



Betriebliche Lohnanalyse mit Hilfe der multiplen Regression – Teil 1

Methodische Anmerkungen zur Revision des Gleichstellungsgesetzes

Autoren: Fritz Dolder / Matthias Moelleney

Beitragsarten: Beiträge

Rechtsgebiete: Arbeitsrecht

Zitiervorschlag: Fritz Dolder / Matthias Moelleney, Betriebliche Lohnanalyse mit Hilfe der multiplen Regression – Teil 1, in: Jusletter 22. Mai 2017

Die Vernehmlassungsvorlage für die geplante Revision des Gleichstellungsgesetzes schlägt unter anderem vor, dass die Betriebe verpflichtet werden, ihr Lohnsystem im Hinblick auf die Lohndiskriminierung (Art. 3 GIG) mit Hilfe von strukturierten multi-criteria Instrumenten zu beurteilen. Die für diese Aufgabe diskutierte Methode der multiplen Regressionsanalyse OLS weist verschiedene technische Problembereiche auf, welche auf mathematische Vorgaben, auf den beschränkten Informationswert der Methode und auf rechtliche Rahmenbedingungen zurückzuführen sind.

Inhaltsverzeichnis

1. Multiple Regression (OLS-Methode) in der betrieblichen Lohnanalyse
2. Probleme des Einsatzes der multiplen Regression zur betrieblichen Lohnanalyse im Rahmen des GIG
 - 2.1. Einschränkungen und Vorgaben der Mathematik
 - 2.2. Beschränkter Erkenntniswert der Regressionsanalyse
 - 2.3. Institutionelle Schranken

[Rz 1] In der auf empirischen Erkenntnissen aufgebauten Entscheidungslehre steht es seit Jahrzehnten ausser Frage, dass strukturierte multi-criteria Methoden bei Entscheidungen über Phänomene mit einer Vielzahl von Eigenschaften (multi-attributive Objekte) im Vergleich zu intuitiven Methoden Resultate von überlegener Qualität sowohl hinsichtlich Reliabilität als auch Validität hervorbringen. Für die Rechtsanwendung dürfte es deshalb schon im Ansatz *unethisch* sein, Entscheidungen mit Hilfe von unstrukturierten und auf der Intuition der Entscheidungsträger basierenden Methoden zu treffen, falls für das betreffende Problem strukturierte multi-criteria Methoden zur Verfügung stehen, deren überlegene Qualität ausser Frage steht. Auf der individuellen Ebene entspricht dies der Vorstellung, dass der Rechtssuchende einen individuellen Anspruch darauf hat, dass seine Anliegen mit denjenigen Methoden beurteilt werden, welche dem neuesten Stand der wissenschaftlichen Erkenntnisse der Entscheidungslehre entsprechen.¹

Unter diesen Voraussetzungen war schon vor dem Erlass des Gleichstellungsgesetzes am 24. März 1995 (Inkrafttreten am 1. Juli 1996) erkennbar, dass die aus vielen Einzelattributen zusammengesetzten Begriffe der «gleichwertigen Arbeit» und des «Benachteiligten» im Einzelfall nur mit geeigneten, strukturierten multi-criteria Methoden und Algorithmen auf sinnvolle und reproduzierbare Weise umgesetzt werden können. Nur durch derartige multi-criteria Methoden kann das zentrale Postulat der *Rechtssicherheit* auch im Bereich der Lohngleichheit wenigstens ansatzweise erfüllt werden.²

- [Rz 2] Bei der Umsetzung von Art. 3 des Gleichstellungsgesetzes (GIG; SR 151.1) steht bis heute die *analytische Arbeitsplatzbewertung* im Zentrum der Rechtsprechung; sie baut im Wesentlichen auf arbeitspsychologischen Erkenntnissen und Zusammenhängen auf und verkörpert formal eine Methode des *Simple Additive Weighting* (SAW, gewichtete Addition). In der Rechtsprechung wurde diese analytische Arbeitsplatzbewertung bisher regelmässig mit Hilfe der Beurteilungsraster der «Vereinfachten Funktionsanalyse» (VFA) (ab 1991) und/oder von ABAKABA (ab 1996) umgesetzt.³ Die Vereinfachte Funktionsanalyse (VFA) ist von verschiedenen Kantonen zur Gesetzesnorm für die Arbeitsplatzbewertung im öffentlichen Personalrecht erhoben worden.⁴
- [Rz 3] Zur Beurteilung der *Gleichwertigkeit* im Rahmen des Gleichstellungsgesetzes wird – soweit ersichtlich – seit BGE 130 III 145 (vom 22. Dezember 2003) *neben* der analytischen Arbeitsplatzbewertung auch das statistische Verfahren der *multiplen Regressionsanalyse* eingesetzt, mit der das Lohnsystem ganzer Betriebe kollektiv beurteilt werden kann. In der Begründung dieses Urteils wurde indessen nur ausgeführt, dass die Feststellung des Sachverhalts durch die Vorinstanz (Cour civile du Tribunal cantonal vaudois), basierend auf zwei Gutachten von Sachverständigen für das Bundesgericht verbindlich sei und *nicht überprüft werden* dürfe (E. 3.1.2 und 3.2). Auch die Vorinstanz des Kantons Waadt hat in ihrem Urteil das fragliche Gutachten nicht inhaltlich überprüft, sondern ohne Kommentar als richtig übernommen.⁵
- [Rz 4] Seit 2001 hat das Büro BASS im Auftrag des Bundes des Eidgenössisches Büro für die Gleichstellung von Frau und Mann (EBG) ein «Standard-Analysemodell» für das *Screening* von Anbietern in der öffentlichen Beschaffung (ÖB) des Bundes entwickelt, welches ebenfalls auf der Methode der multiplen Regressionsanalyse beruht.⁶ Eine derartige informelle Steuerung der *Methode* der Rechtsanwendung durch die Exekutive ist zwar ungewöhnlich, aber auch kein völlig isolierter Einzelfall. Zum Vergleich: In der öffentlichen Beschaffung von Bund und Kantonen werden Submissionsofferten mit Hilfe einer Mehrzahl, teilweise sogar gesetzmässig vorgegebener Beurteilungskriterien im Rahmen der erwähnten gewichteten Addition (Simple additive weighting, SAW) beurteilt.⁷
- [Rz 5] Als Ergebnis dieser Arbeiten stellt der Bund seit 2010 durch das Eidg. Büro für Gleichstellung von Frau und Mann (EBG) mit dem «Standard-Analysemodell» und der dazugehörigen Software Logib der Wirtschaft ein Instrument zur Beurteilung von betrieblichen Lohnsystemen gratis zur Verfügung, welches auf der statistischen Methode der *multiplen Regressionsanalyse* beruht. Es wurde auf der Grundlage von empirischen Daten der periodischen Lohnstrukturerhebungen LSE des Bundesamtes für Statistik (BFS) entwickelt.⁸

Es ist (zunächst noch) ausschliesslich für Betriebe bestimmt, welche im Rahmen des öffentlichen Beschaffungswesens dem Bund Leistungen anbieten und durch Gesetz (Art. 8 Abs. 1 Bst. c und Art. 8 Abs. 2 des Bundesgesetzes über das öffentliche Beschaffungswesen; BöB; SR 172.056.1) zur Respektierung der Lohngleichheit

verpflichtet sind.⁹ Eine ausdrückliche gesetzliche Verpflichtung dieser Betriebe zur Benutzung des Modells oder einer anderen individuell bezeichneten statistischen Methode besteht zwar nicht. Der Bundesrat sieht aber anscheinend mindestens eine indirekte Verpflichtung der Anbieter zur Benutzung des «Standard-Analysemodells» EBG und diese kommen aufgrund situationspezifischer Überlegungen wohl kaum um dieses «Standard-Analysemodell» herum.¹⁰

[Rz 6] Der Bund hat, wie erwähnt, zurzeit *nur* in Betrieben von Anbietern im öffentlichen Beschaffungswesen eine Möglichkeit zur Steuerung der Methode betrieblicher Lohnanalysen zur Kontrolle der Lohngleichheit von Frau und Mann. Die geplante Revision des GIG geht weiter und strebt eine Verpflichtung *aller* privatrechtlich organisierten Betriebe, also der gesamten privaten Wirtschaft, zur periodischen betrieblichen Lohnanalyse nach den methodischen Vorgaben des Bundes an. In der Vernehmlassungsvorlage des Bundesrats vom 18. November 2015 zur Revision des Gleichstellungsgesetzes scheint das Standard-Analysemodell EBG bzw. dessen Software Logib und damit die Methode der *multiplen Regressionsanalyse* im Zentrum der angestrebten zukünftigen Rechtsanwendung zu stehen. Zudem ist es offenbar ein Ziel der geplanten Revision, auch ausserhalb des öffentlichen Beschaffungswesens des Bundes *rechtliche Sanktionen* an die mit Hilfe der statistischen Regressionsanalyse erarbeiteten Resultate betrieblicher Lohngleichheitskontrollen zu knüpfen.

Die auf die Vorlage des Bundesrats eingegangenen Vernehmlassungen haben erwartungsgemäss in methodischer Hinsicht nichts Nennenswertes gebracht.¹¹ An einer Medienkonferenz vom 26. Oktober 2016 hat der Bundesrat von der Vorsteherin des Justizdepartements erneut seine Absicht zur Revision des Gesetzes verkünden lassen und eine Gesetzesvorlage und Botschaft an das Parlament *für Mitte 2017* angekündigt.

[Rz 7] Anwendungsbereiche: Neben der dargestellten betrieblichen Ebene im Rahmen des GIG existiert ein zweiter Anwendungsbereich von multi-criteria Instrumenten zur Lohnanalyse auf *nationaler Ebene*. Seit Jahrzehnten werden auf dieser Ebene soziologisch orientierte Lohnanalysen auf der Basis riesiger Stichproben mit Tausenden von Arbeitnehmer/innen mehrheitlich mit Hilfe der Regressionsanalyse ausgewertet. Diese Ebene ist *nicht Gegenstand der vorliegenden Studie*; der aktuelle Stand der wissenschaftlichen Erkenntnisse für die Schweiz (November 2016) wird dargestellt von FLURINA SCHMID (2016) auf der Basis von Stichproben des Schweizerischen Haushalts-Panels (SHP) des FORS (SHP 2004 und 2012). Diese Studie ist umfassend angelegt und vorsichtig in der Beurteilung von Kausalitäten.¹²

1. Multiple Regression (OLS-Methode) in der betrieblichen Lohnanalyse

[Rz 8] Die *multiple Regressionsanalyse* ermöglicht es, die Übereinstimmung (*goodness of fit*) zwischen beobachteten empirischen Daten und einem Modell mit mehreren, vom Analysten ausgewählten unabhängigen Variablen x_i zu beurteilen. Dies geschieht dadurch, dass eine abhängige Variable y in einer Linearkombination zu einer Mehrzahl von als relevant vermuteten unabhängigen Variablen x_i in Beziehung gesetzt wird, denen die Koeffizienten b_i zugeordnet werden:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_j x_j + \dots + b_J x_J + u$$

wobei y abhängige Variable

b_0	konstantes Element der Regressionsfunktion
b_j	Regressionskoeffizienten ($i = 1, 2, \dots, J$)
x_i	unabhängige Variable ($i = 1, 2 \dots J$)
u	Störgrösse

In der betrieblichen Lohnanalyse bildet der Lohn (in irgendeiner Form pro Zeiteinheit) die abhängige Variable y und die als lohnbestimmend vermuteten Parameter, wie (*notwendige oder tatsächliche?*) Ausbildung, Führungsverantwortung, seelische / körperliche Belastung, Dienstalter, Lebensalter usw. die unabhängigen Variablen x_i . Die Koeffizienten b_i (Regressionskoeffizienten) und das konstante Element b_0 (Schnitt der y -Achse) werden dabei anhand des Kriteriums der kleinsten Quadrate (KQ-Kriterium, OLS-Methode, engl.: *Ordinary Least Square Method*) berechnet bzw. geschätzt. Neben dieser OLS-Methode existieren methodische Varianten (Doubly Robust Regression und Quartilsreduktion), welche indessen bisher nicht in das Standardmodell des Bundes aufgenommen worden sind.¹³

[Rz 9] Auswertungsverfahren: Die Resultate einer multiplen Regressionsanalyse werden mit verschiedenen Verfahren ausgewertet. Zum einen wird die Qualität der Übereinstimmung (*KQ-Kriterium, goodness of fit, Modellkonformität*) der tatsächlich beobachteten empirischen Daten und dem Modell mit dem Bestimmtheitsmass R^2 (Determinationskoeffizient) beurteilt. Dieses beruht auf einer Zerlegung der Gesamtstreuung der beobachteten (y) und der nach dem Modell berechneten y -Werte $[\hat{y}]$ nach dem Ansatz: Gesamtstreuung = erklärte Streuung + nicht erklärte Streuung.

Daraus wird das Bestimmtheitsmass R^2 (Determinationskoeffizient) aus dem Verhältnis (Division) von erklärter Streuung zur Gesamtstreuung berechnet:

$$R^2 = \text{erklärte Streuung} \div \text{Gesamtstreuung}$$

R^2 hat in normierter Form einen Wertebereich zwischen 0 und 1; wird die gesamte Streuung erklärt, wird $R^2 = 1$, wird nichts erklärt, wird $R^2 = 0$. Dies ergibt eine andere Darstellungsweise von R^2 als

$$R^2 = 1 - [(\text{nicht erklärte Streuung}) \div \text{Gesamtstreuung}].^{14}$$

Die «Stunde des Glücks» schlägt dem Modellbauer, wenn sein $R^2 > 0.6$ d.h. sein Modell die beobachteten Daten zu 60% oder mehr erklärt. Daraus erklärt sich die «Jagd nach dem $R^2 > 0.6$ ». Dieses R^2 ist aber auf vielfache Weise anfällig für Störungen; es kann insbesondere durch die Auswahl der unabhängigen Variablen x_i beeinflusst werden.¹⁵ Über die Ursachen der nicht erklärten Streuung und damit über den Inhalt und die Zahl der vermuteten weiteren Einflüsse (unabhängige Variablen x_i bzw. Störgrösse u) liefert ein Regressionsmodell *keine Erkenntnisse*.

[Rz 10] Ein anderes Auswertungsverfahren verwendet dichotome Dummyvariablen (0 / 1-Variablen), um den «Einfluss» einzelner Parameter auf das Resultat (hier: Lohnhöhe) zu beurteilen. Im Ursprungsland derartiger Lohnstudien (USA) stand dabei naturgemäss zunächst die Dummyvariable der Rasse der Arbeitnehmer im Vordergrund (weiss / anders), während in der Schweiz bisher – soweit ersichtlich – nur die Dummyvariable FM (Frau = 1, Mann = 0) eingesetzt worden ist. Der dafür errechnete Regressionskoeffizient β_5 (Standard-Analysemodell EBG) wird als Entscheidungsgrenze (cut-off value) unter der Bezeichnung «Diskriminierungskoeffizient» für das «Aussortieren» von im Hinblick auf die Lohnungleichheit «verdächtigen» Betrieben verwendet.

Methodenbeschrieb EBG (2016), S. 6/7:

$$\ln(Y_i) = .0 + .1 * AUSB_i + .2 * ERWERB_i + .3 * ERWERB_i^2 + .4 * DIENSTJ_i + .5 * GESCHL_i + .i$$

wobei: [...]

$GESCHL_i$ = Dummy-Variable für das Geschlecht $ist = 1$, wenn Person i eine Frau ist; $ist = 0$, wenn Person i ein Mann ist. [...]

.5 = Koeffizient der Variable für das Geschlecht, GESCHL («Diskriminierungskoeffizient»)

Bsp.: .5 = -0.1 bedeutet: Die Frauen verdienen bei gleichwertiger Qualifikation (unter Berücksichtigung von Dienstalter, Ausbildung und potentieller Erwerbserfahrung) rund 10% weniger als die Männer.

[Rz 11] Bei der Lohnanalyse auf *betrieblicher Ebene* zur Beurteilung bzw. zur Rechtfertigung behaupteter geschlechtsspezifischer Lohndiskriminierungen im Rahmen des GIG wird die *multiple Regressionsanalyse* nach dem OLS-Verfahren u.a. folgendermassen eingesetzt:¹⁶ Sie wird im Standard-Analysemodell EBG empfohlen zur Bewertung des *kollektiven* Lohnsystems ganzer Betriebe im Hinblick auf die Lohnungleichheit Frauen/Männer. Mit Hilfe einer Dummyvariablen Frauen (1) / Männer (0) wird dabei der erwähnte Regressionskoeffizient β_5 ermittelt. Dessen Höhe (Beispiel $\beta_5 = - .111$) bildet die Grundlage für das *Aus-Sortieren* der untersuchten Betriebe in verdächtige / nicht-verdächtige im Hinblick auf die Lohnungleichheit im Sinne von Art. 3 GIG. Die Effizienz der Regressionsanalyse zum Nachweis von Verstössen gegen Art. 3 GIG wird im Folgenden zu untersuchen sein (unten Abschnitt 2.2).¹⁷

[Rz 12] Auf der *individuelle Ebene* kann mit den berechneten Koeffizienten b_i bzw. β_i und den beobachteten Daten der unabhängigen Variablen x_i eines/einer individuellen Arbeitnehmers/in k ein *modellkonformer Lohn* \hat{y}_k für diese(n) einzelne(n) Arbeitnehmer/in k berechnet werden. Dieser kann sodann mit dem tatsächlich beobachteten Lohn y_k dieses/dieser Arbeitnehmers/in verglichen und die gefundene Differenz $d = (\hat{y}_k - y_k)$ bewertet werden. Das rechtliche Gewicht einer derartigen auf Statistik beruhenden Differenz d im Hinblick auf einen Verstoß gegen Art. 3 GIG wird im Folgenden zu untersuchen sein (unten Abschnitt 2.2).¹⁸ In diesem Zusammenhang wäre allerdings daran zu erinnern, dass im privatrechtlichen Arbeitsverhältnis Lohnungleichheitsklagen nach GIG meistens erst *nach* Beendigung des Arbeitsverhältnisses eingeleitet und beurteilt werden. So war es auch in dem mehrfach erwähnten Musterfall für die Regressionsanalyse BGE 130 III 145.¹⁹

[Rz 13] Wer auch immer in der empirischen Sozialforschung die *multiple Regressionsanalyse* einsetzt, sollte sich bewusst

sein, dass er mit einem *Modell* arbeitet, dessen Resultate in vielfacher Hinsicht durch die Autoren des Modells («Modellbauer») beeinflusst, im schlimmsten Fall auch *manipuliert* werden können. Der Nachbenutzer eines derartigen Modells ist dem fachlichen Verständnis dieses Modellbauers in vielfacher Hinsicht *ausgeliefert*, insbesondere bei der Auswahl der unabhängigen Variablen x_i oder der Definition des Inhalts der einzelnen Variablen x_i . Vereinfacht ausgedrückt misst die Methode nur das, was der Modellbauer wissen bzw. demonstrieren möchte: Sie gibt nur «*Resultate im Sinne des Erfinders*». ²⁰ Dies hat dazu geführt, dass eine veritable *Giffliste* bestehend aus 10 einzelnen Warnungen S1 ... S10 formuliert worden ist, welche Verhaltensmuster beschreibt, welche die Resultate verändern können und die Absurdität der «*Jagd nach einem möglichst hohen R^2 -Wert*» deutlich machen. ²¹

2. Probleme des Einsatzes der multiplen Regression zur betrieblichen Lohnanalyse im Rahmen des GIG

2.1. Einschränkungen und Vorgaben der Mathematik

Bevor ein Modell der multiplen Regressionsanalyse zur Beurteilung von Arbeitslohn und Lohnsystemen eingesetzt wird, muss geklärt werden, ob die mathematischen Voraussetzungen für eine derartige Beurteilung gegeben sind. Für diese mathematischen Rahmenbedingungen sind verschiedene Listen von Kriterien («Annahmen») A1 bis A5 bzw. bis A7 formuliert worden. ²² Mehrere der formulierten Kriterien erscheinen für eine Anwendung in der Lohnanalyse einzelner Betriebe als problematisch und es ist *prima facie* zweifelhaft, ob die Voraussetzungen für eine Modellierung der betrieblichen Lohnanalyse mit Hilfe einer multiplen Regressionsanalyse OLS durchwegs erfüllt sind.

[Rz 14] Stichprobengrösse: Da die multiple Regressionsanalyse als *statistische* Methode eine Vielzahl gleichartiger Daten einer Stichprobe voraussetzt, benötigt sie für sinnvolle und zuverlässige Resultate eine minimale Datenmenge (Stichprobengrösse). Für die betriebliche Lohnanalyse wird beispielsweise vorgeschlagen, dass die minimale Stichprobengrösse des Betriebs 50 Arbeitnehmer/innen umfassen müsse, wobei jedes Geschlecht *mindestens 10 Personen* umfassen müsse. Diese Werte dürften vermutlich zu niedrig sein; die für verwertbare Resultate erforderliche minimale Stichprobengrösse sollte eher zwischen 100 bis 150 Mitarbeitern liegen mit einer einigermaßen *gleichmässigen Verteilung* auf beide Geschlechter. Selbst die vom Bund finanzierte Studie FELFE et al (2015) musste für ihre Überlegungen zur betrieblichen Lohnanalyse die Stichproben entsprechend eingrenzen. ²³

[Rz 15] Linearitätsannahme: Die multiple Regressionsanalyse baut darauf auf, dass zwischen den einzelnen unabhängigen Variablen x_i (Ausbildung, Führungsverantwortung, körperliche / psychische Belastung, Dienstalter usw.) und der abhängigen Variablen y (Lohnhöhe) *lineare Beziehungen* bestehen. Es sollte daher vor Durchführung einer multiplen Regressionsanalyse abgeklärt werden, ob die Annahme einer Linearität dieser Beziehungen sinnvoll und sachgerecht erscheint. Ist es beispielsweise sinnvoll, dass ein(e) Arbeitnehmer/in mit einer m -fach grösseren Zahl von Untergebenen einen m -fach höheren Lohn bezieht als ein(e) Arbeitnehmer/in mit einer m -fach geringeren Zahl von Untergebenen? Oder sollte vielleicht eher der Wert von x_i logarithmiert werden? Nicht sachgerechte Linearitätsannahmen führen zu falschen Schlussfolgerungen. Insgesamt ist nicht ohne Weiteres ersichtlich, ob die Linearitätsannahmen zwischen der Lohnhöhe y und allen der im Standard-Analysemodell des Bundes verwendeten unabhängigen Variablen x_i sachlogisch vertretbar sind.

Als Alternativen zu Linearitätsannahmen bieten sich verschiedene Formen der «*Linearisierung*» von y und/oder x_i an, wodurch die Variablen vor der Regressionsanalyse derart transformiert werden, dass sie eine lineare Beziehung generieren.²⁴ Das Standard-Analysemodell EBG arbeitet anscheinend mit *logarithmierten* Werten der abhängigen Variablen $\ln(Y_i)$ (Lohn):

Methodenbeschrieb EBG (2016), S. 6/7:

$$\ln(Y_i) = .0 + .1 * AUSB_i + .2 * ERWERB_i + .3 * ERWERB_i^2 + .4 * DIENSTJ_i + .5 * GESCHL_i + .i$$

wobei:

$\ln(Y_i)$ = Logarithmierter, auf Vollzeit standardisierter Bruttolohn von Person i . Die erklärende Variable «Lohn» wird logarithmiert aufgrund **rechtsschiefer** Lohnverteilung und für eine direktere Interpretation der Ergebnisse (=> prozentualer Einfluss von AUSB etc). [...]

Diese Linearisierung erscheint zwar vom mathematischen Standpunkt aus als *zulässig*, aber sie wird bei der praktischen Rechtsanwendung des GIG vor Ort vom betroffenen Arbeitgeber/in bzw. der betroffenen Arbeitnehmer/in wohl *kaum verstanden*, sondern eher als mathematischer Kunstgriff empfunden werden: Warum wird mein Lohn nicht in Franken berechnet? Sondern mit diesem rätselhaften natürlichen Logarithmus $\ln(Y_i)$? Wenn mein Lohn *tel quel* in Franken eingesetzt worden wäre, so wäre die Verletzung der Lohngleichheit natürlich sofort erkannt worden, bzw. wäre niemals eine Lohnungleichheit berechnet worden usw. usw.

[Rz 16] Normalverteilung der Residuen:

Die Kriterienlisten der Regressionsanalyse setzen ausnahmslos voraus, dass die sog Residualwerte (Residuen) der abhängigen Variablen y (hier: Lohnhöhe) eine Normalverteilung aufweisen.²⁵ Diese Residuen entstehen bei der Zerlegung der Gesamtabweichung einer Beobachtung y_k vom Mittelwert \bar{y} aller Beobachtungen nach dem Zusammenhang

Gesamtabweichung = erklärte Abweichung + Residuum

$$y_k - \bar{y} = [\hat{y}_k - \bar{y}] + [y_k - \hat{y}_k],$$

wobei

k Kennzeichen einer einzelnen Beobachtung bei einem einzelnen Arbeitnehmer,

y_k tatsächlich beobachteter Wert von y in der Beobachtung k ,

\hat{y}_k geschätzter / berechneter Wert von y_k gemäss Regressionsmodell,

\bar{y} Mittelwert von y aus allen Beobachtungen k

Am einfachsten kann das Residuum als Abweichung des aus dem Regressionsmodell prognostizierten Wertes \hat{y}_k von dem tatsächlich beobachteten Wert y_k umschrieben werden. Oder als Differenz zwischen den «vorhergesagten» und den tatsächlich «beobachteten» Werten der abhängigen Variablen. Eine Normalverteilung dieser Residuen ist erforderlich, um die Ergebnisse der Regressions-schätzung, namentlich die Werte der Regressionskoeffizienten b_i und von R^2 mit statistischen Tests (z.B. F- oder t-Statistik) auf Signifikanz zu überprüfen.

Aufgrund der Kenntnisse der betrieblichen Verhältnisse erscheint es *prima facie* als *ziemlich unwahrscheinlich*, dass eine betriebliche Lohnanalyse in der Praxis eine Normalverteilung dieser Residualwerte von y ergeben wird, da erfahrungsgemäss die Einkommen y_k (= Arbeitslohn der Beobachtung k) eine ausgeprägt *rechtsschiefe* und keine symmetrische Verteilung aufweist.²⁶ Die niedrigen Arbeitslöhne werden häufiger auftreten als die höheren Arbeitslöhne. Daraus ergibt sich mit hoher Wahrscheinlichkeit eine Residuenverteilung, welche *keine* Normalverteilung aufweist.²⁷

[Rz 17] Form der Verteilungsfunktion: Für ein erfolgreiches Regressionsmodell sollte die Form der Häufigkeitsverteilung der beobachteten Werte sowohl der abhängigen Variablen y (Lohn) als auch der einzelnen unabhängigen Variablen x_i für beide Geschlechter getrennt betrachtet einigermassen vergleichbar und dabei mehr oder weniger *gleichmässig über den untersuchten Wertebereich verteilt* sein. Die Methode der multiplen Regressionsanalyse OLS stösst aber hier an ihre Grenzen: Die Verteilung dieser Parameter wird in der betrieblichen Praxis insbesondere bei KMUs regelmässig eine stark unterschiedliche Form für Frauen und Männer aufweisen, ohne dass daraus bereits auf eine Verletzung von Art. 3 GIG geschlossen werden könnte. An der Voraussetzung dieser zumindest ähnlichen und gleichmässigen Form der Verteilung dürfte es in den meisten Betrieben des Handwerks, insbesondere in der Bauindustrie und im Transportgewerbe, fehlen und eine erfolgreiche Regressionsanalyse OLS dort bereits an der stark unterschiedlichen Häufigkeitsverteilung der Parameter für Frauen und Männer scheitern. Selbst die vom Bund bezahlte Studie FELFE et al. (2015) räumt dies ein und empfiehlt, eine Prüfung der Verteilung der Daten von Frauen und Männern «*vorzuschalten*», was immer dies im Einzelnen bedeuten mag.²⁸ Der Bundesrat hat in seinem Schlussbericht vom November 2015 diese Schlussfolgerungen übernommen ohne allerdings inhaltlich darauf einzugehen.²⁹

Als vorläufiges Teilergebnis kann an dieser Stelle festgehalten werden: Die multiple Regressionsanalyse OLS unterliegt einer Vielzahl von Einschränkungen aufgrund mathematischer Vorgaben. Es bestehen begründete Zweifel, ob einzelne dieser Vorgaben bei der Lohnanalyse auf Betriebsebene immer eingehalten werden können.

2.2. Beschränkter Erkenntniswert der Regressionsanalyse

[Rz 18] *Post hoc ergo propter hoc*

Ein Gespenst geht um in den empirischen Sozialwissenschaften, welche auf Korrelationen zwischen verschiedenen Grössen vertrauen: Es ist das Gespenst des *post hoc ergo propter hoc*. Auch unter diesem Blickwinkel könnten die vorsichtigen Schlussfolgerungen von FLURINA SCHMID (2016) vielleicht zu einer veränderten Wahrnehmung der Kausalitäten im Bereich des *gender wage gap* und des Erkenntniswerts der Regressionsanalyse auf betrieblicher Ebene führen.³⁰

[Rz 19] Regressionskoeffizient β_5 als Ausgangspunkt (cut-off value) für das Aus-Sortieren von «verdächtigen» Betrieben?

Das Standard-Analysemodell des Bundes versucht, den geschlechtsspezifischen Einfluss auf den Lohn durch den Einsatz der dichotomen Dummyvariablen für FM (Frau = 1, Mann = 0) zu bewerten und mit dem zugehörigen Regressionskoeffizienten β_5 zu messen. Im Methodenbeschrieb des Standard-Analysemodells EBG³¹ lautet die allgemeine Form der Regressionsgleichung

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n + u_i$$

Wird diese Gleichung auf die persönlichen Qualifikationsmerkmale eingeschränkt, so lautet sie in der softwaremässigen Fassung von LOGIB:

$$\ln(Y_i) = \beta_0 + \beta_1 * AUSB_i + \beta_2 * ERWERB_i + \beta_3 * ERWERB_i^2 + \beta_4 * DIENSTJ_i + \beta_5 * GESCHL_i + u_i$$

und der Regressionskoeffizient β_5 wird dabei folgendermassen umschrieben:

GESCHL_i = Dummy-Variable für das Geschlecht ist = 1, wenn Person i eine Frau ist; ist = 0, wenn Person i ein Mann ist.

β_0 = Konstante (entspricht «Grundlohn», wenn Ausstattungsmerkmale gleich 0 sind).

β_1 , ..., β_4 = Koeffizienten (zeigen den Einfluss der entsprechenden Faktoren auf den Lohn an).

β_5 = Koeffizient der Variable für das Geschlecht, GESCHL («Diskriminierungskoeffizient»)

Bsp.: $\beta_5 = -0.1$ bedeutet:

Die Frauen verdienen bei gleichwertiger Qualifikation (unter Berücksichtigung von Dienstalter, Ausbildung und potentieller Erwerbserfahrung) rund 10% weniger als die Männer.

u_i = Störgrösse

Für die praktische Anwendung in der betrieblichen Lohnanalyse wird diese *personenbezogene* Form der Regressionsgleichung durch die beiden *arbeitsplatzbezogenen* Variablen «Betriebliches Kompetenzniveau» mit vier Kategorien (Stufen) und «Berufliche Stellung» mit fünf Kategorien (Stufen) erweitert.

[Rz 20] Schon der Inhalt des fraglichen Regressionskoeffizienten β_5 («Diskriminierungskoeffizient») der Dummyvariablen FM ist nur schwer deskriptiv zu erfassen. Misst der Regressionskoeffizient β_5 nun einen prozentualen, linearen *Einfluss* des Geschlechts auf den Arbeitslohn *jeder einzelnen* Arbeitnehmerin? Kann von diesem Regressionskoeffizienten β_5 auf eine prozentuale Häufigkeit von Verstössen gegen Art. 3 GIG in einem untersuchten Einzelbetrieb geschlossen werden?

Also: Gemessen wird $\beta_5 = - .111$, daher verdienen alle in dem fraglichen Betrieb beschäftigten Frauen 11% Lohn weniger, als ihnen gemäss Art. 3 GIG zustehen würde? Und können deshalb *tel quel* 11% mehr Lohn verlangen?

Oder: Gemessen wird $\beta_5 = - .111$, daher beinhalten 11% aller Arbeitsverhältnisse in diesem Betrieb einen Verstoß gegen Art. 3 GIG?

Eine präzise Umschreibung des Inhalts von $\beta_5 = - .111$ wäre wahrscheinlich: Der logarithmierte Mittelwert *aller* in dem untersuchten Betrieb bezahlten Frauenlöhne \bar{y}_f ist um 11.1% niedriger als der logarithmierte Mittelwert *aller* in dem Betrieb bezahlten Männerlöhne \bar{y}_m .³²

[Rz 21] β_5 – ein effizienter Indikator für Verstöße gegen Art. 3 GIG?

Der *Erkenntniswert* des Regressionskoeffizienten β_5 für die Entdeckung von Verstößen gegen Art. 3 GIG in einem untersuchten einzelnen Betrieb ist schwierig zu beurteilen. Dies dürfte unter anderem damit zusammenhängen, dass die Berechnung von Regressionskoeffizienten β_i nach dem Kriterium der kleinsten Quadrate (KQ-Kriterium = OLS Ordinary Least Square) auf *statistischen* Berechnungen beruht, also auf einer Vielzahl gleichartiger Beobachtungen. Demgegenüber beruht der normative Begriff Einhalten / Verletzen von Art. 3 GIG *nicht auf statistischen Zusammenhängen*, (was gelegentlich vergessen wird), sondern auf einer Berechnung einer Differenz $d = y_m - y_f$ zwischen einem (evtl. mehreren) *individuellen* Männerlohn y_m und einem (evtl. mehreren) *individuellen* Frauenlohn y_f in einem Einzelfall. Vom mathematischen Standpunkt aus leuchtet es nicht ohne Weiteres ein, warum eine Grösse wie ein Regressionskoeffizient β_i , welcher auf statistischen Grössen wie Mittelwert, Varianz (bzw. Standardabweichung) und Kovarianz beruht, Erkenntnisse über Differenzen zwischen zwei *individuellen* y-Werten (Männerlohn, Frauenlohn) eines Einzelfalles generieren soll – und nur derartige Differenzen können Verstöße gegen Art. 3 GIG verkörpern.³³

Ob ein hoher Betrag von β_5 ausnahmslos *für* die Häufigkeit von Verstößen gegen Art. 3 GIG in dem betreffenden untersuchten Betrieb spricht, erscheint nicht als gesichert. Es sind Gegenbeispiele vorstellbar, bei denen ein einzelner untersuchter Betrieb zwar einen grossen Betrag von β_5 erzielt, aber weniger (oder gar keine) Verstöße gegen Art. 3 GIG aufweist als andere Betriebe, mit einem geringeren Betrag von β_5 . So wird ein einzelner Betrieb zwar einen hohen Betrag von β_5 , jedoch kaum Verstöße gegen Art. 3 GIG generieren, wenn er ausschliesslich Frauen (aber keine Männer) im Niedriglohnsegment in typischen Frauenberufen mit verhältnismässig geringer Arbeitsplatzproduktivität beschäftigt, und alle (seltenen) Männer als administrative und technische Kader zu höheren Löhnen in der Geschäftsleitung des Betriebs arbeiten. In einem derartigen Betrieb wird es keine (oder nur sehr selten) Arbeitsplatzbewertungen geben, welche gleichzeitig auf Männer und Frauen zutreffen und damit Ausgangspunkte für einen direkten Lohnvergleich und die Berechnung der Differenz d und allenfalls von Verstößen gegen Art. 3 GIG bilden könnten. Im Ergebnis wird der Mittelwert aller Frauenlöhne \bar{y}_f in diesem Betrieb relativ niedrig, derjenige aller Männerlöhne \bar{y}_m aber hoch sein. Der Regressionskoeffizient β_5 als Differenz zwischen den beiden Mittelwerten wird daher ebenfalls einen relativ hohen Betrag ergeben. Beispiele: Die bereits genannten frauendominierten Betriebe wie Grosswäschereien, Reinigung von Gebäuden und Räumen und teilweise Gastronomie, Hotellerie und Detailhandel. Ähnliches dürfte für die männerdominierten Bereiche wie die Bauindustrie, insbesondere Untertagebau, allenfalls auch für das Transportgewerbe gelten, wo nur sehr wenige Frauen an administrativen Arbeitsplätzen, die Männer aber durchwegs in technischen Funktionen mit hoher Arbeitsplatzbewertung arbeiten.

Um die eingangs formulierten Fragen zu beantworten: Falls in einem einzelnen Betrieb ein $\beta_5 = -0.111$ gemessen wird, können sicher *nicht* alle in diesem Betrieb beschäftigten Frauen *tel quel* 11% mehr Lohn und entsprechende Lohnnachzahlungen einfordern. Und ebenso wenig arbeiten 11% der beschäftigten Frauen in einem Arbeitsverhältnis, welches gegen Art. 3 GIG verstösst.

Diese Überlegungen zeigen, dass die Häufigkeit von Verstössen gegen Art. 3 GIG in einem einzelnen Betrieb von einer Reihe variabler Umstände des Einzelfalles innerhalb dieses Betriebes abhängt und mit β_5 allein nicht auf adäquate Weise «eingefangen» werden kann. Ob ein individueller Verstoss gegen Art. 3 GIG vorliegt, entscheidet allein die «lokale» Differenz $d = y_m - y_f$ zwischen den involvierten individuellen Frauen- und Männerlöhnen mit gleicher Arbeitsplatzbewertung. Im Ergebnis erscheint es darnach als fraglich, ob die Höhe des Regressionskoeffizienten β_5 einen *ein-eindeutigen und zuverlässigen Indikator* für die Entdeckung von Verstössen gegen Art. 3 GIG in einem einzelnen Betrieb darstellt.

[Rz 22] Und wie weiter?

Wie auch immer man schliesslich den *Erkenntniswert* von β_5 für das Auffinden von Verstössen gegen Art. 3 GIG qualifizieren sollte, für das Aussortieren «verdächtiger» Betriebe müssen *normative* Entscheidungsgrenzen (cut-off values) für β_5 festgelegt werden. Diese Entscheidungsgrenzen können höchstens teilweise kognitiv begründet werden und es müssen daher auch rechtspolitische Gesichtspunkte berücksichtigt werden. Und dieses Festlegen von Entscheidungsgrenzen führt deshalb unvermeidlich zu *Kontroversen* um deren Höhe. Und wenn diese Kontroverse einmal abgeschlossen ist, so beginnt die nächste Kontroverse um *Toleranzbeträge*: Um welchen Betrag darf β_5 von der einmal festgelegten Entscheidungsgrenze *abweichen*, ohne dass der untersuchte Betrieb als verdächtig qualifiziert wird? Dazu unten Abschnitt 2.4 (Standard-Analysemodell EBG).

Und wenn man schliesslich glücklich einen Betrieb als «verdächtig» aussortiert hat, so ist damit noch nicht viel gewonnen. Denn die vermuteten Verstösse gegen Art. 3 GIG müssen noch im Einzelnen *lokalisiert*, also innerhalb des untersuchten Betriebs aufgefunden und einzelnen Arbeitnehmer/innen zugeordnet werden. Zu diesem Zweck müssen *tatsächlich* ausbezahlte, «beobachtete» Löhne y_m von Männern gefunden werden, welche bei gleicher Bewertung des Arbeitsplatzes (in Punkten) höher sind als derjenige der betreffenden Frau(en) y_f . Auch nach einer Regressionsanalyse eines einzelnen untersuchten Betriebes muss der Nachweis eines Verstosses gegen Art. 3 GIG «manuell» durch *direkte Konfrontation* der einzelnen beteiligten Frauen- und Männerlöhne erfolgen.

Dieser manuelle Nachweis wird regelmässig mit einem *erheblichen Arbeitsaufwand* verbunden sein, welcher auch durch die Beweislast erleichterung von Art. 6 GIG *nicht substantiell vermindert* wird.³⁴ Und wenn schliesslich die höheren Männerlöhne y_m bei identischer Arbeitsplatzbewertung glücklich aufgefunden sind, so ist damit noch nicht alles gewonnen: Denn schon beginnt die nächste Kontroverse um *Toleranzwerte* bei der Lohnhöhe: Wie viele Prozente darf der Männerlohn y_m vom Frauenlohn y_f abweichen, ohne dass ein Verstoss gegen Art. 3 GIG vorliegt?

[Rz 23] Andere Dichotomien?

Zudem sollte im vorliegenden Zusammenhang nicht übersehen werden, dass aus jeder anderen dichotomen

Eigenschaft im Kontext der Lohnfindung ebenfalls eine Dummyvariable gebildet werden könnte, welche einen ähnlichen Regressionskoeffizienten β_i erzeugen würde wie die Dummyvariable Frauen/Männer. Diese anderen Dummyvariablen könnten als *Referenzvariablen* verwendet und mit dem geschlechtsspezifischen β_5 der Dummyvariablen FM verglichen werden. Übrigens stehen viele dieser dichotomen Eigenschaften mit Auswirkungen auf die Lohnfindung nach Art. 8 der Bundesverfassung der Schweizerischen Eidgenossenschaft (BV; SR 101) ebenfalls unter einem Diskriminierungsverbot; dies scheint aber die Wissenschaft weniger zu interessieren als der *gender wage gap*. Es darf beispielsweise vermutet werden, dass eine Dummyvariable für die Dichotomien Schweizer / Ausländer, Europäer / Aussereuropäer, Muttersprache Deutsch / andere oder Tätigkeit technisch / kaufmännisch ein vergleichbares β_i generieren würde wie die Dummyvariable Frauen/Männer.³⁵

[Rz 24] Wie bereits an anderer Stelle gezeigt worden ist,³⁶ erscheint es zudem als zweifelhaft, ob die multiple Regressionsanalyse sich dafür eignet (sei es mit Hilfe von β_5 , R^2 oder anderer Indikatoren), eine *mittelbare oder indirekte* Lohndiskriminierung von Frauen (Beschäftigungsdiskriminierung) zu entdecken. Diese wird bekanntlich dadurch verursacht, dass Frauen bei der *Vergabe* betrieblichen Funktionen mit hohem Anforderungsniveau, hoher hierarchischer Stellung und hohem Lohnniveau benachteiligt werden. Der individuelle Datensatz der durch eine derartige *Zuteilung von Arbeit* («Aufgabenzuteilung», «Beförderung» gemäss Art. 3 Abs. 2 GIG) benachteiligten Frauen erscheint regelmässig als *modellkonform* und dürfte daher die verwendeten Indikatoren (β_5 usw.) kaum beeinflussen: Niedrige betriebliche Stellung – niedrige Anforderungen – niedriger Lohn.

Als vorläufiges Teilergebnis kann an dieser Stelle festgehalten werden: Strukturierte multi-criteria Methoden sind zwar unentbehrlich für eine reproduzierbare und damit rechtssichere Anwendung von vielschichtigen Gesetzesbegriffen wie «gleichwertig» (Art. 8 Abs. 3 BV) oder «benachteiligen» (Art. 3 Abs. 1 GIG). Die multiple Regressionsanalyse OLS hat indessen nur einen beschränkten Erkenntniswert für die rechtssichere Entdeckung individueller Verstösse gegen Art. 3 GIG innerhalb eines Betriebs.

2.3. Institutionelle Schranken

[Rz 25] Vergleiche mit ausserbetrieblichen Parametern?

Die Freiheit zur Gestaltung des Inhalts von Verträgen regiert als zentrales Gestaltungsinstrument des *Privatrechts* auch den Arbeitsvertrag und erfasst damit auch die Festlegung des Arbeitslohnes. Dieser wird in erster Linie durch *individuelle Vereinbarung* zwischen den unmittelbar Beteiligten (Arbeitgeber/ innen, Arbeitnehmer/innen) festgelegt; deren Gestaltungsfreiheit kann aber durch zwingendes Gesetzesrecht oder durch Gesamtarbeitsvertrag (GAV) eingeschränkt werden.³⁷ Dementsprechend sind die Diskriminierung und das Diskriminierungsverbot des Art. 3 GIG jedenfalls im *Bereich des Obligationenrechts* (Art. 2 bis 10 GIG) *inner-betriebliche* Begriffe und entfalten keine kollektiven, betriebs-übergreifenden Wirkungen. Ein privatrechtlicher Arbeitgeber A darf aufgrund von Art. 3 GIG nur innerhalb seines Betriebs A nicht diskriminieren, er darf aber sehr wohl eine Frau schlechter bezahlen als der/die Arbeitgeber/in X eine Frau im Betrieb X, oder als der/die Arbeitgeber/in Y einen Mann im Betrieb Y.³⁸

Daraus ergibt sich für die *Methode* der Lohnanalyse im Rahmen des GIG, dass ausserbetriebliche Parameter zum Nachweis oder zur Rechtfertigung einer Lohnungleichheit nur dann verwendet werden dürfen, wenn dies zweiseitig vereinbart und zum Vertragsinhalt erhoben worden ist.³⁹ Zwar verkörpert Art. 3 GIG fraglos *zwingendes*

Gesetzesrecht, doch bildet er zumindest im Bereich des Privatrechts nach der hier vertretenen Ansicht noch *keine ausreichende Rechtsgrundlage* für die Verwendung *ausser-betrieblicher* Parameter bei der Beurteilung der «Gleichwertigkeit» einer Arbeit. Wenn nach der projektierten Revision des GIG ein Obligatorium für betriebliche Lohnanalysen eingeführt werden und dem Betrieb dadurch eine fremde, nicht von den direkt Beteiligten konzipierte statistische Methode der Lohnanalyse (Regressionsmodell) *aufgezwungen* werden sollte, so würde dies einen massiven und tiefgreifenden *Eingriff in die Vertragsfreiheit* der Beteiligten und gegebenenfalls in die Tarifautonomie der Sozialpartner darstellen. Als besonders problematisch erscheint es, durch die geplante Revision den direkt Beteiligten die *Auswahl der Beurteilungskriterien x_i (unabhängige Variable) zu entziehen*, nach denen ihr Lohnsystem analysiert werden soll.

[Rz 26] Eine Erweiterung einer Lohnanalyse auf Daten von «faktischen» Arbeitgeber/innen, welche zwar formell nicht Lohnschuldner/innen oder Arbeitgeber/innen sind, aber im Hintergrund bei der Lohnfindung mitwirken, ist zwar u.U. bei gesellschaftsrechtlichen Verflechtungen vorstellbar. Dann müsste gegebenenfalls geprüft werden, wer die Fäden der Lohnfindung im Hintergrund tatsächlich gezogen hat und das Lohnsystem des fraglichen Betriebs *tatsächlich* festgelegt hat. Dann dürften u.U. auch Daten aus dem Betrieb des «faktischen» Arbeitgebers für die Lohnanalyse beigezogen werden.⁴⁰ Das absichtsvolle Teilen von Betrieben in juristisch selbständige Einheiten zum Zwecke der Umgehung des Verbots der Lohndiskriminierung wird sich allerdings nie (vollständig) verhindern lassen.⁴¹ Dagegen dürfte der Vergleich eines betrieblichen Lohnsystems mit ausserbetrieblichen Erkenntnissen aus *kollektiven* Untersuchungen (Branchen, Regionen, sog. «Referenzmarkt») jedenfalls im Privatrecht *de lege lata* nicht zulässig sein.

[Rz 27] Untereinheiten: Aufgrund der Vertragsfreiheit in der Lohnfindung dürfte es ebenso wenig zulässig sein, dass die Stichprobe aller Arbeitnehmer/innen eines privatrechtlichen Betriebs zur Berechnung von bereichsspezifischen, «modellkonformen» Löhnen \hat{y} aufgeteilt wird, also isolierte Stichproben für einzelne Untereinheiten eines Betriebes mit unterschiedlichen «modellkonformen» Löhnen gebildet werden. Einmal abgesehen davon, dass gar nicht feststeht, ob derartigen bereichsspezifischen modellkonformen Löhnen \hat{y} überhaupt eine rechtliche Bindungswirkung zukommen würde, welche die Grundlage für Lohnforderungen bilden könnte. Diese Problematik wurde in [BGE 130 III 154](#) zwar kurz erwähnt, aber nicht beurteilt.⁴²

[Rz 28] *Ausserbetriebliche* Kriterien x_i und Regressionskoeffizienten b_i ? Darf die Zahl und der Inhalt der unabhängigen Variablen (Kriterien) x_i ohne Mitwirkung der unmittelbar Beteiligten Vertragsparteien in die betriebliche Lohnanalyse eingeführt werden, beispielsweise durch ein Expertengutachten im Rechtsstreit? Dürfen einem Betrieb A *ausserbetriebliche* Kriterien x_i für die Lohnanalyse *aufgezwungen* werden, auch wenn sie nach Ansicht des Arbeitgebers/in A oder der Arbeitnehmer/innen A nicht relevant sind und sachlich besser geeignete innerbetriebliche Kriterien für die Lohnkontrolle des Betriebs A zur Verfügung stehen? Dürfen *ausser-betriebliche* Kriterien x_i , welche zur Regressionsanalyse einer ganzen *Branche* oder eines *Referenzmarktes* eingesetzt worden sind, für die Regressionsanalyse in Betrieb A verwendet werden? Dürfen Regressionskoeffizienten b_i , welche durch Regressionsanalyse aus den Lohndaten einer Branche oder Region berechnet worden sind, auf einen Betrieb A übertragen werden, um dort einen modellkonformen Lohn \hat{y} zu berechnen und nach Verstössen gegen Art. 3 GIG zu fahnden?⁴³

Alle diese Fragen sind *de lege lata* und für den Bereich des Privatrechts *negativ* zu beantworten. Diskriminierung ist

im Bereich des Privatrechts – wie ausgeführt – ein inner-betrieblicher Begriff. Im privatrechtlichen Arbeitsverhältnis dürfen daher ohne Mitwirkung der direkt Beteiligten keine ausser-betrieblichen Variablen x_i oder Koeffizienten b_i aus anderen Betrieben, oder *Mittelwerte* einer Branche oder Region eingesetzt werden.⁴⁴

[Rz 29] Im mehrfach erwähnten [BGE 130 III 145 ff](#) (2003) sind zumindest in einem der beiden Gutachten vor dem Zivilgericht Waadt anscheinend auch ausser-betriebliche Koeffizienten (Mittelwerte einer ganzen Region Waadt und Genf) eingesetzt worden: Eine Rechtsgrundlage für einen derartigen Einsatz betriebs-übergreifender regionaler oder berufsspezifischer (Vergleichs-) Parameter in einem Gutachten zur Lohnanalyse war und ist bis heute *de lege lata* für den Bereich des Privatrechts *nicht* erkennbar; leider hat aber anscheinend die Zivilkammer des Kantonsgerichts Waadt diese Frage nicht untersucht. Dies blieb ohne Folgen, da im waadtländischen Verfahren tatsächlich nur betriebs-interne Parameter (Koeffizienten) den Fall entschieden haben, und das Bundesgericht dies als Sachverhaltsfeststellung gewertet und nicht materiell überprüft hat.⁴⁵

[Rz 30] Sachverhaltsfeststellung oder Rechtsanwendung?

Die Rechtsprechung ist schon Jahre vor dem Einsatz der multiplen Regressionsanalyse in der Lohnanalyse zu dem Ergebnis gelangt, dass die Bewertung und Einstufung einer bestimmten Arbeitsleistung weder eine reine Sach-, noch Rechts- noch Ermessensfrage darstelle, sondern Elemente von allen dreien enthalte.⁴⁶ In dem eben zitierten [BGE 130 III 145](#) wurden die statistischen Verfahren der Regressionsanalyse der Sachverhaltsfeststellung zugerechnet und das Gericht konnte sich dadurch eine inhaltliche Beurteilung der Methode und ihrer Resultate ersparen.⁴⁷ Demgegenüber überprüft das Bundesgericht im öffentlichen Personalrecht die methodischen Einzelheiten der analytischen Arbeitsplatzbewertung (VFA, ABAKABA) und deren Anwendung auf den Einzelfall und behandelt diese Anwendung der methodischen Kriterien daher vermutlich als *Frage der Rechtsanwendung*.⁴⁸

[Rz 31] Indessen weisen die diskutierten statistischen Methoden (multiple Regression, aber auch ABAKABA und alle anderen Verfahren der analytischen Arbeitsplatzbewertung nach dem Modell der gewichteten Addition SAW) durchwegs einen Inhalt auf, welcher *generell-abstrakter* Art ist und zudem eine normative Funktion erfüllt. Sie sind mathematische *Algorithmen*, welche auf eine unbekannte Zahl von Einzelfällen auf immer die gleiche, unveränderte Art angewendet werden können. Der Anwender verändert diese Algorithmen bei der Anwendung *nicht*. Damit verkörpern diese Algorithmen indessen «*bewährte Lehre*» (*Literatur*) im Sinne von Art. 1 Abs. 3 des Schweizerischen Zivilgesetzbuches ([ZGB; SR 210](#)), sind also rechtlich zu beachten, auch wenn sie von privaten Autoren ausserhalb der staatlichen Behördenorganisation geschaffen worden sind. Dies würde umso mehr gelten, wenn eine Bundesbehörde eine «Anerkennung» dieser Methoden aussprechen würde, wie dies im Entwurf Art. 13–b vorgesehen ist. Eine derartige «Anerkennung» durch eine hoheitlich handelnde Behörde des Bundes würde die Methode normativ erheblich aufwerten.⁴⁹

[Rz 32] Die Anwendung derartiger Algorithmen bzw. statistischer Methoden mit konstantem, generell-abstraktem Inhalt auf den Einzelfall ist indessen Rechtsanwendung *und nicht Sachverhaltsfeststellung*. Sie sollte daher schon *de lege lata* vom Bundesgericht überprüft werden, und nicht nur bei *offensichtlich* unrichtiger «Feststellung des Sachverhalts».⁵⁰ Es ist zu hoffen, dass derartige statistische Algorithmen in Zukunft in der Rechtsanwendung nicht nur immer wichtiger werden, sondern auch, dass sie als «*bewährte Lehre*» im Sinne von Art. 1 Abs. 3 ZGB anerkannt und damit vom

Bundesgericht uneingeschränkt überprüft werden können. Nur so werden sie dazu beitragen können, die *Rechtssicherheit* in ihrem Anwendungsbereich zu erhöhen.⁵¹

Als vorläufiges Teilergebnis kann an dieser Stelle festgehalten werden: Strukturierte multi-criteria Methoden sind zwar unentbehrlich für eine reproduzierbare und damit rechtssichere Anwendung von vielschichtigen Gesetzesbegriffen wie «gleichwertig» (Art. 8 Abs. 3 BV) oder «benachteiligen» (Art. 3 Abs. 1 GIG). Die multiple Regressionsanalyse OLS unterliegt indessen bei der Beurteilung von Lohndiskriminierungen auf Betriebsebene gewichtigen Einschränkungen institutioneller Art, namentlich einem Ausschluss ausserbetrieblicher Parameter.

FORTSETZUNG TEIL 2 FOLGT.

FRITZ DOLDER, Prof.Dr.iur. Dr.sc.techn. ETH, (em) Universität Basel.

MATTHIAS MOELLENEY, Center for Human Resources Management & Leadership, HWZ Hochschule für Wirtschaft Zürich.

Interessenkonflikte: Keine / unabhängig. Die Autoren sind nicht am gewerblichen Geschäft mit betrieblichen Lohnanalysen beteiligt. Der vorliegende Beitrag wurde weder durch den Bund noch durch Verbände oder Unternehmungen der Privatwirtschaft angeregt oder finanziell gefördert. M. Moelleney ist zurzeit Präsident der Zürcher Gesellschaft für Personalmanagement (ZGP).

-
- 1 DANIEL KAHNEMAN (Nobelpreis für Wirtschaft 2002), *Thinking, Fast and Slow*, London 2012, S. 229: «In contrast, Meehl and other proponents of algorithms have argued strongly that it is *unethical* to rely on intuitive judgements for important decisions if an algorithm is available that will make fewer mistakes».
 - 2 Art. 8 Abs 3 BV 1998 / 1999, bzw. Art. 4 Abs. 2 aBV gemäss Revision 14. Juni 1981, Art. 3 GIG «weder direkt noch indirekt benachteiligt»; zur qualitativen Überlegenheit von multi-criteria Methoden über unstrukturierte, intuitive one-reason Entscheidungen: KAHNEMAN, a.a.O., S. 227; Kapitel 21 (Intuition vs. Formulas) ist voll von erfolgreichen Anwendungsbeispielen von multi-criteria Methoden: S. 222–233.
 - 3 Anwendungsbeispiel der SAW-Methode in BGE 126 II 217; in BGE 131 II 393 wurde im *öffentlichen* Personalrecht das Verfahren der «Vereinfachten Funktionsanalyse» (VFA), diskutiert und im Ergebnis als nicht diskriminierend qualifiziert (S. 400/401 Erw. 6.1 und 6.4, Kanton Solothurn). Literatur: CHRISTIAN KATZ / CHRISTOF BAITSCH (1996), *Lohnleichheit für die Praxis*, 1. A. Zürich 1996, mit Fragebogen ABAKABA S. 70 ff, und Formularteil S. 101–114; CHRISTIAN KATZ / CHRISTOF BAITSCH (2006), *Arbeit bewerten – Personal beurteilen, Lohnsysteme mit Abakaba*, 2. A. Zürich 2006.
 - 4 Beispiel Kanton Zürich: Personalverordnung vom 16. Dezember 1998, 177.11:
 - § 8. ¹ Der Regierungsrat und die obersten kantonalen Gerichte legen in der Vollzugsverordnung den Einreichungsplan fest. Dieser enthält die Richtpositionen, die nach 29 Lohnklassen geordnet sind.
 - ² Die Richtpositionen werden gemäss dem Verfahren der «Vereinfachten Funktionsanalyse» eingereiht. Massgebend sind die *vorausgesetzte* Ausbildung und Erfahrung, die mit der Stelle verbundenen geistigen *Anforderungen*, die Verantwortung, die psychischen und körperlichen *Anforderungen* und Belastungen, die *Beanspruchung* der Sinnesorgane und die besonders äussern Arbeitsbedingungen, denen die Inhaberin oder der Inhaber der Stelle ausgesetzt ist.
 - 5 Es ist daher nicht präzise, wenn etwa ausgeführt wird:

Erläuternder Bericht des Bundesamtes für Justiz vom 15. November 2015, S. 13, [...] 2.1.2 Artikel 13b Methode der Lohnanalyse [...] «Die Verwendung der Regressionsanalyse im Bereich der Lohnleichheit wurde vom Bundesgericht zugelassen».

CHRISTINA FELFE / JUDITH TRAGESER / ROLF ITEN, Studie zu den statistischen Analysen der Eidgenossenschaft betreffend die Lohnleichheit von Frau und Mann, Auftraggeber: Eidgenössisches Büro für die Gleichstellung von Frau und Mann EBG, Letzte Überarbeitung: 2. November 2015, zitiert als FELFE et al. (2015), S. 54 «Zuletzt möchten wir noch einmal daran

erinnern, dass das Bundesgericht bisher als statistische Methode nur die OLS-Methode *geprüft* und *zugelassen* hat». Ebenso YVES FLÜCKIGER / ROMAN GRAF, *Lohndiskriminierung und Entschädigung vor Gericht*, in: *Human Resource Management Jahrbuch 2012*, Weka Verlag, Zürich, S. 75–96, S. 87: «Im Entscheid ... wird die angewandte Methode als *zuverlässiges*, dem Schweizer Recht entsprechendes Vorgehen *anerkannt*, das erlaubt, eine mögliche Lohndiskriminierung festzustellen, sowie die Höhe des Schadens, welcher der diskriminierten Person entstanden ist, zu berechnen». OLIVER SCHRÖTER / CLAUDIO MARTI WHITEBREAD, *Die Messmethoden des Bundes zur Lohnungleichheit sind aussagekräftig*, *Die Volkswirtschaft*, 24. Februar 2016, «Auch die verwendete statistische Methode des EBG-Analysenmodells ist wissenschaftlich und *juristisch anerkannt* und sollte deshalb gemeinsam mit der in der Praxis bewährten *Toleranzschwelle von fünf Prozent* beibehalten werden» .

Das Bundesgericht hat sich bisher glücklicherweise (noch) nicht in die *alles dominierende Diskussion* über Zahl und Inhalt der verwendeten unabhängigen Variablen x_i eingeschaltet.

- 6 Eidgenössisches Büro für die Gleichstellung von Frau und Mann EBG, Standard-Analysenmodell zur Überprüfung der Lohnungleichheit zwischen Frau und Mann im Beschaffungswesen des Bundes (Methodenbeschreibung), Anleitung zur Durchführung der standardisierten Überprüfung, Bern, Juni 2016, zitiert als «Methodenbeschreibung 2016», S. 1: «Das Büro für arbeits- und sozialpolitische Studien (BASS) hat im Auftrag des EBG und der BKB das Standard-Analysenmodell zur Überprüfung der Einhaltung der Vorgaben entwickelt und in einer Pilotphase (2001–2003) getestet». (BKB = Beschaffungskonferenz des Bundes)
- 7 FRITZ DOLDER, [Rechtsanwendung mit Hilfe von multi-criteria Methoden](#), in: Jusletter 1. Juni 2015, Rz. 74, 81, 106.
- 8 EDI/EBG, Standard-Analysenmodell zur Überprüfung der Lohnungleichheit zwischen Frau und Mann im Beschaffungswesen des Bundes (Methodenbeschreibung), Anleitung zur Durchführung der standardisierten Überprüfung, Bern, Stand Juni 2016; zitiert als «Methodenbeschreibung 2016»; EBG, Testen Sie die Lohnungleichheit in Ihrem Unternehmen (Anleitung logib), Bern August 2014.
- 9 Art. 8 Abs. 1 Bst. c BöB. «Sie vergibt den Auftrag nur an Anbieter oder Anbieterinnen, welche für jene Arbeitnehmer und Arbeitnehmerinnen, die Leistungen in der Schweiz erbringen, die Gleichbehandlung von Frau und Mann in Bezug auf die Lohnungleichheit gewährleisten».
- 10 Sehr bezeichnend insofern die Antwort des Bundesrats vom 9. November 2016: [16.3657 Motion Franz Grüter](#), Einreichungsdatum: 15. September 2016; Lohnungleichheit im Beschaffungswesen. Aber fair und korrekt; Stand der Beratungen: Im Rat noch nicht behandelt: «Die Verwendung von Logib ist weder im Bundesbeschaffungsrecht noch im üblicherweise von Bundesbeschaffungsstellen eingesetzten Selbstdklarationsformular der Beschaffungskonferenz des Bundes vorgegeben. Der Nachweis der Lohnungleichheitseinhaltung ist auch mit anderen Instrumenten möglich und zulässig, *vorausgesetzt, dass diese Lohnanalysen nach dem Standard-Analysenmodell des Bundes erfolgen, auf dem auch Logib basiert*».
- 11 Erläuternder Bericht des Bundesamtes für Justiz vom 15. November 2015, S. 13, 2.1.2 Artikel 13b Methode der Lohnanalyse [...] Mit dem Standard-Analysenmodell des Bundes (Instrument Logib) verfügt der Bund bereits über eine wissenschaftlich und *juristisch anerkannte* Methode zur Durchführung der Lohnanalyse durch die Arbeitgeberinnen und Arbeitgeber.
Mit der *statistischen Regressionsanalyse* (z.B. *Standard-Analysenmodell des Bundes*) kann die geschlechtsspezifische Lohnsituation in einem Unternehmen überprüft werden. *Es lässt sich damit ermitteln, ob ein Geschlecht systematisch benachteiligt wird.* [FD:??] Bestehende Lohnunterschiede werden mittels der persönlichen Qualifikationsmerkmale (Ausbildung, Berufserfahrung, Dienstjahre) sowie der arbeitsplatzbezogenen Merkmale (Anforderungsprofil, berufliche Stellung) erklärt. Es ist jedoch grundsätzlich nicht ausgeschlossen, im Rahmen der statistischen Regressionsanalyse allenfalls auch weitere persönliche oder arbeitsplatzbezogene Merkmale zu berücksichtigen. Die statistische Regressionsanalyse ist auf nationaler (Universitäten, Bundesamt für Statistik) wie internationaler Ebene ein anerkannter und etablierter Standard. *Die Verwendung der Regressionsanalyse im Bereich der Lohnungleichheit wurde vom Bundesgericht zugelassen.* [...] Voraussetzung für die Durchführung einer *Regressionsanalyse* sind genügende Fallzahlen: Erforderlich sind mindestens 50 Beschäftigte im Unternehmen, wobei beide Geschlechter mit mindestens zehn Personen vertreten sein müssen. [...]
- 12 Beispiele früherer Studien in den USA: ALAN S. BLINDER, *Wage discrimination: reduced form and structural estimates*, *Journal of Human Resources* 8, 1973, S. 436–455; RONALD OAXACA, *Male-female wage differentials in urban labor markets*, *International Economic Review* 14(3), 1973, S. 693–709; Aktuell in der Schweiz: FLURINA SCHMID, (FORS, Uni Lausanne), *The gender gap in Switzerland over time*, *Schweizerische Zeitschrift für Soziologie* 42 (3), 2016, 443–467 (November 2016) S. 461 «There are still major unexplained wage differences between women and men, but according to the data it seems that this gap gets smaller from one cohort to the other». S. 463 «Selection into gender-specific jobs plays a key role in the emergence of the gender pay gap. Most women work in female-dominated occupations, which unfortunately often go with a lower wage level. Changing from a female-dominated to an integrated or male-dominated occupation appears to be financially advantageous for women. [...]
- Die Botschaft ist nicht zu übersehen: Schickt Eure Töchter zum Studium an die ETH oder in technische Berufslehren!
- 13 Die Methode der kleinsten Quadrate (GAUSS 1835) wurde 1908 von PEARSON zur Beurteilung multivariater Phänomene eingeführt. Zur Einführung in die multiple Regressionsanalyse sind auch für Juristen geeignet. KLAUS BACKHAUS / BERND ERICHSON / WULFF PLINKE / ROLF WEIBER, *Multivariate Analysemethoden*, 13. A. Berlin 2011; zum Bestimmtheitsmass R^2 S. 72–76, zur Beurteilung der Signifikanz von R^2 mittels F-Statistik S. 76–79; LUDWIG FAHRMEIR / THOMAS KNEIB / STEFAN LANG /

BRIAN D. MARX, Regression; Models, Methods and Applications, Berlin / Heidelberg 2013; DIETER URBAN / JOCHEN MAYERL, Regressionsanalyse, 4. A. Wiesbaden 2011.

- 14 BACKHAUS ET AL., a.a.O., S. 74–76: umso grösser, je höher der Anteil der erklärten Streuung an der Gesamtstreuung ist. Extremfall, wenn die gesamte Streuung erklärt wird ist $R^2 = 1$. Ursprünglich: Korrelation zwischen den beobachteten und den geschätzten Y-Werten (deshalb R^2); R^2 wird in seiner Höhe durch die *Zahl* der Variablen Xi beeinflusst. URBAN/MAYERL, a.a.O., S. 59 R^2 : «... scheint er auch in idealer Weise dazu geeignet zu sein, mehrere Modellschätzungen miteinander vergleichbar zu machen». [...] «Dementsprechend ist in der Praxis der Regressionsanalyse oftmals auch eine Jagd nach einem hohen R^2 zu beobachten. Die *Stunde des Glücks* scheint demjenigen geschlagen, dessen Regressionsmodell ein R^2 von *über 0.60* aufweist».
- 15 URBAN/MAYERL, a.a.O., S. 61: Wenn z.B. der R^2 -Koeffizient klein ist, dann kann das daran liegen, dass seine Komponenten negativ beeinflusst werden, und zwar a) durch die fehlerhafte Messung von Variablen, b) durch den fehlerhaften Ausschluss erklärungsbedürftiger Variablen, [und die fehlerhafte Berücksichtigung nicht relevanter Variablen FD] c) durch eine fehlerhafte Modellspezifikation (z.B. falsche Linearitätsannahmen).
- 16 Literatur zur Anwendung der multiplen Regressionsanalyse im Rahmen des Gleichstellungsgesetzes: YVES FLÜCKIGER, Approche économique des discriminations salariales, AJP 2001, 1340 f.; YVES FLÜCKIGER / ROMAN GRAF, Lohndiskriminierung und Entschädigung vor Gericht, Jahrbuch 2012, Human Resources Management, Zürich 2012: WEKA-Verlag, S. 75–96; ELISABETH FREIVOGEL, Kommentar GIG, 2. A. 2009, N. 119 bis 127 zu Art. 3 GIG; MARIANNE SCHÄR MOSER / JÜRIG BAILLOD, Analyse von Lohndiskriminierung, AJP 2006, 1384–1391; SILVIA STRUB, Lohnungleichheit und Lohndiskriminierung, AJP 2006, 1376–1383.
- 17 Eben weil die kollektive Beurteilung mit Hilfe von β_5 bzw. R^2 (oder anderen Indikatoren) nur eine *prima facie* Einschätzung des Lohnsystems eines einzelnen untersuchten Betriebs, aber *keine Lokalisierung* auf individuelle Arbeitnehmer/innen liefert.
- 18 ZGP Zürcher Gesellschaft für Personalmanagement, Vernehmlassung Änderung GIG, 23. Februar 2016, S. 3: «Die Problematik geht aber noch weiter: Mittels statistischer Analysen werden auf der Basis der (standardisierten) Ausbildungsjahre, der (potenziellen) Erwerbsjahre, der Dienstjahre, des Anforderungsniveaus und der beruflichen Stellung (hierarchische Position) die Abweichungen zu einem *plausiblen Lohn*, dem *richtigen Lohn*, berechnet, und zwar mit Formeln, die z.T. auf Schätzungen basieren. Den Schätzungen wiederum liegt ein angenommenes Lohnsystem zugrunde».
- 19 BGE 130 III 145, S. 149 «Par lettre du 5 février 1997, A. a été licenciée pour le 31 août 1997; libérée immédiatement de son obligation de travailler, il lui a été toutefois demandé de rester disponible pour répondre à des demandes d'informations relatives aux dossiers dont elle avait la charge. Pour remplacer la demanderesse, X. a engagé en 1997 E., né en 1960, en lui octroyant une rémunération égale à celle qu'avait atteinte l'intéressée en *fin de contrat*».
- 20 MATTHIAS MOELLENEY, Interview im Tages-Anzeiger vom 6. August 2014: «Eine Saläranalyse mit Logib gibt nur Resultate im Sinne des Erfinders.»; vgl. URBAN / MAYERL, a.a.O., S. 61 und OAXACA, R., a.a.O., (Fn. 12) S. 699: «It is clear that the magnitude of the estimated effects of discrimination *crucially depends* upon the *choice* of the control variables for the wage regressions. A researcher's *choice* of control variables *implicitly reveals his or her attitude* toward what constitutes discrimination in the labor market. If it were possible to control for virtually all sources of variation in wages, one could pretty well eliminate labor market discrimination as a significant factor in determining wage differentials by sex (or race)». Ähnlich THOMAS GEISER / MATHIAS MAUCHLE, Rechtsgutachten Marktlohn und konjunkturelle Lage als Rechtfertigungsgrund für geschlechtsspezifische Lohnunterschiede gemäss Art. 3 Abs. 2 GIG, Univ. St. Gallen, 10. Februar 2016, S. 35: «Bei gleichwertiger Arbeit ist in erster Linie auf wissenschaftliche Arbeitsplatzbewertungsmethoden abzustellen. Welche Kriterien in den Bewertungsmethoden berücksichtigt werden, hängt jedoch wiederum von Wertentscheidungen ab, die sich nicht durchwegs wissenschaftlich begründen lassen.» https://www.ebg.admin.ch/dam/ebg/de/dokumente/2016/06/rechtsgutachten_marketlohnundkonjunkturellegealsrechtfertigung (Alle Websites zuletzt besucht am 22. Mai 2016).
- 21 URBAN / MAYERL, a.a.O., S. 109–111 *Giftliste*; Beispiel: S1 Erhöhung der Zahl der unabhängigen Variablen; jede zusätzliche X-Variable kann den Wert von R^2 nur vergrössern. So lässt sich «mit einer entsprechend grossen Anzahl einflussstarker Variablen ein beliebig hoher Determinationskoeffizient erzielen.»
- 22 Übersicht bei BACKHAUS et al., a.a.O., S. 85/86; URBAN / MAYERL, a.a.O., S. 193.
- 23 ZGP Zürcher Gesellschaft für Personalmanagement, Vernehmlassung Änderung GIG, 23. Februar 2016, S. 4: «Logib baut – wie die Lohnstrukturhebung des BFS – auf einer statistischen Methode (Regression) auf, was den Nachteil hat, dass es nur in Unternehmen mit mindestens 100 bis 150 Mitarbeitenden und einer einigermaßen gleichen Verteilung der Geschlechter verwertbare Ergebnisse liefert». Dazu auch FELFE et al. S. 98, Einschränkung für die «betriebliche Ebene»: «... unter Verwendung der bereinigten Stichproben der LSE 2008, LSE 2010 oder LSE 2012, welche auf Beschäftigte in Unternehmen mit *mindestens 50 Beschäftigten sowie mindestens je zehn männlichen Beschäftigten und weiblichen Beschäftigten begrenzt wurden*, und der drei alternativen statistischen Methoden». Entwurf der deutschen Bundesregierung für ein «Lohnungleichheitsgesetz» vom 11. Januar 2017: § 12 Reichweite (1) Der Anspruch nach § 10 besteht für Beschäftigte nach § 5 Absatz 2 in Betrieben mit in der Regel *mehr als 200 Beschäftigten* bei demselben Arbeitgeber.
- 24 Für eine derartige «Linearisierung» von y und/oder x kann eine beliebige nichtlineare Funktion f eingesetzt werden: $X = f(x)$, so können beispielsweise die Werte von x potenziert (z.B. quadriert) oder *logarithmiert* werden. BACKHAUS et al., a.a.O., 86/87, Tabelle 1.21 und S. 96, Tabelle 1.25. Auch durch zweckmässige *Skalierung* einzelner Variablen kann eine

Linearisierung erzielt werden, ohne dass dies notwendigerweise zu einer Manipulation führt. Eine weniger seriöse Methode zur «Behandlung» von Daten vor einer Regressionsanalyse beschreibt Giftliste S3: URBAN / MAYERL, S. 110: «Schliesse die Minderheitsgruppe von Beobachtungsfällen aus der Regressionschätzung aus, bei denen sich deutlich abweichende Formen von Variablenbeziehungen erkennen lassen (z.B. *andere* als Linearbeziehungen). Diese Fälle ... können somit den R^2 -Wert reduzieren. Ausschluss entsprechender Fälle ... kann der R^2 -Wert beträchtlich ansteigen».

- 25 URBAN / MAYERL, a.a.O., S. 193 (Annahme A 5), BACKHAUS et al., a.a.O., S. 85/86 (Annahmen A1 bis A7: A7. Die Störgrössen u_k sind normalverteilt.)
- 26 Für die Einzelheiten dieser Überprüfung: URBAN / MAYERL, a.a.O., S. 193–196: Annahme A5 (Normalverteilung), BACKHAUS et al., a.a.O., S. 77–80. BACKHAUS 96: Nicht-Normalverteilung der Störgrössen.
- 27 URBAN/ MAYERL, a.a.O., S. 194. «So weist beispielsweise das Nettoeinkommen in aller Regel eine *rechtsschiefe* Verteilung auf. [...] In solchen Fällen geschieht es nicht selten, dass bei Berücksichtigung dieser Variablen in Regressionschätzungen empirische Residuenverteilungen entstehen, die vom Ideal einer Normalverteilung *stark abweichen*».
- 28 FELFE et al. (2015), S. 100: «muss auch auf Unternehmensebene garantiert werden, dass sowohl Frauen als auch Männer in den einzelnen beruflichen Stellungen und den verschiedenen Kompetenzniveaus vertreten sind. Falls dies nicht der Fall ist, müsste die Datenbasis auf Frauen und Männer *begrenzt werden*, die in den Ausprägungen der Erklärungsfaktoren *übereinstimmen*. Dies könnte man vor der Kontrolle der unternehmensspezifischen Lohngleichheit deskriptiv analysieren. Alternativ sollte man die Verwendung der «Doubly Robust Regression» in Erwägung ziehen».
- S. 125: «Die *hinreichend ähnliche Verteilung* von Frauen und Männern in einzelnen Ausprägungen der berücksichtigten Erklärungsfaktoren wird nicht thematisiert. Diese Thematik kann vor allem dann zu einem Problem führen, wenn ein Unternehmen sich dadurch charakterisiert, dass Frauen und Männer in sehr unterschiedlichen beruflichen Stellungen arbeiten und sehr unterschiedliche Tätigkeiten verrichten. In diesem Fall kann die OLS-Methode zu *verzerrten Ergebnissen* führen. [...] Unternehmen, in welchen sich Männer und Frauen hinsichtlich der Ausprägungen der Erklärungsfaktoren stark unterscheiden, *sollten jedoch mit Vorsicht behandelt werden*. Zusätzlich zu den bereits durchgeführten Plausibilisierungen im Rahmen der Kontrollen im öffentlichen Beschaffungswesen könnte eine Prüfung der Verteilung von Frauen und Männern *vorgeschaltet* werden. Bei Fällen in denen *Verzerrungen* auftreten könnten, können Sensitivitätsanalysen durchgeführt werden, um zu überprüfen, ob die Toleranzschwelle aufgrund dieser Problematik *signifikant überschritten* werden kann».
- S. 129: «Empfehlung 6: Die Methode der OLS-Regression ist eine anerkannte wissenschaftliche Methode und sollte aus praktischen Gründen beibehalten werden. Dabei ist aber der Gefahr Rechnung zu tragen, dass bei Unternehmen, in denen eine *hinreichend ähnliche Verteilung* von Frauen und Männern über die einzelnen Ausprägungen der Erklärungsfaktoren *nicht gegeben* ist, *Verzerrungen* auftreten können. Im Rahmen der Plausibilisierung bei den Kontrollen im öffentlichen Beschaffungswesen werden diese Gefahren bereits teilweise aufgefangen. Zusätzlich könnte der Analyse eine Prüfung der Verteilung von Frauen und Männern *vorgeschaltet* werden. Bei Fällen, in denen *Verzerrungen* auftreten könnten, können Sensitivitätsanalysen durchgeführt werden, um zu überprüfen, ob die Toleranzschwelle aufgrund dieser Problematik *statistisch signifikant überschritten werden kann*».
- 29 Überprüfung der statistischen Methoden des Bundes betreffend die Lohngleichheit von Frau und Mann, Bericht des Bundesrates in Erfüllung des Postulats 14.3388 Noser vom 2. Juni 2014, vom 18. November 2015, S. 12: «Die Studie empfiehlt jedoch, der Analyse eine Prüfung der Verteilung von Frauen und Männern im Unternehmen *vorzuschalten*. Bei Fällen, in denen *Verzerrungen* auftreten könnten, könnten zusätzliche Analysen durchgeführt werden, um zu überprüfen, ob die Toleranzschwelle aufgrund dieser Problematik *statistisch signifikant überschritten* würde. Diesbezüglich ist das EBG bereits daran, das bestehende Verfahren weiterzuentwickeln, um diese Problematik noch verstärkt zu berücksichtigen».
- 30 A.S. BLINDER, a.a.O. (Fn. 12), S. 441 «But correlation does not imply causation, and if we are interested in the latter, we ought to specify a more precise model». FLURINA SCHMID, a.a.O. S. 463 (oben Fn. 12) «Selection into gender-specific jobs plays a key role in the emergence of the gender pay gap. Most women work in female-dominated occupations, which unfortunately often go with a lower wage level. *Changing from a female-dominated to an integrated or male-dominated occupation appears to be financially advantageous for women*. [...]».
- 31 Methodenbeschrieb EBG (2016), S. 6/7.
- 32 In diesem Sinn FELFE et al., a.a.O., S. 34: «Differenz der *logarithmierten* Mittelwerte in %». EBG Methodenbeschrieb (2016), S. 6: «Bei semi-logarithmischen Regressionsgleichungen lässt sich der Koeffizient einer Dummy-Variable *näherungsweise* als prozentualer *Einfluss* des erklärenden Faktors auf die abhängige Variable (Prozentsatz-differenz) interpretieren. Diese Interpretation wird jedoch ungenauer, je höher der Koeffizient ist. Die exakte Berechnungsformel lautet: $100 \cdot (e^\beta - 1)$ ». STRUB, Methodisches Vorgehen zur Überprüfung der Lohngleichheit zwischen Frau und Mann im Beschaffungswesen des Bundes, Anleitung zur Durchführung der standardisierten Überprüfung Bern, Juni 2005, S. 7: «Der «Diskriminierungskoeffizient» β_5 misst den *Einfluss* des Geschlechtes auf den Lohn: [...] Bei semi-logarithmischen Regressionsgleichungen lassen sich die Koeffizienten *näherungsweise* als prozentualen Einfluss des erklärenden Faktors auf die abhängige Variable interpretieren».
- FELFE et al. S. 99, N. 54: «Koeffizienten in einem semi-logarithmischen Modell entsprechen *näherungsweise* einer prozentualen Veränderung des Lohnes (ein Koeffizient von 0,01 entspricht einer Veränderung von näherungsweise einem Prozent, ein Koeffizient von 1 entspricht einer Veränderung von näherungsweise 100 Prozent)».
- Interview Sommaruga im Tages-Anzeiger vom 10. Dezember 2016 (S. 3): «Eine Frau verdient im Durchschnitt für *dieselbe* Arbeit fast 9 Prozent weniger, *ohne dass es dafür eine Erklärung gibt*. Was für eine Botschaft ist das an die Frauen? Für mich ist es inakzeptabel, dass man immer noch um gleiche Löhne kämpfen muss».

Demgegenüber empfehlen Mathematiker Zurückhaltung bei der Interpretation von Regressionskoeffizienten: BACKHAUS et al., a.a.O., S. 70: «Die Grösse eines Regressionskoeffizienten darf allerdings *nicht* als Mass für die Wichtigkeit der betreffenden Variablen angesehen werden. ...denn der numerische Wert b_j ist abhängig von der Skala auf der die Variable X_i gemessen wurde. ... Eine andere Möglichkeit, die Regressionskoeffizienten miteinander vergleichbar zu machen besteht darin, sie zu standardisieren (= Beta-Werte)».

URBAN / MAYERL, a.a.O., S. 104: «Die vergleichende Interpretation von standardisierten Regressionskoeffizienten ist jedoch *umstritten*, und es gibt gute Gründe, dabei sehr vorsichtig vorzugehen, oder solche Vergleiche überhaupt gänzlich zu *unterlassen*».

- 33 Einzelheiten der Berechnung von Regressionskoeffizienten beispielsweise bei BACKHAUS et al., a.a.O. S. 114–115.
- 34 Auch eine Differenz $d = (\hat{y} - y)$ zwischen theoretisch berechnetem, modellkonformem Lohn und tatsächlich bezahltem Lohn einer einzelnen Arbeitnehmerin indiziert noch *keinen Verstoß* gegen Art. 3 GIG: Es besteht *de lege lata* keine rechtliche Verpflichtung zur Zahlung eines Lohnes, welcher einem bestimmten, auf Statistik aufgebauten Regressionsmodell entspricht: «*modellkonformer*» Lohn. Dies wird solange gelten, als ein (mit zahlreichen Willküreelementen behaftetes!) Regressionsmodell weder vom Gesetzgeber (als zwingendes Gesetzesrecht) erlassen noch von den direkt Beteiligten in einem einzelnen Betrieb *zweiseitig* anerkannt und zum Vertragsinhalt erhoben worden ist. Auch nach der Beweislastleichterung von Art. 6 GIG müssen individuelle Mitarbeiter oder mindestens individuelle Bereiche / Abteilungen / Arbeitsgruppen innerhalb des untersuchten Betriebs benannt werden, aus deren Löhnen sich allenfalls ein Verstoß gegen Art. 3 GIG ergeben könnte. Der Entwurf der deutschen Bundesregierung für ein «Lohnungleichheitsgesetz» vom 11. Januar 2017 verwendet für diese lokale Fokussierung der Untersuchung «Vergleichstätigkeiten», welche bei einem Gesuch einer Arbeitnehmerin um Auskunftserteilung anzugeben sind: «§ 10 Individueller Auskunftsanspruch (1) [...] Dazu haben die Beschäftigten in zumutbarer Weise eine gleiche oder gleichwertige Tätigkeit (Vergleichstätigkeit) zu benennen».
- ZGP Zürcher Gesellschaft für Personalmanagement, in DENARIS 01/2016, S. 18: «Des Weiteren ist zu berücksichtigen, dass eine Logib-Analyse lediglich *Hinweise* liefern kann, ob und wenn ja, *in welchem Bereich* eines Unternehmens Lohnungleichheiten bestehen. [...]»
- ZGP Zürcher Gesellschaft für Personalmanagement, Vernehmlassung zur Änderung GIG, 23. Februar 2016, S. 4: «(...) Wenn ein Unternehmen statistische Lohnungleichheitsanalysen durchführt, erhält es, völlig unabhängig von der eingesetzten Methode, bestenfalls Hinweise, *ob irgendwo* eine Lohnungleichheit besteht. Es kann sie aber wegen der statistisch bedingten Mindestgrösse der Mitarbeitergruppe *nicht genau lokalisieren* und es erhält auch *keine Hinweise auf deren Gründe*».
- 35 Art. 8 BV erfasst Herkunft, Rasse, Geschlecht, Alter, Sprache, soziale Stellung, Lebensform, religiöse, weltanschauliche oder politische Überzeugung, körperliche, geistige, psychische Behinderung. Bei der Lohnfindung dürften vor allem die Dichotomien Schweizer – Ausländer; Europäer – andere, Familienväter – andere, Techniker – Administrative, Muttersprache Deutsch – andere, Vollzeit – Teilzeit eine Rolle spielen; dazu auch die vielen von FLURINA SCHMID auf der nationalen Ebene verwendeten Variablen, a.a.O. S. 454, 456 und 460, sowie BLINDER, a.a.O., (Fn. 12), S. 441. Beispiele von Regressionskoeffizienten unter den vielen dichotomen Dummyvariablen bei RONALD OAXACA, a.a.O., (Fn. 12), S. 700–704: Die höchsten Regressionskoeffizienten β_5 erzielten die unterschiedlichen Industriezweige Bergbau ($\beta_5 = .3604$) bis Landwirtschaft ($\beta_5 = .1285$) und «personal services» ($\beta_5 = -.0618$), alle Zahlen für weisse Männer, alles bei einem $R^2 = .43$. URBAN / MAYERL a.a.O., S. 284, empfehlen, immer eine *Referenzvariable* zu verwenden, um Regressionskoeffizienten verschiedener Dummyvariablen einzustufen.
- 36 DOLDER, a.a.O. Rz. 57.
- 37 Selbst FLÜCKIGER / GRAF anerkennen die Wirkung der Vertragsfreiheit auf die Lohnfindung (a.a.O., S. 83): «Bei Verfahren, welche die Lohnpolitik des Unternehmens untersuchen, wird oft die Vertragsfreiheit des Arbeitgebers ins Spiel gebracht. So soll der Arbeitnehmer [?] im liberalen Wirtschafts- und Rechtssystem der Schweiz bei der Festlegung der Löhne eine gewisse Freiheit haben. *Es ist schwierig, dieses Argument zu widerlegen*. [...] dass die Unternehmen in einer liberalen Wirtschaftsordnung frei sind, die Löhne über dem Branchendurchschnitt anzusetzen, um beispielsweise die produktivsten Arbeitskräfte anzuziehen. [...]».
- 38 In diesem Sinne wohl BGE 133 III 545, S. 552, E. 4.3 «Le principe constitutionnel de l'égalité salariale entre l'homme et la femme (cf. art. 8 al. 3, dernière phrase, Cst.) est fondé sur la notion de *travail de valeur égale* (...). Autrement dit, *après d'un même employeur*, la travailleuse a droit à un salaire égal à celui que touche le travailleur s'ils accomplissent tous deux, dans des conditions égales, des tâches semblables ou des travaux, certes de nature différente, mais *ayant une valeur identique*».
- FREIVOGEL N. 90 zu Art. 3: «Die Arbeitgeberschaft ist also grundsätzlich nicht verpflichtet, gleiche Löhne zu bezahlen wie andere Arbeitgebende. [...] Entsprechend kann der Bewertungsvergleich (...) in der Regel nur zwischen Tätigkeiten verlangt werden, die in verschiedenen Anstellungsverhältnissen, aber *bei ein und derselben* Arbeitgeberschaft erbracht werden.» (mit Hinweisen auf BGE 131 III 393, 408, 130 III 145). N. 91 «Kantons-, gemeinde- oder *firmenübergreifende* Vergleiche (...) können nicht verlangt werden, wenn Arbeitgeber bezüglich Lohnsystem, Festsetzung der Löhne und Bezahlung der Löhne nicht nur formellrechtlich, sondern *tatsächlich* voneinander völlig unabhängige Entitäten und alleinverantwortlich sind». Entwurf der deutschen Bundesregierung für ein «Lohnungleichheitsgesetz» vom 11. Januar 2017: § 12 Reichweite [...] (2) Die Auskunftspflicht nach § 10 umfasst 1. nur Entgeltregelungen, die in *demselben Betrieb* und bei *demselben Arbeitgeber* Anwendung finden, [...]
- 39 Alternative Instrumente für die Einführung eines strukturierten multi-criteria Lohn-Analysesystems bieten neben dem

Individualvertrag (z.B. Anhang Betriebsreglement) oder dem GAV (ebenfalls Anhang) allenfalls eine Betriebsordnung nach Art. 38 ArG, ein zweiseitiger Änderungsvertrag *ad hoc* zwischen den Beteiligten, dagegen kaum das einseitige Weisungsrecht (Organisationsrecht) des Arbeitgebers nach Art. 321d OR. Letzteres würde vermutlich zumindest eine Informationspflicht des Arbeitgebers nach Art. 9 / 10 des MWG auslösen. **BGE 130 III 145 S. 159** «La détermination de la valeur comparative des activités différentes doit se faire *auprès d'un seul et même employeur*, car il ne saurait être question, sauf circonstances spéciales, de procéder à des comparaisons *entre plusieurs entreprises, communes ou cantons*, qui ont des systèmes de rémunération totalement indépendants les uns des autres (ELISABETH FREIVOGEL, Commentaire de la loi sur l'égalité, n. 104 et 105 ad art. 3 LEg; cf. également KATHRIN KLETT, op. cit., ch. 2 p. 1296).» Anmerkung: Das Zitat bezieht sich auf den Aufsatz KATHRIN KLETT, Richterliche Prüfungspflicht und Beweiserleichterung, AJP 2001, 1293 ff.

- 40 FREIVOGEL, N. 91 zu Art. 3: «Dann kann oder muss der Vergleich ausgeweitet werden, und dann können auch andere juristische Personen (...) als die formellrechtliche Arbeitgeberschaft belangt werden».
- 41 Es liegt z.B. nahe, das Tieflohnsegment mit überwiegender Frauenarbeit auszulagern und rechtlich wie tatsächlich selbständig zu organisieren, um dem Diskriminierungsverbot zu entgehen. Rundschau SRF vom 16. November 2016: Der Bund geht auch hier mit dem guten Beispiel voran und lässt die Schuhe der Schweizerarmee von der RUAG in der Region Cluj in Rumänien *ausschliesslich von Frauen* zu Monatslöhnen von etwa 300 Euro produzieren. Bestätigung: CHRISTA LUGINBÜHL, Packt die Politik ihre Chance? Public Eye Magazin Nr. 3 (Januar 2017), S. 22–23: Stundenlohn unter zwei Franken.
- 42 **BGE 130 III 145 S. 154**, «L'expert Q. a procédé à une seconde analyse pour déterminer les salaires des cadres de la défenderesse. Se basant cette fois-ci sur les *données internes* afférentes aux années 1994 à 1997 que X. lui avait communiquées, il a estimé l'équation des salaires annuels bruts d'abord pour l'ensemble du personnel de la recourante, puis *pour son seul personnel administratif*».
- 43 Dieses Verbot der Verwendung ausserbetrieblicher Parameter schimmert auch in **BGE 130 III 145 S. 159/160** durch: «3.2 En l'occurrence, le Professeur Q., dans sa seconde analyse, a estimé l'équation salariale régissant les salaires chez la recourante à partir des *données internes* que celle-ci lui avait fournies. Du moment que les magistrats de la Cour civile n'avaient pas de compétences spéciales en sciences du travail et en économétrie, ils n'avaient évidemment pas à discuter les variables et les coefficients que l'expert a intégrés dans sa formule. On ne voit donc pas en quoi, à considérer les principes susrappelés, la méthode d'évaluation en cause, effectuée par un professeur spécialiste en la matière et *fondée sur des éléments intrinsèques à la défenderesse elle-même*, ne serait pas conforme au droit fédéral».
- 44 ZGP, Vernehmlassung Änderung GIG, 23. Februar 2016, S. 3: «Logib basiert auf der Datenstruktur der oben diskutierten Lohnstrukturerhebung des Bundes, die Berechnungen erfolgen aber *lediglich mit den Lohndaten des zu analysierenden Unternehmens*. Das effektive Lohnsystem des zu analysierenden Unternehmens wird bei der Auswertung allerdings nicht berücksichtigt, sondern *nur die in Logib vorgesehenen Parameter*».
Dies bedeutet u.a., dass auch wenn ein Betrieb für seine Zwecke bestimmte Variablen für besonders relevant ansieht, dies nicht berücksichtigt wird. Anders anscheinend FLÜCKIGER/GRAF, a.a.O., S. 85, zum «Referenzmarkt» als Erkenntnisquelle: «... kann die gleiche ökonomische Schätzung mit den Daten des *Referenzmarktes* durchgeführt werden, ...».
- 45 **BGE 130 III 145 S. 153 B.c.** ... «Une seconde expertise a été confiée à Q., Professeur à l'Université de Genève, qui a déposé son rapport le 17 juin 1999 et un rapport complémentaire le 11 février 2000. L'expert s'est proposé, dans un premier temps, pour mettre en évidence la politique salariale appliquée par la défenderesse, de comparer le comportement de X. à l'égard de son personnel féminin avec la pratique prévalant *dans le marché genevois et vaudois de la finance*».
- 46 **BGE 125 II 385**, E. 5b, 391; Kommentar dazu bei KATHRIN ARIOLI / BIBIANE EGG, Arbeitswissenschaftliche Gutachten im Lohngleichheitsprozessen, AJP 2001, 1299–1307, zur zentralen Frage S. 1301/02, Kritik an der «Vermischung» der drei Kategorien S. 1303/1304.
- 47 **BGE 130 III 160 E. 3.1.2** «La question de la *fiabilité de l'expertise* Q. concerne *l'appréciation des preuves*. Le grief à ce propos est donc irrecevable dans la présente instance (**ATF 127 III 543** consid. 2c)».
- 48 Beispiele: In **BGE 124 II 436** wurden als zulässige und diskriminierungsfreie Kriterien – wenn richtig angewendet – aufgezählt: Ausbildung, Alter, Dienstalter, Qualifikation, Erfahrung, Aufgabenbereich, Leistung; in **BGE 131 II 393**, 399 (Erw. 5.2 und 6, wurde die Methode «Vereinfachte Funktionsanalyse (VFA)» mit ihren 6 Kriterien (= unabhängige Variablen x_i) im Einzelnen untersucht und als nicht diskriminierend qualifiziert (Krankenpflegeberufe und Polizei im Kanton Solothurn). Vgl. auch § 8 der Personalverordnung des Kantons Zürich, **177.11** (vom 16. Dezember 1998), welche die VFA ausdrücklich anerkennt: Oben Fn. 4.
- 49 Die Verteilung der Zuständigkeiten bei dieser mehrstufigen «Anerkennung» ist indessen nach dem Text des Revisionsentwurfs noch nicht eindeutig geregelt: *Art. 13b* Methode der Lohnanalyse
1 Die Lohnanalyse ist nach einer *anerkannten* Methode durchzuführen.
2 Die zuständige Behörde veröffentlicht eine Liste der *anerkannten* Methoden.
3 Der Bundesrat legt die Kriterien fest, die bei der *Anerkennung* einer Methode zu berücksichtigen sind. Er hört vorgängig die Arbeitgeber- und Arbeitnehmerverbände an.
- 50 Dass die aktuelle Situation unbefriedigend ist und auf eine Willkürprüfung hinausläuft, betonen auch GEISER / MAUCHLE, a.a.O., S. 34, allerdings aus anderem Blickwinkel: «Allerdings scheint aus Arbeitnehmersicht unbefriedigend, dass das Bundesgericht im Hinblick auf die angewendete Bewertungsmethode den kantonalen Gerichten viel Ermessen belässt, indem es bisher regelmässig sowohl die vereinfachte Funktionsanalyse als auch die statistische Lohnanalyse (Regressionsanalyse) für *zulässig* erachtete. Die Überprüfung der angewendeten Methode durch das BGer läuft damit im

Ergebnis auf eine *Willkürprüfung* hinaus. Es wäre in diesem Zusammenhang zu fordern, dass das BGer konsequent jenen wissenschaftlichen Instrumenten zur Analyse von Lohndiskriminierungen den Vorzug geben würde, welche sich geschlechterbenachteiligenden Wertungen enthalten». https://www.ebg.admin.ch/dam/ebg/de/dokumente/2016/06/rechtsgutachten_marktlohnundkonjunkturrellelagealsr

- 51 Vgl. zur Unentbehrlichkeit von multi-criteria Methoden in der Rechtsanwendung für andere Probleme (Rückfallrisiko und Bewertung von Submissionsofferten in der öffentlichen Beschaffung) bei DOLDER (2015), a.a.O., Abschnitt 2, Rz. 8 ff. Oder etwas hemdsärmelig ausgedrückt: Die in einem Rechtsstreit unterliegende Partei wird vermutlich wesentlich «*lieber*» aufgrund eines wissenschaftlich anerkannten, systematisch strukturierten Algorithmus verlieren, als aufgrund einer (wie auch immer gearteten) unstrukturierten, persönlichen «Intuition» eines Richters.