
VERDIENSTUNTERSCHIEDE ZWISCHEN MÄNNERN UND FRAUEN

Eine Ursachenanalyse auf Grundlage der
Verdienststrukturerhebung 2014

Claudia Finke, Florian Dumpert, Martin Beck

↳ **Schlüsselwörter:** geschlechtsspezifische Verdienstunterschiede – Verdienststrukturerhebung – Oaxaca-Blinder-Dekomposition – Support Vector Machine – Random Forest

ZUSAMMENFASSUNG

Anknüpfend an Analysen auf Basis der Verdienststrukturerhebungen 2006 und 2010 wird im vorliegenden Beitrag mit aktuellen Daten für 2014 die Höhe des bereinigten Gender Pay Gap berechnet. Ergänzt werden die Untersuchungen um experimentelle Berechnungen zur verbesserten Schätzung der potenziellen Berufserfahrung mithilfe von Machine-Learning-Verfahren. Die Analysen zeigen, dass der bereinigte Gender Pay Gap im Zeitverlauf leicht rückläufig ist und die Ausweitung der Verdienststrukturerhebung 2014 um bislang ausgeschlossene Betriebe keine nennenswerten Effekte auf die zentralen Ergebnisse hat. Die approximative Berücksichtigung von Erwerbsunterbrechungen aufgrund von Mutterschaft wirkt sich nur wenig auf die Höhe des bereinigten Gender Pay Gap aus.

↳ **Keywords:** *gender pay gap – structure of earnings survey – Oaxaca-Blinder decomposition – Support Vector Machine – Random Forest*

ABSTRACT

Building on previous analyses based on the structure of earnings surveys of 2006 and 2010, this article describes the calculation of the adjusted gender pay gap using current data for 2014. The research is supplemented by experimental calculations to improve the estimation of potential professional experience by means of machine learning methods. The analyses indicate that the adjusted gender pay gap has slightly declined over time. Furthermore, the results show that the integration of previously excluded companies in the structure of earnings survey 2014 has no significant effect on the essential findings. The approximate consideration of career interruptions related to motherhood has only a minor effect on the adjusted gender pay gap.

Claudia Finke

hat Sozialwissenschaften mit dem Schwerpunkt angewandte Sozialforschung an der Ruhr-Universität Bochum studiert und ist im Statistischen Bundesamt als Referentin im Bereich „Tarifstatistiken, Vierteljährliche Verdiensterhebung, Arbeitskostenindex, Überschuldung“ tätig. Ihre Schwerpunkte liegen in der konzeptionellen und methodischen Weiterentwicklung der Vierteljährlichen Verdiensterhebung sowie in der Analyse des Gender Pay Gap.

Florian Dumpert

ist wissenschaftlicher Mitarbeiter am Lehrstuhl für Stochastik der Universität Bayreuth. Der Diplom-Mathematiker forscht im Bereich maschineller statistischer Lernverfahren, insbesondere Support Vector Machines.

Martin Beck

ist Diplom-Ökonom und leitet seit 2007 im Statistischen Bundesamt die Gruppe „Unternehmensregister, Klassifikationen, Verdienste, übergreifende Unternehmensstatistiken“. Er befasst sich derzeit unter anderem mit der effizienteren Gestaltung der Datengewinnung und -analyse durch die Einführung neuer statistischer Methoden.

1

Einleitung

Lohnunterschiede zwischen Männern und Frauen sind auch gegenwärtig ein relevantes politisches Thema. So wird die Umsetzung aktueller politischer Maßnahmen von der Bundesregierung unter anderem mit Differenzen im Verdienst von Arbeitnehmern und Arbeitnehmerinnen begründet (Bundesministerium für Familie, Senioren, Frauen und Jugend, 2017). Hierzu zählen etwa der Ausbau der Kinderbetreuung und die Einführung des ElterngeldPlus als Anreize für weniger und kürzere familienbedingte Erwerbsunterbrechungen sowie die Einführung des gesetzlichen Mindestlohns, von der Frauen in besonderem Maße profitieren. Auch das am 11. Januar 2017 im Bundeskabinett beschlossene Gesetz zur Förderung der Transparenz von Entgeltstrukturen soll dazu beitragen, „unmittelbare und mittelbare Entgeltdiskriminierung wegen des Geschlechts zu beseitigen“.¹ Zudem plant die Bundesregierung eine Reform der Pflegeberufe mit dem Ziel einer Aufwertung des Berufsfelds, das insbesondere von Frauen ausgeübt wird. Vor dem Hintergrund der bestehenden Relevanz des Gender Pay Gap ist es umso wichtiger, Anhaltspunkte zur aktuellen Höhe und Entwicklung der Lohnlücke bereitzustellen.

Im Zentrum der vorliegenden Auswertungen steht die Ermittlung des bereinigten Gender Pay Gap sowie eine detaillierte Ursachenanalyse auf Basis der Verdienststrukturerhebung 2014. Die Analysen knüpfen an Auswertungen zum bereinigten Gender Pay Gap auf Grundlage der Verdienststrukturerhebungen 2006 und 2010 an.² Ergänzt werden die beschriebenen Untersuchun-

1 Bundesrats-Drucksache 8/17 vom 12. Januar 2017: Entwurf eines Gesetzes zur Förderung der Transparenz von Entgeltstrukturen.

2 Eine ausführliche Beschreibung der Methodik und der Ergebnisse der Berechnung des bereinigten Gender Pay Gap auf Grundlage der Verdienststrukturerhebung 2006 findet sich in Finke (2010) und Finke (2011). Ergebnisse zum bereinigten Gender Pay Gap auf Basis der Verdienststrukturerhebung 2010 enthält Joachimiak (2013). Bereits die Vorgängererhebung der Verdienststrukturerhebung wurde für Analysen der Verdienstunterschiede von Männern und Frauen herangezogen. So führte beispielsweise Kulmiz (2001) Untersuchungen auf Basis der Gehalts- und Lohnstrukturerhebung (GLS) 1990 durch. Boll und andere (2016a) haben zudem auf Grundlage des Structure of Earnings Survey 2010, das sämtliche nationale Verdienststrukturerhebungen der EU-Mitgliedstaaten umfasst, Auswertungen zum Gender Pay Gap in der Europäischen Union (EU) und in den einzelnen Mitgliedstaaten durchgeführt.

gen um experimentelle Berechnungen zur verbesserten Schätzung der potenziellen Berufserfahrung mithilfe von Support-Vector-Machine- und Random-Forest-Modellen.

2

Rückblick: Bisheriges Vorgehen und Ergebnisse

Zur Identifikation geschlechtsspezifischer Verdienstunterschiede werden in der Regel zwei Indikatoren herangezogen: der unbereinigte und der bereinigte Gender Pay Gap. Der unbereinigte Gender Pay Gap vergleicht den Durchschnittsverdienst aller Arbeitnehmer beziehungsweise Arbeitnehmerinnen in allgemeiner Form miteinander. Mithilfe des unbereinigten Gender Pay Gap wird auf diese Weise auch der Teil des Verdienstunterschieds erfasst, der durch schlechtere Zugangschancen von Frauen hinsichtlich bestimmter Berufe oder Karriere-stufen verursacht wird, die möglicherweise ebenfalls das Ergebnis benachteiligender Strukturen sind. Der bereinigte Gender Pay Gap hingegen misst den Verdienstabstand von Männern und Frauen mit vergleichbaren Qualifikationen, Tätigkeiten und Erwerbsbiografien.

2.1 Methodik bisheriger Analysen

Berechnungen zum unbereinigten Gender Pay Gap werden vom Statistischen Bundesamt seit 2006 auf Basis der Verdienststrukturerhebung durchgeführt. Fortgeschrieben werden die Ergebnisse in den Jahren, in denen keine Verdienststrukturerhebung stattfindet³, mit Ergebnissen der Vierteljährlichen Verdiensterhebung. Methodisch stützt sich das Statistische Bundesamt bei der Berechnung des unbereinigten Gender Pay Gap aufgrund der Bedeutung von internationalen Vergleichen auf europaweit einheitliche Vorgaben von Eurostat, dem Statistischen Amt der Europäischen Union. Hiernach wird der unbereinigte Gender Pay Gap als Differenz zwischen den durchschnittlichen Bruttostundenverdiensten männlicher und weiblicher Beschäftigter in Prozent der durchschnittlichen Bruttostundenverdienste männlicher Beschäftigter definiert (Eurostat, 2008):

3 Während die Vorgängererhebungen der Verdienststrukturerhebung in unregelmäßigen Abständen stattfanden, wird die seit 2006 bestehende Verdienststrukturerhebung alle vier Jahre durchgeführt.

$$(1) \text{ GPG}_{\text{unadj}} = \frac{\bar{y}_m - \bar{y}_f}{\bar{y}_m} \cdot 100.$$

Dabei sind:

\bar{y}_m = durchschnittlicher Bruttostundenverdienst ohne Sonderzahlungen von Männern;

\bar{y}_f = durchschnittlicher Bruttostundenverdienst ohne Sonderzahlungen von Frauen.

Zur Ermittlung des bereinigten Gender Pay Gap nutzt das Statistische Bundesamt die Oaxaca-Blinder-Dekomposition (Oaxaca, 1973; Blinder, 1973). Die Komponentenzerlegung gilt als die am häufigsten verwendete Methode zur Ermittlung von Verdienstunterschieden (Hübler, 2003, hier: Seite 557; Neumark, 2004, hier: Seite 8). Die Oaxaca-Blinder-Dekomposition kann nicht nur genutzt werden, um Anhaltspunkte zur Höhe des bereinigten Gender Pay Gap zu erhalten, sondern ermöglicht auch eine Ursachenanalyse des Verdienstunterschieds. Grundlage des Dekompositionsverfahrens bilden zwei nach dem Geschlecht differenzierte, semilogarithmische Regressionsfunktionen, mit denen der Einfluss verschiedener Strukturmerkmale auf den logarithmierten Bruttostundenverdienst bestimmt wird. Durch Subtraktion der Lohnfunktionen für Männer und Frauen sowie durch Umformungen ergibt sich folgende Gleichung:

$$(2) \quad \overline{\ln Y_M} - \overline{\ln Y_F} = \underbrace{(\beta_0^M - \beta_0^F) + \sum_j \bar{x}_j^F (\beta_j^M - \beta_j^F)}_{\text{unexplained gap}} + \underbrace{\sum_j \beta_j^M (\bar{x}_j^M - \bar{x}_j^F)}_{\text{explained gap}}.$$

Dabei sind:

$\ln Y$ = logarithmierter Bruttostundenverdienst;

β_j = Regressionskoeffizient eines Merkmals j ;

β_0 = Regressionskonstante;

\bar{x}_j = arithmetisches Mittel eines Merkmals j ;

M/F = Männer/Frauen.

Mithilfe von Formel (2) wird die Lohndifferenz in zwei Bestandteile zerlegt: Mit dem erklärten Gender Pay Gap ist der Teil der Verdienstdifferenz gemeint, der auf

geschlechtsspezifische Unterschiede in den erklärenden Variablen zurückzuführen ist. Dieser Effekt wird daher auch Merkmals- oder Ausstattungseffekt genannt. Der unerklärte Teil des Gender Pay Gap, das heißt der Teil des Verdienstunterschieds, der sich nicht mit Ausstattungsunterschieden von Männern und Frauen erklären lässt, untergliedert sich wiederum in zwei Bereiche: Einerseits enthält der unerklärte Gender Pay Gap einen Teil, der auf Unterschieden hinsichtlich der Koeffizienten – also auf der unterschiedlichen Bewertung der erklärenden Variablen – basiert (zweiter Term des “unexplained gap”). Hiermit ist beispielsweise gemeint, dass ein Masterabschluss bei Männern und Frauen monetär unterschiedlich honoriert wird oder etwa beide Geschlechtergruppen unterschiedlich stark von der Ausübung einer Leitungsposition profitieren. Andererseits umfasst der unerklärte Gender Pay Gap auch einen Teil, der nicht auf die berücksichtigten erklärenden Variablen zurückgeführt werden kann (erster Term des “unexplained gap”). Mit anderen Worten bedeutet dies: Selbst wenn sowohl die für die Lohnfindung relevanten Ausstattungsmerkmale als auch deren Honorierung für Männer und Frauen identisch wären, könnte zwischen Arbeitnehmern und Arbeitnehmerinnen ein Verdienstunterschied bestehen, der nicht auf die Modellvariablen zurückgeführt werden kann. Dieser schlägt sich empirisch in der Differenz der Konstanten der Regressionsgleichungen nieder.

Der unerklärte Gender Pay Gap – auch bereinigter Gender Pay Gap genannt – wird vom Statistischen Bundesamt als Obergrenze für Lohndiskriminierung interpretiert, da nicht sämtliche lohnrelevanten Merkmale für Analysezwecke zur Verfügung stehen.⁴ So liegen beispielsweise in der Verdienststrukturerhebung keine Angaben zum individuellen Verhalten in Lohnverhandlungen und zu Erwerbsunterbrechungen vor. Hierauf wird in Kapitel 5 näher eingegangen.

2.2 Datengrundlage

Grundlage der seit 2006 vom Statistischen Bundesamt vorgenommenen Analysen zum unbereinigten und bereinigten Gender Pay Gap bildet die Verdienststrukturerhebung. Bei der europaweit harmonisierten Erhebung

⁴ Wichtige Grenzen der Gleichsetzung des bereinigten Gender Pay Gap mit Diskriminierung werden in Finke (2011) diskutiert.

handelt es sich um eine repräsentative Stichprobe, die in Deutschland dezentral von den Statistischen Ämtern der Länder durchgeführt wird.¹⁵ Die Erhebungseinheiten sind Betriebe sowie deren Beschäftigungsverhältnisse. Zur Auskunft verpflichtet sind die Arbeitgeber. Sie entnehmen die geforderten Daten der Lohnabrechnung und den Personalstammdaten.

Im Rahmen der Verdienststrukturerhebung werden neben Angaben zu den Bruttoverdiensten der Beschäftigten eine Vielzahl lohdeterminierender Merkmale erfragt. Dadurch ist die Erhebung besonders gut für eine Ursachenanalyse des Verdienstunterschieds von Männern und Frauen geeignet. So werden arbeitnehmerbezogene Angaben wie beispielsweise das Geschlecht und das Geburtsjahr, der Ausbildungsabschluss, die Leistungsgruppe¹⁶, der Beruf sowie das Eintrittsdatum in das Unternehmen und die Art der Beschäftigung (zum Beispiel befristet/unbefristet) erfragt. Schließlich liegen auch Angaben zum Betrieb vor, wie etwa der Standort, der Wirtschaftszweig, der Tarifvertrag und die Betriebsgröße. Da die erforderlichen Merkmale weitgehend im Rechnungswesen der Betriebe vorhanden sind, findet die Übermittlung der Daten in vielen Fällen direkt über das Abrechnungssystem der Betriebe statt. Dies führt zu einer hohen Datenqualität, da beispielsweise sehr genaue Verdienstangaben in die Erhebung eingehen. Einschränkungen, die mit der Datengrundlage verbunden sind, betreffen die mangelnde Übertragbarkeit auf bestimmte Beschäftigtengruppen. So werden in der Verdienststrukturerhebung generell keine Selbstständigen erfasst. Darüber hinaus wurden bis einschließlich 2010 Beschäftigte des Wirtschaftsabschnittes „A – Land- und Forstwirtschaft, Fischerei“ der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), nicht berücksichtigt.¹⁷ Zudem waren die Ergebnisse bis 2010 nicht auf Betriebe mit weniger als zehn sozialversicherungspflichtig Beschäftigten übertragbar. Für EU-weite Ver-

gleiche des Verdienstunterschieds von Männern und Frauen stellen die beschriebenen Restriktionen keine Einschränkung dar. Dies hängt damit zusammen, dass etwa die Nichtabdeckung bestimmter Branchen und Betriebsgrößen nicht nur Deutschland betrifft, sondern alle EU-Mitgliedstaaten. Denn die EU-weit einheitliche Erhebung “Structure of Earnings Survey” sieht diesbezüglich keine verpflichtende Datenlieferung an Eurostat vor.

2.3 Ergebnisse

Für das Jahr 2006 ermittelte das Statistische Bundesamt für Deutschland einen unbereinigten Gender Pay Gap von 23 %. Im früheren Bundesgebiet und Berlin (24 %) fiel der Wert wesentlich höher aus als in den neuen Ländern (6 %). Der bereinigte Gender Pay Gap lag 2006 in Deutschland insgesamt bei 8 %, das heißt, Frauen verdienten auch bei gleicher Qualifikation und Tätigkeit je Stunde durchschnittlich 8 % weniger als Männer, wobei dieser Wert eine Obergrenze der nicht erklärten Verdienstunterschiede darstellt. Aus dem Ergebnis für den bereinigten Gender Pay Gap folgt, dass sich in Deutschland im Jahr 2006 zwei Drittel des unbereinigten Gender Pay Gap auf unterschiedliche, für die Entlohnung relevante Eigenschaften von Männern und Frauen zurückführen ließen. Die wichtigsten Unterschiede waren den Berechnungen zufolge die zwischen weiblichen und männlichen Arbeitnehmern ungleiche Besetzung von Leistungsgruppen sowie die zwischen den Geschlechtern unterschiedlich ausfallende Berufs- beziehungsweise Branchenwahl. Darüber hinaus gingen Frauen eher einer Teilzeitbeschäftigung nach und waren tendenziell schlechter ausgebildet.

Auswertungen differenziert nach beiden Landesteilen ergaben folgendes, eher überraschendes Bild: Während der bereinigte Gender Pay Gap in den neuen Ländern (12 %) den des früheren Bundesgebiets und Berlin (8 %) überstieg, lag der unbereinigte Gender Pay Gap im Westen über dem für den Osten. Das Phänomen, dass der bereinigte Gender Pay Gap in den neuen Ländern höher als der unbereinigte ausfiel, lässt sich damit erklären, dass in den neuen Ländern beschäftigte Frauen lohnrelevante Merkmale aufwiesen, die rechnerisch einen höheren Durchschnittsverdienst als den der Männer rechtfertigen würden. Zum Beispiel arbeiten Frauen gegenüber Männern in Ostdeutschland vermehrt in

5 Ausführliche Informationen zur Verdienststrukturerhebung lassen sich dem Qualitätsbericht der Erhebung entnehmen (Statistisches Bundesamt, 2016a).

6 Eine detaillierte Beschreibung der Leistungsgruppen kann dem Glossar der Fachserie der Verdienststrukturerhebung 2014 entnommen werden (Statistisches Bundesamt, 2016b).

7 Der Wirtschaftszweig „0 – Öffentliche Verwaltung, Verteidigung; Sozialversicherung“ der WZ 2008 wurde bereits 2010 in die Verdienststrukturerhebung integriert (Günther, 2013, hier: Seite 127). Die Branche wird jedoch entsprechend der EU-Vorgaben nicht in die Berechnung des EU-weit vergleichbaren unbereinigten Gender Pay Gap einbezogen und ist laut EU-Verordnung auch kein Bestandteil der Verdienststrukturerhebung.

Tabelle 1

Unbereinigter und bereinigter Gender Pay Gap

	2006		2010	
	unbereinigter Gender Pay Gap	bereinigter Gender Pay Gap	unbereinigter Gender Pay Gap	bereinigter Gender Pay Gap
	%			
Deutschland	23	8	22	7
Früheres Bundesgebiet und Berlin	24	8	24	7
Neue Länder	6	12	7	9

großen Unternehmen und verfügen über ein höheres Dienstalder.

Die vier Jahre später auf Grundlage der Verdienststrukturhebung 2010 ermittelten Ergebnisse lassen darauf schließen, dass es sich beim Gender Pay Gap um einen vergleichsweise persistenten Indikator handelt. Sowohl der unbereinigte als auch der bereinigte Gender Pay Gap verringerten sich bundesweit im Vergleich zu 2006 um jeweils einen Prozentpunkt. Im Westen lässt sich beim unbereinigten Verdienstunterschied keine Veränderung im betrachteten Zeitraum feststellen, während sich der bereinigte Indikator analog zum Bundestrend um einen Prozentpunkt reduziert hat. Für den Osten kann eine leichte Erhöhung des unbereinigten geschlechtsspezifischen Verdienstunterschieds von einem Prozentpunkt konstatiert werden. Beim bereinigten Gender Pay Gap lässt sich dagegen eine Verringerung von 3 Prozentpunkten feststellen. [↪ Tabelle 1](#)

3

Weiterentwicklungspotenzial und aktuelles Vorgehen

In die Verdienststrukturhebung 2014 konnten wichtige Beschäftigtengruppen, die bislang nicht repräsentiert wurden, integriert werden. Hierzu zählen Beschäftigte im Wirtschaftsabschnitt „A – Land- und Forstwirtschaft, Fischerei“. Bereits seit 2010 liegen Angaben zu den Arbeitnehmerinnen und Arbeitnehmern im Bereich „O – Öffentliche Verwaltung, Verteidigung; Sozialversicherung“ vor. Auch Beschäftigte, die in Kleinstbetrieben tätig sind, werden ab 2014 abgedeckt. Auf diese Weise ist es möglich, die Berechnungen zum geschlechtsspezifischen Verdienstunterschied für alle abhängig Beschäftigten der Land- und

Forstwirtschaft, der Fischerei, des Produzierenden Gewerbes und des Dienstleistungsbereichs durchzuführen. Für EU-Vergleiche und Analysen im Zeitverlauf können die Daten jedoch alternativ so eingegrenzt werden, dass ausschließlich die Grundgesamtheit der für die Jahre 2006 und 2010 durchgeführten Auswertungen zum Gender Pay Gap betrachtet wird. In Kapitel 4 werden für beide Abgrenzungen der Verdienststrukturhebung – das heißt mit beziehungsweise ohne die Berücksichtigung von Beschäftigten in den Wirtschaftsabschnitten A und O sowie in Kleinstbetrieben – Ergebnisse zum bereinigten Verdienstunterschied von Männern und Frauen dargestellt. Vor dem Hintergrund, auch für 2014 mit den bisherigen Berechnungen vergleichbare Ergebnisse zu generieren, sind die für 2006 und 2010 berechneten Regressionsmodelle hinsichtlich der berücksichtigten Variablen soweit wie möglich übernommen worden. Infolge grundsätzlicher Änderungen bestimmter Merkmale in der Erhebung waren jedoch geringfügige Modifikationen bei der Spezifikation der Modelle notwendig. Aus [↪ Übersicht 1](#) lassen sich sämtliche in das Regressionsmodell und anschließend in die Oaxaca-Blinder-Zerlegung einbezogenen Variablen entnehmen.

Die in der Übersicht 1 als Berufserfahrung benannte Variable wird in [↪ Tabelle 2](#) näher erläutert. Hier werden die rechnerisch möglichen Jahre der entgangenen Berufserfahrung (Ausbildungsjahre) aufgrund von Bildungsphasen dargestellt.

In der [↪ Übersicht 2](#) werden die nach der International Standard Classification of Occupations bezeichneten Berufshauptgruppen, eine weitere Variable für die Regressionsanalyse, aufgeführt.

Eine zusätzliche Möglichkeit zur Weiterentwicklung der bisher durchgeführten Analysen besteht darin, weitere,

Übersicht 1

In die Regressionsanalysen einbezogene Variablen

Variablen	Definition/Ausprägung
Unabhängige Variablen	
Höchster beruflicher Ausbildungsabschluss	Dummy-Codierung, Referenzkategorie: Abschluss einer anerkannten Berufsausbildung Weitere Kategorien: ohne beruflichen Ausbildungsabschluss, Meister-/Techniker- oder gleichwertiger Fachschulabschluss, Bachelor, Diplom/Magister/Master/Staatsexamen und Promotion
Berufserfahrung ¹ Berufserfahrung quadriert ²	Berufserfahrung = Alter – Ausbildungsjahre – 6
Dienstalter	Dienstalter = Berichtsjahr – Eintrittsjahr
Leistungsgruppe	Dummy-Codierung, Referenzkategorie: angelernte Arbeitnehmer Weitere Kategorien: Arbeiter in leitender Stellung, herausgehobene Fachkräfte, Fachangestellte, ungelernete Arbeitnehmer, geringfügig Beschäftigte, Auszubildende
Berufshauptgruppe (nach ISCO-08) ³	Dummy-Codierung, Referenzkategorie: ISCO 4 (Bürokräfte und verwandte Berufe) Weitere Kategorien: ISCO 0, ISCO 1, ISCO 2, ISCO 3, ISCO 5, ISCO 6, ISCO 7, ISCO 8, ISCO 9 (siehe Übersicht 2)
Art des Arbeitsvertrags	Dummy-Codierung, Referenzkategorie: unbefristet Weitere Kategorie: befristet
Beschäftigungsumfang	Dummy-Codierung, Referenzkategorie: vollzeitbeschäftigt Weitere Kategorie: teilzeitbeschäftigt
Altersteilzeit	Dummy-Codierung, Referenzkategorie: keine Altersteilzeit Weitere Kategorie: Altersteilzeit
Tarifbindung des Betriebes ⁴	Dummy-Codierung, Referenzkategorie: keine Tarifbindung Weitere Kategorie: mit Tarifbindung
Zulagen für Schicht-, Wochenend-, Feiertags- und Nachtarbeit	Dummy-Codierung, Referenzkategorie: keine Zulagen Weitere Kategorie: Zulagen
Gebietsstand	Dummy-Codierung, Referenzkategorie: früheres Bundesgebiet und Berlin Weitere Kategorie: neue Länder
Differenzierter Regionstyp ⁵	Dummy-Codierung, Referenzkategorie: ländlicher Raum Weitere Kategorie: städtischer Raum
Unternehmensgröße ⁶	Dummy-Codierung, Referenzkategorie: 1 000 Arbeitnehmer und mehr Weitere Kategorien: 1 bis 9 Arbeitnehmer, 10 bis 49 Arbeitnehmer, 50 bis 249 Arbeitnehmer, 250 bis 499 Arbeitnehmer, 500 bis 999 Arbeitnehmer
Einfluss der öffentlichen Hand auf die Unternehmensführung	Dummy-Codierung, Referenzkategorie: kein oder eingeschränkter Einfluss der öffentlichen Hand auf die Unternehmensführung durch Kapitalbeteiligung (50 % und weniger), Satzung oder sonstige Bestimmungen Weitere Kategorie: beherrschender Einfluss der öffentlichen Hand auf die Unternehmensführung durch Kapitalbeteiligung (mehr als 50 %), Satzung oder sonstige Bestimmungen
Wirtschaftsgruppe ⁷	Dummy-Codierung, Referenzkategorie: Wirtschaftsgruppe Q 86.1 (Krankenhäuser) Weitere Kategorien: Wirtschaftsgruppen A 01.1 bis S 96.0, ohne Q 86.1
Abhängige Variable	
ln(Bruttolohnverdienst)	Logarithmierter Bruttolohnverdienst Bruttolohnverdienst = (Bruttomonatsverdienst – Sonderzahlungen)/bezahlte Stunden

1 Um die Berufserfahrung als erklärende Variable auch in Untersuchungen berücksichtigen zu können, in denen das Merkmal nicht direkt abgefragt wurde, erfolgt in der Regel eine näherungsweise Ermittlung dieser Größe über das Alter und die Ausbildungsdauer (Achatz und andere, 2005, hier: Seite 474). Den Arbeitnehmern wird im Rahmen dieses Vorgehens ein ununterbrochener Erwerbsverlauf unterstellt. Da jedoch vor allem bei Frauen Erwerbsunterbrechungen zu beobachten sind, führt dies insbesondere bei weiblichen Arbeitnehmern zu einer Überschätzung der Berufserfahrung (Hinz/Gartner, 2005, hier: Seite 26). Zur Umrechnung des höchsten (Aus-)Bildungsabschlusses in Jahren siehe Tabelle 2. Aufgrund von Änderungen in Bezug auf die Erfassung der Bildungsinformationen in der Verdienststrukturerhebung ergeben sich Unterschiede zwischen den für 2006 und 2010 durchgeführten Berechnungen und den vorliegenden Analysen.

2 An dieser Stelle wird auch der quadrierte Term in die Regressionsanalyse einbezogen, da zwischen der Berufserfahrung und dem Verdienst ein kurvilinearer Zusammenhang besteht.

3 ISCO (International Standard Classification of Occupations) ist eine von der Internationalen Arbeitsorganisation (ILO) vorgenommene Klassifikation der Berufe (International Labour Organization, 2007; siehe Übersicht 2). Die Kategorie ISCO 0 (Angehörige der regulären Streitkräfte, Soldaten) ist nur für die Regressionsanalysen relevant, bei denen sämtliche im Rahmen der Verdienststrukturerhebung 2014 erfassten Wirtschaftsabschnitte einbezogen wurden.

4 Im Rahmen der Berechnung des bereinigten Gender Pay Gap 2006 und 2010 wurden Angaben zur Tarifbindung des Arbeitnehmers herangezogen. Für die Analysen der Verdienststrukturerhebung 2014 erfolgte die Berücksichtigung der Tarifbindung des Betriebes. Die Betriebsangabe ist die Angabe, die üblicherweise für Analysen der Tarifbindung genutzt wird. Vergleichsrechnungen, in denen – ceteris paribus – die Tarifbindung des Arbeitnehmers beziehungsweise alternativ die des Betriebs herangezogen wurde, zeigen keine relevanten Auswirkungen auf die Höhe des bereinigten Gender Pay Gap. Auch der Betrag beider Merkmale auf den erklärten Anteil des Gender Pay Gap unterscheidet sich nur unwesentlich.

5 Während unter „Ballungsraum“ für die 2006 und 2010 durchgeführten Analysen Kreistypen subsummiert wurden, die entsprechend der Typologie des Bundesamtes für Bauwesen und Raumordnung als Kernstädte in Agglomerationen und (hoch-)verdichtete Kreise in Agglomerationen definiert wurden, fasste die Ausprägung „kein Ballungsraum“ alle übrigen Kreistypen zusammen. Da die Kreistypen in der beschriebenen Form nicht mehr existieren, wurde im Zusammenhang mit den vorliegenden Untersuchungen auf den in der Verdienststrukturerhebung 2014 vorliegenden differenzierten Regionstyp zurückgegriffen. Dieser kategorisiert die Regionen in städtische und ländliche Räume. Durch die Nutzung dieser Information soll die Vergleichbarkeit mit den Ergebnissen der bisherigen Untersuchungen gewährleistet werden. Für Informationen zur aktuellen Kategorisierung der siedlungsstrukturellen Kreistypen siehe Bundesamt für Bauwesen und Raumordnung (www.bbsr.bund.de).

6 Die Kategorie 1 bis 9 Arbeitnehmer ist ausschließlich für die Regressionsanalysen relevant, in die auch Angaben der Kleinbetriebe einfließen.

7 Die Wirtschaftszweige A 01.1 bis A 03.2 und O 84.1 bis O 84.3 der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), sind ausschließlich für die Regressionsanalysen relevant, bei denen sämtliche im Rahmen der Verdienststrukturerhebung 2014 erfassten Wirtschaftsabschnitte einbezogen wurden.

Verdienstunterschiede zwischen Männern und Frauen

Tabelle 2

Entgangene Berufserfahrung aufgrund von Bildungsphasen

Höchster beruflicher Ausbildungsabschluss	Höchster allgemeinbildender Schulabschluss				
	ohne Abschluss	Hauptschulabschluss	Mittlere Reife	Abitur	unbekannt
	Jahre				
Ohne Abschluss	9	9	10	13	10
Berufsausbildung	12	12	13	16	13
Meister/Techniker/ Fachschulabschluss	12	12	13	16	13
Bachelor	12,5	12,5	13,5	16,5	13,5
Diplom/Master und ähnliche Abschlüsse	15	15	16	19	16
Promotion	18	18	19	22	19
Unbekannt	11,5	11,5	12,5	15,5	12,5

die Lohnhöhe bestimmende Informationen in die Schätzmodelle einzubeziehen. Hierauf wird in Kapitel 5 mit Blick auf Erwerbsunterbrechungen und die tatsächliche Berufserfahrung näher eingegangen.

Übersicht 2

Berufshauptgruppen der ISCO-08

Code	Bezeichnung
ISCO 1	Angehörige der gesetzgebenden Körperschaft, leitende Verwaltungsbedienstete und Führungskräfte
ISCO 2	Akademische Berufe/Wissenschaftler
ISCO 3	Techniker und gleichrangige nichttechnische Berufe
ISCO 4	Bürokräfte und verwandte Berufe
ISCO 5	Dienstleistungsberufe und Verkäufer
ISCO 6	Fachkräfte in Land- und Forstwirtschaft und Fischerei
ISCO 7	Handwerks- und verwandte Berufe
ISCO 8	Bediener von Anlagen und Maschinen und Montageberufe
ISCO 9	Hilfsarbeitskräfte
ISCO 0	Angehörige der regulären Streitkräfte, Soldaten

ISCO: International Standard Classification of Occupations.

4

Ergebnisse für das Berichtsjahr 2014

Um Anhaltspunkte für die Entwicklung des bereinigten Gender Pay Gap zu erhalten, wurde zunächst eine Oaxaca-Blinder-Zerlegung durchgeführt, die auf Regressionsanalysen basiert, die mit dem für 2006 und 2010 berechneten Modell hinsichtlich der einbezogenen

nen Variablen und der zugrunde gelegten Abgrenzung der Beschäftigten vergleichbar sind.

Die getrennt für Männer und Frauen berechneten Regressionsmodelle zeigen, dass mithilfe der in den Analysen berücksichtigten Variablen rund 76% der Varianz des Bruttostundenverdienstes der Frauen erklärt werden. Bei den Männern liegt die Erklärungskraft bei etwa 79%. Die Güte

der Modelle ist somit verglichen mit ähnlichen Studien hoch.¹⁸ ↘ Tabelle 3 auf Seite 50 f.

In beiden Regressionsanalysen geht von fast allen Variablen ein signifikanter Einfluss aus. Ausnahmen bilden lediglich einige Branchen (Wirtschaftsgruppen).¹⁹ Die Koeffizienten der Dummy-Variablen „Handwerks- und verwandte Berufe“ (ISCO 7) sowie „Dienstleistungsberufe und Verkäufer“ (ISCO 5) sind zudem lediglich im Modell für die Frauen signifikant.

Anhand der Ergebnisse lässt sich zudem feststellen, dass – unter Konstanthaltung aller anderen in die Analysen einbezogenen Merkmale – beispielsweise Arbeitnehmerinnen beziehungsweise Arbeitnehmer in leitender Stellung, herausgehobene Fachkräfte und Fachangestellte über ein verglichen mit angelernten Beschäftigten höheres Gehalt verfügen. Demgegenüber erzielen Ungelernte einen gegenüber Angelernten geringeren Verdienst. Während etwa der Stundenlohn von ungelerten Arbeitnehmern um fast 7% unter dem der angelernten männlichen Beschäftigten liegt, ist der

- Geisberger und Glaser erhalten im Rahmen von Analysen zum geschlechtsspezifischen Lohnunterschied auf Basis der für Österreich durchgeführten Verdienststrukturerhebung ebenfalls hohe Werte von jeweils rund 68% (Geisberger/Glaser, 2014, hier: Seite 8). Strub und andere (2016) kommen in ihren auf Grundlage der Lohnstrukturerhebung 2012 durchgeführten Analysen, in denen sie die geschlechtsspezifischen Verdienstunterschiede von Männern und Frauen in der Schweiz analysieren, auf einen korrigierten Determinationskoeffizienten von 62% beziehungsweise 69%.
- Die Regressionskoeffizienten und die dazugehörigen beobachteten Signifikanzniveaus der Wirtschaftsgruppendifferenzen können auf Anfrage bei den Autoren bezogen werden. Aus Übersichtsgründen wurde auf eine Darstellung der entsprechenden Werte verzichtet.

Tabelle 3

Ergebnisse der Regressionsanalyse auf Basis der Verdienststrukturerhebung 2014

	Ohne die Wirtschaftsabschnitte A und O ¹ sowie ohne Kleinstbetriebe				Einschließlich der Wirtschaftsabschnitte A und O ¹ sowie der Kleinstbetriebe			
	Männer		Frauen		Männer		Frauen	
	Koeffizient	p	Koeffizient	p	Koeffizient	p	Koeffizient	p
Einflussfaktoren								
Neue Länder (Referenz: früheres Bundesgebiet und Berlin)	- 0,217	< 0,0001	- 0,176	< 0,0001	- 0,210	< 0,0001	- 0,178	< 0,0001
Städtischer Raum (Referenz: ländlicher Raum)	0,039	< 0,0001	0,046	< 0,0001	0,040	< 0,0001	0,048	< 0,0001
Berufshauptgruppen nach ISCO-08 (Referenz: ISCO 4)								
ISCO 0	-	-	-	-	0,096	< 0,0001	0,039	0,001
ISCO 1	0,384	< 0,0001	0,266	< 0,0001	0,328	< 0,0001	0,244	< 0,0001
ISCO 2	0,171	< 0,0001	0,115	< 0,0001	0,160	< 0,0001	0,115	< 0,0001
ISCO 3	0,111	< 0,0001	0,045	< 0,0001	0,082	< 0,0001	0,027	< 0,0001
ISCO 5	- 0,005	0,016	- 0,103	< 0,0001	- 0,024	< 0,0001	- 0,110	< 0,0001
ISCO 6	- 0,032	< 0,0001	- 0,100	< 0,0001	- 0,042	< 0,0001	- 0,125	< 0,0001
ISCO 7	0,002	0,306	- 0,113	< 0,0001	- 0,014	< 0,0001	- 0,118	< 0,0001
ISCO 8	- 0,032	< 0,0001	- 0,126	< 0,0001	- 0,048	< 0,0001	- 0,125	< 0,0001
ISCO 9	- 0,072	< 0,0001	- 0,155	< 0,0001	- 0,073	< 0,0001	- 0,142	< 0,0001
Dienstalter	0,004	< 0,0001	0,005	< 0,0001	0,004	< 0,0001	0,005	< 0,0001
Unternehmensgröße (Referenz: 1 000 Arbeitnehmer und mehr)								
1 bis 9 Arbeitnehmer	-	-	-	-	- 0,205	< 0,0001	- 0,157	< 0,0001
10 bis 49 Arbeitnehmer	- 0,148	< 0,0001	- 0,109	< 0,0001	- 0,155	< 0,0001	- 0,113	< 0,0001
50 bis 249 Arbeitnehmer	- 0,095	< 0,0001	- 0,072	< 0,0001	- 0,098	< 0,0001	- 0,073	< 0,0001
250 bis 499 Arbeitnehmer	- 0,046	< 0,0001	- 0,043	< 0,0001	- 0,047	< 0,0001	- 0,044	< 0,0001
500 bis 999 Arbeitnehmer	- 0,012	< 0,0001	- 0,014	< 0,0001	- 0,013	< 0,0001	- 0,014	< 0,0001
Höchster beruflicher Ausbildungsabschluss (Referenz: Abschluss einer anerkannten Berufsausbildung)								
ohne beruflichen Ausbildungsabschluss	- 0,044	< 0,0001	- 0,025	< 0,0001	- 0,042	< 0,0001	- 0,025	< 0,0001
Meister-/Techniker- oder gleichwertiger Fachschulabschluss	0,039	< 0,0001	0,034	< 0,0001	0,038	< 0,0001	0,033	< 0,0001
Bachelor	0,074	< 0,0001	0,059	< 0,0001	0,059	< 0,0001	0,042	< 0,0001
Diplom/Magister/Master/ Staatsexamen	0,160	< 0,0001	0,130	< 0,0001	0,137	< 0,0001	0,112	< 0,0001
Promotion	0,310	< 0,0001	0,334	< 0,0001	0,300	< 0,0001	0,303	< 0,0001
Leistungsgruppe (Referenz: angelernte Arbeitnehmer)								
Arbeitnehmer in leitender Stellung	0,535	< 0,0001	0,500	< 0,0001	0,537	< 0,0001	0,505	< 0,0001
herausgehobene Fachkräfte	0,337	< 0,0001	0,328	< 0,0001	0,333	< 0,0001	0,327	< 0,0001
Fachangestellte	0,130	< 0,0001	0,130	< 0,0001	0,128	< 0,0001	0,128	< 0,0001
ungelernte Arbeitnehmer	- 0,068	< 0,0001	- 0,048	< 0,0001	- 0,073	< 0,0001	- 0,057	< 0,0001
geringfügig Beschäftigte	- 0,181	< 0,0001	- 0,137	< 0,0001	- 0,136	< 0,0001	- 0,105	< 0,0001
Auszubildende	- 0,908	< 0,0001	- 0,816	< 0,0001	- 0,898	< 0,0001	- 0,821	< 0,0001
Tariffindung des Betriebes (Referenz: keine Tariffindung)	0,069	< 0,0001	0,078	< 0,0001	0,070	< 0,0001	0,081	< 0,0001
Potenzielle Berufserfahrung								
Berufserfahrung	0,014	< 0,0001	0,010	< 0,0001	0,012	< 0,0001	0,009	< 0,0001
Berufserfahrung (quadratiert)	- 0,0002	< 0,0001	- 0,0002	< 0,0001	- 0,0002	< 0,0001	- 0,0002	< 0,0001

Verdienstunterschiede zwischen Männern und Frauen

Tabelle 3 (Fortsetzung)

Ergebnisse der Regressionsanalyse auf Basis der Verdienststrukturerhebung 2014

	Ohne die Wirtschaftsabschnitte A und O ¹ sowie ohne Kleinbetriebe				Einschließlich der Wirtschaftsabschnitte A und O ¹ sowie der Kleinbetriebe			
	Männer		Frauen		Männer		Frauen	
	Koeffizient	p	Koeffizient	p	Koeffizient	p	Koeffizient	p
Befristeter Arbeitsvertrag (Referenz: unbefristeter Arbeitsvertrag)	- 0,079	< 0,0001	- 0,066	< 0,0001	- 0,078	< 0,0001	- 0,069	< 0,0001
Teilzeitbeschäftigt (Referenz: vollzeitbeschäftigt)	- 0,049	< 0,0001	- 0,009	< 0,0001	- 0,064	< 0,0001	- 0,010	< 0,0001
Beherrschender Einfluss der öffentlichen Hand (Referenz: kein oder eingeschränkter Einfluss der öffentlichen Hand)	- 0,020	< 0,0001	- 0,014	< 0,0001	- 0,015	< 0,0001	- 0,011	< 0,0001
Altersteilzeit (Referenz: keine Altersteilzeit)	0,347	< 0,0001	0,300	< 0,0001	0,361	< 0,0001	0,298	< 0,0001
Zulagen für Schicht-, Wochenend-, Feiertags- und Nachtarbeit (Referenz: keine Zulagen)	0,065	< 0,0001	0,074	< 0,0001	0,062	< 0,0001	0,079	< 0,0001
Wirtschaftsgruppen-Dummys	ja ²	ja ²	ja ²	ja ²	ja ²	ja ²	ja ²	ja ²
Konstante	2,472	< 0,0001	2,440	< 0,0001	2,511	< 0,0001	2,463	< 0,0001
R ²	0,791		0,756		0,784		0,745	
Korrigiertes R ²	0,791		0,755		0,784		0,745	
n (Stichprobenumfang)	419 794		368 124		530 808		503 353	
Fallzahl hochgerechnet	15 186 465		13 474 845		19 060 496		18 092 029	
Abhängige Variable: logarithmierter Brutto- stundenverdienst								

- 1 Klassifikation der Wirtschaftszweige, Ausgabe 2008. Wirtschaftsabschnitt A: Land- und Forstwirtschaft, Fischerei. Wirtschaftsabschnitt O: Öffentliche Verwaltung, Verteidigung; Sozialversicherung.
- 2 Die Regressionskoeffizienten und die dazugehörigen beobachteten Signifikanzniveaus der Wirtschaftsgruppendifferenzdummys können auf Anfrage bei den Autoren bezogen werden. Aus Übersichtsgründen wurde auf eine Darstellung der entsprechenden Werte verzichtet.

entsprechende Wert bei den Frauen knapp 5%. Die mit Abstand größten Lohnabschläge im Vergleich zur Referenzgruppe der Angelernten müssen weibliche und männliche geringfügig Beschäftigte sowie Auszubildende hinnehmen.

Beschäftigte mit einem Hochschulabschluss oder einer Promotion verdienen deutlich mehr als ihre Kollegen, die den Abschluss einer anerkannten Berufsausbildung aufweisen. Dies gilt für beide Geschlechtergruppen. Arbeitnehmerinnen und Arbeitnehmer ohne Abschluss einer anerkannten Berufsausbildung beziehen wie erwartet jeweils einen geringeren Bruttostundenlohn als Beschäftigte mit abgeschlossener Berufsausbildung.

Anhand der Ergebnisse lässt sich darüber hinaus feststellen, dass zwischen der Berufserfahrung und dem Verdienst ein umgekehrt U-förmiger Zusammenhang besteht. Dies trifft sowohl für männliche als auch für weibliche Beschäftigte zu.

Hinsichtlich des Gebietsstandes zeigt sich, dass im Osten tätige Arbeitnehmerinnen und Arbeitnehmer

einen geringeren Stundenverdienst erzielen als Beschäftigte im Westen, wobei der Effekt bei den Männern etwas stärker ausgeprägt ist als bei den Frauen.

Wenig überraschend ist auch, dass beispielsweise „Angehörige der gesetzgebenden Körperschaft sowie leitende Verwaltungsbedienstete und Führungskräfte“ (ISCO 1) verglichen mit Bürokräften und verwandten Berufen (ISCO 4) einen Lohnzuschlag erhalten. Demgegenüber beziehen „Hilfsarbeitskräfte“ (ISCO 9) einen im Vergleich zu den Bürokräften und verwandten Berufen geringeren Bruttostundenverdienst.

Der Effekt, der von der Dummy-Variablen ausgeht, die zur Abbildung des Einflusses der Altersteilzeit auf den Verdienst generiert wurde, fällt in beiden Modellen positiv aus. Beschäftigte in Altersteilzeit verfügen demnach verglichen mit den übrigen Arbeitnehmern über einen höheren Bruttostundenverdienst. Begründen lässt sich dies damit, dass Altersteilzeitbeschäftigte ihre Arbeitsstunden zwar um die Hälfte reduzieren, jedoch weiterhin zumindest 70% des ursprünglichen Nettogehalts beziehen.

Männliche Arbeitnehmer in Unternehmen mit 10 bis 49 Angestellten erzielen gegenüber Beschäftigten in Unternehmen mit 1000 und mehr Arbeitskräften einen um etwa 15 % geringeren Lohn. Frauen in kleinen Unternehmen müssen im Vergleich zu den in großen Firmen tätigen Kolleginnen Lohneinbußen in Höhe von lediglich 11 % hinnehmen. Generell gilt, dass das Lohnniveau mit der Unternehmensgrößenklasse ansteigt.

Hinsichtlich der Wirtschaftsgruppen zeigt sich beispielsweise, dass Beschäftigte im Bereich Fondsmanagement (Wirtschaftszweig K 66.3) einen höheren Lohn erhalten als Männer und Frauen, die in Krankenhäusern tätig sind. Demgegenüber weisen Arbeitskräfte, die in Call Centern (Wirtschaftszweig N 82.2) beschäftigt sind, einen gegenüber dem Krankenhauspersonal geringeren Verdienst auf. Die beschriebenen Resultate ergeben sich sowohl für Männer als auch für Frauen.

Die Ergebnisse der für beide Geschlechtergruppen separat geschätzten Lohnfunktionen bilden die Grundlage der im Folgenden mithilfe der Oaxaca-Blinder-Dekomposition vorgenommenen Zerlegung der Lohn Differenz. Das Dekompositionsverfahren bestimmt zum einen den Teil des Verdienstabstands, der auf strukturelle Unterschiede der Beschäftigten zurückzuführen ist, und zum anderen den Teil der Lohnlücke, der nicht mithilfe dieser Unterschiede erklärt werden kann.

↘ **Tabelle 4** zeigt, dass sich der unbereinigte Gender Pay Gap in Deutschland im Jahr 2014 auf 22,3 % beläuft. Insgesamt 16,5 Prozentpunkte des Gender Pay Gap lassen sich auf Unterschiede zwischen Männern und Frauen hinsichtlich der in die Analysen einbezogenen Merkmale zurückführen. Damit beträgt der erklärte Anteil am Verdienstunterschied 73,9%. Die restlichen 5,8 Prozentpunkte beziehungsweise 26,1 % des Lohnabstands

Tabelle 4

Zerlegung des Gender Pay Gap

Ergebnisse der Verdienststrukturerhebung 2014 ohne die Wirtschaftsabschnitte A und O¹ sowie ohne Kleinbetriebe

	Verdienstunterschied	Anteil
	%	
Unbereinigter Gender Pay Gap	22,3	100
Erklärter Teil	16,5	73,9
Unerklärter Teil (= bereinigter Gender Pay Gap)	5,8	26,1

¹ Klassifikation der Wirtschaftszweige, Ausgabe 2008. Wirtschaftsabschnitt A: Land- und Forstwirtschaft, Fischerei. Wirtschaftsabschnitt O: Öffentliche Verwaltung, Verteidigung; Sozialversicherung.

können nicht durch Differenzen in den berücksichtigten Merkmalen erklärt werden. Dies bedeutet, dass der Gender Pay Gap um rund 16 Prozentpunkte geringer wäre, wenn Männer und Frauen dieselben im Rahmen der Analyse betrachteten Charakteristika aufweisen würden. Dennoch bliebe auch dann ein Verdienstunterschied von etwa 6 % bestehen. Da nicht alle lohdeterminierenden Merkmale in die Analyse einbezogen werden konnten, ist dieser Wert als Obergrenze der nicht erklärten Verdienstunterschiede zu interpretieren.

↘ **Tabelle 5** gibt Anhaltspunkte zum Beitrag einzelner Faktoren auf den erklärten und unerklärten Teil des Gender Pay Gap. In Bezug auf den **erklärten** Teil des Gender Pay Gap zeigt sich, dass etwa 7 Prozentpunkte der Lohnlücke durch den höheren Anteil von Frauen in eher niedrig bezahlten Tätigkeiten zurückzuführen sind.¹⁰ Zudem lassen sich fast 5 Prozentpunkte des Lohnabstands mit Unterschieden in den Leistungsgruppen erklären.

Mit deutlichem Abstand wirken sich neben der Leistungsgruppe und der Tätigkeit auch der Beschäftigungsumfang und die geringfügige Beschäftigung auf den erklärten Teil des Gender Pay Gap aus. Dass Frauen häufiger teilzeitbeschäftigt sind, erklärt 2,1 Prozentpunkte des Unterschieds in den Stundenlöhnen und dass sie zudem eher einer geringfügigen Beschäftigung nachgehen weitere 1,3 Prozentpunkte.

Die mangelnde Relevanz des Faktors Berufserfahrung deutet darauf hin, dass Männer und Frauen hinsichtlich dieses Merkmals nur marginale Unterschiede aufweisen. Bei der Interpretation des Faktors Berufserfahrung sollte jedoch beachtet werden, dass im Rahmen der an dieser Stelle betrachteten Analysen lediglich die potenzielle Berufserfahrung der Beschäftigten ohne jegliche Berücksichtigung etwaiger Erwerbsunterbrechungen berücksichtigt wurde. Dieser Aspekt wird in Kapitel 5 noch einmal aufgegriffen.

¹⁰ Unter Tätigkeit wird in diesem Zusammenhang der gemeinsame Erklärungsbeitrag der Faktoren Beruf und Wirtschaftszweig verstanden. Im Rahmen der für 2006 und 2010 durchgeführten Analysen wurde die Zusammenfassung beider Merkmale vorgenommen, da beide Beiträge inhaltlich kaum voneinander zu trennen sind. So sind bestimmte Berufe in der Regel nur in gewissen Wirtschaftszweigen vorzufinden. Aus diesem Grund ist es nicht möglich, sowohl den Wirtschaftszweig als auch den Beruf auf einer tief gegliederten Ebene in die Analysen einzubeziehen. Um einen Vergleich mit den für 2006 und 2010 erzielten Ergebnissen zu gewährleisten, wird im Rahmen der vorliegenden Analysen analog zur Vorgehensweise der Vorgängeruntersuchungen verfahren.

Verdienstunterschiede zwischen Männern und Frauen

Tabelle 5

Detaillierte Zerlegung des unbereinigten Gender Pay Gap

Ergebnisse der Verdienststrukturerhebung 2014 ohne die Wirtschaftsabschnitte A und O¹ sowie ohne Kleinbetriebe

Erklärungsfaktoren	Erklärter Teil des Gender Pay Gap	Anteil des erklärten Teils am gesamten Gender Pay Gap	Unerklärter Teil des Gender Pay Gap: bereinigter Gender Pay Gap	Anteil des unerklärten Teils am Gender Pay Gap insgesamt
	Prozentpunkte	%	Prozentpunkte	%
Insgesamt	16,5	73,9	5,8	26,1
Konstante	X	X	5,4	24,1
Gebietsstand	0,2	1,0	1,7	7,4
Differenzierter Regionstyp	0,2	0,7	-0,1	-0,6
Tätigkeit	6,7	30,1	-3,2	-14,4
Dienstalter	0,2	1,1	-1,1	-4,8
Unternehmensgröße	-0,2	-1,1	-0,5	-2,1
Ausbildungsabschluss	0,6	2,8	-0,0	-0,0
Leistungsgruppen 1 bis 5	4,7	21,2	1,6	7,3
Geringfügige Beschäftigung	1,3	5,8	-0,5	-2,3
Auszubildende	-0,3	-1,4	-0,3	-1,3
Tarifbindung	0,1	0,3	0,0	0,2
Potenzielle Berufserfahrung	0,0	0,1	4,7	21,2
Art des Arbeitsvertrags	0,3	1,2	0,4	2,0
Beschäftigungsumfang	2,1	9,4	-0,3	-1,3
Einfluss der öffentlichen Hand auf die Unternehmensführung	0,2	0,7	0,2	0,9
Altersteilzeit	0,1	0,4	-2,5	-11,4
Zulagen	0,3	1,5	0,3	1,4

¹ Klassifikation der Wirtschaftszweige, Ausgabe 2008. Wirtschaftsabschnitt A: Land- und Forstwirtschaft, Fischerei. Wirtschaftsabschnitt O: Öffentliche Verwaltung, Verteidigung; Sozialversicherung.

In Bezug auf den **unerklärten** Teil des Gender Pay Gap bildet die Konstante den bedeutendsten Faktor. Der positive Wert der Konstante verdeutlicht, dass Frauen unabhängig von den berücksichtigten Einflussfaktoren im Durchschnitt einen um 5,4 Prozentpunkte geringeren Verdienst erhalten würden. Der positive Wert ist jedoch nicht allein mit Diskriminierung weiblicher Arbeitnehmer gleichzusetzen. So enthält der unerklärte Gender Pay Gap auch Effekte, die auf nicht im Modell abgebildete, lohnrelevante Unterschiede zwischen Männern und Frauen, wie beispielsweise das individuelle Verhalten in Lohnverhandlungen, zurückzuführen sind. Diese Effekte werden von der Konstanten aufgefangen (Achatz und andere, 2005, hier: Seite 478). Der Rest des unerklärten Teils des Gender Pay Gap kommt dadurch zustande, dass bei gleicher Ausstattung von Männern und Frauen die Bewertung, das heißt die Entlohnung auf dem Arbeitsmarkt, für einige Merkmale diskriminierend gegenüber Frauen und die Bewertung mancher Merkmale wiederum benachteiligend gegen-

über Männern wirkt. Die positive Summe dieser Effekte bedeutet, dass die berücksichtigten Merkmale insgesamt zuungunsten von Frauen wirken.

Ein Merkmal, das deutlich zulasten von Frauen wirkt, ist die Berufserfahrung. Ein zusätzliches Jahr an Berufserfahrung wird bei Frauen nach den Ergebnissen der Analyse in geringerem Maße honoriert als bei Männern. Allerdings gilt auch hier die bereits erwähnte Einschränkung, dass lediglich die potenzielle und nicht die tatsächliche Berufserfahrung in die Analysen einbezogen wurde. Auch der Gebietsstand und die Leistungsgruppe wirken zulasten von Frauen. Vor allem im Zusammenhang mit den einzelnen Leistungsgruppen

sollte jedoch deren Heterogenität beachtet werden, da diese unter Umständen zu Verzerrungen bei der Ermittlung des Gender Pay Gap bei gleicher Leistungsgruppe führt. Dies hängt damit zusammen, dass auch innerhalb der jeweiligen Leistungsgruppe eine gewisse vertikale Segregation besteht. Diese Unterschiede werden im Rahmen der vorliegenden Analyse nicht erfasst und schlagen sich dementsprechend im unerklärten Teil des Gender Pay Gap nieder. Zugunsten von Frauen, das heißt diskriminierend für Männer, wirken demgegenüber insbesondere die Faktoren Tätigkeit, Altersteilzeit und Dienstalter (siehe Tabelle 5, Spalte 3).

In einem weiteren Schritt erfolgte die Berechnung der Oaxaca-Blinder-Dekomposition unter Rückgriff auf sämtliche in der Verdienststrukturerhebung 2014 abgedeckten Beschäftigtengruppen, das heißt abweichend von den zuvor dargestellten Analysen auch unter Einbezug von Beschäftigten der Wirtschaftsabschnitte A und O sowie der Arbeitnehmerinnen und Arbeitneh-

mer in Kleinstbetrieben. Auf Grundlage der beschriebenen Datenabgrenzung ergibt sich für im Rahmen der für Arbeitnehmerinnen durchgeführten Regressionsanalysen ein korrigierter Determinationskoeffizient von 74 %. Der entsprechende Wert bei den Männern liegt mit 78 % sogar noch etwas höher.

Mit wenigen Ausnahmen sind sämtliche Variablen der an dieser Stelle betrachteten Schätzungen signifikant. Insgesamt lassen sich hinsichtlich der Effektstärke und der Richtung des Zusammenhangs einzelner Merkmale kaum Unterschiede zu den Regressionsmodellen feststellen, die auf Basis der zuvor dargestellten Abgrenzung der Verdienststrukturerhebung 2014 berechnet wurden. Der Effekt der ausschließlich in den vorliegenden Modellen relevanten Dummy-Variablen für die Kleinstbetriebe signalisiert, dass männliche Beschäftigte in Unternehmen mit lediglich ein bis neun Angestellten gegenüber Arbeitnehmern in Firmen mit mindestens vierstelligem Mitarbeiterstab einen um etwa 20 % geringeren Lohn aufweisen. Bei den in Kleinstbetrieben tätigen Arbeitnehmerinnen beläuft sich der Wert auf knapp 16 %. Von der Dummy-Variablen „Angehörige der regulären Streitkräfte, Soldaten“ (ISCO 0), die erst durch die Ausweitung der Abgrenzung der Verdienststrukturerhebung 2014 auf den Wirtschaftsabschnitt O „Öffentliche Verwaltung, Verteidigung; Sozialversicherung“ relevant wird, geht sowohl bei den Männern als auch bei den Frauen ein schwacher positiver Einfluss auf die abhängige Variable aus.

Die Oaxaca-Blinder-Dekomposition ergibt ein mit den Ergebnissen der für die Verdienststrukturerhebung 2014 vorgenommenen Abgrenzung entsprechend der Arbeitnehmerabdeckung der Verdienststrukturerhebung 2006 vergleichbares Bild: 16,2 Prozentpunkte beziehungsweise etwa drei Viertel des unbereinigten Verdienst-

abstands zwischen Männern und Frauen in Höhe von 21,9 % lassen sich auf geschlechtsspezifische Unterschiede in den berücksichtigten Variablen zurückführen. Und 5,6 Prozentpunkte beziehungsweise ein Viertel des Verdienstunterschieds kann nicht auf derartige Ausstattungsunterschiede zurückgeführt werden. [↘ Tabelle 6](#)

Auch die detaillierte Zerlegung zeigt überwiegend vergleichbare Ergebnisse mit der Dekomposition auf Basis der Verdienststrukturerhebung 2006, die zu Beginn dieses Kapitels vorgestellt wurde. Zu den wichtigsten Faktoren zählen die Tätigkeit und die Leistungsgruppe. So sind 5,3 Prozentpunkte der Lohndifferenzen auf Unterschiede hinsichtlich der Verteilung von Männern und Frauen nach Tätigkeiten zurückzuführen. Weitere 4,9 Prozentpunkte lassen sich durch geschlechtsspezifische Unterschiede bei der Besetzung der Leistungsgruppen erklären. Schließlich erklärt auch hier der jeweils höhere Anteil von Frauen, die einer Teilzeitbeschäftigung oder einer geringfügigen Beschäftigung nachgehen, einen Teil der Differenzen in den Stundenverdiensten. Der Faktor „potenzielle Berufserfahrung“ hat in der vorliegenden Arbeitnehmerabgrenzung der Verdienststrukturerhebung ebenfalls keine relevante Bedeutung. [↘ Tabelle 7](#)

Anhand des Ergebnisses der Dekomposition des unerklärten Teils des Lohnabstands zwischen Männern und Frauen lässt sich feststellen, dass die Konstante nicht nur den größten Anteil ausmacht, sondern auch den gesamten unerklärten Teil des Verdienstunterschieds übersteigt. Daraus kann geschlossen werden, dass die berücksichtigten Merkmale in der Summe zulasten von Männern wirken. Die Betrachtung einzelner Merkmale zeigt ferner, dass insbesondere die Faktoren Altersteilzeit und Tätigkeit benachteiligend für Männer sowie die potenzielle Berufserfahrung, die Leistungsgruppe und der Gebietsstand zulasten von Frauen wirken.

Zusammenfassend kann festgestellt werden, dass der bereinigte Verdienstunterschied zwischen Männern und Frauen unabhängig von der Abgrenzung der in die Analyse einbezogenen Arbeitnehmerinnen und Arbeitnehmer 2014 bei rund 6 % lag.

Tabelle 6

Zerlegung des Gender Pay Gap

Ergebnisse der Verdienststrukturerhebung 2014 einschließlich der Wirtschaftsabschnitte A und O¹ sowie der Kleinstbetriebe

	Verdienstunterschied	Anteil
	%	
Unbereinigter Gender Pay Gap	21,9	100
Erklärter Teil	16,2	74,2
Unerklärter Teil (= bereinigter Gender Pay Gap)	5,6	25,8

¹ Klassifikation der Wirtschaftszweige, Ausgabe 2008. Wirtschaftsabschnitt A: Land- und Forstwirtschaft, Fischerei. Wirtschaftsabschnitt O: Öffentliche Verwaltung, Verteidigung; Sozialversicherung.

Tabelle 7

Detaillierte Zerlegung des unbereinigten Gender Pay Gap

Ergebnisse der Verdienststrukturerhebung 2014 einschließlich der Wirtschaftsabschnitte A und O¹ sowie der Kleinbetriebe

Erklärungsfaktoren	Erklärter Teil des Gender Pay Gap		Unerklärter Teil des Gender Pay Gap: bereinigter Gender Pay Gap	
	Prozentpunkte	%	Prozentpunkte	%
Insgesamt	16,2	74,2	5,6	25,8
Konstante	X	X	6,5	29,9
Gebietsstand	0,2	0,9	1,3	5,7
Differenzierter Regionstyp	0,1	0,6	-0,1	-0,5
Tätigkeit	5,3	24,1	-2,9	-13,5
Dienstalter	0,2	1,0	-0,8	-3,6
Unternehmensgröße	0,6	2,5	-0,3	-1,4
Ausbildungsabschluss	0,6	2,9	-0,4	-1,7
Leistungsgruppen 1 bis 5	4,9	22,3	1,2	5,6
Geringfügige Beschäftigung	1,2	5,5	-0,4	-1,9
Auszubildende	-0,3	-1,5	-0,3	-1,2
Tarifbindung	0,2	0,8	0,1	0,3
Potenzielle Berufserfahrung	-0,1	-0,3	4,6	20,9
Art des Arbeitsvertrags	0,1	0,6	0,4	1,6
Beschäftigungsumfang	2,8	12,7	-0,6	-2,6
Einfluss der öffentlichen Hand auf die Unternehmensführung	0,1	0,4	0,1	0,7
Altersteilzeit	0,1	0,3	-3,5	-15,8
Zulagen	0,3	1,5	0,7	3,1

¹ Klassifikation der Wirtschaftszweige, Ausgabe 2008. Wirtschaftsabschnitt A: Land- und Forstwirtschaft, Fischerei. Wirtschaftsabschnitt O: Öffentliche Verwaltung, Verteidigung; Sozialversicherung.

5

Weiterentwicklung der Schätzmodelle

5.1 Berücksichtigung von Erwerbsunterbrechungen

Analysen zum bereinigten Gender Pay Gap sind umso aussagekräftiger, je umfassender und genauer lohnbestimmende Variablen bei der Schätzung der Lohnfunktion berücksichtigt werden können. Allerdings verfügt keine der üblicherweise herangezogenen Datenquellen über den vollständigen relevanten Datenkranz. Bereits in der ersten Untersuchung zum bereinigten Gender Pay Gap auf Basis der Verdienststrukturerhebung 2006 wurde daher ausdrücklich darauf hingewiesen, dass es sich bei dem empirisch ermittelten Wert von 8 % eher um

eine „Obergrenze“ des Ausmaßes von Lohndiskriminierung handelt, da in die Modellberechnung nicht alle lohndeterminierenden Variablen einfließen können (Finke, 2010, hier: Seite 10). Beispielsweise liegen in der Verdienststrukturerhebung keine Angaben zum individuellen Verhalten in Lohnverhandlungen, zum Familienstand oder auch zur tatsächlichen Berufserfahrung vor, da diese vom nationalen und vom EU-Gesetzgeber nicht als Erhebungsmerkmale vorgesehen wurden.

Die tatsächliche Berufserfahrung wird in den Lohnfunktionen und damit auch im Oaxaca-Blinder-Dekompositionsmodell durch die potenzielle Berufserfahrung (siehe Übersicht 1) approximiert. Diese wiederum

erklärt in den empirischen Modellen einen erheblichen Teil des bereinigten Gender Pay Gap. Das Analyseergebnis besagt also, dass ein zusätzliches Jahr potenzieller Berufserfahrung bei Frauen geringer honoriert wird als bei Männern. Wenn die Arbeitgeber jedoch die tatsächliche Berufserfahrung kennen und diese bei der Entlohnung berücksichtigen, wären die Schätzergebnisse verzerrt, wenn die Differenz von potenzieller und tatsächlicher Berufserfahrung von Frauen größer als die der Männer ist. Dann würde bei einer Berücksichtigung der tatsächlichen anstelle der potenziellen Berufserfahrung in den Schätzmodellen der bereinigte Gender Pay Gap niedriger ausfallen. Beide Variablen unterscheiden sich durch die Dauer der Erwerbsunterbrechungen. Zu berücksichtigen wären unter anderem Arbeitslosigkeit sowie Unterbrechungen der Erwerbstätigkeit aufgrund der Geburt und Erziehung von Kindern. Da letztere ganz überwiegend die tatsächliche Berufserfahrung von Frauen mindern und somit hinsichtlich

Männern und Frauen asymmetrisch wirken, wäre deren Berücksichtigung in den Schätzmodellen zum bereinigten Gender Pay Gap anzustreben, um Verzerrungen zu vermeiden. Zu erwarten wäre dann ein niedrigerer bereinigter Gender Pay Gap.

Da die Dauer der Erwerbsunterbrechungen wegen Mutterschaft in der Verdienststrukturerhebung nicht vorliegt und auch aus der Literatur hierfür kein Schätzmodell bekannt ist, wurde die Fragestellung dahingehend vereinfacht, ob eine erwerbstätige Frau Mutter von mindestens einem Kind ist. Auch diese Information liegt in der Verdienststrukturerhebung nicht vor. Um die Mutterschaft dennoch in der Lohnfunktion berücksichtigen zu können, wurden zwei Ansätze verfolgt, nämlich die Schätzung der Eigenschaft „Mutter“ (a) auf der Basis von Merkmalen, die in der Verdienststrukturerhebung vorliegen, beziehungsweise (b) mithilfe von Daten aus externen Quellen.

Der erste Ansatz nutzt die Tatsache aus, dass in der Verdienststrukturerhebung neben dem Bruttomonatsverdienst auch die an das Finanzamt abgeführte Lohnsteuer einschließlich Solidaritätszuschlag erhoben wird. Mittels dieser Informationen und den Berechnungsformeln der Einkommensteuer wurde versucht, auf die Steuerklasse und von dieser auf einkommensteuerlich zu berücksichtigende Kinder zu schließen. Dies war jedoch nicht in hinreichendem Maße erfolgreich, da die Zuordnung zu den Steuerklassen unter anderem wegen individueller Abzugsbeträge oft nicht zweifelsfrei möglich war und zudem der Rückschluss auf die steuerlich zu berücksichtigenden Kinder häufig erfolglos blieb. Der Ansatz musste letztlich verworfen werden.

Daraufhin wurden als externe Quellen die Einkommensteuerstatistik und der Mikrozensus näher untersucht. Die Grundidee bestand darin, Datenbestände zu finden, die für erwerbstätige Frauen die Information über die Geburt mindestens eines Kindes enthalten, und diese geeignet auf die Verdienststrukturerhebung zu „übertragen“. Dafür war es zusätzlich erforderlich, dass die externe Quelle auch Variablen enthielt, die die Mutterschaft statistisch erklären können und die in der Verdienststrukturerhebung ebenfalls zur Verfügung stehen. Für die Einkommensteuerstatistik war dies nicht der Fall. Es konnten nur steuerlich relevante, das heißt in der Regel Kinder unter 25 Jahren indirekt ihren Müttern zugeordnet werden, was insbesondere bei älteren

erwerbstätigen Frauen zu einer Unterschätzung der Muttereigenschaft geführt hätte. Darüber hinaus war die Schnittmenge der Variablen mit der Verdienststrukturerhebung zu gering. Den weiteren Analysen wurden daher die Daten des Mikrozensus 2012 zugrunde gelegt.

Um die potenzielle Berufserfahrung verbessert schätzen zu können, wird also in einem ersten Schritt die Information benötigt, ob eine Arbeitnehmerin überhaupt Mutter¹¹ wenigstens eines Kindes ist (und somit eine Erwerbsunterbrechung angenommen werden kann¹²) oder nicht. Mit dieser Minimalanforderung beginnt die weitere Untersuchung. Es handelt sich daher um ein (binäres) Klassifikationsproblem, also um die Frage, welcher der beiden Gruppen (im Folgenden abgekürzt durch „Mutter“ oder „nicht Mutter“) eine Arbeitnehmerin zugehörig ist. Da a priori unklar ist, welche Merkmale und Merkmalskombinationen der Verdienststrukturerhebung für eine gute Schätzung der Mutterschaft einer Arbeitnehmerin herangezogen werden sollten und welche Verteilung diese Merkmale aufweisen, wurden statistische Verfahren in Erwägung gezogen, welche auf derartige A-priori-Annahmen nicht angewiesen sind. In die engere Auswahl kamen Random Forests (Feuerhake/Dumpert, 2016; Hastie und andere, 2008) und Support Vector Machines (SVMs) (Dumpert und andere, 2016; Feuerhake/Dumpert, 2016; Hamel, 2009), zwei nichtparametrische Machine-Learning-Methoden, die in der Regel mindestens die gleiche Güte erreichen wie andere Klassifikationsverfahren (Bennett/Campbell, 2000; Caruana/Niculescu-Mizil, 2006; Kotsiantis, 2007; Caruana und andere, 2008; Fernández-Delgado und andere, 2014; Wainberg und andere, 2016). Random Forests sind im Vergleich zu Support Vector Machines schneller im Lernen und späteren Klassifizieren. Support Vector Machines weisen hingegen einen deutlich reichhaltigeren theoretischen Hintergrund hinsichtlich ihrer statistischen Eigenschaften auf. Aus diesem Grund wurden zunächst Support Vector Machines als Klassifikationsmethode gewählt. Das Ziel des Einsatzes

11 Die Untersuchung wurde auf Arbeitnehmerinnen beschränkt. Selbstverständlich wird die potenzielle Berufserfahrung auch bei Vätern überschätzt, die aufgrund der Geburt eines Kindes längere Zeit nicht im Arbeitsleben standen. Aufgrund der betrachteten Altersjahrgänge ist jedoch davon auszugehen, dass die Zahl dieser Fälle im Vergleich zu der von Müttern, die dem Arbeitsmarkt eine gewisse Zeit nicht zur Verfügung standen, zu vernachlässigen ist.

12 Es ist zu beachten, dass es nicht darum geht, die Dauer der Erwerbsunterbrechung zu schätzen. Jedoch wird angenommen, dass Frauen, die wenigstens ein Kind geboren haben, eine Erwerbsunterbrechung vorweisen.

zes der Support Vector Machines bestand darin, Muster im Zusammenspiel erklärender Merkmale aus der Verdienststrukturerhebung zu erkennen, welche eine korrekte Vorhersage der Mutterschaft einer Arbeitnehmerin erlauben. Ohne weitere Informationen muss ein solches Vorhaben jedoch scheitern, da die Verdienststrukturerhebung kein Merkmal „Mutterschaft“ enthält. Von keiner Arbeitnehmerin in der Stichprobe der Verdienststrukturerhebung ist bekannt, ob sie jemals ein Kind geboren hat oder nicht. Obwohl die Muttereigenschaft für Arbeitnehmerinnen in der Verdienststrukturerhebung geschätzt werden soll, gibt es dort keine Möglichkeit für die statistische Methode, das zugrunde liegende Muster zu erlernen. Es bedarf daher einer weiteren Datenquelle, die sowohl Merkmale enthält, die auch in der Verdienststrukturerhebung vorhanden sind, als auch das binäre Merkmal „Mutterschaft“. Der Mikrozensus 2012 erfüllt

Übersicht 3

Relevante Merkmale des Mikrozensus 2012

- › Erwerbsstatus
 - › Alter
 - › Geschlecht
 - › gegenwärtiger Beruf
 - › Berufsschlüssel
 - › Stellung im Beruf
 - › Führungs-/Aufsichtskraft
 - › Öffentlicher Dienst
 - › Wirtschaftszweig
 - › Größenklasse des Betriebes
 - › Bundesland der Arbeitsstätte
 - › Befristung des Arbeitsvertrages
 - › Jahr des Beginns der jetzigen Tätigkeit
 - › Vollzeit-/Teilzeittätigkeit
 - › geringfügige Beschäftigung
 - › Vertrag mit Zeitarbeitsvermittlung
 - › Wochenarbeitszeit
 - › Überstunden der letzten Woche
 - › Samstagsarbeit, Sonntagsarbeit, Feiertagsarbeit
 - › Abendarbeit, Nachtarbeit
 - › Schichtarbeit
 - › höchster allgemeinbildender Schulabschluss
 - › höchster Berufs- oder Hochschulabschluss
 - › Einkommen aus Erwerbstätigkeit
 - › Nettoeinkommen des letzten Monats
-

diese Voraussetzungen und wurde deshalb als Datenquelle ausgewählt.¹³ ↗ **Übersicht 3** zeigt die Merkmale einer Person im Mikrozensus 2012, die auch in der Verdienststrukturerhebung 2014 vergleichbar vorliegen und sinnvoll auswertbar sind.

Um die für die Klassifikation „Mutter/nicht Mutter“ relevanten erklärenden Merkmale zu finden, wurden Testrechnungen mit verschiedenen Kombinationen der genannten Merkmale durchgeführt. Dabei stellte sich heraus, dass das Alter der Arbeitnehmerin sowie die Größe des Betriebes, in welchem sie beschäftigt ist, die wichtigsten Merkmale darstellen. Die Hinzunahme weiterer Merkmale zu diesen beiden verbesserte das Ergebnis nur noch marginal. Im Weiteren wurde daher stets mit den erklärenden Merkmalen Alter der Arbeitnehmerin und Größe des Betriebes gearbeitet, die übrigen Merkmale wurden nicht mehr herangezogen.

Jedes auf statistischem Lernen basierende Klassifikationsverfahren führt in der Praxis unvermeidlich auch zu Missklassifikationen. Deren Ausmaß und Verteilung lassen sich schätzen, indem man die mittels der Trainingsdaten gelernte Support Vector Machine auf Testdaten anwendet, für die das korrekte Klassifikationsergebnis „Mutter/nicht Mutter“ bereits bekannt ist. Die mittlere Missklassifikationsrate über mehrere verschiedene, zufällige Aufteilungen des Datensatzes (in Trainings- und Testdatensatz) sowie über Tests mit verschiedenen Unterstichproben des zur Verfügung stehenden Mikrozensus-Datenmaterials betrug etwa 20 %. Gleichzeitig wurde eine starke Disparität zwischen den auftretenden Fehlern festgestellt. Der Fehler, dass eine Mutterschaft geschätzt wurde, obwohl eine solche nicht vorlag, trat ungefähr dreimal so häufig auf wie der entgegengesetzte Fehler (irrtümliches Schätzen, dass keine Mutterschaft vorliegt). In den derzeitigen Modellen zur Erklärung des Verdienstunterschiedes zwischen Männern und Frauen werden alle Frauen so behandelt, als läge keine Erwerbsunterbrechung aufgrund einer Mutterschaft vor. Das irrtümliche Unterstellen einer Mutterschaft ist daher der schwerer wiegende Fehler, wenn man einen konservativen Ansatz zur Verbesserung des bisherigen Verfahrens verfolgt. Aus diesem Grund wurden weitere Anpassungen vorgenommen, um einen Ausgleich zwischen den beiden beschriebenen Arten

13 Hintergrundinformationen zum Mikrozensus liefert die Internetseite des Statistischen Bundesamtes (www.destatis.de).

der Fehlklassifikation zu erreichen (SVM-Modell „Parität“). Diese Anpassungen waren erfolgreich, gingen allerdings zu Lasten der Gesamtmissklassifikationsrate, sodass folgende (über mehrere verschiedene, zufällige Aufteilungen des Datensatzes) gemittelte Ergebnisse erzielt wurden. [↪ Tabelle 8](#)

Tabelle 8

Ergebnisse des Tests der Güte der Klassifikation des Support-Vector-Machine-Modells „Parität“

	%
Missklassifikationsrate	26
Irrtümlich als „nicht Mutter“ klassifiziert	14
Irrtümlich als „Mutter“ klassifiziert	12
Anteil der „Mütter“, die richtigerweise als „Mütter“ klassifiziert wurden	78
Anteil der „nicht Mütter“, die richtigerweise als „nicht Mütter“ klassifiziert wurden	68

Es zeigte sich, dass die SVM-Methodik ihre eigentliche Stärke, nämlich die Mustererkennung (unter Inkaufnahme langer Rechenzeiten), nicht ausspielen konnte. Random Forests lieferten bei deutlich kürzeren Berechnungsdauern vergleichbar gute Ergebnisse hinsichtlich der Gesamtmissklassifikationsraten. Allerdings weisen die Random Forests eine deutlich höhere Disparität zwischen den auftretenden Fehlern als die Support Vector Machines auf. Für die weitere Analyse wurde daher das SVM-Modell „Parität“ verwendet.

5.2 Ergebnisse

Mittels des SVM-Modells „Parität“ wurden die Arbeitnehmerinnen in der Verdienststrukturerhebung 2014 hinsichtlich der Eigenschaft „Mutter/nicht Mutter“ klassifiziert. Um den Effekt der Erwerbsunterbrechungen wegen der Geburt und Erziehung mindestens eines Kindes auf den bereinigten Gender Pay Gap messen zu können, muss die Information über die Mutterschaft in den Regressionsanalysen und im Dekompositionsmodell berücksichtigt werden. Naheliegender wäre die Bildung einer Dummy-Variablen. Da diese bei den Männern allerdings definitionsgemäß immer die Ausprägung „nicht Mutter“ hat, führt dies dazu, dass die Dummy-Variable in der Regression für die Männer nicht berücksichtigt wird. Dies wiederum hat zur Folge, dass die Dekomposition nicht durchgeführt werden kann. Daher bleibt methodisch nur die Möglichkeit, die Mutterschaft über einen

Abzug bei der Berechnung der potenziellen Berufserfahrung zu berücksichtigen. Hierfür sind Informationen über die (durchschnittliche) Dauer der Erwerbsunterbrechungen wegen der Geburt und Erziehung von Kindern notwendig. Da hierzu weder Erkenntnisse aus der Verdienststrukturerhebung noch Schätzergebnisse aus den Modellen vorliegen, wurde auf Angaben aus der Literatur zurückgegriffen. In Anlehnung an Boll und andere (2016b, hier: Seite 88 ff.) wird eine durchschnittliche Erwerbsunterbrechung durch Mutterschaft von acht Jahren unterstellt und die potenzielle Berufserfahrung von Müttern entsprechend modifiziert.

Mit dieser Annäherung an die tatsächliche Berufserfahrung wurde der bereinigte Lohnunterschied zwischen Männern und Frauen für 2014 mit dem oben geschilderten Verfahren und ansonsten gleichen Daten neu berechnet, und zwar sowohl ausschließlich (Abgrenzung der Verdienststrukturerhebung 2006) als auch einschließlich (Abgrenzung der Verdienststrukturerhebung 2014) der Beschäftigten in Kleinstbetrieben und in den Wirtschaftsabschnitten A und O. Dabei war zu beachten, dass das zugrunde liegende SVM-Modell „Parität“ mithilfe einer Zufallsstichprobe aus den zur Verfügung stehenden Mikrozensusdaten trainiert wurde. Um Zufallseffekte auszuschließen, wurden insgesamt zehn solcher Trainingsdatensätze zufällig erzeugt und zum Lernen des SVM-Modells „Parität“ verwendet. Somit standen auch zehn Klassifikationen der Eigenschaft „Mutter/nicht Mutter“ zur Verfügung, die in die Regressionen und die Dekompositionen einfließen. Letztlich wurden auf diese Weise zehn Werte für den bereinigten Verdienstabstand zwischen Männern und Frauen

Tabelle 9

Bereinigter Gender Pay Gap 2014 unter Berücksichtigung kindbedingter Erwerbsunterbrechungen

	Abgrenzung Verdienststrukturerhebung 2006	Abgrenzung Verdienststrukturerhebung 2014
Kennzahl		
Arithmetisches Mittel	5,62	5,33
Standardabweichung	0,10	0,10
Minimum	5,38	5,10
Maximum	5,77	5,48
Spannweite	0,39	0,38
Bereinigter Gender Pay Gap ohne Berücksichtigung von Erwerbsunterbrechungen	5,83	5,63
Differenz in Prozentpunkten	0,21	0,30

geschätzt, die sich jedoch nur marginal unterscheiden.¹⁴ [↪ Tabelle 9](#)


Die Berücksichtigung der Erwerbsunterbrechungen durch Mutterschaft führt also unter den gemachten Annahmen zu einem um rund 0,2 bis 0,3 Prozentpunkte niedrigeren bereinigten Gender Pay Gap.

6

Fazit und Ausblick

Mit den Daten der Verdienststrukturerhebung 2014 konnte für Deutschland zum dritten Mal nach 2006 und 2010 der bereinigte Gender Pay Gap ermittelt und eine detaillierte Ursachenanalyse durchgeführt werden. Für 2014 war es dabei möglich, auch die Beschäftigten in Kleinstbetrieben und in den Wirtschaftsabschnitten A und O einzubeziehen und so Aussagen für die gesamte Wirtschaft zu treffen. Zu Vergleichszwecken wurde der bereinigte Gender Pay Gap auch in der Abgrenzung der Beschäftigten der Verdienststrukturerhebung 2006 berechnet. Unabhängig von dieser Abgrenzung lag der bereinigte Lohnunterschied zwischen Männern und Frauen 2014 bei 6%. Seit 2006 lässt sich ein kontinuierlicher Rückgang feststellen (2006: 8%; 2010: 7%) bei annähernd gleichen Werten für den unbereinigten Gender Pay Gap (22% bis 23%).

Durch den Einsatz von Machine-Learning-Verfahren wurde erstmals versucht, Erwerbsunterbrechungen durch Mutterschaft approximativ in den Schätzmodellen zu berücksichtigen und so eine seit langem bekannte Schwachstelle der Modelle auszuräumen oder zumindest deren Relevanz abzuschätzen. Unter den gemachten Annahmen ergab sich für 2014 eine Abnahme des bereinigten Gender Pay Gap um rund 0,3 Prozentpunkte.

Wegen der vierjährlichen Periodizität der Verdienststrukturerhebung werden für 2018 das nächste Mal Daten zur Berechnung des bereinigten Gender Pay Gap vorliegen. Bis dahin ist geplant, die Modellrechnungen weiter zu verbessern. 

14 Um die Stabilität und Validität der Ergebnisse zu überprüfen, wurden außerdem zehn Random-Forest-Modelle mit ebenfalls zufällig gezogenen Trainingsdatensätzen gerechnet. Diese bestätigten die auf Basis des SVM-Modells „Parität“ ermittelten Ergebnisse.

LITERATURVERZEICHNIS

- Achatz, Juliane/Gartner, Hermann/Glück, Timea. *Bonus oder Bias? Mechanismen geschlechtsspezifischer Entlohnung*. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie. Jahrgang 57. Heft 3. 2005, Seite 466 ff.
- Bennett, Kristin P./Campbell, Colin. *Support Vector Machines: Hype or Hallelujah?* In: SIGKDD Explorations Newsletter. Band 2. Ausgabe 2. 2000, Seite 1 ff.
- Blinder, Alan S. *Wage Discrimination: Reduced Form and Structural Estimates*. In: The Journal of Human Resources. Band 8. Ausgabe 4. 1973, Seite 436 ff.
- Boll, Christina/Rossen, Anja/Wolf, André. *The EU Gender Earnings Gap: Job Segregation and Working Time as Driving Factors*. In: Institut für Arbeitsmarkt- und Berufsforschung. IAB-Discussion Paper. Band 36. 2016a.
- Boll, Christina/Jahn, Malte/Lagemann, Andreas/Puckelwald, Johannes. *Dauerhaft ungleich – berufsspezifische Lebenserwerbseinkommen von Frauen und Männern in Deutschland*. In: Hamburgisches WeltWirtschaftsinstitut HWWI Policy Paper 98. Hamburg 2016b.
- Bundesamt für Bauwesen und Raumordnung. *Siedlungsstrukturelle Kreistypen*. [Zugriff am 28. Februar 2017]. Verfügbar unter: www.bbsr.bund.de
- Bundesministerium für Familie, Senioren, Frauen und Jugend. *Lohnungerechtigkeit*. Hintergrundmeldung vom 11. Januar 2017.
- Caruana, Rich/Karampatziakis, Nikos/Yessenalina, Ainur. *An Empirical Evaluation of Supervised Learning in High Dimensions*. In: Proceedings of the 25th International Conference on Machine Learning. Helsinki 2008, Seite 96 ff.
- Caruana, Rich/Niculescu-Mizil, Alexandru. *An Empirical Comparison of Supervised Learning Algorithms*. In: Proceedings of the 23rd International Conference on Machine Learning. Pittsburgh 2006, Seite 161 ff.
- Dumpert, Florian/Eschwege, Katja von/Beck, Martin. *Einsatz von Support Vector Machines bei der Sektorzuordnung von Unternehmen*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2016, Seite 87 ff.
- Eurostat. *Working Group Labour Market Statistics. State of the Play on the SES 2006 and Gender Pay Gap 2007*. Arbeitsunterlage. Luxemburg 2008.
- Fernández-Delgado, Manuel/Cernadas, Eva/Barro, Senén. *Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?* In: Journal of Machine Learning Research. Band 15. 2014, Seite 3133 ff.
- Feuerhake, Jörg/Dumpert, Florian. *Erkennung nicht relevanter Unternehmen in den Handwerksstatistiken*. In: WISTA Wirtschaft und Statistik. Ausgabe 2/2016, Seite 79 ff.

LITERATURVERZEICHNIS

Finke, Claudia. *Verdienstunterschiede zwischen Männern und Frauen 2006*. Projektbericht. Statistisches Bundesamt. Oktober 2010. [Zugriff am 1. März 2017]. Verfügbar unter: www.destatis.de

Finke, Claudia. *Verdienstunterschiede zwischen Männern und Frauen*. In: *Wirtschaft und Statistik*. Ausgabe 1/2011, Seite 36 ff.

Geisberger, Tamara/Glaser, Thomas. *Geschlechtsspezifische Verdienstunterschiede. Analysen zum „Gender Pay Gap“ auf Basis der Verdienststrukturerhebung 2010*. In: *Statistische Nachrichten*. Ausgabe 3/2014, Seite 1 ff.

Günther, Roland. *Methodik der Verdienststrukturerhebung 2010*. In: *Wirtschaft und Statistik*. Ausgabe 2/2013, Seite 127 ff.

Hamel, Lutz. *Knowledge Discovery with Support Vector Machines*. Hoboken 2009.

Hammermann, Andrea/Schmidt, Jörg. *Facetten des Gender Pay Gap: Empirische Evidenz auf Basis des Sozio-oekonomischen Panels*. In: *Institut der deutschen Wirtschaft IW policy papers*. Ausgabe 8/2015.

Hastie, Trevor/Tibshirani, Robert/Friedman, Jerome. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second Edition. New York 2009.

Hinz, Thomas/Gartner, Hermann. *Geschlechtsspezifische Lohnunterschiede in Branchen, Berufen und Betrieben*. In: *Zeitschrift für Soziologie*. Jahrgang 34. Ausgabe 1/2005, Seite 22 ff.

Hübler, Olaf. *Geschlechtsspezifische Lohnunterschiede*. In: *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*. Jahrgang 36. Ausgabe 4/2003, Seite 539 ff.

International Labour Organization. *ISCO International Standard Classification of Occupations*. [Zugriff am 24. Februar 2017]. Verfügbar unter: www.ilo.org

James, Gareth/Witten, Daniela/Hastie, Trevor/Tibshirani, Robert. *An Introduction to Statistical Learning with Applications in R*. New York 2013.

Joachimiak, Walter. *Frauenverdienste – Männerverdienste: Wie groß ist der Abstand wirklich?* In: *STATmagazin*. Wiesbaden 2013. [Zugriff am 29. Januar 2017]. Verfügbar unter: www.destatis.de

Kotsiantis, Sotiris B. *Supervised Machine Learning: A Review of Classification Techniques*. In: *Informatica*. Band 31. Heft 3/2007, Seite 249 ff.

Kulmiz, Leontine von. *Lohndiskriminierung von Frauen*. In: *Wirtschaft und Statistik*. Ausgabe 5/2001, Seite 406 ff.

Neumark, David. *Sex Differences in Labor Markets*. London, New York 2004.

Oaxaca, Ronald. *Male-Female Wage Differentials in Urban Labour Markets*. In: *International Economic Review*. Band 14. Ausgabe 3/1973, Seite 693 ff.

LITERATURVERZEICHNIS

Statistisches Bundesamt. *Verdienststrukturerhebung*. Qualitätsbericht. Wiesbaden 2016a. [Zugriff am 1. März 2017]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Fachserie 16, Heft 1, Verdienststrukturerhebung 2014*. Wiesbaden 2016b. [Zugriff am 1. März 2017]. Verfügbar unter: www.destatis.de

Strub, Silvia/Abrassart, Aurélien/Bannwart, Livia/Oesch, Thomas. *Analyse der Löhne von Frauen und Männern anhand der Lohnstrukturerhebung 2012*. Schlussbericht. Büro für arbeits- und sozialpolitische Studien. Bern 2016.

Wainberg, Michael/Alipanahi, Babak/Frey, Brendan J. *Are Random Forests Truly the Best Classifiers?* In: *Journal of Machine Learning Research*. Band 17. 2016, Seite 1 ff.

Herausgeber

Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung

Dieter Sarreither, Präsident des Statistischen Bundesamtes

Redaktionsleitung: Kerstin Hänsel

Redaktion: Ellen Römer

Ihr Kontakt zu uns

www.destatis.de/kontakt

Erscheinungsfolge

zweimonatlich, erschienen im April 2017

Das Archiv aller Ausgaben ab Januar 2001 finden Sie unter www.destatis.de/publikationen

Print

Einzelpreis: EUR 18,- (zzgl. Versand)

Jahresbezugspreis: EUR 108,- (zzgl. Versand)

Bestellnummer: 1010200-17002-1

ISSN 0043-6143

ISBN 978-3-8246-1061-7

Download (PDF)

Artikelnummer: 1010200-17002-4, ISSN 1619-2907

Vertriebspartner

IBRo Versandservice GmbH

Bereich Statistisches Bundesamt

Kastanienweg 1

D-18184 Roggentin

Telefon: +49 (0) 382 04 / 6 65 43

Telefax: +49 (0) 382 04 / 6 69 19

destatis@ibro.de

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

© Statistisches Bundesamt (Destatis), 2017

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.