

Winker, Peter

Working Paper

Zur Messung der Lohndifferenzierung mit Entropiemaßen

Diskussionsbeiträge - Serie II, No. 158

Provided in Cooperation with:

Department of Economics, University of Konstanz

Suggested Citation: Winker, Peter (1991) : Zur Messung der Lohndifferenzierung mit Entropiemaßen, Diskussionsbeiträge - Serie II, No. 158, Universität Konstanz, Sonderforschungsbereich 178 - Internationalisierung der Wirtschaft, Konstanz

This Version is available at:

<https://hdl.handle.net/10419/101613>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

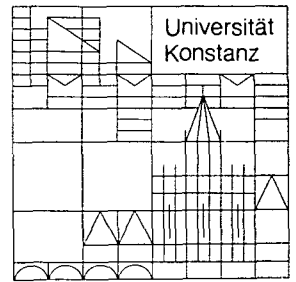
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Sonderforschungsbereich 178
„Internationalisierung der Wirtschaft“

Diskussionsbeiträge

Juristische
Fakultät

Fakultät für Wirtschafts-
wissenschaften und Statistik

Peter Winker

**Zur Messung der Lohndifferenzierung
mit Entropiemaßen**

Postfach 5560
D-7750 Konstanz

Serie II — Nr. 158
Oktober 1991

24. JAN. 1992

W 113 - 158 ~~11~~

**Zur Messung der Lohndifferenzierung
mit Entropiemaßen**

Peter Winker

Serie II - Nr. 158

Oktober 1991

Ich danke B. Fitzenberger, W. Franz, G. Heidbrink, U. Oser und W. Smolny für die anregende Diskussion einer ersten Fassung dieser Arbeit, sowie W. Eggert für die Unterstützung bei der Vorbereitung der Simulationen in Abschnitt 6.

Zusammenfassung

In dieser Arbeit wird die Entwicklung der intersektoralen Lohnstruktur in der BRD untersucht. Als Ziel weiterer Arbeiten steht dabei auch die Erklärung des Mismatchparameters für den Arbeitsmarkt in Mengentransportierungsmodellen im Hintergrund.

Hier wird zunächst axiomatisch ein Dispersionsmaß hergeleitet, das im wesentlichen auf der Entropiefunktion beruht. Es werden nichtparametrische Verfahren getestet und angewandt, die eine Beurteilung der Güte der berechneten Maßzahlen erlauben. Mit diesem Instrumentarium wird die Entwicklung der sektoralen Lohnstruktur in der BRD von 1970 bis 1989 analysiert. Insbesondere wird dabei die durch Betrachtung des Variationskoeffizienten induzierte Hypothese einer stetig zunehmenden intersektoralen Lohndifferenzierung verworfen.

1 Einführung

In der Diskussion um die Persistenz von Arbeitslosigkeit in den 80er Jahren gewann neben der absoluten Lohnhöhe auch die Lohnstruktur an Bedeutung.¹ Hierbei wurde insbesondere nach Zusammenhängen zwischen der Entwicklung der sektoralen, regionalen und qualifikatorischen Lohnstruktur und der gesamtwirtschaftlichen Beschäftigungsentwicklung gesucht, wie sie sich etwa im Mismatchparameter von Mengenrationierungsmodellen² darstellen.

Zwar fällt es verhältnismäßig leicht, anhand der Daten des Statistischen Bundesamtes – insbesondere aufgrund der seit 1976 vorliegenden Beschäftigtenstatistik – qualitative Aussagen über Veränderungen von Lohnstrukturen zu machen; doch ergeben sich erhebliche Schwierigkeiten, will man die Veränderungen von Lohnstrukturen als quantitatives Merkmal einsetzen. Dies liegt unter anderem daran, daß die benutzten Maßzahlen häufig nur das Verhältnis von jeweils zwei Gruppen beschreiben. Als Beispiel mögen die bei Bellmann und Buttler (1989) dargestellten Dispersionsindizes dienen.³ Im Extremfall wird sogar nur die Lohn Differenz zwischen den extremalen Lohngruppen betrachtet.⁴ Eine solche Betrachtung erlaubt keine Aussagen über die Veränderungen der anderen Lohngruppen, so daß diese Maßzahlen nur beschränkt als Indikatoren für Nivellierungstendenzen geeignet sind.⁵

Als globale Maß der Lohnstruktur findet sich bei Hardes (1988) und Gundlach (1986) für die Entwicklung der intersektoralen Lohnstruktur der Rangkorrelationskoeffizient, der aufgrund seines ordinalen Charakters ebenfalls nur begrenzt aussagekräftig scheint. Häufiger wird der Variationskoeffizient verwandt,⁶ dessen quantitative Interpretation jedoch problematisch ist.⁷ Insbesondere berücksichtigt der Variationskoeffizient die Stellung einzelner Gruppen in der Lohnstruktur nur indirekt.⁸ Diese Information ist jedoch essentiell, um festzustellen, ob die Veränderung der Lohnstruktur geeignet ist, Ungleichgewichte auf den beobachteten Arbeitsmärkten abzu-

¹Vgl. etwa Franz (1989), OECD (1985) oder Sengenberger (1986).

²Vgl. Smolny (1991).

³Vgl. auch Franz (1991), S. 321.

⁴Vgl. Soltwedel (1980).

⁵Vgl. Thiehoff (1987), S. 122.

⁶Er findet sich etwa in Franke (1985) für intersektorale und qualifikatorische Lohnstrukturen und in Koller (1987) für regionale Lohnstrukturen.

⁷Vgl. auch das Beispiel in Abschnitt 4.

⁸Dies gilt auch für das von Lampert (1963), S. 13, diskutierte Maß der summierten absoluten Abweichung vom Mittelwert.

bauen.

In den folgenden Abschnitten soll daher ein axiomatisch⁹ begründetes, globales Maß der Lohndifferenzierung für die o.g. Lohnstrukturen eingeführt werden. Zunächst wird dazu ein Anforderungskatalog an ein aussagekräftiges und leicht zu interpretierendes Maß aufgestellt. Danach wird das Maß zunächst für zwei Gruppen abgeleitet und eine Verallgemeinerung für n Gruppen vorgeschlagen.

Dem Dispersionsmaß liegt die Entropiefunktion als Maß der Information einer Verteilung zugrunde. Ich werde daher auch kurz auf die informationstheoretische Begründung dieser Funktion eingehen.

Für dieses Maß gilt wie für jedes andere globale Lohnstrukturmaß, daß sich in den beobachteten Werten verschiedene Ursachen vermischen, die nur theoretisch analytisch voneinander geschieden werden können.¹⁰ Allerdings setzt sich das unten definierte Maß aus drei Komponenten zusammen, die einzeln ausgewiesen eine genauere Analyse von Entwicklungen ermöglichen. Es handelt sich dabei um Größen für die Entwicklung der Beschäftigtenstruktur, des Lohngefüges im engeren Sinne und der relativen Lohnspanne.

Schließlich muß darauf hingewiesen werden, daß die Betrachtung von Lohnstrukturen einen Vergleich von Verteilungen für die einzelnen Lohngruppen impliziert. Die gängigen Statistiken ebenso wie die hier vorgestellte in der vorliegenden Form arbeiten jedoch nur mit den Mittelwerten bzw. den Medianen dieser Verteilungen. In einer späteren Arbeit wird daher eine Verallgemeinerung auf den Vergleich von Verteilungen versucht werden.¹¹

Ich werde das zunächst abzuleitende Maß an den Daten der Beschäftigtenstatistik für die intersektorale Lohnstruktur der Bundesrepublik Deutschland für den Zeitraum 1970-1989 erproben, um zu testen, inwieweit man damit zu ähnlichen Resultaten kommt wie mit gängigen Methoden.¹² Um die berechneten Maßzahlen statistisch bewerten zu können, ist ein Rückgriff auf nichtparametrische Schätzverfahren notwendig. Diese werden abschließend vorgestellt und Resultate für eine theoretische Verteilung ausgewiesen, die eine Anwendung im vorliegenden Fall als legitimiert erscheinen lassen.

In dieser Arbeit wird lediglich eine deskriptiv statistische Analyse der intersektoralen Lohnstruktur versucht. Auf die möglichen Bestimmungsgründe derselben wird nicht eingegangen. Dies bleibt ebenso weiteren Ar-

⁹Vgl. Ramser (1987) für eine Übersicht über Dispersionsmaße und ihre Klassifizierung.

¹⁰Vgl. Thiehoff (1987), S. 35, und Lampert (1963), S. 7.

¹¹Ein Ansatz zur Analyse von Lohnstrukturen auf der Basis von Individualdaten findet sich in Burda (1991).

¹²Vgl. hierzu auch Thiehoff (1987) und Gundlach (1986)

beiten überlassen wie die Analyse regionaler und qualifikatorischer Lohnstrukturen.¹³

2 Anforderungen für zwei Lohngruppen

Die im folgenden zu definierende Maßzahl soll das Ausmaß der Lohndifferenzierung zwischen zwei gegebenen Lohngruppen i und j ausdrücken. Hierbei können die Gruppen nach verschiedenen Kriterien gebildet werden, etwa sektoral, regional oder nach Qualifikation. Ich verwende als Symbolik W_i für den Lohnsatz der Lohngruppe i , B_i für die Anzahl der Beschäftigten in Lohngruppe i und $g_i := \frac{B_i}{B_i + B_j}$ für den relativen Beschäftigtenanteil in Lohngruppe i . Es gilt also insbesondere stets $g_i + g_j = 1$. Die Veränderungen der Maßzahl sollen eine Aussage über Nivellierungstendenzen ermöglichen.

Das von nun an mit $WD(W_i, W_j, B_i, B_j)$ oder kurz mit $WD(W_i, W_j, g_i)$ bezeichnete Maß muß also insbesondere den folgenden Anforderungen genügen:

- Bei zwei Lohngruppen soll das Maß der Lohndifferenzierung bei gegebener Lohnspanne und Beschäftigtenverteilung auf die Lohngruppen unabhängig davon sein, welche Lohngruppe den höheren Lohnsatz erhält. Dies führt zu der Forderung nach Symmetrie in i und j , d.h. $WD(W_i, W_j, g_i) = WD(W_j, W_i, g_j)$.
- Wenn für $W_i > W_j$ W_i ceteris paribus wächst, muß $WD(W_i, W_j, g_i)$ ebenfalls ansteigen. Analoges gilt für die anderen möglichen Bewegungen von W_i bzw. W_j .
- Nähert sich g_i dem Wert $\frac{1}{2}$, so entspricht dies ceteris paribus einer gleichmäßigeren Verteilung der Beschäftigten auf die beiden Lohngruppen, d.h. einer höheren Lohndifferenzierung. $WD(W_i, W_j, g_i)$ muß steigen. Die Umkehrung gilt, falls sich g_i von $\frac{1}{2}$ entfernt.
- Eine gleiche absolute Lohndifferenz muß bei hohen Lohnsätzen als vergleichsweise geringere Lohndifferenzierung gewertet werden, d.h. die Maßzahl sollte mit Steigen von $\max\{W_i, W_j\}$ fallen.
- Linearität in der absoluten Lohndifferenz erleichtert die Interpretation der Zahlen, d.h. verdoppelt sich $|W_i - W_j|$ bei Konstanz von

¹³Vgl. Franz (1991), S. 315ff.

$\max\{W_i, W_j\}$ und festem g_i , so soll sich dies in einer Verdopplung der Maßzahl widerspiegeln.

- Schließlich ermöglicht eine Normierung auf $[0, 1]$ den Vergleich von verschiedenen Verteilungen (etwa zwischen unterschiedlichen Staaten).

Zusätzlich zu diesen Kriterien sollte die Maßzahl natürlich auch den üblichen Mindestanforderungen¹⁴ an Verteilungsmaße genügen, d.h. es soll Unabhängigkeit vom Einkommensmaßstab und von der Gesamtzahl der Beschäftigten gewährleistet sein.¹⁵ Es wird jedoch keine Invarianz gegenüber beliebigen affinen Transformationen gefordert, da dies dem Argument zuwiderlaufen würde, daß eine gegebene, konstante, absolute Lohndifferenz eine umso geringere Lohndifferenzierung darstellt je höher das allgemeine Lohnniveau ist.

3 Ein Entropiemaß für zwei Lohngruppen

Die aufgestellten Anforderungen an eine Maßzahl führen dazu, daß in deren Definition die Gewichte g_i in einer funktionalen Form eingehen müssen, die symmetrisch zu $\frac{1}{2}$ ist und für den Wert $g_i = \frac{1}{2}$ ihr Maximum erreicht. Diese Funktion $H(\cdot)$ beschreibt dann den Informationsgehalt der Beschäftigtenverteilung auf die Lohngruppen.

Daher bietet sich für $H(\cdot)$ die aus der Informationstheorie bekannte Entropiefunktion an:¹⁶ Für ein zufällig mit Wahrscheinlichkeit p eintretendes Ereignis E möchte man den Informationsgehalt der Nachricht bestimmen, daß E tatsächlich eingetreten ist. Offenbar ist dieser umso geringer je höher die a priori Wahrscheinlichkeit des Eintretens war. Dies führt auf eine monoton fallende Funktion $h(p)$. Beobachtet man nun nacheinander zwei Ereignisse E und F mit Eintrittswahrscheinlichkeiten $P[E]$ bzw. $P[F]$, dann gilt für die Wahrscheinlichkeit des Auftretens von E und F :

$$P[E \text{ und } F] = P[E] P[F | E] = P[F] P[E | F]$$

Das Maß des Informationsgehaltes sollte der Nachricht, daß E und F eingetreten sind, die Summe der Informationsgehalte der Nachricht, daß E

¹⁴Vgl. Ramser(1987), S. 23f.

¹⁵D.h. weder die Multiplikation aller Lohnsätze mit einem konstanten Faktor noch die identische Reduplikation der Beschäftigten sollen zu einer Veränderung des Dispersionsindex führen.

¹⁶Zur Entropie als Informationsmaß vgl. auch Rényi (1966), S. 446, Jaglom (1984), S. 57, Theil (1971), S. 636ff oder Guiaşu (1977).

eingetreten ist, und der anschließenden Information, daß auch F eingetreten ist, zuweisen. D.h. man fordert die Additivität des Informationsmaßes, also $h(P[E]) + h(P[F | E]) = h(P[E \text{ und } F])$. Aus dieser Forderung und einigen Axiomen¹⁷ folgt, daß gelten muß: $h(p) = \log_2(\frac{1}{p})$

Der Informationsgehalt der Nachricht, daß E eingetreten ist, ist nicht gleich dem der komplementären Nachricht, daß E nicht eingetreten ist. Für diese gilt: $h(1 - p) = \log_2(\frac{1}{1-p})$. Für den vor dem Eintreffen der Nachricht erwarteten Informationsgehalt ergibt sich somit:

$$\begin{aligned} H(p) &:= ph(p) + (1 - p)h(1 - p) \\ &= p \log_2 \frac{1}{p} + (1 - p) \log_2 \frac{1}{1-p} \end{aligned}$$

Diese Funktion wird als *Entropie*¹⁸ der Verteilung bezeichnet, die zwei Punkten die Wahrscheinlichkeiten p und $1 - p$ zuordnet. Die Betrachtung dieser Funktion als Maß des Informationsgehaltes einer Verteilung geht auf die für die moderne Informationstheorie bahnbrechende Arbeit von Shannon(1948) zurück. Die in Schaubild 1 dargestellte Funktion läßt sich wegen $\lim_{p \rightarrow 0} p \log_2 \frac{1}{p} = 0$ durch 0 stetig für $p = 0$ und $p = 1$ fortsetzen.

Sie ist beliebig oft differenzierbar in $]0, 1[$, symmetrisch in p und $1 - p$ und es gilt:

$$\max_{0 \leq p \leq 1} H(p) = H(\frac{1}{2}) = 1$$

Anstelle der „schönen“ Funktion H könnte zunächst auch jede andere Funktion $f : [0, 1] \rightarrow [0, 1]$ verwandt werden, die symmetrisch um $\frac{1}{2}$ und streng monoton wachsend auf dem Intervall $]0, \frac{1}{2}[$ ist. Allerdings wäre dann weder die Interpretation als Informationsmaß der Verteilung gegeben noch die direkte Verallgemeinerung auf den Fall von n Lohngruppen möglich.¹⁹

Man kann nun die Maßzahl wie folgt definieren:

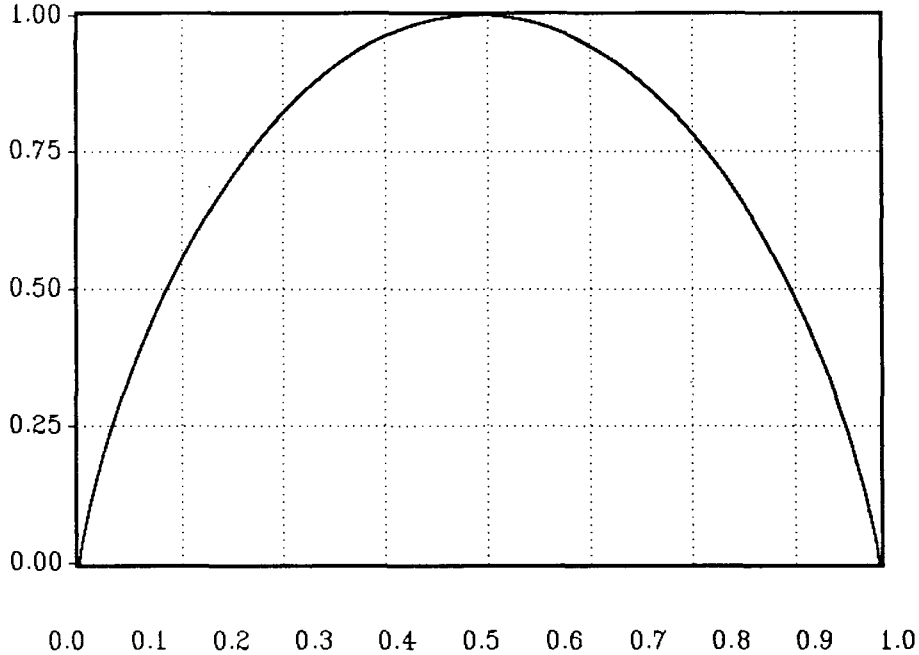
$$WD(W_i, W_j, g_i) := H(g_i) \frac{|W_i - W_j|}{\max\{W_i, W_j\}} \quad (1)$$

¹⁷Es läßt sich zeigen, daß folgende Postulate genügen, um zwangsweise zu der folgenden funktionalen Form zu gelangen: $H(\cdot)$ hängt nur von der Wahrscheinlichkeitsverteilung ab, ist stetig in p mit Maximum für $p = 1/2$ und erfüllt die Anforderung der Additivität (Vgl. Rényi, S. 441.).

¹⁸Der Begriff der Entropie stammt ursprünglich aus der theoretischen Physik.

¹⁹Vgl. Abschnitt 4.

Abbildung 1: Entropiefunktion für $n = 2$



Wegen $H(x) = H(1 - x)$ für $x \in]0, 1[$ gilt

$$\begin{aligned} WD(W_i, W_j, g_i) &= H(g_i) \frac{|W_i - W_j|}{\max\{W_i, W_j\}} \\ &= H(g_j) \frac{|W_j - W_i|}{\max\{W_j, W_i\}} = WD(W_j, W_i, g_j). \end{aligned}$$

D.h. WD ist symmetrisch in i und j . Für $W'_i > W_i \geq W_j$ und festes g_i erhält man²⁰

$$\frac{|W'_i - W_j|}{\max\{W'_i, W_j\}} = \frac{|W'_i - W_j|}{W'_i} \geq \frac{|W_i - W_j|}{W_i} = \frac{|W_i - W_j|}{\max\{W_i, W_j\}}$$

also $WD(W'_i, W_j, g_i) > WD(W_i, W_j, g_i)$. Bei gegebener Beschäftigtenverteilung und Ansteigen des Lohnsatzes in der oberen Lohngruppe steigt also auch WD wie gefordert. Für $\delta > 0$ gilt offenbar

$$\frac{1}{\max\{W_i + \delta, W_j + \delta\}} < \frac{1}{\max\{W_i, W_j\}}$$

²⁰ Da $n \leq k \Rightarrow \frac{n}{k} \leq \frac{n+\delta}{k+\delta}$ für $\delta > 0$.

und somit $WD(W_i + \delta, W_j + \delta, g_i) < WD(W_i, W_j, g_i)$. Steigt also der Lohnsatz in allen Lohngruppen gleichmäßig an, so verringert sich die ausgewiesene Lohndifferenzierung aufgrund des dann höheren allgemeinen Lohnniveaus. Schließlich überzeugt man sich durch Einsetzen in (1), daß für alle $\lambda \in]0, +\infty[$ gilt:

$$WD(\lambda W_i, \lambda W_j, g_i) = WD(W_i, W_j, g_i)$$

und

$$WD(W_i, W_j, \lambda B_i, \lambda B_j) = WD(W_i, W_j, B_i, B_j).$$

$WD(W_i, W_j, g_i)$ erfüllt also alle oben aufgeführten Anforderungen und kann als Maßzahl für Lohnnivellierungstendenzen ins Auge gefaßt werden.

4 Verallgemeinerung auf n Lohngruppen

Durch die Verallgemeinerung der oben eingeführten Entropiefunktion ist es möglich, die Lohnstrukturmaßzahl WD für n Lohngruppen zu definieren.²¹ Für einen Gewichtsvektor²² (x_1, \dots, x_n) (d.h. $\sum_{i=1}^n x_i = 1$ und $x_i \in]0, 1[$ für $i = 1, \dots, n$) sei die Entropie H_n wie folgt definiert:²³

$$H_n(x_1, \dots, x_n) := \sum_{i=1}^n x_i \log_n \frac{1}{x_i}$$

H_n kann wie im Fall von zwei Variablen stetig in den Rand des Einheits-simplex fortgesetzt werden.²⁴ H_n ist also insbesondere stetig, beliebig oft differenzierbar im Inneren des Polyquaders und es gilt:

$$\max H_n(x_1, \dots, x_n) = H_n\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = 1$$

Ferner hat die allgemeine Entropiefunktion die Symmetrieeigenschaft

$$H_n(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = H_n(x_1, \dots, x_j, \dots, x_i, \dots, x_n),$$

²¹Zur Herleitung der allgemeinen Entropiefunktion vgl. Feinstein (1958) oder Theil(1971), S. 640f.

²²In der Informationstheorie spricht man eher von Wahrscheinlichkeitsvektoren.

²³Diese Definition unterscheidet sich von der Standardfassung durch die verwendete Basis n des Logarithmus, um eine Skalierung auf $[0, 1]$ zu gewährleisten.

²⁴Hierzu verwendet man $\lim_{x \rightarrow 0} x \log \frac{1}{x} = 0$.

und es gilt schließlich noch:

$$H_n(0, \dots, 0, 1, 0, \dots, 0) = 0$$

Wie im Fall für zwei Lohngruppen bezeichne W_i den Lohn und B_i die Anzahl der Beschäftigten in Lohngruppe i . Der Anteil der Beschäftigten g_i in Lohngruppe i ergibt sich dann aus der Gleichung:

$$g_i := \frac{B_i}{\sum_{j=1}^n B_j}$$

Als natürliche Verallgemeinerung des Entropiemaßes der Beschäftigtenverteilung auf zwei Lohngruppen $H(g_i)$ erhält man nun das Entropiemaß der Beschäftigtenverteilung auf n Lohngruppen $BV := H_n(g_1, \dots, g_n)$, welches wiederum den Informationsgehalt der Beschäftigtenverteilung widerspiegelt.

Der Ausdruck $|W_i - W_j|$ aus (1) kann nicht unmittelbar auf den Fall von n Lohngruppen übertragen werden, da hierin die Information über das Lohngefüge und die Lohnspanne vereinigt sind. Für mehr als 2 Lohngruppen lassen sich diese Größen getrennt ausweisen. Ohne Beschränkung der Allgemeinheit seien die Lohngruppen in steigender Folge angeordnet, d.h. es gelte $W_1 \leq W_2 \leq \dots \leq W_n$. Für das Lohngefüge betrachtet man zunächst die $n - 1$ relativen Lohndifferenzen²⁵

$$d_i := \frac{W_{i+1} - W_i}{W_n - W_1}.$$

Es läge nun nahe, den Informationsgehalt des Lohngefüges durch die Entropie der Verteilung der relativen Lohndifferenzen $H_{n-1}(d_1, \dots, d_{n-1})$ zu messen. Dieses Maß wäre jedoch stark abhängig von n , da mit einer wachsenden Anzahl von Lohngruppen W_n , d_{n-1} und d_1 immer größer und W_1 immer kleiner würde, d.h. die Entropie würde abnehmen. Da damit ein Vergleich von Verteilungen mit einer unterschiedlichen Anzahl von Lohngruppen nicht mehr möglich wäre, wähle ich eine leichte Abwandlung, die eine geringere Abhängigkeit von n und mehr Robustheit gegenüber „Ausreißern“ erwarten läßt.

Hierzu schließt man an beiden Enden der Verteilung von W_1, \dots, W_n ein vorgegebenes Quantil (etwa 5%), zumindest aber einen Lohnsatz aus und erhält eine gekürzte Folge $W_{n_1} \leq W_{n_1+1} \leq \dots \leq W_{n_2}$.²⁶ Sei $m := n_2 - n_1$

²⁵D.h. es gilt $\sum_{i=1}^n d_i = 1$.

²⁶Dies hat den zusätzlichen Vorzug, daß die in Abschnitt 6 benutzten nichtparametrischen Schätzverfahren für Quantile asymptotisch konsistent sind im Gegensatz zu entsprechenden Schätzern für Maximalwerte. Siehe hierzu Bickel und Freedman (1981) und Efron und Tibshirani (1986), S. 61.

und berechnet man nun analog zum obigen Vorgehen für $1 \leq i \leq m$

$$d_i := \frac{W_{i+1} - W_i}{W_{n_2} - W_{n_1}},$$

so ist $LG := H_m(d_1, \dots, d_m)$ ein Schätzer des Informationsgehaltes der Lohnsatzverteilung.

Wünschenswert wäre natürlich ein zumindest asymptotisch von n unabhängiges Maß des Informationsgehaltes der Beschäftigten- bzw. Lohnsatzverteilung.

Ein solches Maß wäre etwa für die durch ihre Dichte $f(x)$ gegebene Lohnsatzverteilung wie folgt konstruierbar: Man teilt die Menge der möglichen Ausprägungen (hier etwa das Intervall $]0, +\infty[$) in äquidistante Intervalle der Länge $\frac{1}{N}$ ein. Für $k \in \mathbb{N}$ sei $p_{Nk} := P(W \in [\frac{k}{N}, \frac{k+1}{N}[$ die Wahrscheinlichkeit für W im Intervall $[\frac{k}{N}, \frac{k+1}{N}[$) zu liegen. Diese wird durch die relativen Häufigkeiten für die gezogene Stichprobe approximiert. Es gilt dann folgendes asymptotische Resultat:²⁷

$$\lim_{N \rightarrow \infty} \left(\sum_{k=1}^{\infty} p_{Nk} \log_2 \frac{1}{p_{Nk}} - \log_2 N \right) = \int_0^{\infty} f(x) \log_2 \frac{1}{f(x)} dx,$$

falls das Integral auf der rechten Seite der Gleichung existiert. Dieses Maß wäre also insbesondere asymptotisch von n unabhängig. Um vernünftige Approximationen zu erhalten müßte aber einerseits N groß sein und andererseits die Größe n der gezogenen Stichprobe wiederum ein Vielfaches von N betragen, um gute Schätzer für p_{Nk} zu erhalten. Diese Bedingungen sind bei Untersuchung der Lohnstruktur auf Grundlage von aggregierten Daten nicht gegeben. Dagegen ist an eine Anwendung dieses Maßes bei der Betrachtung von Lohnstrukturen auf der Basis von Individualdaten zu denken. Dies bleibt weiteren Arbeiten zum Thema vorbehalten.

Der Skalierungsfaktor $LS := \frac{W_{n_2} - W_{n_1}}{W_{n_2}}$ ist schließlich ein asymptotisch von n unabhängiger Schätzer für die relative Lohnspanne, die zusammen mit LG dem Term $\frac{|W_i - W_j|}{\max\{W_i, W_j\}}$ in (1) entspricht.

Durch die Normierung aller drei Komponenten auf das Intervall $[0, 1]$ wird ein Vergleich für verschiedene Lohnstrukturen möglich, solange diese auf einer vergleichbaren Klasseneinteilung beruhen und die mögliche Verzerrung durch unterschiedliche Stichprobengrößen berücksichtigt wird. Allerdings muß darauf hingewiesen werden, daß der Verzicht auf die Berücksich-

²⁷Vgl. Rényi (1966), S. 476.

tigung extremer Quantile dazu führt, daß nur Lohnstrukturen mit mindestens etwa 10 Lohngruppen sinnvoll untersucht werden können.

Zusammenfassend ergibt sich nun mit den Bezeichnungen von oben die Maßzahl für n Lohngruppen wie folgt:

$$\begin{aligned} WD(W_1, \dots, W_n, B_1, \dots, B_n) &:= WD(W_1, \dots, W_n, g_1, \dots, g_n) \\ &:= H_n(g_1, \dots, g_n) \cdot H_m(d_1, \dots, d_m) \cdot \frac{W_{n_2} - W_{n_1}}{W_{n_2}} \quad (2) \\ &:= BV \cdot LG \cdot LS \end{aligned}$$

Diese verallgemeinerte Maßzahl der Lohndifferenzierung genügt den im Abschnitt 2 aufgeführten Anforderungen, soweit diese auf den Fall von n Lohngruppen übertragbar sind. Man erhält in der Maßzahl WD somit einen Indikator, der drei wesentliche Bestimmungsgrößen von Lohnstrukturen erfaßt. In $BV := H_n(g_1, \dots, g_n)$ wird die Information über die Besetzung der einzelnen Lohngruppen gemessen, $LG := H_m(d_1, \dots, d_m)$ mißt die Information²⁸ über die relativen Lohnpositionen und der Skalierungsfaktor $\frac{W_{n_2} - W_{n_1}}{W_{n_2}}$ drückt die relative Spannweite der Lohnpositionen aus. Für empirische Untersuchungen lassen sich diese drei Einflußgrößen natürlich auch getrennt ausweisen und ermöglichen damit eine detailliertere Interpretation des Gesamtindikators WD . Dann entspricht dem Konzept eines von der Beschäftigtenverteilung unabhängigen Lohnstrukturmaßes, wie es sich etwa im bereinigten Variationskoeffizienten²⁹ findet, der Term LG aus dem WD -Maß. Im Gegensatz zu diesem Entropiemaß lassen sich im Variationskoeffizienten die Einflüsse der einzelnen Komponenten nicht direkt ablesen. Außerdem genügt der Variationskoeffizient nicht der Forderung nach Symmetrie und nach Anwachsen mit einer zunehmenden Gleichverteilung der Beschäftigten auf die einzelnen Lohngruppen.

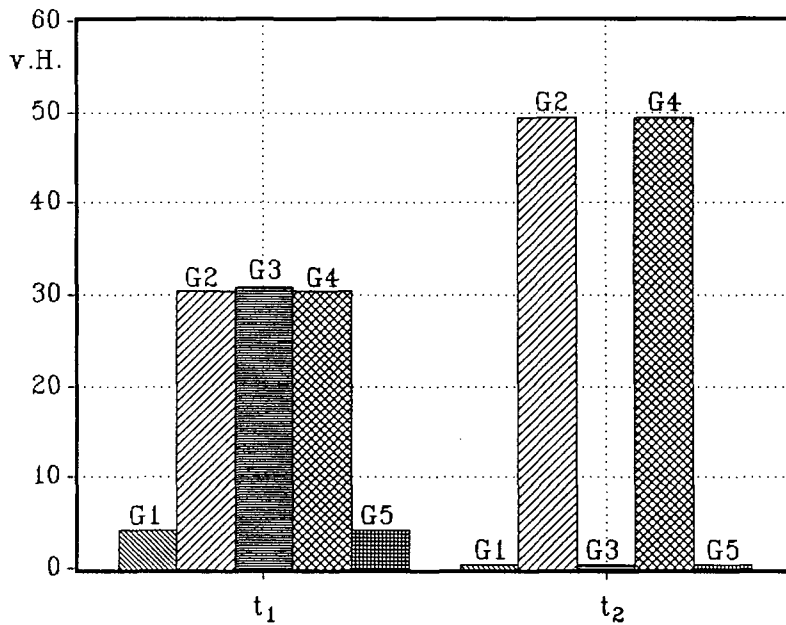
Im folgenden hypothetischen Beispiel für fünf Leistungsgruppen (G1 – G5) weisen die hier eingeführte Maßzahl WD und der Variationskoeffizient ein gegenläufiges Verhalten auf. Im Beispiel wurden für die beiden Zeitpunkte t_1 und t_2 gleiche Lohnsätze in Höhe von 1DM für G1, 2DM für G2, ..., 5DM für G5 unterstellt. D.h. sowohl das Lohngefüge als auch die absolute Lohnspanne blieb konstant. Lediglich die Verteilung der Beschäftigten hat sich wie in Schaubild 2 ausgewiesen verändert. Während WD klei-

²⁸Zur informationstheoretischen Interpretation und Herleitung der Entropiefunktion vgl. auch Topsøe(1974).

²⁹D.h. der Variationskoeffizient wird für eine unterstellte konstante Beschäftigtenstruktur berechnet. Vgl. Thiehoff (1987), S. 183 und Gerfin (1978).

ner wird, also dem optischen Eindruck einer zunehmenden Nivellierung entspricht, steigt der Variationskoeffizient an. Die Teilmaßzahlen für das Lohngefüge und die relative Lohnspanne in WD bleiben dabei konstant, da die Veränderungen nur aus der Beschäftigtenverteilung herrühren.

Abbildung 2: Zwei hypothetische Beschäftigtenverteilungen



Natürlich läßt sich aus diesem einfachen Beispiel keine Aussage der Art ableiten, daß WD das „bessere“ Maß der Lohndifferenzierung sei. Doch spricht das unterschiedliche Verhalten in diesem Beispiel³⁰ dafür, bei der Analyse von Lohnstrukturen auch Maße auf der Basis von Entropiefunktionen hinzuzuziehen.³¹ Einer solchen Untersuchung, dem Vergleich mit dem

³⁰ Ein hypothetische Beispiel mit ähnlichen Merkmalen läßt sich auch leicht für den Fall einer ausschließlichen Änderung des Lohngefüges bei konstanter Beschäftigtenverteilung generieren.

³¹ Schon in einer der ersten Untersuchungen der Lohnstruktur für die Bundesrepublik Deutschland weist Lampert (1963) darauf hin, „...daß sich zur Darstellung einer Lohnstruktur die Verwendung mehrerer Maße nebeneinander empfiehlt, weil jedes Maß spezifische Vor- und Nachteile hat.“

Variationskoeffizienten und der Frage der statistischen Interpretation der Ergebnisse sind die beiden folgenden Abschnitte gewidmet.

5 Die sektorale Lohnstruktur der BRD 1970–89

Für die Anwendung dieses Lohnstrukturmaßes bot sich aufgrund der vergleichsweise günstigen Datenlage die sektorale Lohnstruktur der BRD an.³² Aufgrund von Quartalsdaten von 1970 bis 1989 für Beschäftigte³³ und für Bruttostundenlöhne männlicher Arbeiter³⁴ wurde das oben eingeführte Informationsmaß berechnet. Zugrundegelegt wurden dabei 13 Sektoren nach der SYPRO-Gliederung.³⁵ Obwohl die verwendeten Primärdaten u.U. zu Verzerrungen der Maßzahl führen können, insbesondere da die tatsächlich geleisteten Arbeitsstunden nicht eingehen, erlauben sie doch eine erste Aussage über die Qualität des neuen Maßes. Zum Vergleich wurde auch die Zeitreihe der Variationskoeffizienten berechnet. Die Entwicklung beider Maße im Zeitablauf von 1970 bis 1989 ist in Schaubild 3 ausgewiesen.³⁶

Der Variationskoeffizient weist über die beobachtete Zeitspanne eine ständige Zunahme der intersektoralen Lohndifferenzierung aus, wenn man einmal von dem leicht rückläufigen Verlauf für die Jahre 1972 und 1973 absieht. Demgegenüber zeigt das *WD*-Maß und insbesondere auch seine drei Komponenten ein deutlich differenzierteres Bild. Zunächst läßt sich hier keine gleichmäßige Zunahme der Differenzierung der Lohnstruktur erkennen. Vielmehr beobachtet man für 1973 eine deutliche Nivellierungstendenz, die – wie die Schaubilder 4 und 5 zeigen – ihren Ursprung sowohl im Lohngefüge wie auch in der relativen Lohnspanne hat. Eine Ursache dafür sind sicher die überdurchschnittlichen Lohnsteigerungen der Niedriglohnbranchen. Bei

³²Einen Überblick über verschiedene Untersuchungen zur intersektoralen Lohnstruktur geben auch Gerlach und Hübler (1985).

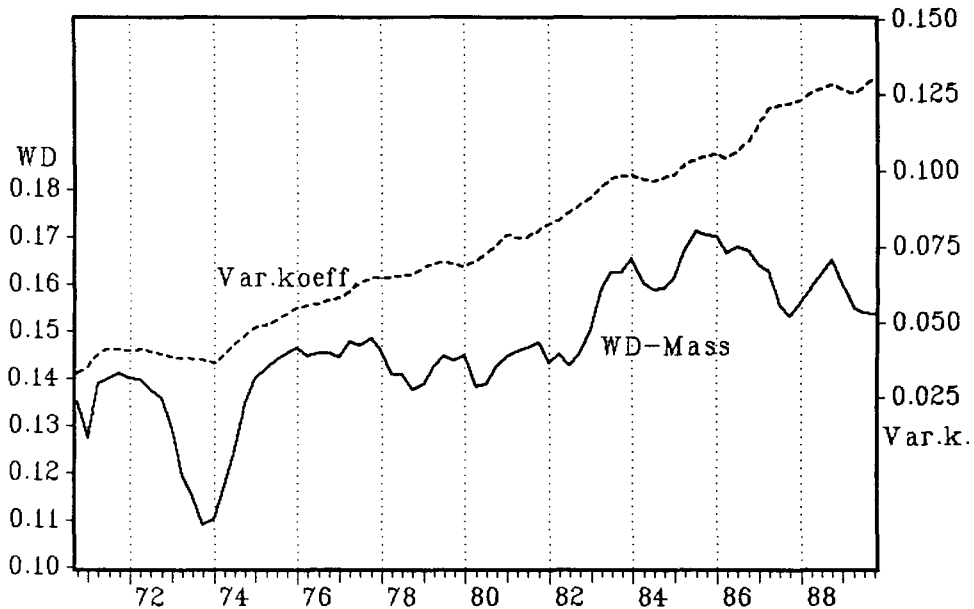
³³Quelle: Statistisches Bundesamt, Wiesbaden, Fachserie 4, Reihe 4.1. Wegen der Änderung der Berichtserstattung zum Januar 1977 wurden für den Zeitraum von 1970 bis 1976 die Daten der Rückrechnung absoluter Zahlen des Monatsberichts im Bergbau und Verarbeitenden Gewerbe des Statistischen Bundesamtes, Wiesbaden 1978, verwandt.

³⁴Quelle: Statistisches Bundesamt, Wiesbaden, Fachserie 16, Reihe 2.1.. Die Daten sind um Überstundenzuschläge bereinigt.

³⁵Im einzelnen waren dies: Bergbau (21), Eisenschaffende Industrie (27), Maschinenbau (32), Straßenfahrzeugbau (33), Elektrotechnik (36), Eisen-, Blech- und Metallwaren (38), Chemische Industrie (40), Holzverarbeitung (54), Papier- und Pappverarbeitung (56), Kunststoffwaren (58), Textilgewerbe (63), Bekleidungsindustrie (64) und Nahrungs- und Genußmittelindustrie (68+69).

³⁶Dargestellt sind jeweils gleitende Durchschnitte vierter Ordnung.

Abbildung 3: Lohnstrukturmaße

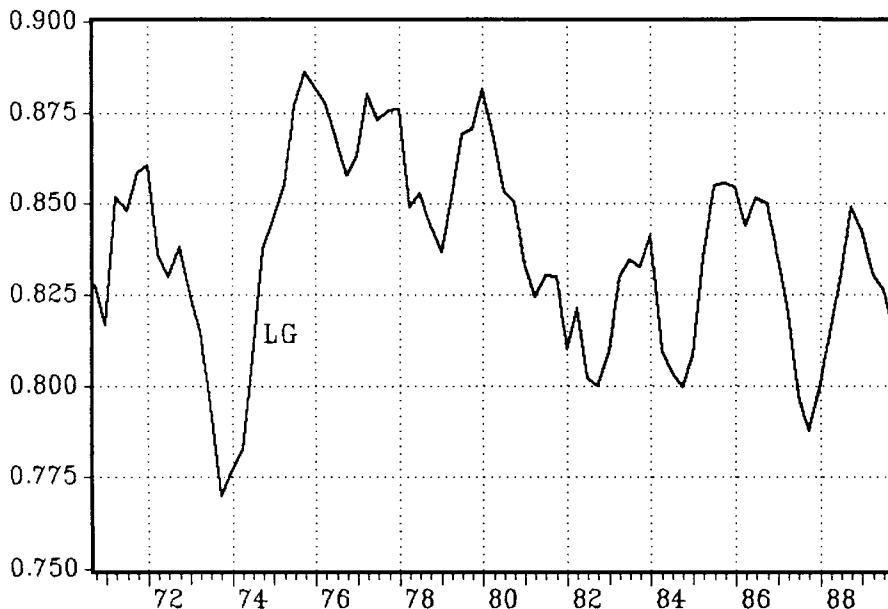


einer steigenden Inflationsrate wie nach dem ersten Ölpreisschock ergibt sich dann die Möglichkeit einer zunehmenden Lohndifferenzierung bei gleichzeitigem starken Ansteigen aller Nominallöhne.

Der Anstieg des *WD*-Maßes 1983 entsteht aufgrund eines umgekehrten Mechanismus. Schaubild 4 zeigt, daß nach einem vorangegangenen Rückgang der Differenzierung des Lohngefüges nun wieder eine stärkere Differenzierung erfolgt. Dieser Effekt wird verstärkt durch eine stark wachsende Lohnspanne, wie sie Schaubild 5 ausweist. Beim Vergleich der beiden Zeitreihen für die Differenzierung des Lohngefüges und die relative Lohnspanne stellt man fest, daß der hier zu zwei Zeitpunkten festgestellte Gleichlauf beider Variablen nicht der Regelfall ist.

Für das Entropiemaß *BV* der Beschäftigtenverteilung auf die einzelnen Sektoren zeigt sich schließlich eine generelle Nivellierungstendenz, d.h. Konzentration auf gewisse Sektoren, die jedoch quantitativ nicht so stark ausgeprägt ist wie die Schwankungen in den beiden anderen Komponenten *LG* und *LS* des *WD*-Maßes.

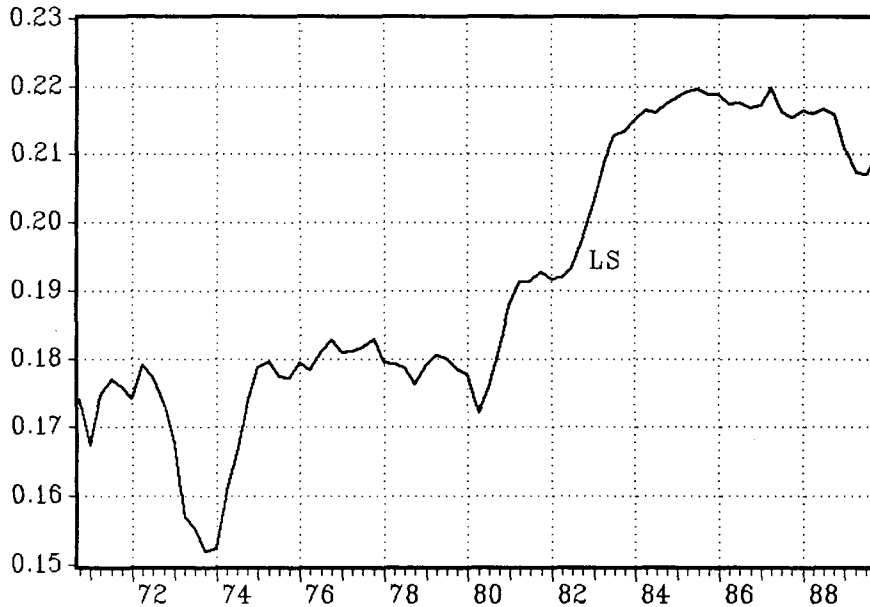
Abbildung 4: Lohngefügekoeffizient



Zusammenfassend läßt sich konstatieren, daß die im Variationskoeffizienten ausgewiesene deutliche Zunahme der intersektoralen Lohndifferenzierung im Zeitablauf durch das *WD*-Maß nicht bestätigt wird.³⁷ Andererseits erlauben die einzelnen Komponenten des *WD*-Maßes eine differenziertere Analyse der Entwicklung der Lohnstruktur. Im folgenden Abschnitt wird nun auf die statistische Bewertung der berechneten Maßzahlen eingegangen. Insbesondere wird ein nichtparametrisches Verfahren vorgestellt, um Standardabweichungen und Konfidenzintervalle für die berechneten Maßzahlen zu bestimmen. Mit diesem Verfahren wird auch ein Test auf signifikante Veränderungen der Maßzahlen konstruiert. Außerdem wird ein Test auf signifikante Veränderungen von reinen Entropiemaßen wie *BV* oder *LG* hergeleitet.

³⁷Die beobachteten qualitativen Unterschiede der Entwicklung findet man auch bei Betrachtung des bereinigten Variationskoeffizienten auf Basis der Beschäftigtenverteilung von 1970.1 und eines bereinigten Entropiemaßes, für das die Komponente *BV* konstant mit ihrem Wert für 1970.1 fortgeschrieben wird.

Abbildung 5: Lohnspannenkoeffizient



6 Zur Bewertung der empirischen Ergebnisse

Die Bewertung und Interpretation der empirischen Ergebnisse – insbesondere der beobachteten Veränderungen – erfordert neben der Zeitreihe der Maßzahl selbst auch eine Aussage über die „Qualität“ dieser Reihe. Üblicherweise wird in der Statistik die „Qualität“ eines Verteilungsmaßes durch seine Standardabweichung bzw. durch die Angabe eines Konfidenzintervalles zu einem vorgegebenen Konfidenzniveau beschrieben.

Selbst wenn man plausible Annahmen über die Verteilung der Grundgesamtheit besitzt, ist es jedoch in der Regel nicht möglich, daraus analytisch die Verteilung der betrachteten Statistik zu bestimmen. Dies gilt insbesondere für das hier diskutierte Lohndifferenzierungsmaß ebenso wie für den gebräuchlicheren Variationskoeffizienten. Daher werden in der Regel für den Variationskoeffizienten auch keine Standardabweichung oder Konfidenzintervalle ausgewiesen. Um trotz der geschilderten Schwierigkeit Schätzwerte für Standardabweichung und Konfidenzintervalle der gemessenen Statisti-

ken angeben zu können, werden nichtparametrische Verfahren verwandt. Es handelt sich dabei um den auf Quenouille (1949, 1956) und Tukey (1958) zurückgehenden Jackknife-Schätzer der Standardabweichung und das von Efron (1979) eingeführte Bootstrap-Verfahren.³⁸

Beide Verfahren beruhen auf folgendem Prinzip: aus den zur Berechnung des betrachteten Parameters verwandten Daten werden Teilmengen benutzt, um für diese den gleichen Parameter zu berechnen. Daraus erhält man eine empirische Verteilung des Parameters, die in den beobachteten Daten begründet ist. Diese Verteilung wird nun benutzt, um Schätzer für die Standardabweichung des berechneten Parameters zu erhalten.

Für den Jackknife-Schätzer erhält man die benutzten Teilmengen dadurch, daß man jeweils einen Datenpunkt unberücksichtigt läßt. In der hier betrachteten Anwendung erhält man somit bei n Lohngruppen neben der Maßzahl selbst eine Verteilung von n Schätzwerten $\hat{\theta}_{(i)}$ für $n - 1$ der n Lohngruppen. Daraus berechnet sich dann der Jackknife-Schätzer für die Standardabweichung des entsprechenden Lohnstrukturmaßes als

$$\hat{\sigma}_{JACK} := \left(\frac{n-1}{n} \sum_{i=1}^n (\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)})^2 \right)^{\frac{1}{2}},$$

wo $\hat{\theta}_{(\cdot)}$ das arithmetische Mittel der $\hat{\theta}_{(i)}$ ist.

Für den Bootstrap-Schätzer wird eine größere Menge von Teilmengen der beobachteten Daten zugelassen. Es wird zufällig eine große Zahl (in der Regel mindestens 1000) von Teilmengen der Größe n aus den Daten gezogen; die Ziehung erfolgt hierbei „mit Zurücklegen“, d.h. einzelne Datenpunkte können in den gezogenen Teilmengen wiederholt auftreten. Für die gezogenen Teilmengen wird wieder der Wert des Parameters berechnet. Man erhält somit eine empirische Verteilung, die aufgrund der großen Zahl von Beobachtungen neben der Schätzung der Standardabweichung auch die Schätzung von Konfidenzintervallen zuläßt.³⁹ Die unmittelbare Schätzung von Konfidenzintervallen aus der Bootstrap-Verteilung läßt sich in einigen Fällen durch die Berücksichtigung von Asymmetrie verbessern. Diese Methode wird von Efron⁴⁰ als „bias-corrected percentile method“ eingeführt. Hall

³⁸ Für eine Einführung in diese und weitere nichtparametrische Schätzverfahren vgl. auch Efron (1982). Eine gute Einführung in die Thematik bietet auch Stine (1989). Amemiya (1986) beschreibt kurz eine Anwendung des Bootstrap-Verfahrens zur Schätzung von Konfidenzintervallen für NLS-Schätzer.

³⁹ Vgl. hierzu auch Efron (1982), S. 78.

⁴⁰ Vgl. Efron (1982), S. 82.

(1988) vergleicht die asymptotischen Eigenschaften verschiedener Bootstrap-Schätzer für Konfidenzintervalle. Er kommt zu dem Ergebnis, daß in einer großen Klasse von Fragestellungen die aus der Bootstrap-Verteilung von $\sqrt{(n)}(\hat{\theta}^* - \hat{\theta})/\sigma^*$ geschätzten t-percentile Konfidenzintervalle bessere Approximationseigenschaften haben als die einfache percentile method, vorausgesetzt, es liegen stabile Schätzer für σ^* vor. Dies ist im Fall des Variationskoeffizienten durch den Jackknife-Schätzer gegeben, so daß hier auch t-percentile Konfidenzintervalle geschätzt werden können.

Läßt sich die untersuchte Maßzahl als Wert eines Funktionals der (unbekannten) Verteilung der Grundgesamtheit darstellen, so liefert Bootstrapping eine Monte Carlo Approximation des nichtparametrischen Maximum Likelihood Schätzers der interessierenden Größe, falls es eine Transformation auf die Standardnormalverteilung gibt. Diese braucht jedoch nicht bekannt zu sein.⁴¹ D.h. gehen sowohl die Größe der gezogenen Stichprobe als auch die Anzahl der Bootstrap-Ziehungen aus der Stichprobe gegen unendlich, so konvergiert der resultierende Bootstrap-Schätzer gegen den tatsächlichen Wert des Parameters. Dieses „Bootstrap-Paradigma“ ist für den Variationskoeffizienten offensichtlich erfüllt, da dieser der Quotient aus Varianz und Mittelwert der entsprechenden Verteilung ist. Das Maß der relativen Lohnspanne LS ist asymptotisch die Differenz zweier Quantilsgrenzen der Verteilung geteilt durch eine Quantilsgrenze, läßt sich also vermöge der Inversen der kumulierten Dichtefunktion der Verteilung als Funktional der Verteilung darstellen. LS genügt also auch dem Bootstrap-Paradigma. Für Entropiemaße gilt dies in der strengen Form nur für die in Abschnitt 4 diskutierte Alternative zu den hier verwandten Maßen BV und LG , die jedoch für kleine Stichproben nicht anwendbar ist. Doch läßt die im folgenden dargestellte Simulation eine Anwendung auf BV und LG im vorliegenden Fall als legitimiert erscheinen.

Die Güte der geschätzten Intervalle wurde für das hier verwandte Entropiemaß der Lohnsatzverteilung LG und den Variationskoeffizienten durch Simulation auf einer Lognormalverteilung geprüft. Die Ergebnisse sind in Tabelle 1 zusammengefaßt.

Für die empirischen Werte wurden 1.000.000 mal jeweils 13 identisch unabhängig lognormalverteilte⁴² Zufallszahlen gezogen und daraus der Va-

⁴¹Vgl. Efron (1982), S. 33. Ein Beweis der asymptotischen Konsistenz der Bootstrap-Schätzer für einige Klassen von Funktionalen findet sich in der Dissertation von Lohse (1984) und der dort zitierten Literatur.

⁴²Mit Mittelwert 12.24 und Varianz 1.51.

Tabelle 1: Simulationsergebnisse

	Variationskoeffizient	Entropiemaß
Erwartungswert		
— empirisch	0.098	0.839
Standardabweichung		
— empirisch	0.020	0.063
— Jackknife	0.019	0.103
— Bootstrap	0.017	0.143
Konfidenzintervall		
— empirisch	[0.066,0.133]	[0.721,0.928]
— Bootstrap	78.00%	98.60%
— bias-corr. Boots.	79.38%	90.36%
— t-percentile	81.34%	n.a.

riationskoeffizient bzw. LG^{43} berechnet. Der empirische Erwartungswert ist dann jeweils das arithmetische Mittel der 1.000.000 Schätzungen. Die ausgewiesenen empirischen Schätzer der Standardabweichung und des Konfidenzintervalls zum Niveau 90% sind die entsprechenden Parameter der generierten diskreten Verteilungen des Variationskoeffizienten bzw. von LG .

Für die ausgewiesenen Jackknife- und Bootstrap-Schätzer wurden 5.000 Stichproben gezogen, aus denen wiederum je 1.000 Bootstrap-Ziehungen vorgenommen wurden. Die ausgewiesenen Werte für die Schätzer der Standardabweichung sind die arithmetischen Mittel aus den Schätzern für die 5.000 gezogenen Stichproben. Für die Konfidenzintervalle ist der Anteil derjenigen Bootstrap-Konfidenzintervalle zum Niveau 90% angegeben, die den geschätzten Erwartungswert enthalten. Für den Variationskoeffizienten wurden auch Konfidenzintervalle nach der „accelerated bias-corrected percentile method“⁴⁴ bestimmt, die jedoch keine nennenswerte Verbesserung gegenüber der bias-corrected percentile method ergaben. Die gleichfalls getestete Methode der kürzesten Konfidenzintervalle⁴⁵ ergab keine brauchbaren Werte.

Analoge Simulationen wurden auch für die anderen Komponenten des WD -Maßes BV und LS sowie für verschiedene theoretische Verteilungen

⁴³ Für LG wurden jeweils die 11 zentralen Werte berücksichtigt.

⁴⁴ Vgl. Efron (1987), S. 178.

⁴⁵ Vgl. Hall (1988), S. 936.

Tabelle 2: Signifikanz von Veränderungen des *WD*-Maßes

Zeitraum	<i>WD</i> -Maß	<i>BV</i>	<i>LG</i>	<i>LS</i>
1972.1-1974.1	-	0	-	-
1974.1-1976.1	+	(-)	0	+
1982.1-1984.1	+	(-)	+	+
1986.1-1987.4	(-)	(-)	0	(+)
1972.1-1989.4	0	0	-	+

der Grundgesamtheit durchgeführt. Die deutliche Überschätzung der Standardabweichung durch Jackknife- und Bootstrap-Schätzer trat hierbei nur für *LG* auf. Für *LS* ist eher eine leichte Unterschätzung festzustellen. Für *LG* und *BV* sind die Standardschätzer der Konfidenzintervalle in der Regel konservativ. Hier führt die bias-corrected percentile method zu besseren Ergebnissen. Für *LS* hingegen sind die geschätzten Bootstrap-Konfidenzintervalle eher zu klein. Dies mag seine Ursache in der geringen Stichprobengröße haben.

Um Aussagen über die Signifikanz von Veränderungen des *WD*-Maßes oder des Variationskoeffizienten machen zu können, bietet sich ebenfalls ein Bootstrap-Schätzer an. Dazu betrachtet man die gemeinsame Verteilung von Lohnsätzen und Beschäftigten für zwei Zeitpunkte. Durch Bootstrap-Ziehungen⁴⁶ erhält man eine Verteilung der Differenzen für das betrachtete Maß. Hieraus lassen sich wieder Konfidenzintervalle bestimmen, aus denen Aussagen über die Signifikanz von Veränderungen ableitbar sind.

Diese Untersuchung wurde mit jeweils 2.000 Bootstrap-Ziehungen für die Entwicklung des *WD*-Maßes und seiner Komponenten in fünf Zeitintervallen vorgenommen. Daraus wurden mit der bias-corrected percentile method Konfidenzintervalle zum Niveau 90% bestimmt. Tabelle 2 weist die betrachteten Zeitintervalle und die Signifikanz der Veränderungen für das *WD*-Maß und seine Komponenten aus. Eine zum gegebenen Niveau signifikante Zunahme der Maßzahl ist dabei mit +, eine signifikante Abnahme mit - ausgewiesen. Ist die Veränderung nicht signifikant, so wird eine 0 eingetragen. Tendenzen, die zum gegebenen Niveau nicht signifikant waren, sind in Klammern ausgewiesen.

⁴⁶Hierbei werden immer Paare von Variablen für die beiden Zeitpunkte gezogen, um der intertemporalen Abhängigkeit der beobachteten Größen Rechnung zu tragen.

Die Abnahme der im WD -Maß ausgedrückten Lohndifferenzierung im Zeitraum 1972.1-1974.1 wird somit als signifikant zum Niveau 90% ausgewiesen. Noch deutlicher wird die in Schaubild 3 dargestellte Entwicklung des WD -Maßes für die Zunahmen von 1974.1-1976.1 und 1982.1-1984.1 bestätigt, während die Abnahme 1986.1-1987.4 zum gegebenen Niveau nicht signifikant war. Für den großen Zeitraum 1972.1-1989.4 kann keine signifikante Zunahme der intersektoralen Lohndifferenzierung bestätigt werden. Allerdings weisen die einzelnen Komponenten des WD -Maßes teilweise differierende Entwicklungen auf. So ist für den Zeitraum 1972.1-1989.4 eine signifikante Zunahme der relativen Lohnspanne LS neben einem signifikant rückgängigem Maß der Differenzierung des Lohngefüges LG festzustellen.

Für die Betrachtung der Veränderung der Entropiemaße BV und LG im WD -Maß bietet sich noch ein anderer Ansatz an. Ebenso wie man den erwarteten Informationsgehalt einer Verteilung durch die Entropie messen kann, läßt sich auch der erwartete Informationsgehalt der Veränderung einer Verteilung messen.⁴⁷ Dieses Maß ist für den Übergang von der durch den Wahrscheinlichkeitsvektor $p = (p_1, \dots, p_n)$ gegebenen Verteilung auf die durch $q = (q_1, \dots, q_n)$ gegebene durch

$$I(q : p) := \sum_{i=1}^n q_i \log \frac{q_i}{p_i}$$

definiert. Sind die Wahrscheinlichkeitsvektoren p und q elementweise identisch, so nimmt $I(q : p)$ erwartungsgemäß den Wert 0 an. Andererseits zeigt man unter der Voraussetzung, daß die p_i und q_i paarweise verschieden sind, daß $I(q : p)$ positiv ist.

Seien hierzu zunächst alle $q_i > 0$. Für die Funktion $f(x) := x - \log(1+x)$ gilt wegen $f(0) = 0$ und $f'(x) > 0$ für alle $x > 0$, daß für positives x $f(x) > 0$ sein muß. Setzt man nun $a_i := \frac{p_i - q_i}{q_i}$, so gilt somit:

$$I(q : p) = \sum_{i=1}^n q_i f(a_i) \geq 0$$

und $I(q : p) = 0 \Leftrightarrow p_i = q_i$ für alle $i \in \{1, \dots, n\}$. Durch Übergang zu Grenzwerten sieht man, daß diese Resultate auch gültig bleiben, wenn für ein i $p_i > q_i = 0$ gilt.

Es gilt also $I(q : p) = -\sum_{i=1}^n q_i \log(1 + a_i)$. Ersetzt man \log durch den natürlichen Logarithmus \ln , dann liefert die Entwicklung von \ln in seine

⁴⁷Vgl. Theil (1971), S. 641.

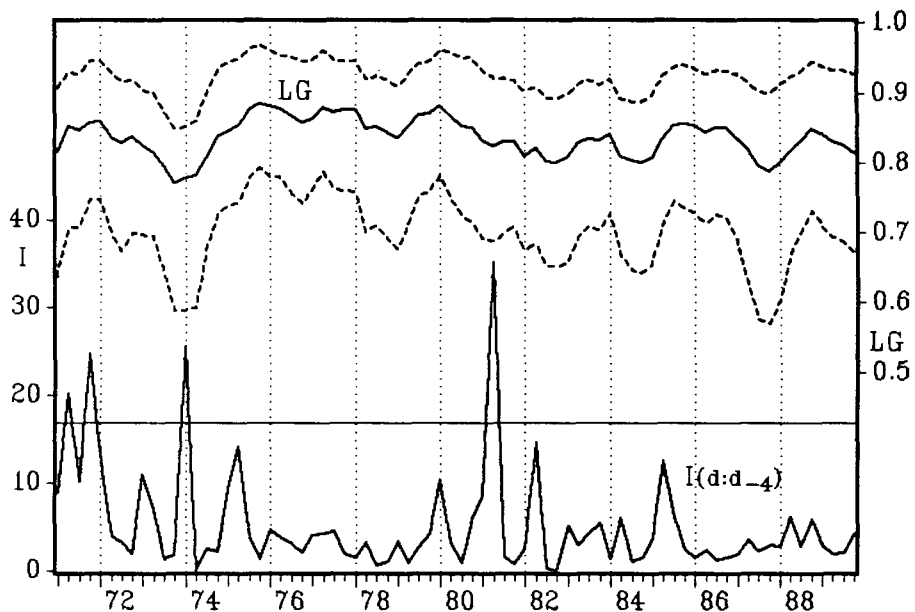
Taylorreihe wegen $\sum_{i=1}^n q_i a_i = 0$ die Approximation

$$I(q : p) \approx \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^2}{q_i}.$$

Diese Approximation ist gültig für $|p_i - q_i| \leq |q_i|$ für alle i . $2n \cdot I(q : p)$ ist also für kleine Veränderungen der Verteilungen χ^2 -verteilt.

In Schaubild 6 sind der Lohngefügekoeffizient LG mit dem bias-corrected Bootstrap-Konfidenzintervall zum Niveau 90% als gleitende arithmetische Mittel aus vier Quartalen, $2n \cdot I(d : d_{-4})$ und der kritische Wert der χ^2 -Verteilung mit 9 Freiheitsgraden zum Niveau 90% dargestellt.

Abbildung 6: Lohngefügekoeffizient und Informationsveränderung



Hieraus ergibt sich, daß der Informationsgehalt der Veränderung der durch die relativen Lohnunterschiede ausgedrückten Lohnsatzverteilung im Verlauf von einem Jahr nur für folgende Zeitpunkte signifikant positiv war: 1971.1, 1971.3, 1974.1 und 1981.1.

7 Schlußbemerkungen und Ausblick

In dieser Arbeit wurde ein auf der Entropiefunktion begründetes Maß der Lohndifferenzierung diskutiert. Im Vergleich zum gebräuchlicheren Variationskoeffizienten zeigt sich ein abweichendes Bild der Entwicklung der intersektoralen Lohnstruktur in der Bundesrepublik Deutschland für den Zeitraum 1970–1989. Insbesondere ist die durch den Variationskoeffizienten ausgewiesene Zunahme der intersektoralen Lohndifferenzierung durch die Betrachtung von Informationsmaßen nicht belegbar.

Wie schon in der Einleitung kurz angesprochen, messen Lohnstrukturmaße Verhältnisse zwischen Verteilungen in den einzelnen Lohngruppen. Damit einher geht einerseits ein Informationsverlust, indem anstelle der tatsächlichen Verteilung einer Lohngruppe nur deren arithmetisches Mittel in die Statistik eingeht. Andererseits ist insbesondere bei einer geringen Anzahl von Lohngruppen die Annahme einer stetigen Verteilung der beobachteten Statistik äußerst fragwürdig, so daß eine Übertragung der oben eingeführten Qualitätsmaße etwa für qualitative Lohnstrukturen mit nur drei Gruppen wenig ratsam erscheint, während die Simulation mit jeweils 13 normalverteilten Zufallsvariablen die oben ausgewiesenen Werte als zuverlässig belegen.

In einer weiteren Arbeit möchte ich daher versuchen, daß hier eingeführte Dispersionsmaß explizit als Dispersionsmaß von Verteilungen zu verallgemeinern, um dann etwa auf der Grundlage der Daten des sozioökonomischen Panels zu dezidierten Aussagen über die Entwicklung der Lohnstruktur in der Bundesrepublik Deutschland gelangen zu können.

Außerdem werde ich die hier eingeführte Methodik auch auf die Untersuchung weiterer Lohnstrukturen anwenden. Hierbei ist insbesondere an regionale und qualifikatorische Lohnstrukturen gedacht. Durch die Normierung des *WD*-Maßes wird aber auch der Vergleich von Lohnstrukturen (etwa der intersektoralen oder qualifikatorischen Lohnstruktur in verschiedenen Ländern der Europäischen Gemeinschaft) ermöglicht.

Schließlich ermöglicht es das durch die Entropiemaße eingeführte informationstheoretische Konzept der Lohnstruktur auch, Fragen der Art zu untersuchen, ob die bestehende Lohnstruktur der zugehörigen Verteilung der Arbeitsproduktivitäten angemessen ist.

Literatur

- [1] T. Amemiya. *Advanced Econometrics*. Basil Blackwell, Oxford 1986
- [2] L. Bellmann und F. Buttler. *Lohnstrukturflexibilität – Theorie und Empirie der Transaktionskosten und Effizienzlöhne*. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 2 (1989), 202-217
- [3] P. J. Bickel und D. A. Freedman. *Some Asymptotic Theory for the Bootstrap*. The Annals of Statistics, 9 (1981), 1196-1217
- [4] M. C. Burda. *Some New Insight on the Interindustry Wage Structure from the German Socioeconomic Panel*. DIW Diskussionspapier Nr. 30, Berlin 1991
- [5] B. Efron. *Bootstrap Methods: Another Look at the Jackknife*. The Annals of Statistics, 7 (1979), 1-26
- [6] B. Efron. *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia, Pennsylvania 1982
- [7] B. Efron. *Better Bootstrap Confidence Intervals*. Journal of the American Statistical Association (1987), 171-185
- [8] B. Efron und C. Stein. *The Jackknife Estimate of Variance*. The Annals of Statistics, 9 (1981), 586-596
- [9] B. Efron und R. Tibshirani. *Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy*. Statistical Science, 1 (1986), 54-77
- [10] A. Feinstein. *Foundations of Information Theory*. McGraw-Hill, New York 1958
- [11] S. F. Franke. *Sektorale, qualifikatorische und regionale Bestimmungsgründe der Lohnstruktur*. Nomos Verlagsgesellschaft, Baden-Baden 1985
- [12] W. Franz. *Beschäftigungsprobleme auf Grund von Inflexibilitäten auf Arbeitsmärkten?* in: H. Scherf (Hrsg.). *Beschäftigungsprobleme hochentwickelter Volkswirtschaften*. Duncker u. Humblot, Berlin 1989, 304-340

- [13] W. Franz. *Arbeitsmarktökonomik*. Springer-Verlag, Berlin Heidelberg 1991
- [14] H. Gerfin. *Die interindustrielle Lohnstruktur in der Bundesrepublik Deutschland*. Jahrbücher für Nationalökonomie und Statistik, 192 (1978), 127-147
- [15] K. Gerlach und O. Hübler. *Lohnstruktur, Arbeitsmarktprozesse und Leistungsintensität in Effizienzlohnmodellen*. in: F. Buttler, J. Kühl und B. Rahmann (Hrsg.). *Staat und Beschäftigung*. Beiträge zur Arbeitsmarkt- und Berufsforschung 88, Nürnberg 1985, 249-290
- [16] S. Guiaşu. *Information Theory with Applications*. McGraw-Hill, New York 1977
- [17] E. Gundlach. *Gibt es genügend Lohndifferenzierung in der Bundesrepublik Deutschland?* Die Weltwirtschaft 1 (1986), 74-88
- [18] P. Hall. *Theoretical Comparison of Bootstrap Confidence Intervals*. The Annals of Statistics, 16 (1988), 927-953
- [19] H.-D. Harges. *Vorschläge zur Differenzierung und Flexibilisierung der Löhne*. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 1 (1988), 52-73
- [20] A. M. Jaglom und I. M. Jaglom. *Wahrscheinlichkeit und Information*. 4. Aufl., Harri Deutsch, Thun und Frankfurt a.M. 1984
- [21] M. Koller. *Regionale Lohnstrukturen*. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, 20 (1987), 30-44
- [22] H. Lampert. *Die Lohnstruktur der Industrie*. Duncker & Humblot, Berlin 1963
- [23] K. Lohse. *Zur Konsistenz des Bootstrap-Verfahrens*. Dissertation, Universität Hamburg 1984
- [24] Organisation for Economic Co-Operation and Development. *Employment Outlook September 1985*. OECD, Paris 1985
- [25] M. H. Quenouille. *Approximate Tests of Correlation in Time-Series*. Journal of the Royal Statistical Society, Ser. B, 11 (1949), 68-84

- [26] M. H. Quenouille. *Notes on Bias in Estimation*. Biometrika, 43 (1956), 353-360
- [27] H. J. Ramser. *Verteilungstheorie*. Springer-Verlag, Berlin Heidelberg 1987
- [28] A. Rényi. *Wahrscheinlichkeitsrechnung*. 2. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1966
- [29] W. R. Schucancy, H. L. Gray und D. B. Owen. *On Bias Reduction in Estimation*. Journal of the American Statistical Association, 66 (1971), 524-533
- [30] W. Sengenberger. *Mangelnde Flexibilität auf dem Arbeitsmarkt als Ursache der Arbeitslosigkeit*. in: H. Krupp, B. Rohwer und K. W. Rothschild (Hrsg.). *Wege zur Vollbeschäftigung*. Rombach Verlag, Freiburg 1986, 91-106
- [31] C. E. Shannon *A Mathematical Theory of Communication*. Bell Systems Technical Journal 27 (1948), 379-423, 623-653
- [32] W. Smolny. *Dynamic Factor Demand in a Rationing Context*. Dissertation, Universität Konstanz, 1991
- [33] R. Soltwedel. *Nivellierungsprobleme in der Tarifpolitik und deren Auswirkungen auf Leistungsbereitschaft, Wirtschaft und Beschäftigung*. Leistung und Lohn (BDA) Nr. 98, Heider-Verlag, Bergisch-Gladbach 1980
- [34] R. Stine. *An Introduction to Bootstrap Methods*. Sociological Methods & Research, 18 (1989), 243-291
- [35] H. Theil. *Principles of Econometrics*. North-Holland, Amsterdam 1971
- [36] R. Thiehoff. *Lohnnivellierung und qualifikatorische Arbeitslosigkeitsstruktur*. Nomos Verlagsgesellschaft, Baden-Baden 1987
- [37] F. Topsøe. *Informationstheorie*. B. G. Teubner, Stuttgart 1974
- [38] J. W. Tukey. *Bias and Confidence in Not-Quite Large Samples*. Annals of Mathematical Statistics (abstract), 29 (1958), 614