die erste verwendet unterjährige Zusatzinformationen in Form von Indikatorreihen, die in sachlogischem Zusammenhang zu der zu disaggregierenden Zeitreihe der Jahresdaten steht. Solche Methoden wurden etwa vorgestellt von Chow und Lin (1971), Fernández (1981), Litterman (1983) oder Polasek und anderen (2010). Sind keine unterjährigen Informationen vorhanden, verbleibt die Möglichkeit rein mathematischer Verfahren. Ein früher Vorschlag für ein solches Verfahren stammt von Lisman und Sandee (1964). Weitere Verfahren wurden von Boot, Feibes und Lisman (1967) sowie Denton (1971) vorgestellt. Weiterhin ist eine Unterscheidung zwischen modellbasierten und nicht modellbasierten Verfahren möglich. Modellbasiert sind etwa die Verfahren von Harvey und Pierse (1984) sowie Wei und Stram (1990), nicht modellbasiert die Verfahren von Lisman und Sandee (1964), Boot, Feibes und Lisman (1967) sowie Denton (1971).

### 2.2 Das Verfahren von Lisman und Sandee und die Hoffersche Formel

Lisman und Sandee (1964) haben ein einfaches Verfahren der temporalen Disaggregation vorgestellt.[2] Sei der Indexwert für das Jahr $t$ mit $x_t$ bezeichnet und die Quar-

2 Für den Hinweis auf dieses Verfahren danke ich Berthold Fischer (ehemals Statistisches Landesamt Baden-Württemberg) und Dr. Udo Vullhorst (Ministerium für Finanzen und Wirtschaft Baden-Württemberg). Für die Ableitung der Formel von Lisman und Sandee siehe auch Vullhorst (2015).

talswerte dieses Jahres mit $y_{t,1}$ bis $y_{t,4}$.[3] Dann lassen sich die Quartale dieses Jahres darstellen als mit den Elementen $m_{ij}$ der Matrix M gewogene Summe des Jahres $t$, des Vorjahres $t-1$ und des Folgejahres $t+1$:

$$(1) \quad \begin{pmatrix} y_{t,1} \\ y_{t,2} \\ y_{t,3} \\ y_{t,4} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \\ m_{41} & m_{42} & m_{43} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ x_t \\ x_{t+1} \end{pmatrix}$$

Sind alle drei Jahre bekannt, dann besteht das Problem nur noch darin, die Elemente der Matrix M zu bestimmen. Lisman und Sandee (1964) führen hierzu mehrere Bedingungen ein. So soll sich etwa der Index für das Jahr als arithmetisches Mittel der vier Quartale ergeben:

$$(2) \quad x_t = \frac{1}{4} \sum_{i=1}^{4} y_{t,i}$$

und die Gewichte für Vorjahr und Folgejahr sollen symmetrisch sein:

$$(3) \quad m_{11} = m_{43},\ m_{21} = m_{33},\ m_{31} = m_{23} \text{ und } m_{41} = m_{13}$$

Diese und weitere Bedingungen führen dann zu der Lösung

$$(4) \quad \begin{pmatrix} y_{t,1} \\ y_{t,2} \\ y_{t,3} \\ y_{t,4} \end{pmatrix} = \left( \frac{1}{16} \begin{pmatrix} 3 & 16 & -3 \\ 1 & 16 & -1 \\ -1 & 16 & 1 \\ -3 & 16 & 3 \end{pmatrix} + \frac{\alpha}{16} \begin{pmatrix} -1 & 2 & -1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \\ -1 & 2 & -1 \end{pmatrix} \right) \begin{pmatrix} x_{t-1} \\ x_t \\ x_{t+1} \end{pmatrix}$$

Bei gegebenen Indexwerten für die Zeitreihe der Jahresdaten ist noch ein geeignetes $\alpha$ zu wählen. Für die Wahl von $\alpha$ gibt es kein mathematisches oder statistisches Verfahren. Nach Lisman und Sandee sollte diese Wahl “in a commonsense way” geschehen, also auf gesundem Menschenverstand beruhen. Sie machen hierzu einen Vorschlag, der durchaus als plausibel angesehen werden kann. Über die Annahme eines sinusförmigen Verlaufs der zerlegten Zeitreihe sowie über das Verhältnis von Amplitude der Reihe der Quartalsdaten zur Differenz zwischen den Jahreswerten ermitteln sie ein $\alpha = -1{,}656$.

Ein weiteres Problem stellen die Ränder der zu zerlegenden Zeitreihe dar. Zu Beginn der Zeitreihe fehlt ein Jahreswert für die Zerlegung des Startjahres. Möglich wäre es, den fehlenden Jahreswert zu schätzen und die Zerlegung auf dessen Grundlage vorzunehmen. Aber die einfachste Möglichkeit besteht sicherlich darin, auf eine Disaggregation des Startjahres zu verzichten. Anders stellt sich die Situation am aktuellen Rand dar, denn hier sind sogar zwei Jahreswerte zu schätzen: das aktuelle Jahr und das Folgejahr. Am aktuellen Rand kann auf eine Disaggregation nicht verzichtet werden. Lisman und Sandee äußern sich nicht zu diesem Problem. Wenn aber der ihrem Verfahren zugrunde liegende Gedanke der Einfachheit beibehalten werden sollte, bieten sich für die Prognose der Jahreswerte am aktuellen Rand ebenfalls einfache Verfahren wie beispielsweise eine Trendextrapolation an.

Bei der Hofferschen Formel handelt es sich um einen Spezialfall des Lisman-Sandee-Ansatzes. Nach Einsetzen von $\alpha = -2$ vereinfacht sich der Ausdruck zu

$$(5) \quad \begin{pmatrix} y_{t,1} \\ y_{t,2} \\ y_{t,3} \\ y_{t,4} \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 5 & 12 & -1 \\ -1 & 20 & -3 \\ -3 & 20 & -1 \\ -1 & 12 & 5 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ x \\ x_{t+1} \end{pmatrix}$$

Diese Formel wird seit Längerem in einigen Bereichen im Statistischen Bundesamt bei der Erstellung der Volkswirtschaftlichen Gesamtrechnungen verwendet, so etwa bei der Berechnung

- der Investitionen in Forschung und Entwicklung (Adler und andere, 2014),
- der Abschreibungen,
- von Versicherungsleistungen sowie
- von schattenwirtschaftlichen Aktivitäten.

Seit wann die Formel bereits verwendet wird und auf welche Weise sie zu ihrem Namen gekommen ist, kann nicht mehr genau rekonstruiert werden. Beschrieben wird das Verfahren in einem internen Vermerk aus dem Jahr 1977. Da das Statistische Bundesamt im Jahr 1978 erstmals Vierteljahresergebnisse der VGR veröffentlicht hat, ist davon auszugehen, dass das Verfahren seitdem auch eingesetzt wird (Hamer/Engelmann, 1978).

3 Im Folgenden wird ausschließlich auf Indexzahlen Bezug genommen. Im Fall von Wertgrößen sind die Berechnungen entsprechend zu normieren.

# 3

# Die Hoffersche Formel im Genauigkeitsvergleich

## 3.1 Verfahren zur Beurteilung der Schätzgenauigkeit

Um im Folgenden die Schätzgenauigkeit der Hofferschen Formel zu beurteilen, werden drei Voraussetzungen benötigt: erstens eine Zeitreihe, die in jährlicher und vierteljährlicher Form vorliegt und welche die zu schätzende Größe beziehungsweise die Zielgröße darstellt, zweitens geeignete Maßzahlen zur Beurteilung der Genauigkeit der einzelnen Verfahren und drittens andere Verfahren der temporalen Disaggregation, die als Vergleichsmaßstab herangezogen werden können.

Als zu zerlegende Reihe wird die Zeitreihe des saison-, kalender- und preisbereinigten Bruttoinlandsprodukts für Deutschland für den Zeitraum von 1991 bis 2014 verwendet. Als Zielgröße für die Zerlegung steht somit die Zeitreihe der Quartalswerte vom ersten Quartal 1991 bis zum dritten Quartal 2015 zur Verfügung (Statistisches Bundesamt, diverse Jahrgänge).

Insgesamt sind 24 Veröffentlichungsstände der Zeitreihe in die Untersuchung einbezogen, und zwar vom vierten Quartal 2009 bis zum dritten Quartal 2015. Der gewählte Zeitraum ergibt sich daraus, dass somit Quartale aus sechs vollständigen Veröffentlichungsjahren zur Verfügung stehen.[4] Die Veröffentlichungsjahre sind jeweils durch eine identische Datensituation gekennzeichnet. ↘ **Tabelle 1 auf Seite 66**

Weiterhin ist festzulegen, mit welchen Verfahren die Hoffersche Formel verglichen werden soll. Hier bietet sich zunächst der Ansatz von Lisman und Sandee an. Ein weiteres mathematisches Verfahren, welches durch die vom Statistischen Amt der Europäischen Union (Eurostat) entwickelte Software ECOTRIM eine gewisse Verbreitung erfahren hat, ist der Ansatz von Denton (1971; Barcellan/Buono, 2002).

4 So wurde im Veröffentlichungsjahr 2015 das Bruttoinlandsprodukt für das vierte Quartal 2014 und die ersten drei Quartale des Jahres 2015 erstmals veröffentlicht.

Entsprechend der Vielzahl von Vorschlägen für Methoden zur temporalen Disaggregation besteht bei der Entscheidung für weitere Verfahren ein gewisser Ermessensspielraum. Deshalb wird ein anderer Weg beschritten: Es wird die Frage gestellt, ob ein deutlich einfacheres Verfahren als die bisher genannten drei zu vergleichbaren Ergebnissen führt beziehungsweise, wie die genannten drei Verfahren im Vergleich zu einem Referenz- oder Benchmarkverfahren abschneiden.

Als ein solches Verfahren kann die „Division durch 4“ angesehen werden. Jahreswerte werden dabei gleichmäßig auf die Quartale aufgeteilt. Im Fall von Indexreihen wird der Wert für das Jahr von den Quartalen des jeweiligen Jahres übernommen. Um die unerwünschten Sprungstellen beim Wechsel von einem zum anderen Jahr zu vermeiden, kann die Reihe geglättet werden. Hier bietet sich zum Beispiel ein (zentrierter) gleitender 4er-Durchschnitt an.

Zur Beurteilung der Schätzgenauigkeit werden mehrere quantitative Maße verwendet (Rinne/Specht, 2002, hier: Seite 133 ff.). Diese beruhen auf dem Schätz- oder Prognosefehler $e$, der als Differenz zwischen einem Schätzwert $\hat{x}$ und einem als „wahr“ angesehenen Wert $x$ berechnet wird. Der Schätzwert ist hier jeweils ein durch eines der Verfahren der temporalen Disaggregation erzeugter Wert, während als „wahrer Wert“ ein saison-, kalender- und preisbereinigter Wert betrachtet wird. Für die $i$-te Beobachtung aus einer Zeitreihe ist der Schätzfehler

(6) $$e_i = x_i - \hat{x}_i.$$

Das erste verwendete Maß für die Beurteilung der Schätzgenauigkeit ist die Mittlere Abweichung beziehungsweise Mittlere Revision (MR). Sie wird berechnet über die Formel

(7) $$MR = \frac{1}{n}\sum_{i=1}^{n} e_i$$

mit $n$ als Zahl der Beobachtungen. Sie ist das arithmetische Mittel der Schätzfehler und gibt einen Hinweis auf die Verzerrung beziehungsweise den bias einer Schätzung. Einzelwerte mit unterschiedlichen Vorzeichen gleichen sich in dieser Formel aus, sodass Werte für die Mittlere Revision, die nahe Null liegen, auf eine im Mittel unverzerrte Prognose hindeuten. Die Mittlere Revision ist als Maß zur Beurteilung nur bedingt geeignet, da die

**Tabelle 1**
**Bruttoinlandsprodukt (preisbereinigt, verkettet) für Deutschland**

| | Saison- und kalenderbereinigte Werte nach Census X-12-ARIMA | | | | Temporal disaggregierte Jahresdaten | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | zentrierter gleitender 4er-Durchschnitt [1] | | Hoffer | | Lisman/Sandee | | Denton | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | 2010 = 100 | % [1] | 2010 = 100 | % [2] | 2010 = 100 | % [2] | 2010 = 100 | % [2] | 2010 = 100 | % [2] | 2010 = 100 | % [2] |
| 2010 | 99,79 | + 3,9 | | | | | | | | | | |
| 1. Vj | | | 97,71 | + 0,8 | 98,37 | + 1,0 | 98,37 | + 1,6 | 98,37 | + 1,8 | 97,66 | + 1,8 |
| 2. Vj | | | 99,69 | + 2,0 | 99,31 | + 1,0 | 99,33 | + 1,0 | 99,33 | + 1,0 | 99,21 | + 1,6 |
| 3. Vj | | | 100,48 | + 0,8 | 100,25 | + 0,9 | 100,27 | + 0,9 | 100,26 | + 0,9 | 100,56 | + 1,4 |
| 4. Vj | | | 101,27 | + 0,8 | 101,18 | + 0,9 | 101,18 | + 0,9 | 101,19 | + 0,9 | 101,71 | + 1,1 |
| 2011 | 103,50 | + 3,7 | | | | | | | | | | |
| 1. Vj | | | 103,16 | + 1,9 | 102,11 | + 0,9 | 102,30 | + 1,1 | 102,36 | + 1,2 | 102,66 | + 0,9 |
| 2. Vj | | | 103,34 | + 0,2 | 103,03 | + 0,9 | 103,61 | + 1,3 | 103,54 | + 1,2 | 103,38 | + 0,7 |
| 3. Vj | | | 103,76 | + 0,4 | 103,58 | + 0,5 | 104,15 | + 0,5 | 104,09 | + 0,5 | 103,85 | + 0,5 |
| 4. Vj | | | 103,73 | + 0,0 | 103,74 | + 0,2 | 103,93 | – 0,2 | 103,99 | – 0,1 | 104,10 | + 0,2 |
| 2012 | 104,13 | + 0,6 | | | | | | | | | | |
| 1. Vj | | | 104,13 | + 0,4 | 103,89 | + 0,2 | 103,91 | + 0,0 | 103,91 | – 0,1 | 104,11 | + 0,0 |
| 2. Vj | | | 104,19 | + 0,1 | 104,05 | + 0,2 | 104,09 | + 0,2 | 104,09 | + 0,2 | 104,12 | + 0,0 |
| 3. Vj | | | 104,36 | + 0,2 | 104,19 | + 0,1 | 104,23 | + 0,1 | 104,22 | + 0,1 | 104,14 | + 0,0 |
| 4. Vj | | | 103,85 | – 0,5 | 104,29 | + 0,1 | 104,30 | + 0,1 | 104,31 | + 0,1 | 104,16 | + 0,0 |
| 2013 | 104,56 | + 0,4 | | | | | | | | | | |
| 1. Vj | | | 103,57 | – 0,3 | 104,40 | + 0,1 | 104,32 | + 0,0 | 104,29 | + 0,0 | 104,18 | + 0,0 |
| 2. Vj | | | 104,51 | + 0,9 | 104,50 | + 0,1 | 104,27 | + 0,0 | 104,30 | + 0,0 | 104,34 | + 0,2 |
| 3. Vj | | | 104,9 | + 0,4 | 104,76 | + 0,2 | 104,53 | + 0,2 | 104,56 | + 0,2 | 104,63 | + 0,3 |
| 4. Vj | | | 105,24 | + 0,3 | 105,17 | + 0,4 | 105,10 | + 0,5 | 105,07 | + 0,5 | 105,06 | + 0,4 |
| 2014 | 106,21 | + 1,6 | | | | | | | | | | |
| 1. Vj | | | 105,99 | + 0,7 | 105,59 | + 0,4 | 105,63 | + 0,5 | 105,65 | + 0,6 | 105,63 | + 0,5 |
| 2. Vj | | | 105,93 | – 0,1 | 106,00 | + 0,4 | 106,14 | + 0,5 | 106,12 | + 0,4 | 106,09 | + 0,4 |
| 3. Vj | | | 106,13 | + 0,2 | 106,32 | + 0,3 | 106,46 | + 0,3 | 106,44 | + 0,3 | 106,44 | + 0,3 |
| 4. Vj | | | 106,78 | + 0,6 | 106,55 | + 0,2 | 106,60 | + 0,1 | 106,61 | + 0,2 | 106,67 | + 0,2 |
| 2015 | – | – | | | | | | | | | | |
| 1. Vj | | | 107,15 | + 0,3 | 106,78 | + 0,2 | 106,78 | + 0,2 | 106,77 | + 0,1 | 106,80 | + 0,1 |
| 2. Vj | | | 107,62 | + 0,4 | 107,01 | + 0,2 | 106,99 | + 0,2 | 107,00 | + 0,2 | 106,98 | + 0,2 |
| 3. Vj | | | 107,96 | + 0,3 | 107,13 | + 0,1 | 107,24 | + 0,2 | 107,24 | + 0,2 | 107,22 | + 0,2 |
| 4. Vj | | | – | – | – | – | – | – | – | – | – | – |

Veröffentlichungsstand: November 2015.
1 Veränderung gegenüber dem Vorjahr.
2 Veränderung gegenüber dem Vorquartal.

untersuchten Verfahren implizit so konstruiert sind, dass sie (abgesehen von rundungsbedingten Abweichungen) den Wert Null annimmt.

Unverzerrte Schätzungen sind jedoch nicht mit „genauen“ Schätzungen gleichzusetzen. Deshalb wird zusätzlich die auch als „Dispersion“ bezeichnete Mittlere absolute Abweichung oder Mittlere absolute Revision (MAR) berechnet. Bei der Mittleren absoluten Revision bleibt das Vorzeichen des Schätzfehlers unberücksichtigt. Sie ist das arithmetische Mittel der Absolutbeträge der Schätzfehler:

$$(8) \qquad MAR = \frac{1}{n}\sum_{i=1}^{n}|e_i|$$

Die Mittlere Revision und die Mittlere absolute Revision gewichten alle Schätzfehler gleich. Anders der Mittlere Quadratwurzelfehler oder der Root mean squared error (RMSE): Hier erfolgt durch das Quadrieren der Fehler eine stärkere Gewichtung oder „Bestrafung" größerer Schätzfehler:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \qquad (9)$$

Der Root mean squared error ist jedoch nicht mehr direkt interpretierbar, wie es die anderen zuvor genannten Größen sind.

Ein viertes Maß ist der Theilsche Ungleichheitskoeffizient (Theilsches U – TU). Mit diesem Maß kann die Schätzgenauigkeit zweier Verfahren direkt miteinander verglichen werden, so zum Beispiel die zwischen dem Hofferschen Verfahren und dem Verfahren „Division durch 4" als Referenzmethode:

$$TU_{Hoffer} = \sqrt{\frac{RMSE^2_{Hoffer}}{RMSE^2_{Division\ durch\ 4}}} \qquad (10)$$

Der Theilsche Ungleichheitskoeffizient ist einfach interpretierbar. Bei Werten nahe Eins sind die verglichenen Verfahren hinsichtlich ihrer Schätzgenauigkeit als gleichwertig anzusehen. Bei Werten über Eins ist das Referenzverfahren vorzuziehen, bei Werten unter Eins das Verfahren im Zähler.

## 3.2 Disaggregation der gesamten Reihe

In einem ersten Schritt wurden die verschiedenen Verfahren der temporalen Disaggregation für die Zeitreihen ab dem zweiten Quartal 1992 überprüft, somit ab dem frühest möglichen Zeitpunkt, zu dem ein Quartal zur Verfügung stand. Gewählt wurden jeweils die Veröffentlichungsstände bis zum zweiten Quartal der Jahre 2010 bis 2015, die jeweils im August des genannten Jahres veröffentlicht worden sind. Diese Veröffentlichungsstände wurden deshalb gewählt, weil im Rahmen der sogenannten Sommerrechnung im August Überarbeitungen für die Quartale der vergangenen vier Jahre möglich sind. Zu den anderen Terminen werden üblicherweise lediglich die Quartale des aktuellen Jahres überprüft und gegebenenfalls überarbeitet (Statistisches Bundesamt, 2015, hier: Seite 8). Weiterhin ist eine Analyse der Genauigkeit der verschiedenen Verfahren nur sinnvoll für sich ändernde Datenkonstellationen beziehungsweise dann, wenn eine gewisse Streuung in den Zeitreihen vorliegt.

In ↘ Tabelle 2 sind die Ergebnisse der Genauigkeitsanalyse dargestellt. Ausgelassen wurden das Startjahr

**Tabelle 2**

**Verfahren der temporalen Disaggregation für das saison-, kalender- und preisbereinigte Bruttoinlandsprodukt für Deutschland**

| | Veröffentlichung bis zum zweiten Quartal … | | | | | | Insgesamt |
|---|---|---|---|---|---|---|---|
| | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | |
| Anzahl der Beobachtungen | 73 | 77 | 81 | 85 | 89 | 93 | 498 |
| | Differenz zur veröffentlichten Veränderungsrate gegenüber dem Vorquartal in Prozentpunkten | | | | | | |
| **MAR (Mittlere absolute Revision)** | | | | | | | |
| Zentrierter gleitender 4er-Durchschnitt | + 0,50 | + 0,51 | + 0,47 | + 0,48 | + 0,46 | + 0,38 | + 0,46 |
| Hoffer | + 0,51 | + 0,55 | + 0,52 | + 0,53 | + 0,51 | + 0,46 | + 0,51 |
| Lisman/Sandee | + 0,48 | + 0,53 | + 0,49 | + 0,50 | + 0,48 | + 0,38 | + 0,47 |
| Denton | + 0,53 | + 0,53 | + 0,50 | + 0,50 | + 0,49 | + 0,40 | + 0,49 |
| **RMSE (Root mean squared error)** | | | | | | | |
| Zentrierter gleitender 4er-Durchschnitt | + 0,68 | + 0,66 | + 0,63 | + 0,64 | + 0,63 | + 0,62 | + 0,64 |
| Hoffer | + 0,70 | + 0,74 | + 0,70 | + 0,71 | + 0,69 | + 0,68 | + 0,70 |
| Lisman/Sandee | + 0,67 | + 0,69 | + 0,64 | + 0,65 | + 0,63 | + 0,62 | + 0,65 |
| Denton | + 0,72 | + 0,71 | + 0,66 | + 0,67 | + 0,65 | + 0,63 | + 0,67 |
| **Theilscher Ungleichheitskoeffizient** | | | | | | | |
| Hoffer | + 1,025 | + 1,108 | + 1,107 | + 1,107 | + 1,099 | + 1,099 | + 1,091 |
| Lisman/Sandee | + 0,974 | + 1,034 | + 1,019 | + 1,025 | + 1,003 | + 1,004 | + 1,010 |
| Denton | + 1,050 | + 1,066 | + 1,046 | + 1,047 | + 1,029 | + 1,030 | + 1,045 |

Vergleich der Veränderungsraten für das vierteljährliche saison-, kalender- und preisbereinigte Bruttoinlandsprodukt mit den Veränderungsraten der temporal disaggregierten Jahresdaten.

1991, für welches kein Vorjahreswert existiert, und das erste Quartal 1992, für welches kein Vorquartalswert vorliegt. Die Zeitreihen beginnen somit mit dem zweiten Quartal 1992.

**Tabelle 3**
**Verfahren der temporalen Disaggregation für das saison-, kalender- und preisbereinigte Bruttoinlandsprodukt für Deutschland am aktuellen Rand bei unbekannten Jahresdaten**

| | Veröffentlichung im Jahr … | | | | | | Insgesamt |
|---|---|---|---|---|---|---|---|
| | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | |
| Anzahl der Beobachtungen | 6 | 6 | 6 | 6 | 6 | 6 | 36 |
| | Differenz zur veröffentlichten Veränderungsrate gegenüber dem Vorquartal in Prozentpunkten | | | | | | |
| **MAR (Mittlere absolute Revision)** | | | | | | | |
| Zentrierter gleitender 4er-Durchschnitt | + 1,17 | + 0,64 | + 0,21 | + 0,53 | + 0,25 | + 0,22 | + 0,50 |
| Hoffer | + 1,30 | + 0,83 | + 0,34 | + 0,54 | + 0,37 | + 0,29 | + 0,61 |
| Lisman/Sandee | + 1,21 | + 0,89 | + 0,33 | + 0,56 | + 0,33 | + 0,28 | + 0,60 |
| Denton | + 1,23 | + 0,83 | + 0,27 | + 0,52 | + 0,32 | + 0,28 | + 0,58 |
| **RMSE (Root mean squared error)** | | | | | | | |
| Zentrierter gleitender 4er-Durchschnitt | + 1,39 | + 0,79 | + 0,24 | + 0,67 | + 0,31 | + 0,24 | + 0,73 |
| Hoffer | + 1,55 | + 0,97 | + 0,41 | + 0,71 | + 0,48 | + 0,33 | + 0,85 |
| Lisman/Sandee | + 1,47 | + 1,04 | + 0,40 | + 0,72 | + 0,44 | + 0,31 | + 0,84 |
| Denton | + 1,43 | + 1,04 | + 0,36 | + 0,64 | + 0,41 | + 0,31 | + 0,81 |
| **Theilscher Ungleichheitskoeffizient** | | | | | | | |
| Hoffer | + 1,111 | + 1,217 | + 1,720 | + 1,065 | + 1,539 | + 1,362 | + 1,160 |
| Lisman/Sandee | + 1,060 | + 1,310 | + 1,668 | + 1,069 | + 1,437 | + 1,289 | + 1,144 |
| Denton | + 1,025 | + 1,310 | + 1,476 | + 0,959 | + 1,316 | + 1,259 | + 1,101 |

Ein Blick auf die Mittlere absolute Abweichung zeigt, dass die einzelnen Verfahren nur geringfügig voneinander abweichen. Am genauesten ist hier der als Referenzverfahren ermittelte zentrierte, gleitende 4er-Durchschnitt, gefolgt von den Verfahren nach Lisman und Sandee sowie nach Denton. Das Verfahren von Hoffer schneidet hier zwar am schlechtesten ab, allerdings zeigen sich die Unterschiede zwischen den Verfahren erst in der zweiten Nachkommastelle. Gleiches gilt für den RMSE und die daraus resultierenden Werten für das Theilsche U. Die Ergebnisse gelten nicht nur für alle Beobachtungen insgesamt, sondern im Wesentlichen auch für die Beobachtungen der einzelnen Veröffentlichungsstände.

Hinsichtlich der Datensituation ist zu berücksichtigen, dass die betrachteten Zeitreihen jeweils aus zwei Teilen bestehen: Weil die temporale Disaggregation jeweils auf Daten über Vorjahr und Folgejahr zugreift, fehlen am Beginn der Zeitreihe und am aktuellen Rand jeweils Jahreswerte. In der „Mitte“ der Zeitreihe dagegen sind die Daten bekannt, die für die Disaggregation benötigt werden. Deshalb wurde die Analyse der einzelnen Verfahren zusätzlich getrennt nach den beiden Datensituationen durchgeführt.

Die Ergebnisse für die Disaggregation bei bekannten Jahreswerten unterscheiden sich, wie es auch zu erwarten war, kaum von denen für die gesamte Reihe. Auf eine Darstellung wird deshalb verzichtet.

↘ Tabelle 3 enthält die Ergebnisse der Analyse der Disaggregation bei unbekannten Jahreswerten am aktuellen Rand der Zeitreihe. Eine Schließung der Datenlücke zu Beginn der Reihe wäre zwar auch möglich. Weil dieser Zeitraum jedoch im Vergleich zum aktuellen Rand nur geringe Relevanz aufweist, wurde darauf verzichtet. Auch hier liefert der zentrierte gleitende 4er-Durchschnitt als Referenzverfahren das genaueste Ergebnis. Die drei anderen Verfahren liegen bei Betrachtung der Mittleren absoluten Abweichung um ungefähr 0,1 Prozentpunkte schlechter als das Referenzverfahren. Abgesehen davon sind die Aussagen identisch mit denen zur Analyse für die gesamte Reihe.

## 3.3 Genauigkeit der Erstveröffentlichung

Das Verhalten der Disaggregationsverfahren am aktuellen Rand ist aus zwei Gründen von Bedeutung. Zum einen erfährt die Erstveröffentlichung von Vierteljahresdaten zum Bruttoinlandsprodukt immer ein besonderes Interesse in Öffentlichkeit, Politik und Medien. Und zum anderen ist eine größere Ungenauigkeit der Verfahren zu vermuten, weil die Disaggregation hier, wie bereits beschrieben, zum Teil auf Grundlage von Prognosedaten durchgeführt wird. Deshalb konzentriert sich die Genauigkeitsanalyse jetzt auf die erste Veröffentlichung von Quartalsdaten. Im Unterschied zu den bisherigen Betrachtungen werden nun alle Quartale eines Jahres in die Untersuchung einbezogen. Weiterhin ist es jetzt auch möglich beziehungsweise sinnvoll, die Mittlere Abweichung zu berechnen.

Die Mittlere Abweichung nimmt bei einer unverzerrten Schätzung den Wert Null an. Da die Werte bei allen Verfahren jedoch bei 0,2 Prozentpunkten liegen, schätzen die Verfahren der temporalen Disaggregation somit leicht verzerrt, das heißt im insgesamt betrachteten Zeitraum systematisch etwas zu niedrig. Ein Blick auf die Ergebnisse für die Quartale einzelner Jahre zeigt hier ein differenzierteres Bild. So gibt es Jahre, in denen die Schätzungen systematisch zu niedrig liegen, während sie in anderen Jahren systematisch zu hoch liegen. Auch sind jetzt deutliche Unterschiede in der Genauigkeit der einzelnen Verfahren zu erkennen. So ist das Referenzverfahren im Jahr 2010 den anderen Verfahren unterlegen, während es im Jahr 2012 als einziges Verfahren nahezu unverzerrt schätzt. ↘ **Tabelle 4**

Der Grund für die systematischen Verzerrungen dürfte in dem gewählten Prognoseverfahren am aktuellen Rand zu suchen sein. Dieses schreibt die aktuelle konjunkturelle Situation (der vergangenen drei Jahre) fort und führt somit bei Änderungen der konjunkturellen Lage zu Verzerrungen.

Die Mittlere absolute Revision und der RMSE beziehungsweise das Theilsche U sind mit den Ergebnissen der bisherigen Untersuchung vergleichbar.

Zusätzlich wurde auch das erste Ergebnis der temporalen Disaggregation mit dem Ergebnis der Berechnung im darauffolgenden Quartal verglichen. Erwartungsgemäß – das Ergebnis der temporalen Disaggregation kann sich nur bei der Änderung von Jahreswerten ändern – gibt es nur äußerst geringe Abweichungen in den Berechnungen. Auf eine Darstellung wird an dieser Stelle deshalb verzichtet.

**Tabelle 4**

**Erstes Ergebnis der temporalen Disaggregation im Vergleich zur ersten Veröffentlichung des saison- und kalenderbereinigten Ergebnisses**

| | Veröffentlichung im Jahr … | | | | | | Insgesamt |
|---|---|---|---|---|---|---|---|
| | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | |
| Anzahl der Beobachtungen | 4 | 4 | 4 | 4 | 4 | 4 | 24 |
| | Differenz zur veröffentlichten Veränderungsrate gegenüber dem Vorquartal in Prozentpunkten | | | | | | |
| **MR (Mittlere Abweichung)** | | | | | | | |
| Zentrierter gleitender 4er-Durchschnitt | + 0,91 | + 0,70 | + 0,02 | − 0,52 | − 0,08 | + 0,20 | + 0,21 |
| Hoffer | + 0,65 | + 0,85 | + 0,23 | − 0,62 | − 0,16 | + 0,24 | + 0,20 |
| Lisman/Sandee | + 0,68 | + 0,83 | + 0,20 | − 0,61 | − 0,15 | + 0,24 | + 0,20 |
| Denton | + 0,68 | + 0,90 | + 0,13 | − 0,61 | − 0,14 | + 0,23 | + 0,20 |
| **MAR (Mittlere absolute Revision)** | | | | | | | |
| Zentrierter gleitender 4er-Durchschnitt | + 0,91 | + 0,98 | + 1,30 | + 0,78 | + 0,70 | + 0,67 | + 0,49 |
| Hoffer | + 0,91 | + 1,01 | + 1,43 | + 0,93 | + 0,85 | + 0,64 | + 0,55 |
| Lisman/Sandee | + 0,90 | + 1,01 | + 1,42 | + 0,91 | + 0,83 | + 0,68 | + 0,54 |
| Denton | + 0,83 | + 0,84 | + 1,26 | + 0,89 | + 0,90 | + 0,89 | + 0,53 |
| **RMSE (Root mean squared error)** | | | | | | | |
| Zentrierter gleitender 4er-Durchschnitt | + 1,250 | + 0,868 | + 0,268 | + 0,710 | + 0,353 | + 0,295 | + 0,719 |
| Hoffer | + 1,201 | + 1,031 | + 0,395 | + 0,839 | + 0,369 | + 0,330 | + 0,776 |
| Lisman/Sandee | + 1,210 | + 1,033 | + 0,422 | + 0,816 | + 0,344 | + 0,317 | + 0,774 |
| Denton | + 1,097 | + 1,113 | + 0,398 | + 0,744 | + 0,353 | + 0,297 | + 0,749 |
| **Theilscher Ungleichheitskoeffizient** | | | | | | | |
| Hoffer | + 0,961 | + 1,188 | + 1,476 | + 1,181 | + 1,045 | + 1,119 | + 1,079 |
| Lisman/Sandee | + 0,968 | + 1,190 | + 1,578 | + 1,149 | + 0,977 | + 1,072 | + 1,076 |
| Denton | + 0,878 | + 1,283 | + 1,486 | + 1,047 | + 1,002 | + 1,004 | + 1,042 |

## 3.4 Robustheit des Ansatzes von Lisman und Sandee

Abschließend soll der Frage nachgegangen werden, inwieweit das Verfahren von Lisman und Sandee sensitiv gegenüber Veränderungen des Parameters α ist. Diese Frage ist insofern von Bedeutung, als dass es keine festen Kriterien für die Bestimmung des Parameters gibt, sondern dass dieser vor der Berechnung festzulegen ist. Zu diesem Zweck wurde die Berechnung der Revisionsmaße für verschiedene Parameter wiederholt. ↘ Tabelle 5

**Tabelle 5**
**Temporale Disaggregation nach Lisman und Sandee (1964) für verschiedene Parameter α**

| | MAR (Mittlere absolute Revision) | RMSE (Root mean squared error) | Theilscher Ungleichheits-koeffizient |
|---|---|---|---|
| Parameter α | | | |
| – 2,8 | 0,63 | 0,856 | 1,331 |
| – 2,4 | 0,56 | 0,774 | 1,204 |
| – 2,0 [1] | 0,51 | 0,701 | 1,091 |
| – 1,8 | 0,49 | 0,669 | 1,042 |
| – 1,656 [2] | 0,47 | 0,649 | 1,010 |
| – 1,6 | 0,47 | 0,641 | 0,998 |
| – 1,2 | 0,44 | 0,598 | 0,931 |
| – 0,8 | 0,42 | 0,576 | 0,896 |
| – 0,4 | 0,41 | 0,575 | 0,896 |
| + 0,0 | 0,43 | 0,598 | 0,931 |
| + 0,4 | 0,46 | 0,641 | 0,998 |
| + 0,8 | 0,51 | 0,701 | 1,090 |

1 Hoffersches Verfahren.
2 Vorschlag von Lisman und Sandee.

Hervorgehoben in der Tabelle 5 sind die Zeilen für den von Lisman und Sandee vorgeschlagenen Wert (α = – 1,656) und für den in der Hofferschen Formel verwendeten Wert (α = – 2,0). Hier wird deutlich, dass das Verfahren im Bereich der beiden Spezialfälle relativ robust ist. Andererseits gibt es ein Intervall für α, in dem das Verfahren von Lisman und Sandee genauere Schätzungen liefert als das Referenzverfahren (Werte zwischen – 1,6 und + 0,6). Somit scheinen die gewählten Parameter zumindest für die Disaggregation der gewählten Reihe nicht die optimale Lösung darzustellen.

Andererseits ist zu bedenken, dass die von Lisman und Sandee gewählte Spezifikation auf ex ante getroffenen Überlegungen beruhen, die sehr wohl plausibel sind.

## 4

## Abschließende Bewertung

Das Hoffersche Verfahren ist ein mathematisches Verfahren der temporalen Disaggregation, welches einen Spezialfall des Verfahrens von Lisman und Sandee (1964) darstellt. Im Genauigkeitsvergleich schneidet das Hoffersche Verfahren bei der temporalen Disaggregation der Zeitreihe des Bruttoinlandsprodukts nur geringfügig schlechter ab als das Verfahren von Lisman und Sandee sowie der Ansatz von Denton (1971). Probeweise durchgeführte Berechnungen mit anderen Zeitreihen ergaben ein ähnliches Bild: Die Qualitätsunterschiede zwischen den Verfahren sind relativ gering, während die Unterschiede zwischen den verschiedenen Zeitreihen deutlich sind. [5] Werden neben der Genauigkeit weitere Kriterien zur Beurteilung herangezogen, wie die Wirtschaftlichkeit oder die einfache Anwendbarkeit, dann ist der Einsatz der Hofferschen Formel zur temporalen Disaggregation aus Kosten-Nutzen-Überlegungen gerechtfertigt.

5 Zusätzliche Berechnungen wurden mit den Konsumausgaben des Staates in jeweiligen Preisen und mit dem Produktionsindex für das Verarbeitende Gewerbe ohne Baugewerbe durchgeführt.

## LITERATURVERZEICHNIS


Adler, Walther/Gühler, Nadine/Oltmanns, Erich/Schmidt, Daniel/Schmidt, Pascal/Schulz, Ingeborg. *Forschung und Entwicklung in den Volkswirtschaftlichen Gesamtrechnungen*. In: Wirtschaft und Statistik. Ausgabe 12/2014, Seite 703 ff.

Barcellan, Roberto/Buono, Dario. *Temporal disaggregation techniques*, *ECOTRIM interface (Version 1.01)*. User manual. Eurostat. 2002.

Boot, J. C. G./Feibes, W./Lisman, J. H. C. *Further Methods of Derivation of Quarterly Figures from Annual Data*. In: Applied Statistics. Jahrgang 16. 1967, Seite 65 ff.

Chen, Baoline. *An Empirical Comparison of Methods for Temporal Distribution and Interpolation at the National Accounts*. Bureau of Economic Analysis. Washington, D. C. 2007. Verfügbar unter: www.bea.govf

Chow, Gregory C./Lin, An-loh. *Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series*. In: Review of Economics and Statistics. Jahrgang 53. Nr. 4/1971, Seite 372 ff.

Denton, Frank T. *Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization*. In: Journal of the American Statistical Association. Jahrgang 66. Nr. 333/1971, Seite 99 ff.
DOI: 10.1080/01621459.1971.10482227

Fernández, R. B. *A methodological note on the estimation of time series*. In: The Review of Economics and Statistics. Jahrgang 63. Nr. 3/1981, Seite 471 ff.

Hamer, Günter/Engelmann, Margot. *Vierteljahresergebnisse der Sozialproduktsberechnung ab 1968*. In: Wirtschaft und Statistik. Ausgabe 1/1978, Seite 15 ff.

Harvey, A. C./Pierse, R. G. *Estimating missing observations in economic time series*. In: Journal of the American Statistical Association. Jahrgang 79. Nr. 385/1984, Seite 125 ff.

Lisman, J. H. C./Sandee, J. *Derivation of Quarterly Figures from Annual Data*. In: Journal of the Royal Statistical Society. Series C (Applied Statistics). Jahrgang 13. Nr. 2/1964, Seite 87 ff.

Litterman, Robert B. *A random walk, Markov model for the distribution of time series*. In: Journal of Business and Economic Statistics. Jahrgang 1. Nr. 2/1983, Seite 169 ff.

Pavía-Miralles, Jose Manuel. *A Survey of Methods to Interpolate, Distribute and Extrapolate time series*. In: Journal of Service Science and Management. Jahrgang 3. Nr. 4/2010, Seite 449 ff. DOI: 10.4236/jssm.2010.34051

Polasek, Wolfgang/Llano, Carlos/Sellner, Richard P. *Bayesian methods for completing data in spatial models*. In: Review of Economic Analysis. Jahrgang 2. Nr. 2/2010, Seite 194 ff.

Rinne, Horst/Specht, Katja. *Zeitreihen*. *Statistische Modellierung, Schätzung und Prognose*. München 2002.

## LITERATURVERZEICHNIS

Statistisches Bundesamt (Herausgeber). *Qualitätsbericht Volkswirtschaftliche Gesamtrechnungen*. Wiesbaden 2015. Verfügbar unter: www.destatis.de

Statistisches Bundesamt (Herausgeber). *Fachserie 18 Volkswirtschaftliche Gesamtrechnungen, Reihe 1.2 Inlandsproduktberechnung. Vierteljahresergebnisse*. Wiesbaden, diverse Jahrgänge.

Vullhorst, Udo. *Zur indikatorgestützten Berechnung des vierteljährlichen Bruttoinlandsprodukts für Baden-Württemberg*. In: Statistisches Monatsheft Baden-Württemberg. Ausgabe 9/2008, Seite 32 ff.

Vullhorst, Udo. *Von Jahres- zu Quartalswerten – ohne unterjährige Information*. In: Statistisches Monatsheft Baden-Württemberg. Ausgabe 5/2015, Seite 35 ff.

Wei, William W./Stram, Daniel O. *Disaggregation of time series models*. In: Journal of the Royal Statistical Society. Series B (Statistical Methodology). Jahrgang 52. Heft 3/1990, Seite 453 ff. DOI: 10.2307/2345669

**Herausgeber**
Statistisches Bundesamt, Wiesbaden
www.destatis.de

**Schriftleitung**
Dieter Sarreither, Präsident des Statistischen Bundesamtes
Redaktionsleitung: Kerstin Hänsel
Redaktion: Ellen Römer

**Ihr Kontakt zu uns**
www.destatis.de/kontakt

**Erscheinungsfolge**
zweimonatlich, erschienen im Oktober 2016
Das Archiv aller Ausgaben ab Januar 2001 finden Sie unter www.destatis.de/publikationen

**Print**
Einzelpreis: EUR 18,– (zzgl. Versand)
Jahresbezugspreis: EUR 108,– (zzgl. Versand)
Bestellnummer: 1010200-16005-1
ISSN 0043-6143
ISBN 978-3-8246-1047-1

**Download (PDF)**
Artikelnummer: 1010200-16005-4, ISSN 1619-2907

**Vertriebspartner**
IBRo Versandservice GmbH
Bereich Statistisches Bundesamt
Kastanienweg 1
D-18184 Roggentin
Telefon: +49 (0) 382 04 / 665 43
Telefax: +49 (0) 382 04 / 669 19
destatis@ibro.de

Papier: Metapaper Smooth, FSC-zertifiziert, klimaneutral, zu 61% aus regenerativen Energien