



Der Einfluss der Kontrollvariablen in Logib Modul 1

Bericht

BEREICH ARBEIT

Eidgenössisches Departement des Innern EDI
Eidgenössisches Büro für die Gleichstellung von Frau und Mann EBG

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra



IMPRESSUM

TITEL

Der Einfluss der Kontrollvariablen in Logib Modul 1

AUTOREN

Boris Kaiser und Lukas Mergele

HERAUSGEBERIN

Eidgenössisches Büro für die Gleichstellung
von Frau und Mann EBG

VERTRIEB

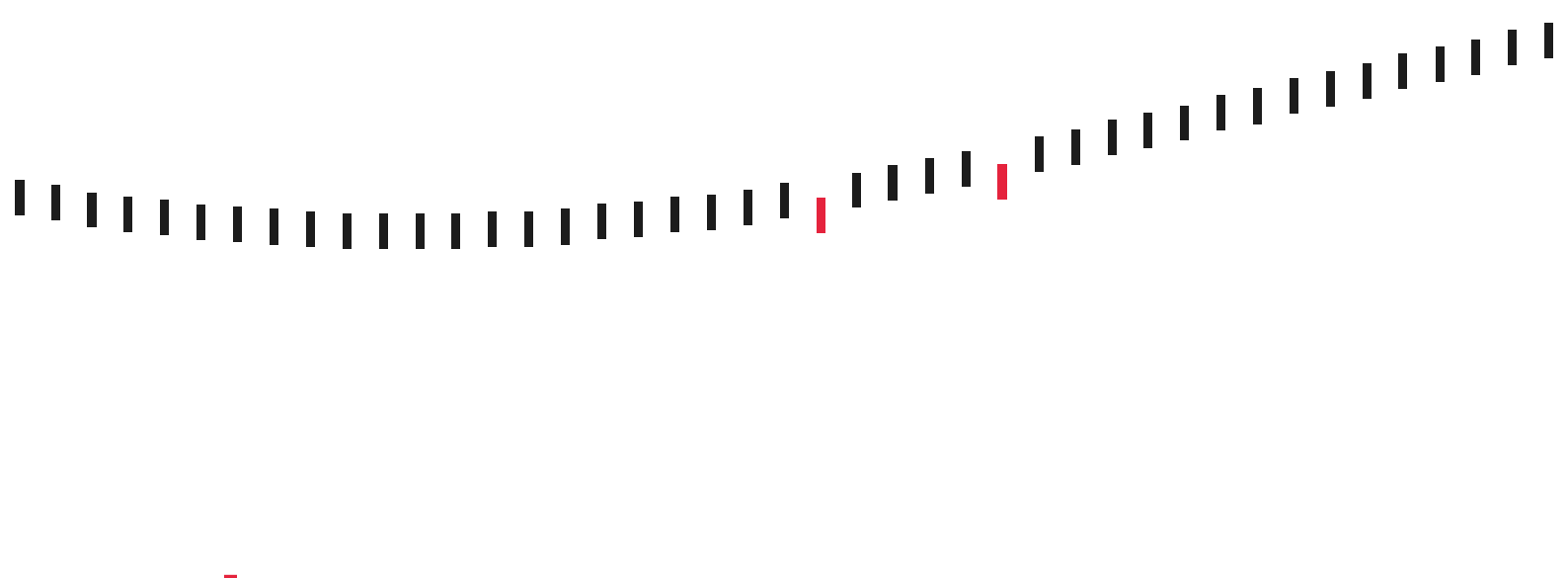
Eidgenössisches Büro für die Gleichstellung
von Frau und Mann EBG
Schwarztorstrasse 51, 3003 Bern
ebg@ebg.admin.ch
www.ebg.admin.ch

Dieser Bericht wurde im Auftrag des EBG verfasst. Die darin enthaltenen Einschätzungen und Interpretationen entsprechen nicht zwingend der Sicht des EBG.

Bericht

Der Einfluss der Kontrollvariablen in Logib Modul 1

Basel | 31.05.2023



Impressum

Der Einfluss der Kontrollvariablen in Logib Modul 1

Schlussbericht

31.05.2023

Auftraggeberin: Eidgenössisches Büro für die Gleichstellung von Frau und Mann EBG

Autoren: Boris Kaiser und Lukas Mergele

Verantwortlich seitens Auftraggeberin: Simon Hilber

Projektleitung seitens Auftragnehmerin: Boris Kaiser

Zitierweise:

Kaiser, B. & Mergele, L. (2023). Der Einfluss der Kontrollvariablen in Logib Modul 1. BSS Volkswirtschaftliche Beratung. Studie im Auftrag des EBG.

BSS Volkswirtschaftliche Beratung AG

Aeschengraben 9

4051 Basel

T +41 61 262 05 55

contact@bss-basel.ch

www.bss-basel.ch

© 2023 BSS Volkswirtschaftliche Beratung AG

Inhalt

Tabellen	iii
Abbildungen	iv
Zusammenfassung	1
1 Einleitung	3
2 Statistische Grundlagen	4
2.1 Identifikation des Geschlechter-Effekts	4
2.2 Schätzung des Geschlechter-Effekts	13
3 Empirisches Vorgehen und Daten	14
3.1 Auswirkungen von Modellmodifikationen	14
3.2 Abschätzung von Omitted Variable Bias	15
3.3 Datengrundlage	16
4 Ergebnisse	18
4.1 Verteilung der Geschlechtereffekte im Standardmodell	18
4.2 Auswirkungen verschiedener Modifikationen	20
4.3 Heterogenitätsanalyse	30
4.4 Einschätzung des Omitted Variable Bias	34
5 Fazit	38
6 Quellenangaben	41
A. Anhang	42
A.1 Deskriptive Statistiken nach Merkmalen	42
A.2 Unterschiede nach Frauenanteil der Firma	48

Tabellen

Tabelle 1: Zahlenbeispiel – Lohn­tabelle mit X1 und X2.....	7
Tabelle 2: Zahlenbeispiel – Durchschnittslöhne und Lohndifferenz	7
Tabelle 3: Zahlenbeispiel – Lohn­tabelle mit X1 und X2, juristische Betrachtung.....	8
Tabelle 4: Zahlenbeispiel – Lohn­tabelle ohne X2.....	9
Tabelle 5: Zahlenbeispiel – Durchschnittslöhne und Lohndifferenz ohne X2	10
Tabelle 6: Zahlenbeispiel – Lohn­tabelle mit endogenem X2.....	12
Tabelle 7: Zahlenbeispiel – Durchschnittslöhne und Lohndifferenz mit endogenem X2	12
Tabelle 8: Relevante Variablen für Logib Modul 1	17
Tabelle 9: Zusätzlich relevante Variablen für Modifikationen	17
Tabelle 10: Modellvarianten mit ausgeschlossenen Variablen, private Firmen	22
Tabelle 11: Modellvarianten mit ausgeschlossenen Variablen, öffentliche Firmen.....	23
Tabelle 12: Modellvarianten mit detaillierteren Variablen, private Firmen	24
Tabelle 13: Modellvarianten mit detaillierteren Variablen, öffentliche Firmen.....	25
Tabelle 14: Modellvarianten mit zusätzlichen Variablen, private Firmen	27
Tabelle 15: Modellvarianten mit zusätzlichen Variablen, öffentliche Firmen.....	28
Tabelle 16: Schätzresultate nach Branchen und kollektivvertraglichen Regelungen	29
Tabelle 17: Logib-Ergebnisse vs. «simulierter» Geschlechtereffekt, private Firmen.....	37
Tabelle 18: Logib-Ergebnisse vs. «simulierter» Geschlechtereffekt, öffentliche Firmen	38
Tabelle 19: Deskriptive Statistiken nach Firmeneigentümerschaft.....	43
Tabelle 20: Deskriptive Statistiken nach Geschlecht der Mitarbeitenden, private Firmen.....	44
Tabelle 21: Deskriptive Statistiken nach Geschlecht der Mitarbeitenden, öffentliche Firmen.....	45
Tabelle 22: Deskriptive Statistiken nach Zahlung von Zulagen.....	46
Tabelle 23: Deskriptive Statistiken nach Firmeneigentümerschaft.....	47

Abbildungen

Abbildung 1: Schematische Darstellung der Einflusskanäle ohne fehlende Variablen.....	6
Abbildung 2: Schematische Darstellung der Einflusskanäle mit fehlenden Variablen.....	9
Abbildung 3: Schematische Darstellung der Einflusskanäle ohne fehlende Variablen.....	11
Abbildung 4: Häufigkeitsverteilung der Geschlechtereffekte mit dem Standardmodell	19
Abbildung 5: Kumul. Wahrscheinlichkeit der Geschlechtereffekte im Standardmodell	20
Abbildung 6: Testergebnisse nach Firmengrösse, ohne berufliche Stellung	31
Abbildung 7: Testergebnisse nach Firmengrösse, mit Berufskategorien.....	32
Abbildung 8: Testergebnisse nach Firmengrösse, mit Zulagen-Dummy	33
Abbildung 9: Kerndichteverteilung der geschätzten Geschlechterkoeffizienten.....	35
Abbildung 10: Geschlechtereffekt und Verzerrung durch unbeobachtete Variablen.....	36
Abbildung 11: Signifikanzrate nach Frauenanteil, ohne berufliche Stellung, alle Firmen	48
Abbildung 12: Signifikanzrate nach Frauenanteil, Zulagen, alle Firmen.....	49
Abbildung 13: Signifikanzrate nach Frauenanteil, Berufskategorien, alle Firmen	50

Zusammenfassung

Für die Analyse der Lohngleichheit zwischen Frauen und Männern innerhalb von Unternehmen stellt der Bund den Arbeitgebenden das Standard-Analyse-Tool «Logib» zur Verfügung. Das Modul 1 von Logib richtet sich an Unternehmen ab 50 Beschäftigten und besteht im Kern aus einer Regressionsanalyse. Anhand von statistischen Tests wird beurteilt, ob in einem Unternehmen ein signifikanter, geschlechterspezifischer Lohnneffekt vorliegt oder nicht.

Das Ziel der vorliegenden Studie besteht darin, zu untersuchen, wie stark die Kontrollvariablen den geschätzten Geschlechtereffekt in Logib Modul 1 beeinflussen. Dabei ist insbesondere die Frage von Interesse, inwieweit der Geschlechtereffekt aufgrund von fehlenden Variablen verzerrt geschätzt wird (sogenannter «Omitted Variable Bias», OVB). Da statistische Lohnanalysen die komplexe Realität von Lohnsetzungsprozessen modellieren müssen, besitzt diese Fragestellung eine hohe empirische und praktische Relevanz.

Das methodische Vorgehen besteht darin, anhand der Schweizerischen Lohnstrukturerhebung (LSE) 2020 verschiedene Analysen auf Unternehmensebene durchzuführen. Erstens untersuchen wir, wie verschiedene Modifikationen der Kontrollvariablen die Ergebnisse von Logib im Vergleich zum heute gültigen Standardmodell beeinflussen. Dabei prüfen wir das Weglassen oder Modifizieren bestehender Variablen sowie die Hinzunahme zusätzlicher Variablen. Zu betonen ist, dass die Analyse alternativer Modellspezifikationen eine rein statistische Betrachtung darstellt; juristische Überlegungen zur Eignung der jeweiligen Merkmale wurden nicht miteinbezogen. Zweitens nutzen wir die Methode von Oster (2019), um mögliche Verzerrungen aufgrund von fehlenden Variablen (OVB) abzuschätzen und deren Einfluss auf die Testergebnisse von Logib zu untersuchen.

Die Lohnanalysen auf Unternehmensebene zeigen: Im Schnitt liefert das Logib-Standardmodell einen Geschlechtereffekt von -5.4%. Bei rund 55% der Unternehmen beträgt der geschätzte Geschlechtereffekt mindestens 5%; in den allermeisten Fällen ist dieser zu Ungunsten der Frauen. Rund 17% der privaten Unternehmen und 14% der öffentlichen Unternehmen bestehen den Test gemäss Logib Modul 1 nicht, das heisst, der Geschlechtereffekt ist statistisch signifikant grösser als 5%. Der durchschnittliche Geschlechtereffekt auf Unternehmensebene ist aus vier Gründen nicht direkt mit der unerklärten Lohndifferenz der nationalen Statistik (vgl. Kaiser & Möhr 2023) vergleichbar: (i) Nur Lohnunterschiede *innerhalb* von Unternehmen werden berücksichtigt, (ii) die verwendeten Kontrollvariablen unterscheiden sich, (iii) nur mittlere und grössere Unternehmen werden einbezogen und (iv) die Ergebnisse sind nicht gewichtet mit der Anzahl Beschäftigten.

Die Ergebnisse zeigen, dass die bestehenden Variablen unterschiedlich wichtig sind, um den OVB zu reduzieren. Dem Merkmal «Berufliche Stellung» kommt insgesamt die grösste Bedeutung zu, während die Variablen «Dienstjahre», «Potenzielle Erwerbsjahre» und «Ausbildungsjahre», zumindest bei privaten Firmen, weniger bedeutsam sind. Wenn man anstelle der Ausbildungsjahre die Ausbildungskategorien verwendet, reduziert sich der Geschlechtereffekt bei den öffentlichen

Firmen merklich, bei den privaten Firmen hingegen kaum. Deutlichere Effekte treten bei der Nutzung von Berufskategorien nach CH-ISCO-19 anstelle des Kompetenzniveaus zutage. Verglichen mit den Ergebnissen von Felder & Wunsch (2021) finden wir nur einen geringfügigen Rückgang der Betriebe, die in den Bereich der kritischen Lohnunterschiede fallen. Bei der Hinzunahme weiterer Kontrollvariablen ins Modell ergibt sich die grösste Veränderung durch die Berücksichtigung eines Indikators für Zulagen für Schicht-, Sonntags- oder Nachtarbeit. Der Lohnunterschied nach Geschlecht sinkt in dieser Spezifikation durchschnittlich um rund einen halben Prozentpunkt im Vergleich zum Standardmodell. Weiterführende Analysen zeigen, dass sich die Zulagen-Variable besonders stark auf Lohnanalysen von grösseren Firmen auswirkt. Weitere geprüfte Merkmale wie Beschäftigungsgrad, Stundenlohn, Zivilstand und Nationalität wirken sich im Durchschnitt nur relativ geringfügig auf den Geschlechtereffekt und die Logib-Testergebnisse von privaten Firmen aus. Demgegenüber fällt auf, dass die Hinzunahme weiterer Variablen den Geschlechtereffekt bei öffentlichen Firmen teilweise bedeutsam senkt. Beispielsweise reduziert sich der Anteil der öffentlichen Unternehmen, die den Logib-Test nicht bestehen, von rund 14% auf 11%, wenn der Indikator «Zulagen» in das Modell inkludiert wird.

Schliesslich bestand ein weiteres Ziel der Studie darin, das Ausmass des OVB indirekt abzuschätzen. Dazu kam die Methode von Oster (2019) zur Anwendung, die es erlaubt, mit Hilfe von statistischen Annahmen eine obere «Schranke» für den OVB zu berechnen. Der um OVB korrigierte Geschlechtereffekt liegt durchschnittlich bei etwa -2,6% über alle Unternehmen. Dieses Ergebnis impliziert, dass die bestehenden Kontrollvariablen im Logib-Standardmodell rund drei Viertel des OVB im Vergleich zum Modell ohne Kontrollvariablen absorbieren. Der Vergleich zwischen dem «simulierten» Geschlechtereffekt anhand der Oster-Methode und dem Geschlechtereffekt aus dem Standardmodell suggeriert, dass die Bereinigung um OVB die Häufigkeit von nicht bestandenen Testergebnissen verringert. Dieser Rückgang ist insbesondere bei den öffentlichen Firmen besonders ausgeprägt.

Zusammengenommen liefern die Untersuchungen wichtige Erkenntnisse, die in die Beurteilung und Weiterentwicklung von Logib Modul 1 einfließen können. Zum einen können die Ergebnisse als Informationsgrundlage hinzugezogen werden zur Frage, welche Modifikationen des Standardmodells als sinnvoll erachtet werden. Zum anderen zeigen die Ergebnisse annäherungsweise auf, wie gross der Einfluss durch verbleibende Verzerrungen ist und wie stark sich diese in den Testergebnissen von Logib niederschlagen.

Abschliessend ist auch auf drei Limitationen dieser Studie hinzuweisen. Erstens beruhen unsere Analysen nur auf Unternehmen, die mit mindestens 50 vollständigen Beobachtungen (d.h. ohne fehlende Werte) in der LSE vertreten sind. Damit dürften insbesondere Unternehmen zwischen 50 und 150 Beschäftigten unterrepräsentiert sein. Zweitens verfügt die LSE nicht über die für das Logib Modul 1 erforderliche Angabe zum firmenspezifisch kodierten Anforderungsniveau der ausgeübten Funktion. Als Ersatz greifen wir auf das in der LSE vorhandene Kompetenzniveau zurück, welches vermutlich einen geringeren Erklärungsgehalt für die innerbetriebliche Lohnstruktur besitzt. Drittens bleibt festzuhalten, dass der Ansatz von Oster (2019) trotz seiner Vorzüge nur eine Annäherung an den tatsächlichen OVB liefert und die quantitativen Ergebnisse basierend auf dieser Methode mit entsprechender Vorsicht zu interpretieren sind.

1 Einleitung

Für die Analyse der Lohngleichheit zwischen Frauen und Männern innerhalb von Unternehmen stellt der Bund den Arbeitgebenden das Standard-Analyse-Tool «Logib» zur Verfügung. Logib besteht aus zwei Modulen: Modul 1 basiert auf einer statistischen Regressionsmethode und richtet sich an mittlere und grössere Unternehmen ab 50 Beschäftigten, während Modul 2 für kleine Unternehmen konzipiert ist. Konkret besteht Modul 1 aus einer linearen OLS-Regression der standardisierten, logarithmierten Löhne auf einen Geschlechter-Indikator und eine Reihe von Kontrollvariablen. Anhand von zwei statistischen t-Tests wird beurteilt, ob in einem Unternehmen ein signifikanter, geschlechterspezifischer Lohneffekt vorliegt oder nicht: In einem ersten Schritt wird geprüft, ob der Geschlechtereffekt signifikant verschieden ist von null. Trifft dies zu, wird in einem zweiten Schritt getestet, ob der Geschlechtereffekt die «Toleranzschwelle» von 0.05 (d.h. 5%) signifikant überschreitet. Erst wenn der Geschlechtereffekt im Betrag signifikant grösser ist als 5%, gilt der Logib-Test als nicht bestanden. Die Toleranzschwelle soll unter anderem dem Umstand Rechnung tragen, dass die Lohnbildung teilweise unternehmensspezifisch ist und dass ein allgemeingültiges Modell nicht in jedem Fall alle lohnrelevanten, exogenen Merkmale als Kontrollvariablen berücksichtigen kann. Fehlende Variablen können dazu führen, dass der Geschlechtereffekt verzerrt geschätzt wird, wenn diese gleichzeitig lohnbestimmend sind *und* mit dem Geschlecht der Beschäftigten korrelieren. In der Statistik bezeichnet man solche Verzerrungen allgemein als «Omitted Variable Bias» (OVB).

Die Frage, wie stark der Einfluss von (fehlenden) Kontrollvariablen auf den geschätzten Geschlechtereffekt ist, wurde im Kontext von Logib Modul 1 bislang nur punktuell untersucht. Felder & Wunsch (2021) führen mithilfe der Lohnstrukturerhebung (LSE) 2018 eine methodische Überprüfung des Moduls 1 durch.¹ Neben einer Analyse der Wahl des statistischen Verfahrens zeigen die Autorinnen, dass die Berücksichtigung der Berufsgruppe im Modul 1 mit einem Rückgang der Betriebe einhergeht, die in den Bereich der kritischen Lohnunterschiede fallen.² Die Hinzunahme weiterer Kontrollvariablen wurde jedoch nicht überprüft.

Das Ziel der vorliegenden Studie ist es, den möglichen Einfluss von (fehlenden) Kontrollvariablen auf die Ergebnisse von Logib Modul 1 einer vertieften Analyse zu unterziehen. Die zentralen Forschungsfragen lauten: Wie stark verändern sich der geschätzte Geschlechtereffekt und die Test-Ergebnisse, wenn zusätzliche Kontrollvariablen in das Modell inkludiert werden? Welche Variablen haben einen grösseren Einfluss, welche nicht? In welcher Grössenordnung ist der potenzielle OVB zu verorten? Nebst fehlenden Variablen ist auch von Interesse, wie stark sich die Ergebnisse von Logib ändern würden, wenn man bestehende Kontrollvariablen weglassen oder diese inhaltlich anders operationalisieren würde.

Das methodische Vorgehen besteht aus einer empirischen Analyse von Schweizer Lohndaten. Dazu nutzen wir die Mikrodaten der Schweizerischen Lohnstrukturerhebung (LSE) 2020 des

¹ Ausgeschlossen in dieser Analyse sind Betriebe mit weniger als 50 Mitarbeitenden, sowie Betriebe mit weniger als 20 Frauen oder 20 Männern.

² Ob Berufsgruppen im Lohnmodell verwendet werden sollen, wird in der arbeitswissenschaftlichen Literatur jedoch kritisch betrachtet (vgl. Klammer et al. 2022).

Bundesamts für Statistik (BFS). Grundlage für die Analyse bilden alle Unternehmen, für die mindestens 50 vollständige Beobachtungen (d.h. ohne fehlende Werte) verfügbar sind.³ Für jedes dieser Unternehmen vergleichen wir die Ergebnisse des bestehenden Logib-Modells («Standardmodell») mit den Ergebnissen alternativer Lohnmodelle, die jeweils eine bestehende Variable ausschliessen, diese anders operationalisieren oder zusätzliche Variablen inkludieren. So lassen sich Erkenntnisse zu Einfluss und Relevanz verschiedener Kontrollvariablen gewinnen. Um die Gröszenordnung des OVB insgesamt abzuschätzen, nutzen wir den sogenannten Bounding-Ansatz von Oster (2019), der in den angewandten Wirtschaftswissenschaften grosse Beachtung gefunden hat. Die Oster-Methode erlaubt es, näherungsweise abzuschätzen, wie sich die Häufigkeit von nicht bestandenen Testergebnissen verändern würde, wenn der OVB umfassend berücksichtigt werden könnte. Die Ergebnisse werden für Unternehmen im privaten und öffentlichen Sektor jeweils separat ausgewiesen.

Dieser Bericht ist wie folgt strukturiert. Zunächst erarbeiten wir in Kapitel 2 statistische Grundlagen, die für das Verständnis von Logib zentral sind. Insbesondere gehen wir auf die Identifikation und Schätzung des Geschlechtereffekts ein. In Kapitel 3 erläutern wir das methodische Vorgehen und beschreiben die Datenaufbereitung. In Kapitel 4 präsentieren wir die Ergebnisse der empirischen Analysen. Kapitel 5 rundet den Bericht mit einem Fazit ab. Ergänzende Informationen, wie eine formale Darstellung der Methoden sowie weitere Ergebnisse sind im Anhang A aufgeführt.

2 Statistische Grundlagen

2.1 Identifikation des Geschlechter-Effekts

Allgemein besteht das Ziel von betrieblichen Lohngleichheitsanalysen darin, zu prüfen, ob Frauen und Männer innerhalb eines Unternehmens für gleichwertige Arbeit gleich entlohnt werden oder ob die Lohngleichheit verletzt ist. Die zentrale Herausforderung dabei ist, zu eruieren, wie der «Wert» einer Arbeitsleistung ermittelt werden kann, da dieser in den allermeisten Fällen nicht direkt gemessen oder beobachtet werden kann.⁴ Aus ökonomischer Sicht wird der Wert der Arbeit vor allem mit dem Beitrag der Beschäftigten zur Wertschöpfung des Unternehmens (Produktivität) in Verbindung gesetzt. Aus rechtlicher Sicht wird hingegen argumentiert, dass sich der Wert an einer arbeitswissenschaftlichen Arbeitsbewertung orientiert. Damit eine Einschätzung erfolgen kann, ob allfällige Lohnunterschiede gerechtfertigt bzw. ungerechtfertigt sind, müssen messbare, produktivitätsrelevante Einflussfaktoren (Kontrollvariablen) hinzugezogen werden. Den

³ Dies entspricht nicht allen Unternehmen in der LSE ab 50 Beschäftigten. Erstens haben Unternehmen die Möglichkeit, nur eine Zufallsstichprobe (Angaben zu jeder 2. oder 3. beschäftigten Person) zu liefern. Zweitens führt der Ausschluss von Beobachtungen mit unplausiblen und/oder fehlenden Werten zu einer Verringerung der Beobachtungen auf Unternehmensebene.

⁴ Hier sind zwei Ausnahmen zu nennen. Die Erste ist der Akkordlohn, bei dem die Beschäftigten direkt anhand der individuellen Produktionsmenge (Stückzahl, Kilogramm) entlohnt werden. Der Lohn lässt sich also vollständig durch die Produktivität erklären. Der Akkordlohn kommt in modernen Arbeitsmärkten aber kaum noch zur Anwendung, war früher aber in der Landwirtschaft verbreitet. Die zweite Ausnahme sind moderne Plattformökonomien, wo der Lohn per Algorithmus bestimmt wird, der ausschliesslich auf beobachtbaren Inputfaktoren basiert.

Lohngleichheitsanalysen liegt in aller Regel eine statistische Methode zugrunde, die das Ziel verfolgt, den Effekt des Geschlechts auf den Lohn von den Kontrollvariablen zu trennen.

Die zugrunde liegende Idee hinter statistischen Methoden ist, für jede Frau einen kontrafaktischen Lohn zu ermitteln, den sie erhalten würde, wenn keine Diskriminierung bei der Lohnfestsetzung vorläge. Die Differenz zwischen dem kontrafaktischen Durchschnittslohn und dem tatsächlichen Durchschnittslohn entspricht der Lohndiskriminierung im statistischen Sinne. In der Ökonometrie spricht man in diesem Kontext auch vom *kausalen Effekt* der Diskriminierung auf die Löhne. Hervorzuheben ist, dass sich diese Definition von der Lohndiskriminierung im juristischen Sinne unterscheiden kann. Letztere entspricht der Lohndifferenz, welche im Durchschnitt vor Gericht eingeklagt werden könnte.⁵

Zu erwähnen ist, dass bei *Lohngleichheitsanalysen* die potenzielle Lohndiskriminierung durch Unternehmen im Fokus steht (Lohndiskriminierung im engeren Sinne). Geschlechterdiskriminierung kann jedoch auch andere Formen annehmen, etwa bei der Anstellung, der Beförderung oder bei der Ausbildungs- und Berufswahl. Diese Diskriminierungsformen werden im vorliegenden Kontext von Logib nicht weiter betrachtet, was aber nicht heisst, dass diesen Formen der Diskriminierung keine hohe gesellschaftliche Relevanz beizumessen ist.

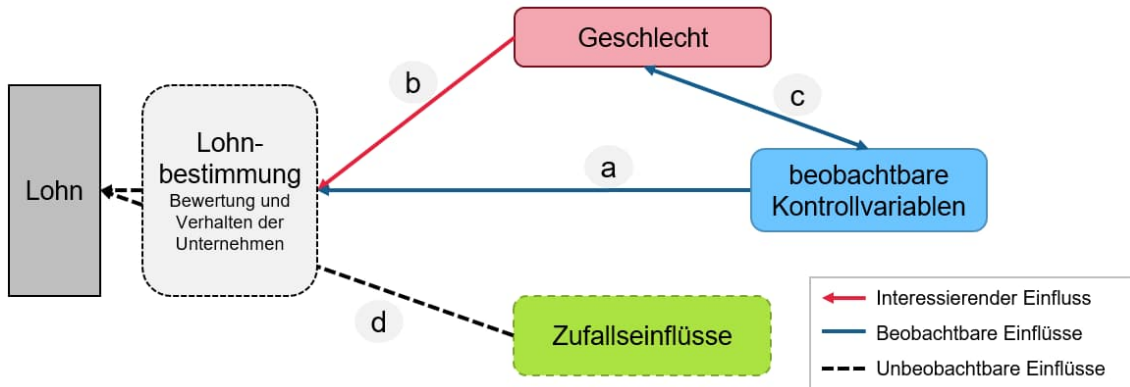
2.1.1 Keine fehlenden Kontrollvariablen

Mithilfe der einfachen schematischen Darstellung in Abbildung 1 illustrieren wir zunächst den Fall, in dem *alle relevanten* Kontrollvariablen beobachtbar sind. Für die Diskussion in diesem Kapitel ist es sinnvoll, zwischen der Lohnbestimmung und dem Lohn zu unterscheiden. Die Lohnbestimmung beschreibt den *Prozess* der Lohnfestsetzung, in welche die Bewertung der Arbeitsleistung der Beschäftigten, und das (potenziell diskriminierende) Verhalten der Unternehmen gegenüber den Beschäftigten einfließen. Der Prozess der Lohnbestimmung ist nicht beobachtbar. Demgegenüber stellt der Lohn das beobachtbare *Ergebnis* der Lohnfestsetzung dar.

In Abbildung 1 bilden die beobachtbaren Kontrollvariablen sämtliche Faktoren ab, die mit der Produktivität der Beschäftigten zusammenhängen (Pfeil a). Dazu gehören unter anderem Merkmale wie Ausbildung, Alter, Dienstjahre und berufliche Stellung. Zweitens beeinflusst das Geschlecht die Lohnbestimmung, wenn Arbeitgebende vergleichbare Frauen und Männer ungleich behandeln (Pfeil b). Diese Lohndiskriminierung führt dazu, dass Frauen und Männer nicht gleich viel verdienen, obwohl sie gleich produktiv sind. Zentral ist, dass die beobachtbaren Merkmale und das Geschlecht in aller Regel korreliert sind. Anders ausgedrückt unterscheiden sich Frauen und Männer in Bezug auf diese Merkmale (Pfeil c). Bestünde kein Zusammenhang zwischen Geschlecht und Kontrollvariablen, wären letztere für die Bestimmung des Geschlechter-Effekts gar nicht relevant. Weitere Einflüsse auf den Lohn sind hier zufällig und stehen mit dem Geschlecht nicht in Zusammenhang (Pfeil d). Dabei kann es sich beispielsweise um Messfehler oder Unschärfen in der Datengrundlage handeln. Zusammenfassend lässt sich sagen, dass in Abbildung 1 der Geschlechter-Effekt und damit das Ausmass der Lohndiskriminierung korrekt identifiziert wird, weil alle relevanten, lohnbestimmenden Kontrollvariablen beobachtbar sind.

⁵ Bei der juristischen Betrachtung stützen wir uns auf Aussagen des Auftraggebers. Wir haben im Rahmen dieses Mandats keine eigenen juristischen Abklärungen vorgenommen.

Abbildung 1: Schematische Darstellung der Einflusskanäle ohne fehlende Variablen



Quelle: Eigene Darstellung.

Um das Verständnis für die Identifikation eines Geschlechtereffekts anhand von Lohndaten zu schärfen, machen wir nachfolgend ein fiktives Zahlenbeispiel, welches der Idee von Kaiser & Möhr (2021, S. 28ff.) folgt. Das Ziel des Zahlenbeispiels ist es, aufzuzeigen, wie die Kontrollvariablen es ermöglichen, den Geschlechtereffekt zu identifizieren, ohne dabei ein statistisches Modell unterstellen zu müssen.⁶ Alle notwendigen Rechenoperationen können mit einem Taschenrechner erfolgen. Im nachfolgenden Zahlenbeispiel gibt es zwei Kontrollvariablen X1 und X2, die je die zwei Ausprägungen «hoch» und «tief» haben. Tabelle 1 zeigt die Beschäftigungsanteile und Durchschnittslöhne für Männer und Frauen in jeder Merkmalskombination. Bei den Männern kommt «hoch» etwas häufiger vor als bei den Frauen. Die Ausprägung «hoch» hat einen positiven Effekt auf den Lohn, wobei der Effekt bei X1 deutlich stärker ist als jener von X2. Es sei angenommen, dass die Lohnfunktion bei den Männern *diskriminierungsfrei* ist und Produktivitätsunterschiede zwischen den Merkmalsausprägungen widerspiegelt. Es wird unterstellt, dass Frauen und Männer bei gleichen Merkmalen gleich produktiv sind und ohne Diskriminierung den gleichen Lohn erhalten würden; folglich gibt es keine weiteren relevanten Merkmale neben X1 und X2. Die letzte Spalte zeigt, dass Arbeitgeber Frauen mit tiefem X2 *diskriminieren* und ihnen im Schnitt 10 Geldeinheiten weniger bezahlen, während Frauen mit hohem X2 gleich entlohnt werden wie Männer.

⁶ «Ohne statistisches Modell» bedeutet im technischen Sinne, dass die Identifikation hier nichtparametrisch erfolgt.

Tabelle 1: Zahlenbeispiel – Lohntabelle mit X1 und X2

X1	X2	Männer		Frauen		Bedingte Lohndifferenz (Lohndiskriminierung)
		Anteil	Lohn	Anteil	Lohn	
tief	hoch	30%	50	25%	50	0
tief	tief	10%	40	25%	30	-10
hoch	hoch	40%	100	25%	100	0
hoch	tief	20%	90	25%	80	-10
alle		100%		100%		

Quelle: Das Zahlenbeispiel ist angelehnt an Kaiser & Möhr (2021, S. 28ff.).

In Tabelle 2 sind die Durchschnittslöhne von Männern (= 77) und Frauen (= 65) ersichtlich. So erhält man im obigen Beispiel eine *gesamte* Lohndifferenz von -12 Geldeinheiten (= 65 – 77), siehe unten. Die Einfachheit des Zahlenbeispiels erlaubt es, den durchschnittlichen Geschlechtereffekt (= Lohndiskriminierung) ohne statistisches Modell und mit einfachen Rechenoperationen zu berechnen: Der «kontrafaktischen Durchschnittslohn» der Frauen erhält man, indem man die Männerlöhne mit den Frauenanteilen multipliziert ($0.25 \cdot 50 + 0.25 \cdot 40 + 0.25 \cdot 100 + 0.25 \cdot 90 = 70$). Das heisst, die Frauen würden im Durchschnitt 70 Geldeinheiten verdienen, wenn keine Diskriminierung bestünde. Der durchschnittliche Geschlechtereffekt ergibt sich dann aus der Differenz zum tatsächlichen Frauenlohn ($65 - 70 = -5$). Alternativ kann man auch nur die bedingte Lohndifferenz in der letzten Spalte betrachten: 50% der Frauen werden nicht diskriminiert, 50% der Frauen werden um -10 Geldeinheiten diskriminiert, sodass der *durchschnittliche* Geschlechtereffekt (d.h. über alle Frauen hinweg) -5 Geldeinheiten beträgt.

Tabelle 2: Zahlenbeispiel – Durchschnittslöhne und Lohndifferenz

Durchschnittslöhne		Lohndifferenz	absolut
Männer	77	Gesamte Lohndifferenz	-12
Frauen	65	Geschlechtereffekt (= Lohndiskriminierung im statistischen Sinne)	-5
Kontrafaktisch	70	Erklärte Lohndifferenz (Selektionseffekt)	-7

Quelle: Das Zahlenbeispiel ist angelehnt an Kaiser & Möhr (2021, S. 28ff.).

Die unterschiedliche Verteilung von Frauen und Männern hinsichtlich der Kontrollvariablen «erklären» somit im Durchschnitt eine Lohndifferenz von -7 Geldeinheiten; diese Differenz bezeichnen wir auch als «Selektionseffekt». Die Berücksichtigung der beiden Kontrollvariablen X1 und X2 erlaubt es somit in diesem Beispiel, den durchschnittlichen Geschlechtereffekt im statistischen Sinne korrekt zu identifizieren. Das heisst: Der gemessene Geschlechtereffekt entspricht dem Ausmass der Lohndiskriminierung im statistischen Sinne.

Aus einer juristischen Perspektive wird Lohndiskriminierung möglicherweise anders beurteilt, weil massgebend ist, welche Löhne benachteiligte Frauen vor Gericht einklagen könnten.⁷ In Tabelle 3 nehmen wir beispielhaft an, dass Arbeitgeber Lohnunterschiede nach Merkmal X2 nicht rechtfertigen dürfen, nach Merkmal X1 hingegen schon. In der Folge können Frauen den Unterschied zwischen ihrem Lohn und dem Vergleichslohn der Männer einklagen. Beim Vergleichslohn nehmen wir an, dass es sich um den höchsten Lohn der Männer in der jeweiligen Gruppe von X1 handelt.⁸ Eine Frau mit tiefem X1 und tiefem X2 kann bei einem Lohn von 30 vor Gericht somit 20 zusätzliche Einheiten geltend machen. Die durchschnittliche Lohndiskriminierung im juristischen Sinne beträgt somit 10 Einheiten und ist damit doppelt so gross wie die durchschnittliche Lohndiskriminierung im statistischen Sinne.

Tabelle 3: Zahlenbeispiel – Lohntabelle mit X1 und X2, juristische Betrachtung

X1	X2	Männer		Frauen		Einklagbare Lohndifferenz
		Anteil	Vergleichs- lohn	Anteil	Lohn	
tief	hoch	40%	50	25%	50	0
tief	tief			25%	30	-20
hoch	hoch	60%	100	25%	100	0
hoch	tief			25%	80	-20
alle		100%		100%		

Quelle: Eigene Darstellung.

2.1.2 Fehlende Kontrollvariablen: Omitted Variable Bias

In Abbildung 2 wird die Annahme gelockert, dass alle relevanten Kontrollvariablen beobachtbar sind. Stattdessen gehen wir nun davon aus, dass unbeobachtete Kontrollvariablen existieren, die sowohl den Lohn beeinflussen (Pfeil e) und gleichzeitig auch mit dem Geschlecht korreliert sind (Pfeil f).⁹ Dabei kann es sich beispielsweise um berufliche Qualifikationen oder Erfahrungen handeln, die in den Daten nicht oder nur unzureichend verfügbar sind. In der Situation von Abbildung 2 lässt sich der Geschlechter-Effekt (Pfeil b) nicht mehr korrekt identifizieren und man spricht allgemein von einem «Omitted Variable Bias» (OVB), was sich auf Deutsch mit «Verzer-

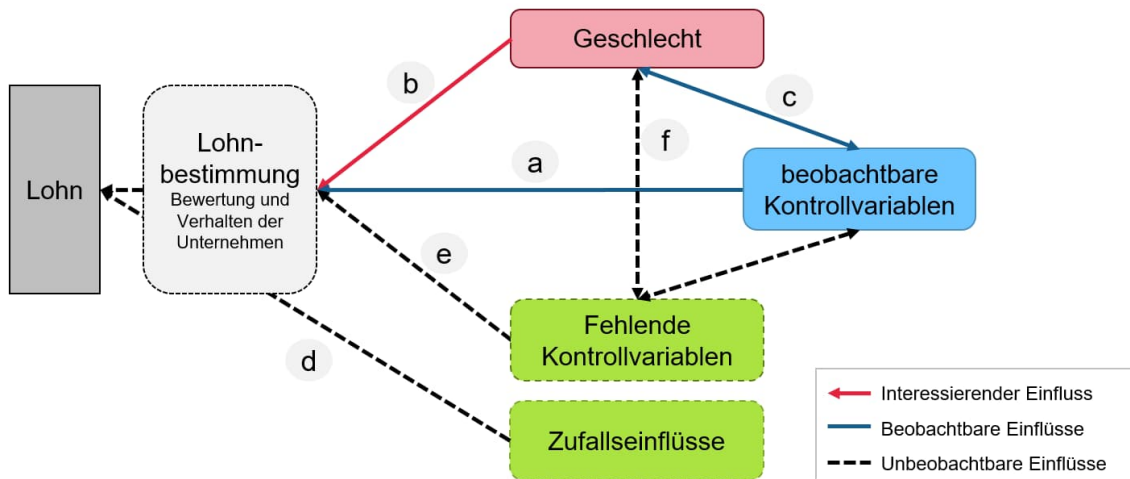
⁷ Dieser Abschnitt stützt sich auf die Aussagen des Auftraggebers. Wir haben keine eigenen juristischen Abklärungen vorgenommen.

⁸ Der Rechtsdienst des EBG liefert hierzu folgende Einschätzung: «Haben zwei Arbeitnehmende unterschiedlichen Geschlechts vergleichbare Positionen bzw. gleiche Arbeit, wird im Falle einer Lohndifferenz vermutet, dass eine geschlechtsbezogene Diskriminierung vorliegt. Es genügt hier eine Vergleichsperson des andern Geschlechts, ohne dass im Einzelfall noch die Lohnstruktur des Unternehmens insgesamt und die Einordnung der fraglichen Positionen in derselben geprüft werden müsste. Diese Aspekte könnten aber allenfalls von Relevanz sein für die Bestimmung der Höhe des diskriminierungsfreien Lohnes und eine fundiertere Analyse, wenn die Gleichwertigkeit in Zweifel gezogen wird bzw. allfällig angeführte Rechtfertigungsgründe zu gewichten sind.» (Antwort vom 26.5.2023)

⁹ Die unbeobachteten Kontrollvariablen können auch mit den beobachteten Kontrollvariablen korreliert sein, was hier aber nicht entscheidend ist.

«... aufgrund fehlender Variablen» übersetzen lässt. Zu betonen ist, dass der OVB nur dann auftritt, wenn beide Kanäle, Pfeil e *und* Pfeil f, existieren. Relevant ist zudem, dass der OVB grundsätzlich bei allen statistischen Verfahren besteht, die auf derselben Datengrundlage basieren. Das Problem des OVB, so wie es hier definiert ist, liegt somit in den Daten und nicht in der Analyse-methode begründet.

Abbildung 2: Schematische Darstellung der Einflusskanäle mit fehlenden Variablen



Quelle: Eigene Darstellung.

Wie gross der OVB ist, wenn eine Kontrollvariable fehlt, ist immer eine empirische Frage. Allgemein aber gilt, dass der OVB umso grösser ist: 1) je grösser der Einfluss der Kontrollvariable auf den Lohn ist; 2) je stärker sich Frauen und Männer in Bezug auf die Kontrollvariable unterscheiden und 3) je weniger die fehlende Kontrollvariable mit den beobachtbaren Kontrollvariablen korreliert.

Um den Einfluss des OVB zu illustrieren, greifen wir wiederum auf das obige Zahlenbeispiel zurück. Dazu nehmen wir nun an, dass das Merkmal X2 nicht verfügbar ist. Somit «beobachten» wir nur die folgende reduzierte Lohntabelle mit den Durchschnittslöhnen und Beschäftigungsanteilen in den Ausprägungen von Merkmal X1 (siehe Tabelle 4).

Tabelle 4: Zahlenbeispiel – Lohntabelle ohne X2

X1	Männer		Frauen		Bedingte Lohndifferenz
	Anteil	Lohn	Anteil	Lohn	
tief	40%	47.50	50%	40.00	-7.50
hoch	60%	96.67	50%	90.00	-6.67
alle	100%		100%		

Quelle: Eigene Darstellung.

Den kontrafaktischen Durchschnittslohn der Frauen erhalten wir wiederum, indem die Männerlöhne mit den Frauenanteilen multipliziert werden ($0.5 \cdot 47.5 + 0.5 \cdot 96.67 = 72.08$), siehe Tabelle 5. Anschliessend bilden wir den geschätzten Geschlechtereffekt anhand der Differenz zum tatsächlichen Durchschnittslohn der Frauen ($65 - 72.08 = -7.08$).¹⁰

Tabelle 5: Zahlenbeispiel – Durchschnittslöhne und Lohndifferenz ohne X2

Durchschnittslöhne		Lohndifferenz	abs.
Männer	77.00	Gesamte Lohndifferenz	-12.00
Frauen	65.00	Geschlechtereffekt (= Lohndiskriminierung)	-7.08
kontrafaktisch	72.08	Erklärte Lohndifferenz	-4.92

Quelle: Eigene Darstellung.

Wie wir sehen, führt das Fehlen von X2 zu einem OVB von $-2.08 (= -7.08 - (-5))$. Mit anderen Worten wird die Stärke des wahren durchschnittlichen Geschlechtereffekts um 2.08 Einheiten überschätzt. Dies liegt erstens daran, dass das Merkmal X2 (von «tief» zu «hoch») einen positiven Effekt auf den Lohn hat. Der zweite Grund ist, dass das Merkmal X2 (von «tief» zu «hoch») negativ mit der Wahrscheinlichkeit, eine Frau zu sein, korreliert ist, selbst innerhalb der Ausprägungen von X1.

Gleichermassen lassen sich auch alternative Zahlenbeispiele generieren, bei denen beim Weglassen von X2 die Stärke des Geschlechtereffekts *unterschätzt* wird, oder bei denen kein OVB auftritt.

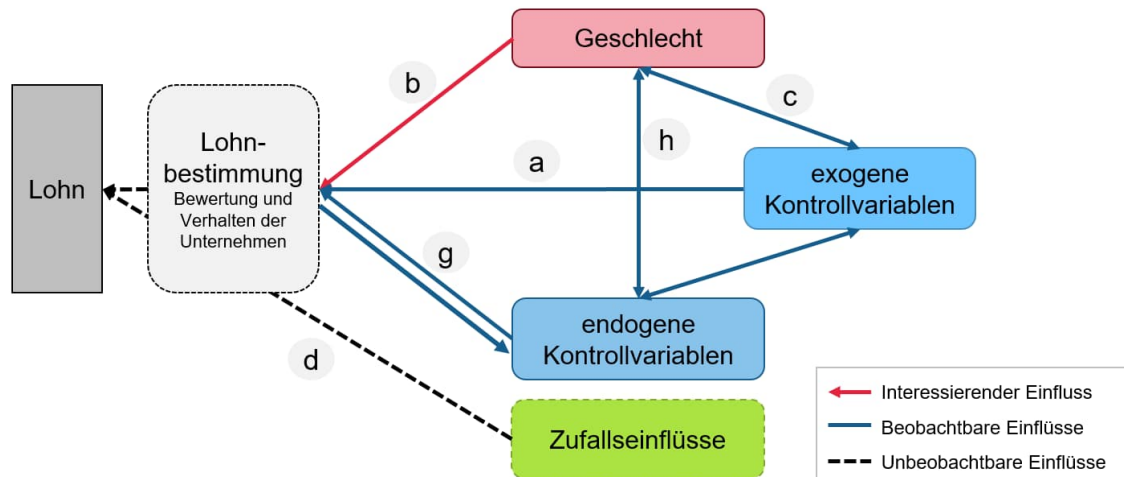
2.1.3 Endogene Kontrollvariablen («Bad Controls»)

In den vorangegangenen Abschnitten sind wir implizit davon ausgegangen, dass alle Kontrollvariablen valide und damit *exogen* sind. Exogenität heisst im vorliegenden Kontext, dass der Lohnbestimmungsprozess und das damit verbundene Verhalten der Unternehmen die Ausprägungen der Kontrollvariablen nicht beeinflusst. Nicht alle potenziellen Kontrollvariablen erfüllen diese Eigenschaft. Wenn eine Kontrollvariable diese Eigenschaft verletzt, ist sie *endogen*, was in der Literatur zur kausalen Inferenz auch als «bad control» bezeichnet wird. In Abbildung 3 ist eine Situation mit exogenen und endogenen Kontrollvariablen dargestellt. Das Problem von endogenen Kontrollvariablen in Zusammenhang mit Geschlechterdiskriminierung besteht dann, wenn das Verhalten des Unternehmens auf eine Kontrollvariable einwirkt (Pfeil g) und die Kontrollvariable gleichzeitig mit dem Geschlecht korreliert ist (Pfeil h). Zur Illustration machen wir ein plakatives Beispiel, wo die Endogenität offensichtlich ist: Eine binäre Kontrollvariable (Dummyvariable), die angibt, ob der Lohn über dem Medianlohn im Unternehmen liegt oder nicht. Der Lohnbestimmungsprozess beeinflusst die Werte dieser Variable per Definition unmittelbar (Pfeil g) und gleichzeitig besteht eine positive Korrelation mit dem Geschlecht (Pfeil h). Folglich «absorbiert» die endogene Kontrollvariable einen Teil des zu identifizierenden Geschlechter-Effekts

¹⁰ Alternativ ergibt sich der geschätzte Geschlechtereffekt durch die Gewichtung der bedingten Lohndifferenzen mit den korrespondierenden Frauenanteilen: $0.5 \cdot (-7.5) + 0.5 \cdot (-6.67) = -7.08$.

(Pfeil b). In diesem Beispiel ist der verzerrte Geschlechter-Effekt also kleiner als der korrekt identifizierte Geschlechter-Effekt. Mit anderen Worten wird das Ausmass der Lohndiskriminierung in diesem Modell unterschätzt, wenn die endogene Kontrollvariable verwendet wird.

Abbildung 3: Schematische Darstellung der Einflusskanäle ohne fehlende Variablen



Quelle: Eigene Darstellung.

Wie lässt sich die Exogenität der Kontrollvariablen beurteilen? Grundsätzlich hat die Beurteilung anhand von inhaltlichen und fachlichen Überlegungen zu erfolgen, da der Lohnbestimmungsprozess und damit potenziell diskriminierendes Verhalten nicht direkt beobachtbar sind. Die Plausibilität der Exogenitätseigenschaft kann mit einem «Gedankenexperiment» eingeschätzt werden. Angenommen es existieren zwei hypothetische Zustände: Ein Zustand ohne Geschlechterdiskriminierung und ein Zustand mit Geschlechterdiskriminierung bei der Lohnfestsetzung, während alle anderen Einflüsse gleich sind. Würden sich die realisierten Werte einer bestimmten Kontrollvariable zwischen den beiden Zuständen unterscheiden? Lautet die Antwort mit grosser Wahrscheinlichkeit nein, handelt es sich um eine exogene Kontrollvariable. Gute Beispiele für exogene Kontrollvariablen sind somit Personenmerkmale, die nach dem Eintritt ins Unternehmen nicht mehr ändern, wie etwa Geburtsjahr, Beruf oder Bildungsabschluss. Naheliegende Beispiele für endogene Kontrollvariablen sind Vergütungsmerkmale, wie beispielsweise die Entrichtung von Bonuszahlungen (Dummyvariable). Im diskriminierungsfreien Zustand ist es wahrscheinlich, dass mehr Frauen Bonuszahlungen erhalten als im diskriminierenden Zustand. Häufig ist die Antwort auf die Frage im Gedankenexperiment jedoch nicht eindeutig. In diesen Fällen ist eine Güterabwägung vorzunehmen. Auf der einen Seite steht der «Nutzen» der Variable, wenn sie exogene Aspekte der Produktivität der Beschäftigten abbildet. Auf der anderen Seite stehen die «Kosten», die dadurch entstehen, dass die Variable gleichzeitig auch Aspekte eines diskriminierenden Verhaltens widerspiegelt. Ob die Inklusion der Variable die Verzerrung des Geschlechter-Effekts reduziert oder erhöht, hängt davon ab, wie wichtig die exogene bzw. endogene Komponente ist.

Bei der obigen Diskussion ist zu beachten, dass sich die Analyse hier auf die Definition «Lohndiskriminierung im engeren Sinne» bezieht. Wenn es darum ginge, auch Diskriminierungsformen miteinzubeziehen, die sich indirekt auf die Löhne auswirken, wie etwa Beförderungsdiskriminierung, verändern sich die Schlussfolgerungen bezüglich Endogenität. In diesem Fall wäre auch die Variable «Berufliche Stellung» als potenziell endogen zu beurteilen, da sich die realisierten Werte

dieser Variable in den beiden hypothetischen Zuständen mit und ohne Beförderungsdiskriminierung unterscheiden.

Um die Konsequenzen von endogenen Kontrollvariablen zu illustrieren, ziehen wir wiederum unser Zahlenbeispiel heran, siehe Tabelle 6. Wir nehmen hier an, dass das Merkmal X2 endogen ist. Das bedeutet: Im diskriminierungsfreien Zustand wären die Anteile der Frauen in den Merkmalskombinationen anders als im tatsächlich beobachteten, diskriminierenden Zustand. Konkret wäre der Frauenanteil in X2 «hoch» ohne Diskriminierung 30% anstatt 25%. Frauen hätten also im diskriminierungsfreien Zustand eine höhere Ausstattung im Merkmal X2.

Tabelle 6: Zahlenbeispiel – Lohntabelle mit endogenem X2

X1	X2	Männer		Frauen			Lohndiskriminierung
		Anteil	Lohn	Anteil im diskriminierungsfreien Zustand	Anteil im diskriminierenden Zustand	Lohn	
tief	hoch	30%	50	30%	25%	50	0
tief	tief	10%	40	20%	25%	30	10
hoch	hoch	40%	100	30%	25%	100	0
hoch	tief	20%	90	20%	25%	80	10
alle		100%		100%	100%		

Quelle: Eigene Darstellung.

Der *geschätzte* kontrafaktische Durchschnittslohn der Frauen beträgt hier 70 Geldeinheiten und der *geschätzte* Geschlechtereffekt beträgt -5 Geldeinheiten (vgl. Tabelle 2). Da X2 endogen ist, ist diese Schätzung aber verzerrt und entspricht *nicht* der Lohndiskriminierung, wie die Ergebnisse in Tabelle 7 zeigen. Der kontrafaktische Lohn im diskriminierungsfreien Zustand folgt aus der Multiplikation der Frauenanteile im diskriminierungsfreien Zustand mit den Löhnen der Männer: $0.3 \cdot 50 + 0.2 \cdot 40 + 0.3 \cdot 100 + 0.2 \cdot 90 = 71$. Der *wahre* Geschlechtereffekt bzw. die wahre Lohndiskriminierung haben somit ein Ausmass von -6 Geldeinheiten. Das heisst, die Endogenität von X2 führt dazu, dass die Grösse des Geschlechtereffekts in diesem Beispiel um 1 Einheit (= -5 - (-6)) unterschätzt wird.

Tabelle 7: Zahlenbeispiel – Durchschnittslöhne und Lohndifferenz mit endogenem X2

Durchschnittslöhne		Lohndifferenz	abs.
Männer	77	Gesamte Lohndifferenz	-12
Frauen	65		
Kontrafaktisch, geschätzt	70	geschätzter Geschlechtereffekt	-5
Kontrafaktisch, wahr	71	wahrer Geschlechtereffekt	-6

Quelle: Eigene Darstellung.

Wie lässt sich nun eine Entscheidung treffen, ob die endogene Variable X2 aus dem Modell entfernt werden soll oder nicht? Dazu kann man den OVB beim Weglassen der Variable mit dem «Endogenitätsbias» beim Inkludieren der Variable gegenüberstellen:

- *Mit X2 im Modell* (vgl. Tabelle 7): Der geschätzte Geschlechtereffekt beträgt -5. Der Bias ist 1 (= -5 – (-6)).
- *Ohne X2 im Modell* (vgl. Tabelle 5): Der geschätzte Geschlechtereffekt beträgt -7.08. Der Bias ist -1.08 (= -7.08 – (-6)).

Mit anderen Worten ist der OVB im Betrag leicht grösser als der «Endogenitätsbias», was dafür sprechen würde, in diesem fiktiven Zahlenbeispiel X2 im Modell zu belassen. In der Realität sind solche Entscheidungen schwierig, weil das Ausmass der Verzerrungen infolge von Endogenität nicht bekannt ist. Stattdessen muss mit Fachwissen und inhaltlichen Überlegungen entschieden werden, wie gross ein potenzielles Endogenitätsproblem ist.

2.2 Schätzung des Geschlechter-Effekts

2.2.1 Methodenübersicht

Die Schätzung des Geschlechter-Effekts kann im Grunde anhand verschiedener statistischer Methoden erfolgen. Dazu gehören parametrische und nichtparametrische Regressionsansätze, Propensity-Score-Methoden, Kontrollvariablen-Matching sowie kombinierte Methoden wie zum Beispiel Doubly-Robust-Regression. Alle diese Methoden besitzen die Gemeinsamkeit, dass die statistischen Einflüsse der beobachtbaren Kontrollvariablen «kontrolliert» werden und der Geschlechtereffekt (d.h. der «kausale Effekt» im ökonomischen Sinne) dadurch aus den Daten geschätzt werden kann. Die Methoden unterscheiden sich aber dadurch, dass unterschiedliche Annahmen getroffen werden, um den Geschlechtereffekt zu identifizieren. Eine Diskussion dieser Annahmen sowie der Vor- und Nachteile einzelner Methoden geht jedoch über diese Studie hinaus.

2.2.2 Umsetzung in Logib

Der mit Abstand gängigste Ansatz zur Analyse von Lohnungleichheit zwischen Frauen und Männern in der wissenschaftlichen Literatur basiert auf einem (log-)linearen Lohnregressionsmodell, das mit Ordinary Least Squares (OLS) geschätzt wird.¹¹ Auch das Standard-Analysemodell des Bundes (Logib Modul 1) für Unternehmen ab 50 Mitarbeitenden basiert auf einem log-linearen Lohnmodell. Formal lässt sich dieses wie folgt aufschreiben:

$$Y = \beta_0 + D\beta + X\beta_1 + u,$$

¹¹ Als Präzisierung ist anzufügen, dass zwei verschiedene Spezifikationen häufig verwendet werden. Die erste ist, dass für beide Geschlechter separate Regressionen geschätzt werden; dabei spricht man von der Oaxaca-Blinder-Zerlegung (Blinder 1973; Oaxaca 1973). Dieses Modell ist aber mathematisch äquivalent zu folgender gemeinsamer Lohnregression: $Y = \beta_0 + D\beta + X\beta_1 + DX\beta_2 + u$, wobei Y der logarithmierte Lohn, D der Geschlechter-Indikator, X der Vektor der Kontrollvariablen und u der Störterm darstellen. Die zweite Spezifikation ist leicht restriktiver, $Y = \beta_0 + D\beta + X\beta_1 + u$, das heisst, der Geschlechter-Effekt wird über eine Dummyvariable gemessen.

wobei Y der logarithmierte standardisierte Lohn, D der Geschlechter-Indikator, X der Vektor der Kontrollvariablen und u der Störterm darstellen. Der Koeffizient β_0 ist eine Konstante, β ist der interessierende Geschlechter-Koeffizient und β_1 ist der Vektor der Koeffizienten der Kontrollvariablen. Die Kontrollvariablen in Logib Modul 1 umfassen Ausbildungsjahre, potenzielle Erfahrung, Dienstjahre, Anforderungsniveau und berufliche Stellung und werden im vorliegenden Kontext auch als «objektive, lohnrelevante und nichtdiskriminierende Faktoren» bezeichnet.¹² Die Standardfehler werden mit dem «konventionellen» Schätzverfahren berechnet, welches unterstellt, dass der Störterm homoskedastisch ist.

Nachdem das Logib-Modell geschätzt ist, wird die Prüfung der Lohnungleichheit wie folgt durchgeführt. Zuerst wird mit einem zweiseitigen t-Test geprüft, ob der Geschlechter-Koeffizient signifikant von null verschieden ist. Wenn ja, wird ein zweiter, einseitiger t-Test durchgeführt, ob der Geschlechter-Koeffizient im Betrag signifikant grösser ist als 0.05. Auf beiden Stufen wird ein Signifikanzniveau von 5% verwendet; dies ist der gängigste Wert in der wissenschaftlichen Literatur. Wird die Nullhypothese des zweiten t-Tests verworfen, gilt der Logib-Test als nicht bestanden. Die Lohnungleichheitsanalyse beruht somit auf einem sequenziellen, zweistufigen Testverfahren.¹³ Dies entspricht nicht der gängigen Praxis in wissenschaftlichen Studien, wo in der Regel nur die Nullhypothese eines Nulleffekts getestet wird. Der Grund für das zweistufige Vorgehen mit der Toleranzschwelle ist aus unserer Sicht, dass das Feststellen einer Lohnungleichheit mit hoher Gewissheit erfolgen soll; anders ausgedrückt wird der Vermeidung von falsch-positiven Ergebnissen eine hohe Bedeutung beigemessen.

Für eine detaillierte Beschreibung von Modell und Methode wird auf EBG (2021, Kapitel 3.2) verwiesen.

3 Empirisches Vorgehen und Daten

3.1 Auswirkungen von Modellmodifikationen

In dieser Studie liegt der Fokus auf der Frage, welchen Einfluss einzelne Kontrollvariablen auf den geschätzten Geschlechter-Effekt im Logib-Modell sowie auf das Ergebnis des t-Tests ausüben. Dabei sind mehrere Experimente von Interesse: Wie verändern sich die Ergebnisse der Lohnungleichheitsanalyse mit Logib, wenn ...

- ... bestehende Kontrollvariablen aus dem Modell entfernt werden?
- ... bestehende Kontrollvariablen inhaltlich modifiziert werden?
- ... zusätzliche Kontrollvariablen in das Modell inkludiert werden?

¹² Vgl. dazu die Konformitätserklärung des Standard-Analyse-Tools Logib (Seite 2). Link: [https://www.ebg.admin.ch/dam/ebg/de/dokumente/lohnungleichheit/logib/konformitaetserklaerung_logib.pdf.download.pdf/Konformit%C3%A4tserkl%C3%A4rung_Standard-Analyse-Tool%20\(Logib\)_V2020.1.pdf](https://www.ebg.admin.ch/dam/ebg/de/dokumente/lohnungleichheit/logib/konformitaetserklaerung_logib.pdf.download.pdf/Konformit%C3%A4tserkl%C3%A4rung_Standard-Analyse-Tool%20(Logib)_V2020.1.pdf)

¹³ Aus methodischer Sicht ist relevant, dass ein p-Wert von unter 0.05 im zweiten Test automatisch bedeutet, dass auch der p-Wert des ersten Tests unter 0.05 liegt. Folglich hat der erste Test keine statistischen Implikationen für den zweiten Test. Statistisch gesprochen besteht kein «Multiples Testproblem».

Das erste Gedankenexperiment liefert Hinweise, wie «wichtig» die einzelnen Kontrollvariablen sind, die zurzeit verwendet werden. Das zweite und dritte Gedankenexperiment liefern Grundlagen zur Frage, wie sich allfällige Weiterentwicklungen des Modells auf die Ergebnisse von Lohngleichheitsanalysen auswirken würden. Dabei sind zwei separate Effekte von zusätzlichen Kontrollvariablen zu erwähnen. Erstens können sich die geschätzten Geschlechter-Effekte verändern, indem ein allfälliger OVB reduziert wird. Zweitens führt die Inklusion von zusätzlichen, lohnbestimmenden Variablen auch dazu, dass die Residualvarianz des Modells sinkt und die Mächtigkeit der t-Tests dadurch steigen kann.¹⁴ Folglich könnten zusätzliche Kontrollvariablen dazu beitragen, dass die Lohngleichheitsanalysen griffiger werden, selbst wenn sie keine Verzerrungen (OVB) beseitigen.

Anhand geeigneter Lohndaten auf Unternehmensebene lassen sich diese Gedankenexperimente relativ einfach durchführen, indem die Ergebnisse der «modifizierten» Modelle mit jenen des Standardmodells verglichen werden. Beim Vergleich zwischen den Modellen sind insbesondere folgende Kennzahlen von Interesse:

- Geschätzter Geschlechter-Koeffizient
- Veränderung des Geschlechter-Koeffizienten
- Präzision des Geschlechter-Koeffizienten: Standardfehler
- p-Werte und Signifikanz bei den beiden t-Tests
 - Häufigkeit der Fälle, wo das Testergebnis signifikant wird
 - Häufigkeit der Fälle, wo das Testergebnis insignifikant wird
- Modellgüte (z.B. adjusted-R²)

Die oben beschriebenen Analysen werden für eine Vielzahl der Unternehmen durchgeführt. Anschliessend werden die Ergebnisse auf die Gesamtwirtschaft und den privaten und öffentlichen Sektor aggregiert, indem die Mittelwerte der Kennzahlen dargestellt werden. Die Ergebnisse zeigen dann beispielsweise, dass bei der Hinzunahme der erklärenden Variable Z in Logib sich der Geschlechter-Koeffizient im privaten Sektor im Durchschnitt um x% verändert und dass sich die Häufigkeit signifikanter Ergebnisse im zweiten t-Test von a% auf b% ändert.

3.2 Abschätzung von Omitted Variable Bias

Unter bestimmten statistischen Annahmen zu den fehlenden relevanten Variablen ist es möglich, das Ausmass des OVB mit sogenannten Schranken («bounds») einzugrenzen. Dazu nutzen wir den Ansatz von Oster (2019). Diese Analysen vermitteln ein Gefühl dafür, wie gross das potenzielle Ausmass des OVB ist.

Die Grundidee des Ansatzes ist, dass der gesamte OVB in zwei Teile gegliedert werden kann. Der erste Teil ist die Verzerrung, welche durch die vorhandenen Kontrollvariablen erfasst und korrigiert wird. Der zweite Teil ist die Verzerrung, welche darüber hinaus noch durch unbeobachtete

¹⁴ Technisch gesprochen steigt die t-Statistik dann, wenn der Effekt der gesunkenen Residualvarianz den Effekt der abnehmenden Freiheitsgrade übersteigt. Die t-Statistik des Geschlechter-Koeffizienten kann also steigen, wenn lohnrelevante Faktoren ins Modell inkludiert werden, die nicht mit dem Geschlecht korreliert sind. Bei der Hinzunahme lohnirrelevanter Variablen verschlechtert sich die t-Statistik des Geschlechter-Koeffizienten hingegen, weil die Freiheitsgrade sinken.

Faktoren besteht. Hängen diese beiden Bestandteile zusammen, lässt sich von der Grösse des kontrollierbaren OVB auf die Grösse des unkontrollierbaren OVB schliessen. Hierzu werden zwei Annahmen benötigt. Erstens benötigt wird eine Annahme über die relative Stärke und Richtung der beiden OVB-Bestandteile, den Parameter δ . In der Literatur hat sich als Standard durchgesetzt, als Obergrenze für diesen Parameter den Wert 1 anzunehmen. Dies entspricht der Annahme, dass die Verzerrung durch beide OVB-Bestandteile gleich stark und gleichläufig ist. Ein δ von zwei hingegen würde der Annahme entsprechen, dass die Verzerrung durch unbeobachtete Faktoren zwei Mal so gross ist wie diese durch beobachtbare Faktoren. Die Annahme der gleichläufigen Verzerrung unterstellt, dass die Richtung der Verzerrung jeweils identisch ist. Zweitens definiert werden muss der Parameter R_{max} , welcher beschreibt bis zu welchem Grad ein vollständiges Modell mit allen beobachtbaren und unbeobachtbaren Variablen in der Lage wäre, bestehende Lohnunterschiede zu erklären. Entsprechend den Empfehlungen von Oster (2019) und der gängigen Praxis in der Literatur setzten wir diesen Wert auf den 1,3-fachen Wert des R^2 s der Lohnregression mit den Logib-Kontrollvariablen. Sollte dieses Produkt den Wert von 1 überschreiten, setzten wir den Wert auf 1, da dies die natürliche Obergrenze des theoretisch möglichen Erklärungsgehaltes ist.¹⁵

3.3 Datengrundlage

Für die empirische Umsetzung der obigen Analysen greifen wir auf Daten der Schweizerischen Lohnstrukturerhebung LSE 2020 zurück. Die Grundgesamtheit der LSE besteht aus allen Unternehmen des sekundären und tertiären Wirtschaftssektors mit mindestens drei Beschäftigten. Die Stichprobe umfasst rund 47'000 Unternehmen.¹⁶ Im vorliegenden Kontext ist hervorzuheben, dass Unternehmen ab 50 Beschäftigten entweder Daten zu allen Beschäftigten (Vollerhebung) oder nur für jeden dritten Beschäftigten (Teilerhebung) übermitteln können.¹⁷ Das bedeutet, dass für ein Unternehmen mit 51 Beschäftigten allenfalls nur 17 Lohnangaben verfügbar sind.

Die Datenaufbereitung erfolgt in drei Schritten. In einem ersten Schritt werden alle benötigten Variablen selektiert bzw. generiert. Tabelle 8 liefert einen Überblick zu den relevanten Variablen für das «Standardmodell» von Logib Modul 1. Zu erwähnen ist, dass die Variable «Anforderungsniveau» in Logib firmenspezifisch codiert wird und sich von der Variable «Kompetenzniveau» in der LSE unterscheidet. Letztere stellt eine Aggregation der Berufsgruppe nach CH-ISCO-19 dar. Tabelle 9 zeigt welche zusätzlichen Variablen aus der LSE im Rahmen der verschiedenen Modifikationen verwendet werden.

¹⁵ Der Ansatz von Oster und seine Annahme ist nicht ohne Kritik geblieben (siehe z.B. Cinelli & Hazlet (2020)). Dennoch hat er in der empirischen Wirtschaftsforschung eine breite Verwendung erreicht.

¹⁶ Siehe Steckbrief der LSE: <https://www.bfs.admin.ch/bfs/de/home/statistiken/arbeit-erwerb/erhebungen/lse.assetdetail.23124233.html>.

¹⁷ Siehe Seite 2 auf dem Fragebogen der LSE: <https://www.bfs.admin.ch/bfs/de/home/statistiken/arbeit-erwerb/erhebungen/lse.assetdetail.23889608.html>.

Tabelle 8: Relevante Variablen für Logib Modul 1

Variable in Logib	Name in R-Paket «logib»	Variablen in der LSE
Geschlecht (Mann=0, Frau = 1)	sex	geschle
Anzahl Bildungsjahre	years_of_training	ausbild (Kodierung gemäss EBG 2021, Kap. 3.2.2)
Potenzielle Berufserfahrung in Jahren	years_of_earning	max(alter – years_of_training – 6, 0)
Anzahl Dienstjahre	years_of_service	dienstja
Betriebliches Anforderungsniveau, firmenspezifisch codiert (4 Stufen)	skill_level	va_ps07, abgeleitet aus der Berufsgruppe nach CH-ISCO-19
Berufliche Stellung (5 Stufen)	professional_position	berufst
Standardisierter Bruttomonatslohn, logarithmiert	log(standardized_salary)	log(mbls)

Quelle: Lohnstrukturerhebung 2020.

Tabelle 9: Zusätzlich relevante Variablen für Modifikationen

Name	Variable in der LSE	Beschreibung
Bildungskategorien	ausbild	Acht Kategorien für die höchste abgeschlossene Ausbildung
Berufskategorien	isco19_2	Berufskategorien nach CH-ISCO-19-Klassifizierung. Einsteller: 10 Kategorien. Zweisteller: 47 Kategorien (effektiv)
Beschäftigungsgrad	ibgr	Fünf Kategorien: Pensum von <50%, 50-69%, 70-79%, 80-89%, >=90%
Stundenlohn	fs04	Bezahlung auf Stundenlohnbasis: Ja / Nein
Zulage	zulagen	Bezahlung von Zulagen (z.B. Schicht-, Sonntags- und Nachtarbeit): Ja / Nein
Zivilstand	zivsta	Drei Kategorien: Ledig, verheiratet, Anderer.
Nationalität oder Aufenthaltsbewilligung	natkat	Sechs Kategorien: Schweizer/innen, Kurzaufenthalter/innen, Aufenthaltler/innen, Niedergelassene, Grenzgänger/innen, Anderer.

Bemerkungen: Kategoriale Variablen werden jeweils in Dummy-Variablen für die einzelnen Kategorien umgewandelt.
Quelle: Lohnstrukturerhebung 2020.

In einem zweiten Schritt werden unplausible oder unvollständige Beobachtungen ausgeschlossen. Erstens werden Beobachtungen entfernt, bei denen die für das Logib-Modell relevanten Variablen fehlenden Werte enthalten. Dies ist insbesondere beim Kompetenzniveau und in geringerem Masse beim Ausbildungsabschluss der Fall. Zweitens wird eine Reihe weiterer Filter gesetzt,

um Beobachtungen mit extremen oder unplausiblen Werten auszuschliessen.¹⁸ Diese Filter entsprechen jenen aus der BFS-Studie zur Analyse der gesamtwirtschaftlichen Lohnunterschiede (vgl. Seite 15 in Kaiser und Möhr 2023). Insgesamt werden durch die Filterung rund 22% der Beobachtungen der LSE 2020 entfernt. Nähere Angaben zur Filterung sind im Anhang aufgeführt.

In einem dritten Schritt behalten wir alle Unternehmen, bei welchen wir über mindestens 50 *gültige* Beobachtungen (Anzahl Beschäftigte) verfügen, weil Logib Modul 1 nur für Lohndatensätze mit mindestens 50 gültigen Beobachtungen zur Anwendung gelangt. Dabei handelt es sich somit nicht um alle Unternehmen ab 50 Beschäftigten, sondern um eine Auswahl davon. Entsprechend sind keine Aussagen für die Grundgesamtheit *aller* Unternehmen ab 50 Beschäftigten möglich. Im Weiteren verzichten wir auf eine Gewichtung der Ergebnisse mit der Anzahl der Mitarbeitenden. Das heisst, kleine und grosse Unternehmen werden gleich gewichtet.

4 Ergebnisse

4.1 Verteilung der Geschlechtereffekte im Standardmodell

Bevor wir die Modifikationen vornehmen, beschreiben wir zunächst die Ergebnisse des Logib-Standardmodells. Dieses dient als Referenzpunkt für alle folgenden Modifikationen. Dazu schätzen wir zunächst das Logib-Standardmodell einzeln für jede Firma und führen das zweistufige Testverfahren durch, um die Signifikanz des geschlechtsspezifischen Effekts zu prüfen. Die daraus gewonnenen Masszahlen aggregieren wir anschliessend über alle Firmen hinweg. So beträgt der durchschnittliche geschätzte Geschlechtereffekt ($\hat{\beta}$) über alle Firmen hinweg -0,054 mit den Standardkontrollen des Logib-Modells. Dies bedeutet, dass nach Bereinigung um die Kontrollvariablen Frauen im Durchschnitt der Firmen etwa 5,4% weniger verdienen als Männer innerhalb desselben Unternehmens. In einem alternativen Modell, das nicht für die mit den Kontrollvariablen erfassten Personenunterschiede kontrolliert, beträgt der Unterschied 12,9% auf Unternehmensebene. Dies gibt einen ersten Hinweis auf die wichtige Rolle von Kontrollvariablen für die Betrachtung von Lohnunterschieden.

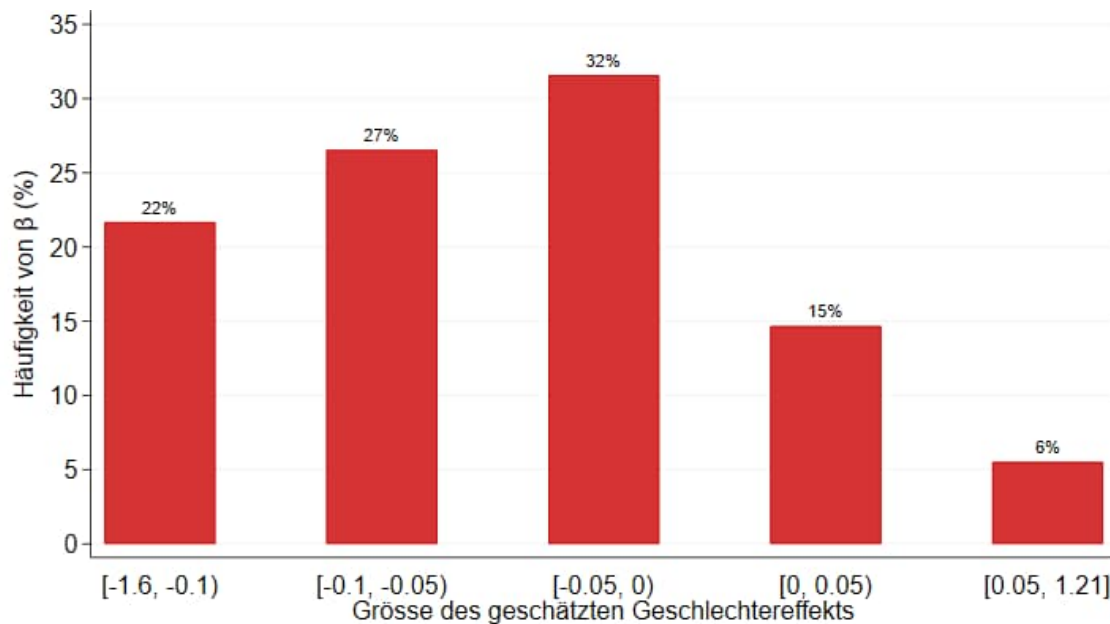
An dieser Stelle ist es angezeigt, die obigen Schätzung auf *Unternehmensebene* mit der Lohndifferenz auf *nationaler Ebene* anhand der LSE 2020 zu vergleichen (Kaiser & Möhr 2023). Die gesamtwirtschaftliche Lohndifferenz auf Basis von logarithmierten Bruttolöhnen beträgt rund 16.4%; die unerklärte Lohndifferenz beläuft sich auf 7.8%. Die Schätzungen auf nationaler Ebene sind jeweils höher. Der Hauptgrund dafür ist, dass auf nationaler Ebene Löhne innerhalb und *zwischen* den Unternehmen miteinander verglichen werden. Wenn Männer in Firmen mit überdurchschnittlich hohem Lohnniveau übervertreten sind, führt dies zu einer Erhöhung der Lohnunterschiede auf nationaler Ebene. Zweitens ist das Set der Kontrollvariablen in der nationalen

¹⁸ Dazu gehören Beobachtungen mit sehr tiefem oder hohem Beschäftigungsgrad von <20% oder >150% (Filter B); standardisierte Bruttolöhne unterhalb 1/3 des Medianlohns oder oberhalb des 15-fachen des Medianlohns (Filter C und D) sowie unplausible Kombinationen von Beschäftigungsgrad und Arbeitsstunden (Filter E).

Statistik grösser als im Logib-Modell. Drittens werden in der vorliegenden Studie nur Unternehmen mit mindestens 50 gültigen Lohnangaben betrachtet. Da die unerklärte Lohndifferenz bei kleinen Unternehmen grösser ist, hat diese Stichprobenrestriktion einen Einfluss auf die Ergebnisse. Viertens betrachten wir in der vorliegenden Studie einen Durchschnitt über Unternehmen, während in Kaiser & Möhr (2023) Durchschnitte über Beschäftigte gebildet werden.

Betrachten wir die Grösse der geschätzten Geschlechtereffekte im Standardmodell etwas genauer. Abbildung 4 zeigt die Häufigkeit, mit der bestimmte Effektgrössen geschätzt wurden. Daraus folgt, dass Unternehmen am häufigsten einen Koeffizienten zwischen $-0,05$ und 0 aufweisen. Für insgesamt rund 80% der Firmen wird ein Koeffizient kleiner als null geschätzt, das heisst, dass Frauen in diesen Firmen durchschnittlich weniger verdienen als vergleichbare Männer. Es verbleibt jedoch umgekehrt ein Anteil von rund 20% der Firmen, in denen der Koeffizient positiv ist, Frauen also bei vergleichbaren Merkmalen durchschnittlich einen höheren Lohn erhalten.

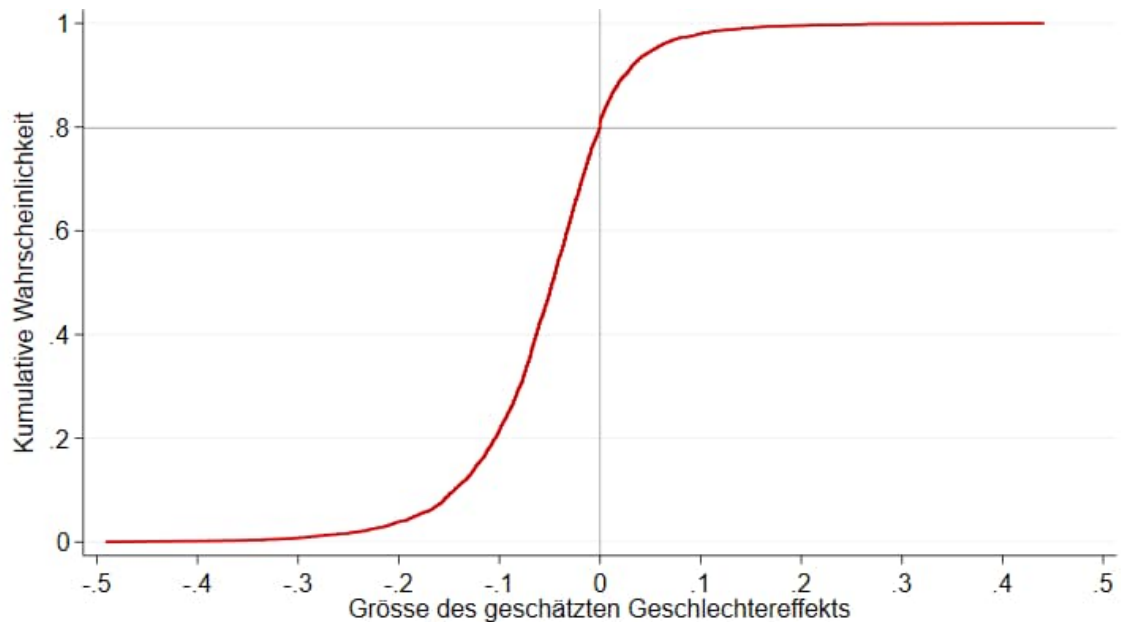
Abbildung 4: Häufigkeitsverteilung der Geschlechtereffekte mit dem Standardmodell



Anmerkungen: Die Balken geben an, wie häufig ein bestimmter Wert für beta geschätzt wird. Zu Grunde liegen Regressionsmodelle mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Die in Abbildung 5 gezeigte kumulative Wahrscheinlichkeitsfunktion der geschätzten Geschlechtereffekte zeigt dieses Muster noch etwas präziser. Je stärker der Anstieg dieser Kurve in einem Bereich ausfällt, desto häufiger wurden Geschlechtereffekte dieser Grösse bei Firmen geschätzt. Besonders stark ist der Anstieg im Bereich zwischen $-0,15$ und $0,05$. In diesem Intervall liegen also die meisten geschätzten Geschlechterkoeffizienten. Diese Aufschlüsselungen erlauben einen Überblick über die Grössenordnungen der mit dem Standardmodell geschätzten Geschlechterunterschiede und helfen damit die Änderungen durch Modifikationen besser einzuordnen.

Abbildung 5: Kumul. Wahrscheinlichkeit der Geschlechtereffekte im Standardmodell



Anmerkungen: Die Kurve zeigt an, wie häufig ein bestimmter Wert für den Geschlechtereffekt geschätzt wird. Ein stärkerer Anstieg ist mit einer grösseren Häufigkeit verbunden. Zu Grunde liegen Regressionsmodelle mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

4.2 Auswirkungen verschiedener Modifikationen

4.2.1 Einfluss der bestehenden Kontrollvariablen

Unsere erste Modifikation ist der Ausschluss einzelner *bestehender* Variablen des Logib-Standardmodells. Der Modellvergleich zwischen dem vollständigen und reduzierten Modell gibt dann Auskunft zur Frage, wie wichtig die einzelnen erklärenden Variablen für die Schätzung der Logib-Lohnregression sind.

Tabelle 10 fasst die Ergebnisse dieser Modifikation für die Gruppe der im Sample enthaltenen privaten Firmen zusammen. Die Spalten zeigen das Ergebnis der Schätzungen mit jeweils einer ausgeschlossenen Kontrollvariable, wobei die erste Spalte das Standardmodell darstellt. Die Zeilen zeigen unterschiedliche Merkmale der jeweiligen Spezifikationen, beginnend mit dem durchschnittlichen Geschlechtereffekt in der ersten Zeile. Bei diesem fällt auf, dass das Ausschliessen einer Variable immer zu einem stärker ausgeprägten Geschlechtereffekt führt. Für Bildungsjahre, Erwerbsjahre und Dienstjahre beträgt der Unterschied zum Standardmodell noch weniger als 0,3 Prozentpunkte (siehe auch absolute Unterschiede in Zeile zwei). Beim Ausschliessen des Kompetenzniveaus beträgt die Differenz absolut bereits 0,7 Prozentpunkte.¹⁹ Noch bedeutendere Folgen hat der Ausschluss der Beruflichen Stellung, was mit einem Anstieg des Geschlechterunterschiedes um 2,5 Prozentpunkte einhergeht und impliziert, dass Frauen in diesem Vergleich durchschnittlich 8% weniger verdienen als Männer innerhalb derselben Firma.

¹⁹ Der enge Zusammenhang zwischen der Variable Bildungsjahre und der in der LSF verfügbaren Kompetenzniveauvariablen (siehe Abschnitt 3.3) verhindert möglicherweise eine stärkere Veränderung.

Die folgende Zeile zeigt, dass mit dem Ausschluss von Kontrollvariablen auch die Standardfehler ansteigen, da der Anteil der erklärten Variation in den Daten sinkt (siehe auch Adjustiertes R^2 in Zeile 5). Die durchschnittlichen p-Werte steigen für die meisten Modellmodifikationen an, was durch den grösseren gemessenen Geschlechterunterschied erklärbar ist. Weiterhin zeigt die Tabelle die Signifikanzraten der beiden t-Tests des Logib-Modells an, welche prüfen, ob im Unternehmen ein Verstoss gegen die Lohngleichheit vorliegt. Im ersten, zweiseitigen t-Test steigt die Signifikanzrate durch das Ausschliessen einzelner Variablen von 38,5% auf bis zu 44,6%. Dieser Wert wird beim Ausschluss der beruflichen Stellung erreicht. Im zweiten, einseitigen t-Test, gegeben der Toleranzschwelle von 5%, spiegelt sich dieser Anstieg entsprechend wider: Während im Standardmodell rund 17% der Firmen den Test nicht bestehen würden, steigt dieser Wert bei Nichtberücksichtigung der beruflichen Stellung auf gut 26% der Firmen im Datensatz.

Die folgenden beiden Zeilen zeigen jedoch, dass diese Modifikationen nicht für alle Unternehmen in dieselbe Richtung gehen. Zwar ist es so, dass im Beispiel der ausgeschlossenen Beruflichen Stellung nun 10,6% der Firmen neuerdings den Test nicht mehr bestehen würden. Für 1,6% der Firmen führt der Ausschluss jedoch dazu, dass diese den Test in dieser Variante nun bestehen würden. Weiter zeigt die Tabelle wie viele Unternehmen im Sample den Test nicht mehr bestehen würden, sollte die Toleranzschwelle auf 2,5% sinken. Im Standardmodell hätte dies einen Anstieg der Signifikanzrate von 17% auf neu 28% der Firmen zur Folge.

Schliesslich beleuchten wir auch den Anteil der nicht identifizierbaren Geschlechtereffekte. Dieser Fall tritt dann ein, wenn im Unternehmen entweder nur Frauen oder nur Männer beschäftigt sind. Zudem ist der Geschlechtereffekt aber auch dann nicht identifizierbar, wenn der Geschlechterunterschied nicht von anderen Unterschieden wie beispielsweise im Kompetenzniveau unterscheidbar ist. Wenn im Extremfall alle Mitarbeiterinnen dem höchsten Kompetenzniveau und alle Mitarbeiter dem niedrigsten Kompetenzniveau angehörten, wäre ein um das Kompetenzniveau bereinigter Lohnvergleich also nicht möglich, weil es keine passenden Überlappungen zwischen den Geschlechtern gibt. Im Falle des Standardmodells und der hier gezeigten Modifikationen ist der Geschlechtereffekt nur in rund 1,5% der Fälle nicht identifizierbar.

Tabelle 10: Modellvarianten mit ausgeschlossenen Variablen, private Firmen

Mittelwert	Standard	Ausgeschlossene Variable				
		Bildungs-jahre	Erwerbs-jahre	Dienst-jahre	Komp.-niveau	Berufli. Stellung
Geschlechter Effekt (beta)	-0.056	-0.058	-0.058	-0.057	-0.063	-0.080
Abs. Veränderung v. beta	.	-0.002	-0.002	-0.001	-0.007	-0.025
Standardfehler	0.045	0.046	0.048	0.045	0.046	0.051
p-Wert	0.272	0.270	0.277	0.269	0.264	0.228
Adj. R2	0.663	0.629	0.606	0.650	0.608	0.547
Signif. 1. t-Test	38.51%	38.74%	37.45%	38.32%	40.19%	44.58%
Signif. 2. t-Test (Tol. 5%)	16.94%	18.77%	17.38%	17.58%	21.13%	25.94%
Wechsel sig. -> insig.	.	1.47%	2.64%	0.85%	2.40%	1.60%
Wechsel insig. -> sig.	.	3.31%	3.08%	1.50%	6.59%	10.60%
Signif. 2. t-Test (Tol. 2,5%)	28.01%	29.33%	28.50%	28.58%	31.60%	37.45%
Anteil n. identifizierbarer betas	1.56%	1.51%	1.56%	1.56%	1.46%	1.46%
N	3917	3917	3917	3917	3917	3917

Anmerkungen: Die erste Spalte beruht auf Regressionen mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die darunter folgenden Merkmale zeigen Mittelwerte über alle Firmen. Die weiteren Spalten stellen die Ergebnisse für die jeweils angezeigten Modifikationen dar. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Tabelle 11 zeigt die entsprechenden Ergebnisse für die Gruppe der im Sample enthaltenen öffentlichen Firmen. In dieser Gruppe sind die geschätzten Geschlechterunterschiede im Durchschnitt niedriger, folglich besteht auch ein höherer Anteil an Firmen das zweistufige Testverfahren von Logib Modul 1. Die Veränderungen hinsichtlich Ausschluss der verschiedenen Variablen ähneln denen der privaten Unternehmen. Den grössten Einfluss auf den geschätzten Geschlechtereffekt hat erneut die berufliche Stellung. Der einzelne Ausschluss der übrigen Variablen führt zu vergleichsweise kleinen Veränderungen der durchschnittlich gemessenen Geschlechterunterschiede. Deutlich geringer ist bei den öffentlichen Firmen der Anteil der nicht identifizierbaren Geschlechtereffekte, da öffentliche Firmen in der Regel über mehr Mitarbeitende verfügen und damit Überlappungen zwischen Männern und Frauen häufiger beobachtbar sind. Eine ausführliche Darstellung der Unterschiede zwischen privaten und öffentlichen Firmen bietet Tabelle 19 im Anhang.

Tabelle 11: Modellvarianten mit ausgeschlossenen Variablen, öffentliche Firmen

Mittelwert	Standard	Ausgeschlossene Variable				
		Bildungs-jahre	Erwerbs-jahre	Dienst-jahre	Komp.-niveau	Berufli. Stellung
Geschlechter Effekt (beta)	-0.042	-0.046	-0.043	-0.045	-0.042	-0.071
Abs. Veränderung v. beta	.	-0.004	-0.000	-0.003	-0.000	-0.030
Standardfehler	0.031	0.033	0.034	0.031	0.033	0.033
p-Wert	0.228	0.219	0.247	0.234	0.245	0.172
Adj. R2	0.698	0.646	0.625	0.684	0.635	0.611
Signif. 1. t-Test	46.15%	43.82%	41.72%	47.32%	45.92%	56.88%
Signif. 2. t-Test (Tol. 5%)	14.22%	17.02%	14.69%	16.78%	13.99%	29.37%
Wechsel sig. -> insig.	.	1.40%	3.26%	0.23%	4.43%	0.23%
Wechsel insig. -> sig.	.	4.20%	3.73%	2.80%	4.20%	15.38%
Signif. 2. t-Test (Tol. 2,5%)	27.04%	28.90%	27.27%	29.60%	28.21%	44.29%
Anteil n. identifizierbarer betas	0.46%	0.46%	0.46%	0.46%	0.46%	0.46%
N	431	431	431	431	431	431

Anmerkungen: Die erste Spalte beruht auf Regressionen mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die darunter folgenden Merkmale zeigen Mittelwerte über alle Firmen. Die weiteren Spalten stellen die Ergebnisse für die jeweils angezeigten Modifikationen dar. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

4.2.2 Erweiterung des Informationsgehalts

Das bestehende Logib-Modell kann angepasst werden, indem der Informationsgehalt bestehender Variablen voll ausgenutzt wird. Hierzu gibt es zwei Möglichkeiten. Erstens verwenden wir Dummyvariablen für alle Ausbildungsstufen anstelle deren Verdichtung auf die Anzahl der Ausbildungsjahre. Dies geschieht vor dem Hintergrund, dass die Ausbildungsjahre je nach Unternehmen nicht zwingend mit einer proportionalen Erhöhung der Löhne einhergehen und die Dummies für die Ausbildungsstufen die Löhne somit besser erklären können. Zweitens nutzen wir anstelle des Kompetenzniveaus – welches eine normative Verdichtung der Berufsgruppen darstellt – Dummyvariablen für die Berufsgruppen selbst. Diese Modifikation unterscheiden wir nochmals nach Verwendung von Ein- oder Zweistellern der Berufsklassifikation (CH-ISCO-19). Die Einsteller unterscheiden 10 verschiedene Berufskategorien, die Zweisteller effektiv 47 Kategorien.²⁰ Durch die Erhöhung der Anzahl an Variablen nimmt der Erklärungsgehalt des Modells für das Lohngefüge innerhalb eines Unternehmens potenziell zu. Durch die Nutzung der Berufsgruppen ist das Modell besser in der Lage, gleichen Lohn für gleiche Arbeit festzustellen. Dies ist jedoch noch nicht dem verfassungsrechtlichen Anspruch gerecht, der auch gleichen Lohn für *gleichwertige* Arbeit fordert. Gleichwertige Arbeit kann bereits vorliegen, wenn wie im Logib-Standardmodell für die Lohnanalyse für das Kompetenz- bzw. Anforderungsniveau kontrolliert wird.

²⁰ Die Hälfte der im Sample enthaltenen Firmen besitzt 13 oder weniger unterschiedliche Berufskategorien (Zweisteller). 90% der Firmen weisen 22 oder weniger unterschiedliche Berufe auf.

In Tabelle 12 sind die Ergebnisse dieser Modifikationen im Vergleich mit dem Standardmodell nach vorigem Schema für private Firmen zusammengefasst. Die Verwendung von Bildungskategorien anstatt Bildungsjahren führt kaum zu einer Änderung mit Blick auf den Geschlechtereffekt. Folglich sind auch kaum Veränderungen in der Zahl der Firmen zu beobachten, die den Logib-Test nicht bestehen würden. Stärker ausgeprägt sind die Veränderungen beim Ersetzen des Kompetenzniveaus durch Dummyvariablen für die Ein- oder Zweisteller der Berufe. Diese Modifikation führt dazu, dass die Grösse des Geschlechtereffekts im Schnitt um rund einen Prozentpunkt steigt. Der Anteil an Firmen, die nun den Logib-Test nicht mehr bestehen würden, bleibt aber stabil bei rund 17%. Der Detaillierungsgrad der Berufsvariable (einstellig vs. zweistellig) scheint nur einen vernachlässigbaren Einfluss zu haben, wie ein Vergleich der beiden Spalten zeigt. Vermutlich enthalten die gröberen Berufskategorien in Kombination mit der Stellung im Beruf bereits den grössten Teil des Informationsgehalts der zweistelligen Berufskategorien. Die Modifikationen unterscheiden sich hauptsächlich im Anteil der nicht identifizierbaren Geschlechterkoeffizienten. Dieser ist bei der Verwendung von Zweistellern auf Grund der höheren Anzahl an Parametern im Modell bei nahezu 4%. Für diesen Anteil an Firmen wäre der Logib-Test mit der gegebenen Modifikation nicht durchführbar.

Tabelle 12: Modellvarianten mit detaillierteren Variablen, private Firmen

Mittelwert	Standard	Ersetzte Variable		
		Bildungskategorien anstatt Bildungsjahre	Berufe (Einsteller) anstatt Kompetenzniveau	Berufe (Zweisteller) anstatt Kompetenzniveau
Geschlechter Effekt (beta)	-0.056	-0.055	-0.064	-0.064
Abs. Veränderung v. beta	.	0.001	-0.009	-0.009
Standardfehler	0.045	0.045	0.052	0.052
p-Wert	0.272	0.275	0.285	0.285
Adj. R2	0.663	0.675	0.708	0.708
Signif. 1. t-Test	38.51%	38.22%	36.26%	36.26%
Signif. 2. t-Test (Tol. 5%)	16.94%	16.34%	16.81%	16.81%
Wechsel sig. -> insig.	.	1.99%	5.53%	5.53%
Wechsel insig. -> sig.	.	1.40%	5.40%	5.40%
Signif. 2. t-Test (Tol. 2,5%)	28.01%	27.85%	27.23%	27.23%
Anteil n. identifizierbarer betas	1.56%	1.74%	2.43%	3.85%
N	3917	3917	3917	3917

Anmerkungen: Die erste Spalte beruht auf Regressionen mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die darunter folgenden Merkmale zeigen Mittelwerte über alle Firmen. Die weiteren Spalten stellen die Ergebnisse für die jeweils angezeigten Modifikationen dar. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Für die in Tabelle 13 gleichermassen analysierten öffentlichen Firmen weist die Modifikation der Bildungskategorien in eine ähnliche Richtung, indem der durchschnittliche Geschlechtereffekt

leicht sinkt. Der Effekt auf die Testergebnisse von Logib ist jedoch markant: Der Anteil der Unternehmen, die den Logib-Test nicht bestehen, sinkt um fast 4 Prozentpunkte.

Die Änderung der Berufskategorien wirkt sich stärker auf das Standardmodell aus als die der Bildungskategorien. Die Auswirkungen auf die Logib-Testverfahren sind wiederum stärker als für private Firmen. Bei den Berufskategorien sinkt der Anteil der öffentlichen Firmen, die den Logib-Test nicht mehr bestehen würden, sogar auf unter 10%.

Tabelle 13: Modellvarianten mit detaillierteren Variablen, öffentliche Firmen

Mittelwert	Standard	Ersetzte Variable		
		Bildungskategorien anstatt Bildungsjahre	Berufe (Einsteller) anstatt Kompetenzniveau	Berufe (Zweisteller) anstatt Kompetenzniveau
Geschlechter Effekt (beta)	-0.042	-0.039	-0.034	-0.034
Abs. Veränderung v. beta	.	0.004	0.008	0.008
Standardfehler	0.031	0.031	0.034	0.034
p-Wert	0.228	0.238	0.267	0.267
Adj. R2	0.698	0.714	0.740	0.740
Signif. 1. t-Test	46.15%	44.52%	39.16%	39.16%
Signif. 2. t-Test (Tol. 5%)	14.22%	10.49%	9.32%	9.32%
Wechsel sig. -> insig.	.	4.43%	6.53%	6.53%
Wechsel insig. -> sig.	.	0.70%	1.63%	1.63%
Signif. 2. t-Test (Tol. 2,5%)	27.04%	26.11%	22.84%	22.84%
Anteil n. identifizierbarer betas	0.46%	0.46%	0.46%	0.70%
N	431	431	431	431

Anmerkungen: Die erste Spalte beruht auf Regressionen mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die darunter folgenden Merkmale zeigen Mittelwerte über alle Firmen. Die weiteren Spalten stellen die Ergebnisse für die jeweils angezeigten Modifikationen dar. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Zusammenfassend bleibt festzuhalten, dass sich die Modifikation der bestehenden Variablen im Durchschnitt nur begrenzt auf die Häufigkeit von signifikanten Testergebnissen bei privaten Unternehmen auswirken. Diese Durchschnittsbetrachtung maskiert aber etwas die Tatsache, dass sich die Testergebnisse bei vielen Unternehmen ändern können. Bei den öffentlichen Unternehmen sieht es anders aus: Hier führen die betrachteten Modifikationen dazu, dass bei deutlich weniger Firmen eine ungleiche Bezahlung detektiert wird.

4.2.3 Hinzunahme zusätzlicher Variablen

Als weiteren Ansatz zur Reduktion des OVB untersuchen wir die Hinzunahme zusätzlicher Variablen. Da bei dieser Studie rein empirische Überlegungen im Vordergrund stehen, wird bei der Hinzunahme zusätzlicher Variablen keine Rücksicht auf den juristischen Diskriminierungsgehalt

genommen. Aus juristischer Sicht dürfen Arbeitgeber z.B. den Zivilstand nicht als Grund angeben, um vor Gericht Lohnunterschiede zu rechtfertigen. Aus diesem Grund darf dieses Merkmal aus juristischen Überlegungen nicht im Modell verwendet werden. Wenn es darum geht den «kausalen» Effekt der Lohndiskriminierung im statistischen Sinne zu messen, kann das Merkmal jedoch potenziell hilfreich sein, weil Zivilstand mit Reservationslöhnen und Berufserfahrung korreliert sein kann. In diesem Sinne kann die Aufnahme solcher Merkmale Hinweise auf den OVB liefern. Ändert sich der Geschlechtereffekt durch die Aufnahme weiterer Variablen, sind lohnbestimmende Faktoren im jetzigen Modell möglicherweise nicht ausreichend berücksichtigt. Wir überprüfen neben Charakteristika der Beschäftigung wie Beschäftigungsgrad, Anstellung im Stundenlohn (ja/nein), Erschwernis-Zulagen (ja/nein) auch individuelle Merkmale wie Zivilstand und Nationalität.

Tabelle 14 zeigt die Schätzergebnisse der verschiedenen Varianten zunächst wieder für private Firmen. Die jeweils einzeln zusätzlich aufgenommenen Variablen haben jedoch oft keinen merklichen Einfluss auf die Grösse des geschätzten Geschlechtereffektes. Dennoch sind Veränderungen im Modell festzustellen. Bei der Hinzunahme des Beschäftigungsgrades verändert sich der mittlere Geschlechtereffekt zwar kaum, das adjustierte R^2 steigt jedoch überdurchschnittlich stark um 0,017. Die Signifikanzrate im 2. t-Test ändert sich insgesamt zwar kaum, es gibt jedoch viele Wechsel von Firmen, die neu ein signifikantes Ergebnis aufweisen oder in die umgekehrte Richtung wechseln. Es ist daher plausibel, dass der Beschäftigungsgrad den Erklärungsgehalt des Lohnmodells verbessert. Eine Erweiterung des Modells um die Anstellung im Stundenlohn führt nur zu geringfügigen Verschiebungen. Die Berücksichtigung des Umstandes, ob eine beschäftigte Person Zulagen erhält oder nicht, senkt hingegen den Geschlechterunterschied um rund einen halben Prozentpunkt. Auch der Anteil der Firmen, der den Logib-Test nicht mehr besteht, verändert sich durch Aufnahme dieser Variable am stärksten und sinkt von knapp 17% im Standardmodell auf etwa 15,7% im erweiterten Modell. Männer üben innerhalb eines Unternehmens häufiger Tätigkeiten mit Zulagenzahlungen aus, was wiederum auch mit unterschiedlichen ausgeübten Berufen zusammenhängen kann (vgl. Tabelle 20). Durch die Berücksichtigung im Modell sinkt dann folglich die gemessene Lohnlücke. Wenn die Modelle zusätzlich den Zivilstand oder die Nationalitätskategorie berücksichtigen, so fallen die Veränderungen wieder geringfügiger aus. Insbesondere beim Zivilstand wären hier auch stärkere Veränderungen erwartbar gewesen, da dieser potenziell auch das mögliche Vorhandensein von Kindern im Haushalt der Arbeitnehmenden erfasst. Dies ist ein in der Literatur vielfach belegter Grund für Geschlechterunterschiede («motherhood penalty»). Der Zivilstand ist auf Ebene der Mitarbeiterschaften zwischen den Geschlechtern nicht gleich verteilt. Tabelle 20 und Tabelle 21 zeigen, dass in den untersuchten privaten und öffentlichen Unternehmen jeweils ein grösserer Anteil der Männer verheiratet ist. Die Berücksichtigung dieser Unterschiede im Modell verändert den geschätzten Geschlechterkoeffizient aber nur geringfügig. Möglicherweise ist auch die Verbindung zwischen dem Zivilstand und Elternschaft auf Unternehmensebene zu gering ausgeprägt, um die Schätzung des Geschlechtereffekts zu beeinflussen.

Tabelle 14: Modellvarianten mit zusätzlichen Variablen, private Firmen

Mittelwert	Standard	Zusätzliche Variable				
		Beschäftigungsgrad	Stundenlohn	Zulagen	Zivilstand	Nationalität
Geschlechter Effekt (beta)	-0.056	-0.058	-0.056	-0.053	-0.055	-0.059
Abs. Veränderung v. beta	.	-0.002	0.001	0.005	0.001	-0.003
Standardfehler	0.045	0.048	0.045	0.045	0.046	0.046
p-Wert	0.272	0.282	0.271	0.278	0.276	0.268
Adj. R2	0.663	0.680	0.674	0.668	0.664	0.668
Signif. 1. t-Test	38.51%	36.38%	38.17%	37.16%	37.45%	38.84%
Signif. 2. t-Test (Tol. 5%)	16.94%	16.14%	16.50%	15.67%	16.24%	17.64%
Wechsel sig. -> insig.	.	3.54%	1.01%	1.78%	1.24%	0.96%
Wechsel insig. -> sig.	.	2.74%	0.57%	0.52%	0.54%	1.66%
Signif. 2. t-Test (Tol. 2,5%)	28.01%	27.08%	27.77%	26.69%	27.62%	29.33%
Anteil n. identifizierbarer betas	1.56%	1.84%	1.68%	1.56%	1.58%	1.58%
N	3917	3917	3917	3917	3917	3917

Anmerkungen: Die erste Spalte beruht auf Regressionen mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die darunter folgenden Merkmale zeigen Mittelwerte über alle Firmen. Die weiteren Spalten stellen die Ergebnisse für die jeweils angezeigten Modifikationen dar. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Tabelle 15 wiederholt die gleichen Modifikationen für die Gruppe der öffentlichen Firmen. Der Beschäftigungsgrad führt durch seine Berücksichtigung im Modell zu einem vergleichbaren Anstieg des Erklärungsgehaltes im Modell (siehe Adj. R²) sowie zu einer spürbaren Reduktion der nicht bestandenen Logib-Tests um rund 3 Prozentpunkte. Auffällig ist auch, dass es zu Verschiebungen zwischen den Firmen kommt, welche eine Ablehnung im Logib-Test erhalten.²¹ Erneut zeigt sich, dass die Berücksichtigung von Zulagen den Geschlechtereffekt am stärksten beeinflusst. Die Berücksichtigung der Zulagen reduziert die Grösse des Geschlechtereffekts im Durchschnitt um rund 0,4 Prozentpunkte auf nunmehr 3,9%. Gleichzeitig würden nur noch 11% der öffentlichen Firmen einen Geschlechterunterschied aufweisen, der statistisch signifikant über der Toleranzschwelle von 5% liegt. Insgesamt beeinflussen die Merkmale Stundenlohn, Zivilstand und Nationalität die Modellergebnisse nur geringfügig.

²¹ Für solche Verschiebungen kann es mehrere Gründe geben. Zunächst einmal wird durch eine zusätzliche Variable der Erklärungsgehalt eines Modells erhöht, was die Standardfehler der geschätzten Koeffizienten typischerweise sinken lässt. Durch kleiner werdende Standardfehler erhöht sich die (mögliche) Signifikanz eines t-Tests des Koeffizienten. Zudem spielen die statistischen Zusammenhänge zwischen Beschäftigungsgrad, Geschlecht und Lohn eine wichtige Rolle. Beides kann sehr firmenspezifisch sein, sodass der Geschlechtereffekt in manchen Fällen steigt und in anderen Fällen sinkt.

Tabelle 15: Modellvarianten mit zusätzlichen Variablen, öffentliche Firmen

Mittelwert	Standard	Zusätzliche Variable				
		Beschäftigungsgrad	Stundenlohn	Zulagen	Zivilstand	Nationalität
Geschlechter Effekt (beta)	-0.042	-0.044	-0.042	-0.039	-0.042	-0.044
Abs. Veränderung v. beta	.	-0.002	-0.000	0.004	0.000	-0.002
Standardfehler	0.031	0.033	0.030	0.031	0.031	0.031
p-Wert	0.228	0.258	0.232	0.246	0.232	0.230
Adj. R2	0.698	0.714	0.712	0.703	0.700	0.700
Signif. 1. t-Test	46.15%	39.63%	45.92%	43.12%	44.99%	45.22%
Signif. 2. t-Test (Tol. 5%)	14.22%	11.19%	13.75%	10.96%	12.82%	13.52%
Wechsel sig. -> insig.	.	5.59%	1.17%	3.96%	1.40%	1.17%
Wechsel insig. -> sig.	.	2.56%	0.70%	0.70%	0.00%	0.47%
Signif. 2. t-Test (Tol. 2,5%)	27.04%	24.24%	27.74%	25.41%	27.27%	28.21%
Anteil n. identifizierbarer betas	0.46%	0.46%	0.46%	0.46%	0.46%	0.46%
N	431	431	431	431	431	431

Anmerkungen: Die erste Spalte beruht auf Regressionen mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die darunter folgenden Merkmale zeigen Mittelwerte über alle Firmen. Die weiteren Spalten stellen die Ergebnisse für die jeweils angezeigten Modifikationen dar. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

4.2.4 Exkurs: Einfluss der Zulagenvariable nach Branche

Da die Zahlung von Zulagen unter allen untersuchten Modellerweiterungen durch neue Kontrollvariablen den stärksten Einfluss auf den Geschlechtereffekt aufweist, untersuchen wir diese Modifikation auch noch einmal separat für verschiedene Branchen sowie für verschiedene kollektivvertragliche Anbindungen. Tabelle 16 zeigt, dass die Zahlung von Zulagen über alle Branchen hinweg gängige Praxis ist. Besonders hoch mit über 80% ist der Anteil an Zulagen zahlenden Firmen in der Energieversorgung sowie im Gesundheits- und Sozialwesen. Dies ist aber nicht mit einem überdurchschnittlich hohen Geschlechterunterschied im Lohn dieser Branchen verbunden. Auch die Anteile der Firmen, welche die Toleranzschwelle signifikant überschreiten, liegt im mittleren Bereich. Durch die Berücksichtigung von Zulagen im geschätzten Modell verringert sich der Anteil der signifikanten zweiten t-Tests beispielsweise von 17% auf rund 13% in der Branche Energieversorgung, während er im Gesundheits- und Sozialwesen konstant bei knapp 8% verbleibt. Am grössten ist die Veränderung in der Kategorie «Verwaltung; Verteidigung; Sozialwesen». Dort fällt die Signifikanzrate im 2. t-Test von rund 26% auf 10% nach Berücksichtigung von Zulagen. Firmen mit und ohne Kollektivverträge unterscheiden sich mit Blick auf die Zahlung von Zulagen und deren Auswirkungen auf die Lohnungleichheit in eher geringem Umfang.

Tabelle 16: Schätzergebnisse nach Branchen und kollektivvertraglichen Regelungen

Branche	Anzahl Firmen	Anteil m. Zulagen	Mittlerer geschätzter Geschlechterkoeffizient		Signifikanzrate 2. t-Test	
			Standardmodell	mit Zulagen	Standardmodell	mit Zulagen
Bergbau; Steine und Erden	7	0.43	0.01	0.01	0.00	0.00
Nahrungsmittel; Getränke	120	0.76	-0.07	-0.06	30.00	25.00
Textil-, Lederwaren	15	0.53	-0.14	-0.14	53.33	46.67
Druckgewerbe	19	0.63	-0.08	-0.07	26.32	21.05
Chemie	103	0.74	-0.07	-0.06	23.30	23.30
Maschinenbau	384	0.61	-0.08	-0.08	27.86	25.00
Übriges Verarbeitendes Gewerbe	284	0.70	-0.08	-0.08	27.46	25.35
Energieversorgung	53	0.87	-0.06	-0.05	16.98	13.21
Wasser; Abfall	25	0.76	-0.01	-0.00	8.00	4.00
Baugewerbe	298	0.65	-0.06	-0.05	16.48	12.45
Handel; Instandhaltung und Rep.	310	0.38	-0.06	-0.06	17.53	15.58
Detailhandel	172	0.56	-0.06	-0.06	16.96	15.79
Verkehr und Lagerei	212	0.58	-0.03	-0.03	10.48	9.52
Gastgewerbe	145	0.43	-0.03	-0.03	4.14	4.14
Information und Kommunikation	178	0.57	-0.06	-0.06	16.29	16.85
Finanzen und Versicherung	192	0.41	-0.07	-0.07	26.56	24.48
Grundstücks- und Wohnungswesen	31	0.39	-0.06	-0.05	12.90	12.90
Freiberuf., wissenschaft. und techn. Dienste	217	0.34	-0.05	-0.05	9.22	10.14
Sonstige wirtschaftl. Dienstleistungen	288	0.66	-0.08	-0.07	27.41	26.30
Verwaltung; Verteidigung; Sozialwesen	70	0.77	-0.05	-0.04	26.09	10.14
Erziehung; Unterricht	131	0.44	-0.04	-0.03	6.11	5.34
Gesundheits- und Sozialwesen	974	0.83	-0.03	-0.03	7.62	7.72
Kunst, Unterhaltung und Erholung	46	0.57	-0.04	-0.04	4.35	2.17

Branche	Anzahl Firmen	Anteil m. Zulagen	Mittlerer geschätzter Geschlechterkoeffizient		Signifikanzrate 2. t-Test	
			Standardmodell	mit Zulagen	Standardmodell	mit Zulagen
Erbringung v. sonst. Dienstleistungen	74	0.38	-0.05	-0.05	14.86	17.57
Firmen ohne Kollektivverträge	2918	0.61	-0.06	-0.06	17.38	16.10
Firmen mit Kollektivverträgen	1430	0.65	-0.05	-0.05	15.19	13.32
N	4348					

Anmerkungen: Die Signifikanzrate 2. t-Test bezeichnet den Anteil von Firmen, welcher für die gegebene Grössenklasse und die gegebene Modellspezifikation den Lohngleichheitstest nicht bestehen würde. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

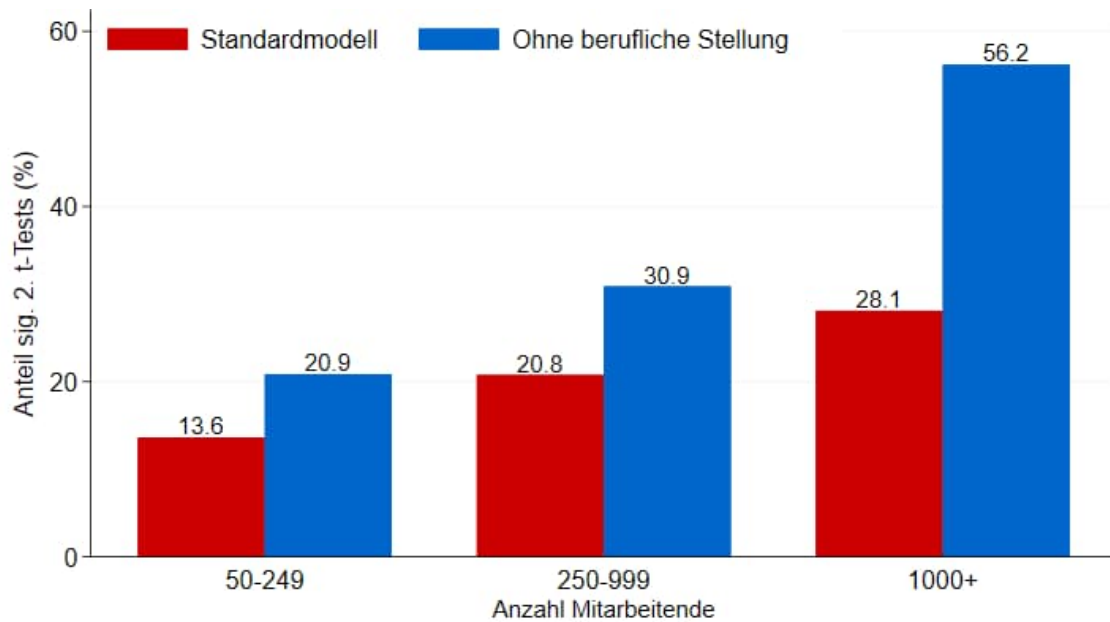
4.3 Heterogenitätsanalyse

Die bisher betrachteten Analysen stellen Durchschnittseffekte für die Logib-Analysen von privaten und öffentlichen Firmen dar. Eine Betrachtung dieser Durchschnitte allein kann daher Veränderungen für bestimmte Unternehmenstypen verbergen. Wichtige Unterscheidungsmerkmale von Firmen aus statistischer Sicht sind vor allem die Unternehmensgrösse sowie der Frauenanteil, da diese eine direkte Auswirkung auf die Präzision des Logib-Testverfahrens haben.

4.3.1 Unternehmensgrösse

Die folgenden Abbildungen unterscheiden daher kleine Firmen mit 50 bis 249 Mitarbeitenden, mittlere Firmen mit 250 bis 999 Mitarbeitenden und Grossunternehmen mit mindestens 1000 Mitarbeitenden. Sie zeigen die Auswirkungen der im Durchschnitt jeweils prägnantesten Modifikationen auf den Anteil der Firmen, die den Lohngleichheitstest der zweiten Stufe nicht bestehen würden. Grundsätzlich ist bei einer niedrigeren Mitarbeitendenzahl weniger wahrscheinlich, dass ein Logib-Test einen gegebenen Geschlechterunterschied im Lohn als signifikant detektiert. Auf Grund der niedrigeren Beobachtungszahl von Löhnen im Unternehmen besteht eine grössere statistische Unsicherheit. Abbildung 6 bestätigt diese Überlegung und zeigt, dass im Logib-Standardmodell der Anteil abgelehnter Tests mit der Unternehmensgrösse steigt (rote Balken). Über alle Unternehmensgrössen hinweg steigt dieser Anteil weiter an, wenn die berufliche Stellung im Modell nicht mehr berücksichtigt wird (blaue Balken). Besonders betroffen sind hiervon allerdings Grossunternehmen. Dort steigt nun der Anteil von Unternehmen im Sample, die den Logib-Test nicht bestehen, von 28% auf 56%. Eine mögliche Erklärung hierfür ist, dass die Hierarchieebenen in einem Grossunternehmen auch in der Lohnstruktur deutlich ausgeprägter sind als in einer Kleinfirma und die berufliche Stellung daher wichtig ist, um diese innerbetriebliche Heterogenität zu erfassen.

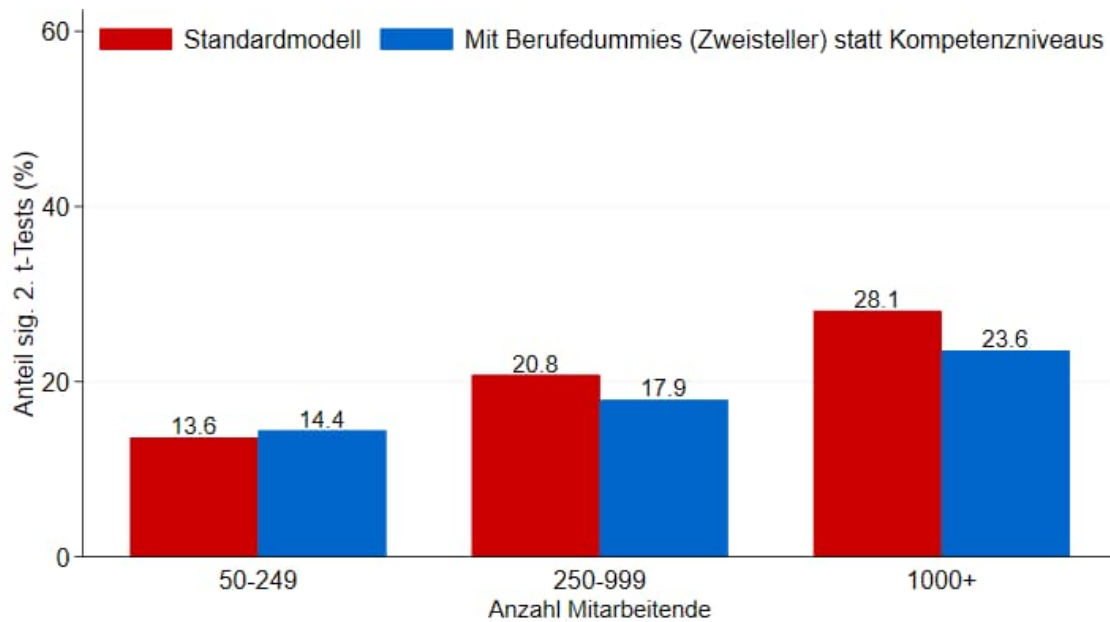
Abbildung 6: Testergebnisse nach Firmengrösse, ohne berufliche Stellung



Anmerkungen: Die Balken geben an, welcher Anteil von Firmen für die gegebene Grösseklasse und die gegebene Modellspezifikation den Lohngleichheitstest nicht bestehen würde, im t-Test der zweiten Stufe also ein signifikantes Ergebnis hätte. Zu Grunde liegen Regressionsmodelle mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturhebung 2020, eigene Berechnungen.

In der Abbildung 7 vergleichen wird das Standardmodell mit der Modifikation, die zweistellige Berufskategorien als Dummyvariablen anstelle der daraus abgeleiteten Kompetenzniveaus verwendet. Während sich für kleinere Firmen kaum eine Veränderung durch diese Modifikation zeigt, sinkt der Anteil nicht bestandener Tests von Grossunternehmen um über vier Prozentpunkte.

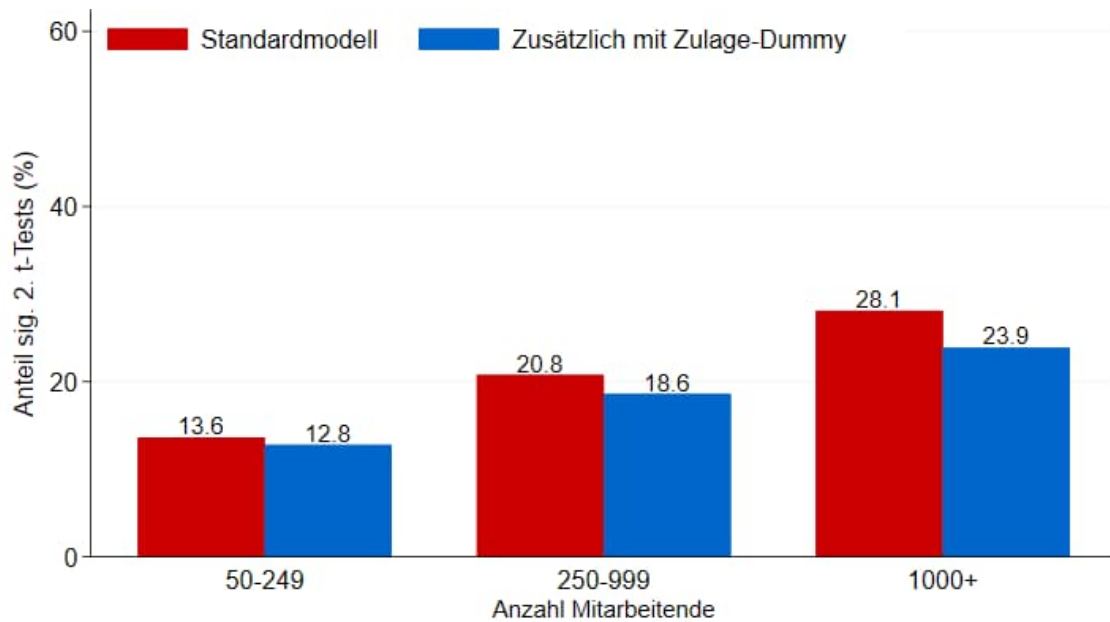
Abbildung 7: Testergebnisse nach Firmengrösse, mit Berufskategorien



Anmerkungen: Die Balken geben an, welcher Anteil von Firmen für die gegebene Grösseklasse und die gegebene Modellspezifikation den Lohngleichheitstest nicht bestehen würde, im t-Test der zweiten Stufe also ein signifikantes Ergebnis hätte. Zu Grunde liegen Regressionsmodelle mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Zu ähnlichen Verschiebungen führt die in Abbildung 8 dargestellte Modifikation, welche das Standardmodell mit einer Zulagen-Dummyvariable als weitere Kontrollvariable ergänzt. Für kleine Firmen verändert sich der Anteil abgelehnter Tests um weniger als 1 Prozentpunkt gegenüber der Standardspezifikation. Dem steht wiederum eine Reduktion von über 4 Prozentpunkten bei den Grossunternehmen gegenüber. Daher sind auch Zulagen bedeutsam, um die Lohnstruktur von Grossunternehmen besser zu erfassen. Insgesamt liegt daher die Folgerung nahe, dass insbesondere für grosse Firmen die Wahl und Zusammensetzung der Kontrollvariablen besonders bedeutsam ist. Dies liegt vermutlich an der grösseren inneren Differenzierung von Positionen und Löhnen in Grossunternehmen, welche mit einem Modell abgebildet werden sollen.

Abbildung 8: Testergebnisse nach Firmengrösse, mit Zulagen-Dummy



Anmerkungen: Die Balken geben an, welcher Anteil von Firmen für die gegebene Grösseklasse und die gegebene Modellspezifikation den Lohngleichheitstest nicht bestehen würde, im t-Test der zweiten Stufe also ein signifikantes Ergebnis hätte. Zu Grunde liegen Regressionsmodelle mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Grosse Unterschiede zwischen den Firmen verschiedener Grösseklassen bestehen neben der Wahrscheinlichkeit eines nicht bestandenen Logib-Tests auch in der Wahrscheinlichkeit, dass der Geschlechterkoeffizient erst gar nicht identifizierbar ist. Für Firmen mit unter 250 Mitarbeitenden beträgt der Anteil für das Standardmodell 2%. Für die mittlere Unternehmensgrösse kommt dies noch in 0,5% der Fälle vor und bei Grossunternehmen überhaupt nie. Auch hier besteht ein direkter Zusammenhang zur Mitarbeitendenzahl. Bei einer hohen Anzahl von Lohnbeobachtungen je Firma ist es wahrscheinlicher, für jede Mitarbeiterin und ihren Lohn einen basierend auf den gewählten Variablen vergleichbaren Mitarbeiter als Referenzpunkt zu finden.

4.3.2 Frauenanteil im Unternehmen

Im Anhang A.2 zeigen wir gleichermassen, dass auch Firmen entsprechend ihres Frauenanteils unterschiedlich von verschiedenen Modifikationen betroffen sind. Generell haben Firmen mit einem Frauenanteil von rund 50% ceteris paribus eine höhere Wahrscheinlichkeit, dass der Lohngleichheitstest nicht bestanden wird. Dies liegt daran, dass die Vergleichsmöglichkeiten zwischen Frauen und Männern besser sind, wenn beide Gruppen etwa gleich gross sind. Die Varianz im Merkmal Geschlecht ist bei einem ausgeglichenen Geschlechterverhältnis am grössten. Damit haben statistische Tests eine höhere Mächtigkeit, das heisst eine höhere Wahrscheinlichkeit, einen Geschlechtereffekt korrekt zu detektieren. Die präsentierten Modifikationen wirken sich jedoch nicht nur auf die Firmen im mittleren Bereich des Frauenanteils aus, sondern haben oft noch stärkere Implikationen für Firmen an den Rändern der Verteilung.

4.4 Einschätzung des Omitted Variable Bias

In diesem Abschnitt ermitteln wir mit dem Ansatz von Oster (2019) eine Obergrenze («Schranke») des Omitted Variable Bias. Die Annahmen und die genaue Funktionsweise des Ansatzes erläutern wir in Abschnitt 3.2. Die Grundidee ist, dass beobachtete und unbeobachtbare Kontrollvariablen miteinander zusammenhängen. Somit können wir aus der Reaktion des Modells auf das Hinzufügen von beobachteten Kontrollvariablen auch Rückschlüsse auf die Folgen des Hinzufügens von unbeobachtbaren Kontrollvariablen ziehen. Diese Analysen erlauben es, auf Basis statistischer Annahmen das potenzielle Ausmass des OVB abzuschätzen. Zu betonen ist, dass dieser Ansatz lediglich eine Annäherung an den tatsächlichen OVB darstellt.

In unserem Fall schätzen wir für jedes Unternehmen im Datensatz drei Grössen:

- den unbereinigten Geschlechtereffekt, also den Lohnunterschied ohne Verwendung von Kontrollvariablen
- den bereinigten Geschlechtereffekt unter Verwendung der Kontrollvariablen im Standardmodell
- den «simulierten» Geschlechtereffekt nach Oster (2019), wenn auch unbeobachtete Kontrollvariablen theoretisch berücksichtigt werden könnten

Abbildung 9 fasst diese Ergebnisse zusammen. Die rote, kurzgestrichelte Kurve zeigt die Häufigkeitsverteilung von geschätzten Geschlechtereffekten ohne jegliche Kontrollvariablen. Im Mittel über alle Firmen liegt der geschätzte Lohnunterschied bei rund -13% (rote vertikale Linie). Die blaue, lang gestrichelte Kurve zeigt die Verteilung der geschätzten Geschlechtereffekte unter Berücksichtigung der Kontrollvariablen im Standardmodell.²² Diese Kurve liegt in ihrem Zentrum bereits wesentlich näher am Nullpunkt. Der durchschnittliche Geschlechtereffekt im Standardmodell über alle berücksichtigten Unternehmen hinweg beträgt -5.4% (vgl. Abschnitt 4.1) und ist durch die blaue vertikale Linie gegeben.

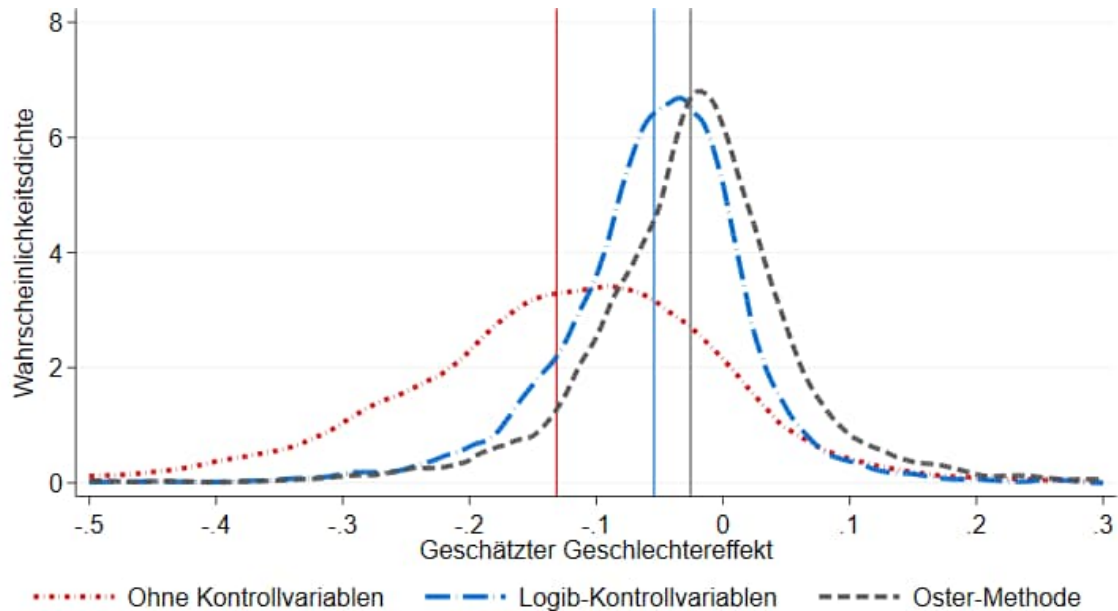
Noch näher am Nullpunkt liegt der Mittelwert der grauen, mittellang gestrichelten Kurve. Diese repräsentiert die Häufigkeitsverteilung die mithilfe des Oster-Ansatzes simulierten Geschlechtereffekte unter vollständiger Berücksichtigung beobachtbarer und unbeobachtbarer Faktoren. Der Lohnunterschied zwischen Männern und Frauen ist hierbei im Durchschnitt über alle Firmen -2,6% (graue vertikale Linie) und damit im Betrag also noch kleiner als im Logib-Standardmodell.

Betrachten wir die Kurven gemeinsam, ergibt sich folgendes Bild: Der Durchschnitt des maximal möglichen OVB ergibt sich gemäss der Oster-Methode durch den horizontalen Abstand zwischen roter und grauer Linie. Er beträgt rund 10 Prozentpunkte. Der grösste Teil dieses maximal möglichen OVB wird aber durch die bestehenden Kontrollvariablen des bestehenden Logib-Standardmodells bereits berücksichtigt (horizontaler Unterschied zwischen roter und blauer Linie). Die verbleibende Differenz zwischen der blauen und grauen Linie repräsentiert den laut Oster-Methode verbleibenden OVB, der nicht durch die im Standardmodell enthaltenen Variablen erfasst wird. Das Logib-Standardmodell überschätzt das Ausmass des Geschlechtereffekts also leicht. Die

²² Die Lohnstrukturerhebung enthält lediglich das Kompetenzniveau und nicht das in den Logib-Tests vorgesehene Anforderungsniveau einer Stelle (siehe Erläuterung in Sektion 3.3). Kann dieses Merkmal in die Analyse einfließen, wäre der Spielraum für unbeobachtbare Faktoren bereits weiter eingegrenzt.

bestehenden Kontrollvariablen scheinen jedoch etwa drei Viertel des gesamten OVB zu berücksichtigen. Der verbleibende OVB beträgt im Durchschnitt aller beobachteten Firmen ca. 2,8 Prozentpunkte.²³

Abbildung 9: Kerndichteverteilung der geschätzten Geschlechterkoeffizienten

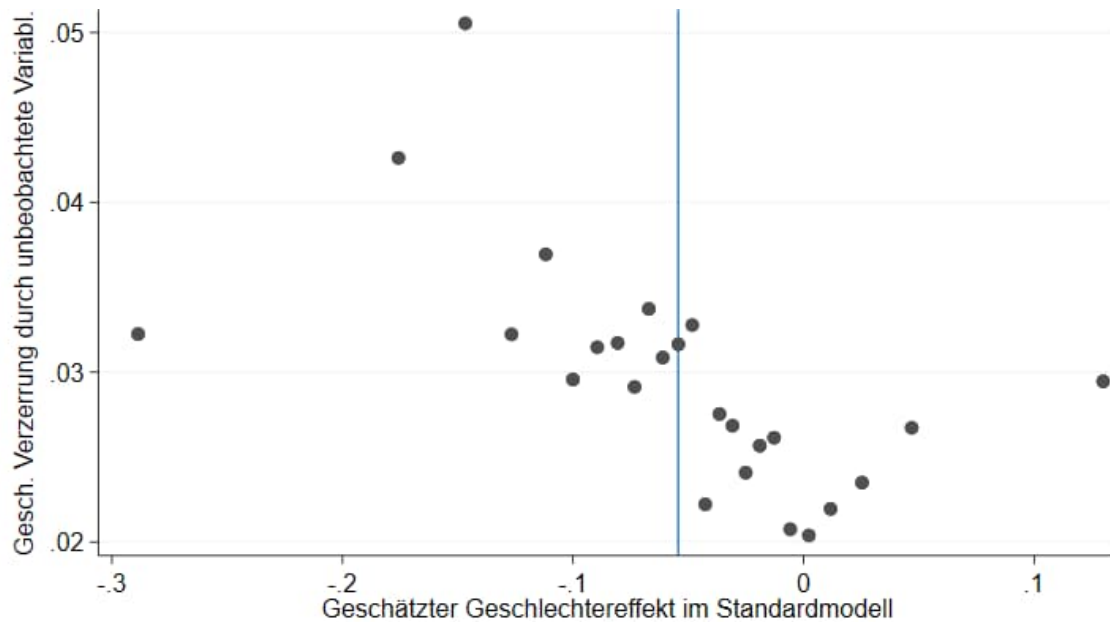


Anmerkungen: Kerndichteverteilung geschätzt mit Hilfe eines Epanechnikov-Kerns. Vertikale Linien zeigen den Mittelwert der jeweils geschätzten Koeffizienten. Annahmen für das Oster-Verfahren: $\delta = 1$, $R_{\max} = \text{Min}(1.3 \times R^2 \text{ des Standardmodells}, 1)$. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Zweitens untersuchen wir, für welche Firmen das Ausmass des verbleibenden OVB besonders gross ist. Abbildung 10 zeigt Gruppen von Firmen entsprechend der Höhe ihres Geschlechtereffekts im Logib-Modell auf der x-Achse und den verbleibenden OVB auf der y-Achse. Dadurch wird sichtbar, dass insbesondere Firmen mit einem besonders stark negativen Geschlechtereffekt eine möglicherweise hohe Verzerrung durch nicht berücksichtigte Variablen aufweisen. Bei Geschlechtereffekten von 15% und mehr hat eine verbleibende Verzerrung von beispielsweise 4 Prozentpunkten allerdings einen geringfügigeren Einfluss als für Firmen, die sich nahe der Toleranzgrenze befinden. Dadurch wird das Testergebnis im 2. t-Test auf Grund einer solchen Verzerrung seltener umgekehrt. Firmen mit gering ausgeprägten Lohnunterschieden nach Geschlecht weisen gemäss der Abbildung hingegen eher geringfügigere Verzerrungen von unter drei Prozentpunkten auf.

²³ Das Logib-Testverfahren berücksichtigt solche Unschärfen auf zwei Arten: Einerseits durch die Toleranzschwelle von 5%. Andererseits stellt der zweite t-Test sicher, dass nicht schon bei sofortiger Überschreitung sondern erst bei statistisch signifikanter Überschreitung ein Verstoß festgestellt wird.

Abbildung 10: Geschlechtereffekt und Verzerrung durch unbeobachtete Variablen



Anmerkungen: Die Grafik zeigt ein Binned Scatterplot. Dieser teilt die Firmen entsprechend ihres geschätzten Geschlechtereffektes in 25 gleich grosse Gruppen ein und zeigt für jede Gruppe die durchschnittliche geschätzte Verzerrung durch unbeobachtete Variablen. Die geschätzte Verzerrung ist berechnet als die Differenz zwischen der Schranke nach dem Oster-Verfahren und dem geschätzten Geschlechterkoeffizienten aus dem Standardmodell nach Logib 1. Annahmen für das Oster-Verfahren: $\delta = 1$, $R_{\max} = \text{Min}(1.3 \times R^2 \text{ des Standardmodells}, 1)$. Vertikale Linie repräsentiert den Durchschnitt des geschätzten Geschlechtereffekts im Standardmodell über alle Firmen. Sample enthält nur Firmen mit mindestens 50 gültigen Lohnbeobachtungen (Beschäftigte). Quelle: Lohnstrukturerhebung 2020, eigene Berechnungen.

Schliesslich ist die Frage von besonderem Interesse, wie häufig die simulierten Geschlechtereffekte mit der Oster-Methode zu anderen Ergebnissen bei den beiden sequenziellen t-Tests führen würden. Da die Oster-Methode eine simulierte Grösse liefert, steht kein Standardfehler zur Verfügung, um einen statistischen Signifikanztest durchzuführen. Um trotzdem Aussagen zu den Test-Ergebnissen abzuleiten, treffen wir nachfolgend die Annahme, dass sich der Standardfehler des simulierten Geschlechtereffekts im Vergleich zum Standardfehler aus dem Standardmodell nicht ändern würde. Diese Annahme wird in der Realität nie genau zutreffen, erscheint mit Blick auf die Analysen in Kapitel 4.2 allerdings nicht unplausibel: Der Standardfehler des Geschlechtereffekts ändert sich häufig nur geringfügig, wenn weitere Variablen in das Modell inkludiert werden.²⁴ Dennoch sind die nachfolgenden Ergebnisse als grobe Annäherung zu interpretieren.

Tabelle 17 zeigt die Häufigkeiten der Testergebnisse für den Geschlechtereffekt aus dem Logib-Standardmodell und den simulierten Geschlechtereffekt gemäss Oster-Methode in einer Kreuztabelle. Die Testergebnisse sind in Form von drei Kategorien dargestellt: 1) keiner der beiden t-Tests führt zu einem statistisch signifikanten Ergebnis, 2) nur der 1. t-Test liefert ein statistisch signifikantes Ergebnis, 3) beide t-Tests führen zu einem statistisch signifikanten Ergebnis. Fällt ein Unternehmen in die dritte Kategorie, gilt der Logib-Test als nicht bestanden. Die Prozentwerte in Tabelle 17 beziehen sich auf die Anzahl der privaten Unternehmen im verwendeten Datensatz. In der Spalte ganz rechts sind die Testergebnisse gemäss Logib-Standardmodell ersichtlich (vgl.

²⁴ Aus theoretischer Sicht kann der Standardfehler steigen oder sinken, wenn weitere Variablen inkludiert werden. Zum einen senken weitere Variablen die Residualvarianz der Regression, was den Standardfehler reduziert. Zum anderen steigt die Kollinearität der Kontrollvariablen mit dem Geschlechter-Indikator, was den Standardfehler erhöht.

Tabelle 10), wonach 16.9% der Unternehmen den Logib-Test nicht bestehen. Die letzte Zeile der Tabelle zeigt die Häufigkeiten der Testergebnisse, wenn der simulierte Geschlechtereffekt gemäss Oster-Methode verwendet wird. Der Anteil der privaten Unternehmen, die den Logib-Test nicht bestehen, beträgt noch 10.9%. Umgekehrt steigt der Anteil der Unternehmen mit einem nicht signifikant von null verschiedenen Geschlechtereffekt von 61.5% auf 68.1%. Die Kreuztabelle zeigt jeweils, wie die Unterschiede zustande kommen. Ein Beispiel: Bei rund 30% der Unternehmen (= 5.1%/16.9%), die den 2. t-Test mit dem Standardmodell nicht bestehen, wird das Ergebnis des 2. t-Tests insignifikant, während der 1. t-Test signifikant bleibt. Insgesamt suggerieren die Resultate der Oster-Methode, dass eine umfassende Berücksichtigung von OVB durch unbeobachtbare Variablen die Häufigkeit von nicht bestandenen Logib-Tests um etwas mehr als einen Drittel reduzieren würde.

Tabelle 17: Logib-Ergebnisse vs. «simulierter» Geschlechtereffekt, private Firmen

		Simulierte Ergebnisse mit Oster-Methode			
		kein t-Test sig.	nur 1. t-Test sig.	2. t-Test sig.	Total
Ergebnisse mit Standardmodell	kein t-Test sig.	55.0%	5.8%	0.7%	61.5%
	nur 1. t-Test sig.	10.4%	10.1%	1.0%	21.6%
	2. t-Test sig.	2.6%	5.1%	9.2%	16.9%
	Total	68.1%	21.0%	10.9%	100.0%

Anmerkungen: Die Tabelle zeigt die Häufigkeitsverteilung der Logib-Testergebnisse mit dem Standardmodell und der Oster-Methode als Kreuztabelle. Die Prozentwerte beziehen sich auf den Anteil der privaten Unternehmen im Datensatz mit dem jeweiligen Testergebnis. Die Annahme ist, dass sich der Standardfehler zwischen den beiden Ansätzen nicht unterscheidet. Quelle: LSE 2020, eigene Berechnungen.

Tabelle 18 zeigt dieselben Ergebnisse wie oben für die öffentlichen Unternehmen im Datensatz. Was stark auffällt ist, dass sich der Anteil der nicht bestandenen Testergebnisse stark reduziert: Mit dem Standardmodell bestehen 14.2% der öffentlichen Firmen den Test nicht, d.h. bei diesen Unternehmen ist der Geschlechtereffekt statistisch signifikant grösser als 5%. Mit dem simulierten Geschlechtereffekt gemäss Oster-Methode sind es noch 5.6%. Das heisst, ein erheblicher Teil der öffentlichen Unternehmen, bei denen mit dem Standardmodell eine Lohndiskriminierung festgestellt wird, würden den Test bestehen, wenn der simulierte Geschlechtereffekt gemäss Oster-Methode verwendet wird.

Tabelle 18: Logib-Ergebnisse vs. «simulierter» Geschlechtereffekt, öffentliche Firmen

		Simulierte Ergebnisse mit Oster-Methode			
		kein t-Test sig.	nur 1. t-Test sig.	2. t-Test sig.	Total
Ergebnisse mit Standardmodell	kein t-Test sig.	46.9%	7.0%	0.0%	53.8%
	nur 1. t-Test sig.	16.1%	15.6%	0.2%	31.9%
	2. t-Test sig.	2.3%	6.5%	5.4%	14.2%
	Total	65.3%	29.1%	5.6%	100.0%

Anmerkungen: Die Tabelle zeigt die Häufigkeitsverteilung der Logib-Testergebnisse mit dem Standardmodell und der Oster-Methode als Kreuztabelle. Die Prozentwerte beziehen sich auf den Anteil der öffentlichen Unternehmen im Datensatz mit dem jeweiligen Testergebnis. Die Annahme ist, dass sich der Standardfehler zwischen den beiden Ansätzen nicht unterscheidet. Quelle: LSE 2020, eigene Berechnungen.

Der Vergleich zwischen den simulierten Testergebnissen mit der Oster-Methode und den Ergebnissen mit dem Standardmodell legen nahe, dass die Häufigkeit nicht bestandener Logib-Tests abnehmen würde, wenn OVB aufgrund fehlender Variablen umfassend berücksichtigt werden könnte. Dieser Rückgang der nicht bestandenen Logib-Tests ist vor allem bei den öffentlichen Firmen sehr ausgeprägt. Gleichzeitig sind die Ergebnisse und die daraus gezogenen Schlussfolgerungen mit Vorsicht zu betrachten: Bei der Oster-Methode und der obigen Simulation des Logib-Tests werden starke statistische Annahmen unterstellt, sodass die quantitativen Ergebnisse stets als grobe Annäherungen zu interpretieren sind.

5 Fazit

Im Rahmen dieser Studie wurden die statistischen Eigenschaften des Logib-Standardmodells mit Blick auf den Einfluss der Kontrollvariablen untersucht. Im Zentrum dieser Betrachtung stand die Frage, ob der Geschlechtereffekt im Logib Modul 1 auf Grund fehlender Variablen verzerrt geschätzt wird. Das Problem des sogenannten Omitted Variable Bias (OVB) entsteht dann, wenn eine im Modell nicht enthaltene Variable gleichzeitig lohnbestimmend und mit dem Geschlecht der Beschäftigten korreliert ist. Wir stützen uns auf Daten der LSE 2020, wobei nur Unternehmen berücksichtigt werden, für die mindestens 50 gültige Beobachtungen ohne fehlende Werte zur Verfügung stehen.

Die Lohnanalysen auf Unternehmensebene zeigen: Im Schnitt liefert das Logib-Standardmodell einen Geschlechtereffekt von -5.4%. Bei rund 55% der Unternehmen beträgt der geschätzter Geschlechtereffekt mindestens 5%; in den allermeisten Fällen ist dieser zu Ungunsten der Frauen. Rund 17% der privaten Unternehmen und 14% der öffentlichen Unternehmen bestehen den Test gemäss Logib Modul 1 nicht, das heisst, der Geschlechtereffekt ist statistisch signifikant grösser als 5%. Der durchschnittliche Geschlechtereffekt auf Unternehmensebene ist aus vier Gründen nicht direkt mit der unerklärten Lohndifferenz der nationalen Statistik (vgl. Kaiser & Möhr 2023) vergleichbar: (i) Nur Lohnunterschiede *innerhalb* von Unternehmen werden berücksichtigt, (ii) die verwendeten Kontrollvariablen unterscheiden sich, (iii) nur mittlere und grössere Unternehmen werden einbezogen und (iv) die Ergebnisse sind nicht gewichtet mit der Anzahl Beschäftigten.

Anhand einer Analyse auf Unternehmensebene versuchten wir durch Modifikationen des Logib-Standardmodells zunächst herauszufinden, wie sensibel die Ergebnisse von Modul 1 auf Veränderungen in der Auswahl und Spezifikation der Kontrollvariablen reagieren. Die Analyse alternativer Modellspezifikationen stellt eine rein statistische Betrachtung dar; juristische Überlegungen zur Eignung der jeweiligen Merkmale wurden nicht miteinbezogen. Die Modifikationen führten wir in drei verschiedenen Formen durch. Die erste Modifikation bestand darin, bestehende Variablen des Logib-Standardmodells jeweils einzeln aus dem Modell auszuschliessen. Dies führte in jedem Fall zu höheren geschätzten Lohnunterschieden nach Geschlecht. Die Ergebnisse zeigen, dass die bestehenden Variablen unterschiedlich wichtig sind, um den OVB zu reduzieren. Dem Merkmal «Berufliche Stellung» kommt insgesamt die grösste Bedeutung zu, während die Variablen «Dienstjahre», «Potenzielle Erwerbsjahre» und «Ausbildungsjahre», zumindest bei privaten Firmen, weniger bedeutsam sind. Als zweite Gruppe an Modifikationen prüften wir eine Erhöhung des Erklärungsgehaltes von bestehenden Merkmalen. Wenn man anstelle der Ausbildungsjahre die Ausbildungskategorien verwendet, reduziert sich der Geschlechtereffekt bei den öffentlichen Firmen merklich, bei den privaten Firmen hingegen kaum. Deutlichere Auswirkungen traten bei der Nutzung von Berufskategorien anstelle der gröberen Kompetenzniveaus zutage. Verglichen mit den Ergebnissen von Felder & Wunsch (2021) fanden wir nur einen geringfügigen Rückgang der Betriebe, die in den Bereich der kritischen Lohnunterschiede fallen. Schliesslich prüften wir auch die Hinzunahme weiterer Variablen in das Modell. Die grösste Veränderung ergab sich hier durch die Aufnahme eines Indikators für Zulagen für Schicht-, Sonntags- oder Nachtarbeit. Der Lohnunterschied nach Geschlecht sinkt in dieser Spezifikation durchschnittlich um rund einen halben Prozentpunkt im Vergleich zum Standardmodell. Weiterführende Analysen zeigen, dass sich die Zulagen-Variable besonders stark auf Lohnanalysen von grösseren Firmen auswirkt. Weitere geprüfte Merkmale wie Beschäftigungsgrad, Stundenlohn, Zivilstand und Nationalität wirken sich im Durchschnitt nur relativ geringfügig auf den Geschlechtereffekt und die Logib-Testergebnisse von privaten Firmen aus. Demgegenüber fällt auf, dass die Hinzunahme weiterer Variablen den Geschlechtereffekt bei öffentlichen Firmen teilweise bedeutsam senkt. Beispielsweise reduziert sich der Anteil der öffentlichen Unternehmen, die den Logib-Test nicht bestehen, von rund 14% auf 11%, wenn der Indikator «Zulagen» in das Modell inkludiert wird.

Schliesslich bestand ein weiteres Ziel der Studie darin, das Ausmass des OVB indirekt abzuschätzen. Dazu kam die Methode von Oster (2019) zur Anwendung, die es erlaubt, mit Hilfe von statistischen Annahmen eine obere «Schranke» für den OVB zu berechnen. Der um OVB korrigierte Geschlechtereffekt liegt bei etwa -2,6% im Durchschnitt über alle Unternehmen. Dieses Ergebnis impliziert, dass die bestehenden Kontrollvariablen im Logib-Standardmodell rund drei Viertel des OVB im Vergleich zum Modell ohne Kontrollvariablen absorbieren. Der Vergleich zwischen dem «simulierten» Geschlechtereffekt anhand der Oster-Methode und dem Geschlechtereffekt aus dem Standardmodell suggeriert, dass die Bereinigung um OVB die Häufigkeit von nicht bestandenen Testergebnissen verringert. Dieser Rückgang ist insbesondere bei den öffentlichen Firmen besonders ausgeprägt. Dies deutet darauf hin, dass die potenziellen Verzerrungen aufgrund von OVB bei öffentlichen Firmen relevanter sind als bei privaten Firmen.

Zusammengenommen liefern die Untersuchungen wichtige Erkenntnisse, die in die Beurteilung und Weiterentwicklung von Logib Modul 1 einfliessen können. Zum einen können die Ergebnisse als Informationsgrundlage hinzugezogen werden zur Frage, welche Modifikationen des Standardmodells als sinnvoll erachtet werden. Zum anderen zeigen die Ergebnisse annäherungsweise auf,

wie gross der Einfluss durch verbleibende Verzerrungen ist und wie stark sich diese in den Testergebnissen von Logib niederschlagen.

Abschliessend ist auch auf drei Limitationen dieser Studie hinzuweisen. Erstens beruhen unsere Analysen nur auf Unternehmen, die mit mindestens 50 vollständigen Beobachtungen (d.h. ohne fehlende Werte) in der LSE vertreten sind. Damit dürften insbesondere Unternehmen zwischen 50 und 150 Beschäftigten unterrepräsentiert sein. Zweitens verfügt die LSE nicht über die für das Logib Modul 1 erforderliche Angabe zum firmenspezifisch kodierten Anforderungsniveau der ausgeübten Funktion. Als Ersatz greifen wir auf das in der LSE vorhandene Kompetenzniveau zurück, welches vermutlich einen geringeren Erklärungsgehalt für die innerbetriebliche Lohnstruktur besitzt. Drittens bleibt festzuhalten, dass der Ansatz von Oster (2019) trotz seiner Vorzüge nur eine Annäherung an den tatsächlichen OVB liefert und die quantitativen Ergebnisse basierend auf dieser Methode mit entsprechender Vorsicht zu interpretieren sind.

6 Quellenangaben

- Blinder, Alan S. 1973. „Wage Discrimination: Reduced Form and Structural Estimates“. *The Journal of Human Resources* 8(4):436–55. doi: 10.2307/144855.
- Cinelli, Carlos, und Chad Hazlett. 2020. „Making Sense of Sensitivity: Extending Omitted Variable Bias“. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82(1):39–67. doi: 10.1111/rssb.12348.
- EBG. 2021. *Standard-Analysemodell des Bundes: Methodenbeschrieb zu Logib Modul 1 und 2*. Bern: Eidgenössisches Büro für die Gleichstellung von Frau und Mann. https://www.ebg.admin.ch/dam/ebg/de/dokumente/lohngleichheit/logib/methodenbeschrieb_logib_m1_m2.pdf.download.pdf/Methodenbeschrieb%20Logib%20Modul%201%20und%202_DE.pdf.
- Felder, Rahel, und Conny Wunsch. 2021. „Evaluation des Analyse- Tools Logib Modul 1 für den Lohnvergleich zwischen Frauen und Männern“. *WWZ Insights No. 3*.
- Kaiser, Boris, und Thomas Möhr. 2021. *Analyse der Lohnunterschiede zwischen Frauen und Männern anhand der Schweizerischen Lohnstrukturerhebung (LSE) 2018*. BSS Volkswirtschaftliche Beratung. Studie im Auftrag des Bundesamts für Statistik.
- Kaiser, Boris, und Thomas Möhr. 2023. *Analyse der Lohnunterschiede zwischen Frauen und Männern anhand der Schweizerischen Lohnstrukturerhebung (LSE) 2020*. BSS Volkswirtschaftliche Beratung. Studie im Auftrag des Bundesamts für Statistik.
- Oaxaca, Ronald. 1973. „Male-Female Wage Differentials in Urban Labor Markets“. *International Economic Review* 14(3):693–709. doi: 10.2307/2525981.
- Oster, Emily. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics* 37.2 (2019): 187-204.

A. Anhang

A.1 Deskriptive Statistiken nach Merkmalen

Die folgenden Tabellen bieten einen Überblick zu den Merkmalen der im verwendeten Datensatz enthaltenen Firmen. Dabei unterscheiden wir zusätzlich auch verschiedene Typen von Firmen und ihren Belegschaften.

Tabelle 19: Deskriptive Statistiken nach Firmeneigentümerschaft

	(1)	(2)	(1)-(2)
	Öffentlich	Privat	
	Mittelwert (Std.abw.)	Mittelwert (Std.abw.)	Mittlere Differenz
Geschätzter Geschlechterkoeffizient (Standardmodell)	-0.042	-0.056	0.014***
	(0.003)	(0.002)	
Signifikanzrate 2. Test	14.219	16.938	-2.719
	(1.688)	(0.603)	
Medianlohn	7210.887	6663.105	547.782***
	(80.247)	(33.509)	
Bezahlt Zulagen	0.761	0.609	0.152***
	(0.021)	(0.008)	
Frauenanteil	0.627	0.418	0.209***
	(0.011)	(0.005)	
50-249 Mitarbeitende	0.480	0.676	-0.196***
	(0.024)	(0.007)	
250-999 Mitarbeitende	0.304	0.263	0.041*
	(0.022)	(0.007)	
1000+ Mitarbeitende	0.216	0.061	0.155***
	(0.020)	(0.004)	
Kollektive Lohnvereinbarung	0.415	0.319	0.096***
	(0.024)	(0.007)	
Mittleres Alter Belegschaft	44.501	42.301	2.200***
	(0.127)	(0.062)	
Mittlere Bildungsjahre	12.949	12.200	0.749***
	(0.071)	(0.028)	
N	431	3917	

Anmerkungen: Die Anzahl der Sterne in der letzten Spalte zeigt das Ergebnis eines paarweisen t-Tests, der prüft ob die mittlere Differenz statistisch von Null unterscheidbar ist. *p<0.1; **p<0.05; ***p<0.01. Quelle: LSE 2020, eigene Berechnungen.

Tabelle 20: Deskriptive Statistiken nach Geschlecht der Mitarbeitenden, private Firmen

	(1)	(2)	(1)-(2)
	Männliche MA	Weibliche MA	
	Mittelwert (Std.abw.)	Mittelwert (Std.abw.)	Mittlere Differenz
Bildungsjahre	12.409	12.153	0.256***
	(0.028)	(0.028)	
Pot. Erwerbserfahrung (Jahre)	24.156	23.518	0.638***
	(0.072)	(0.085)	
Dienstjahre	7.230	6.584	0.645***
	(0.060)	(0.057)	
Anteil Vollzeit	0.824	0.534	0.290***
	(0.003)	(0.004)	
Anteil Stundenbezahlung	0.920	0.902	0.018***
	(0.004)	(0.004)	
Anteil Zulagenerhalt	0.215	0.171	0.044***
	(0.005)	(0.005)	
Zivilstand: Anteil verheiratet	0.511	0.442	0.069***
	(0.002)	(0.003)	
Nationalität: Anteil Schweiz	0.596	0.669	-0.073***
	(0.004)	(0.004)	
N	3912	3842	

Anmerkungen: Fallzahlen nach Spalten unterschiedlich, da für Firmen ohne weibliche oder männliche Mitarbeitende kein Mittelwert berechnet werden kann. Die Anzahl der Sterne in der letzten Spalte zeigt das Ergebnis eines paarweisen t-Tests, der prüft ob die mittlere Differenz statistisch von Null unterscheidbar ist. *p<0.1; **p<0.05; ***p<0.01
Quelle: LSE 2020, eigene Berechnungen.

Tabelle 21: Deskriptive Statistiken nach Geschlecht der Mitarbeitenden, öffentliche Firmen

	(1)	(2)	(1)-(2)
	Männliche MA	Weibliche MA	
	Mittelwert (Std.abw.)	Mittelwert (Std.abw.)	Mittlere Differenz
Bildungsjahre	13.187	12.841	0.346***
	(0.073)	(0.070)	
Pot. Erwerbserfahrung (Jahre)	25.582	25.171	0.411
	(0.185)	(0.178)	
Dienstjahre	8.840	7.887	0.953***
	(0.151)	(0.115)	
Anteil Vollzeit	0.756	0.344	0.412***
	(0.008)	(0.008)	
Anteil Stundenbezahlung	0.982	0.955	0.027***
	(0.002)	(0.004)	
Anteil Zulagenerhalt	0.345	0.302	0.043*
	(0.016)	(0.017)	
Zivilstand: Anteil verheiratet	0.550	0.503	0.047***
	(0.006)	(0.004)	
Nationalität: Anteil Schweiz	0.776	0.807	-0.031***
	(0.009)	(0.006)	
N	429	431	

Anmerkungen: Fallzahlen nach Spalten unterschiedlich, da für Firmen ohne weibliche oder männliche Mitarbeitende kein Mittelwert berechnet werden kann. Die Anzahl der Sterne in der letzten Spalte zeigt das Ergebnis eines paarweisen t-Tests, der prüft ob die mittlere Differenz statistisch von Null unterscheidbar ist. *p<0.1; **p<0.05; ***p<0.01
Quelle: LSE 2020, eigene Berechnungen.

Tabelle 22: Deskriptive Statistiken nach Zahlung von Zulagen

	(1)	(2)	(1)-(2)
	Firmen ohne Zulagenzahlungen	Firmen mit Zulagenzahlungen	
	Mittelwert (Std.abw.)	Mittelwert (Std.abw.)	Mittlere Differenz
Geschätzter Geschlechterkoeffizient (Standardmodell)	-0.056	-0.054	-0.002
	(0.002)	(0.002)	
Signifikanzrate 2. Test	14.233	18.134	-3.902***
	(0.869)	(0.744)	
Medianlohn	6902.806	6605.891	296.915***
	(58.297)	(35.692)	
Frauenanteil	0.627	0.418	0.209***
	(0.011)	(0.005)	
50-249 Mitarbeitende	0.772	0.587	0.185***
	(0.010)	(0.009)	
250-999 Mitarbeitende	0.202	0.306	-0.104***
	(0.010)	(0.009)	
1000+ Mitarbeitende	0.026	0.106	-0.081***
	(0.004)	(0.006)	
Kollektive Lohnvereinbarung	0.306	0.343	-0.036**
	(0.011)	(0.009)	
Privatfirma	0.937	0.879	0.058***
	(0.006)	(0.006)	
Mittleres Alter Belegschaft	41.777	42.966	-1.188***
	(0.104)	(0.068)	
Mittlere Bildungsjahre	12.557	12.104	0.453***
	(0.047)	(0.031)	
N	1633	2715	

Anmerkungen: Die Anzahl der Sterne in der letzten Spalte zeigt das Ergebnis eines paarweisen t-Tests, der prüft ob die mittlere Differenz statistisch von Null unterscheidbar ist. *p<0.1; **p<0.05; ***p<0.01

Quelle: LSE 2020, eigene Berechnungen.

Tabelle 23: Deskriptive Statistiken nach Firmeneigentümerschaft

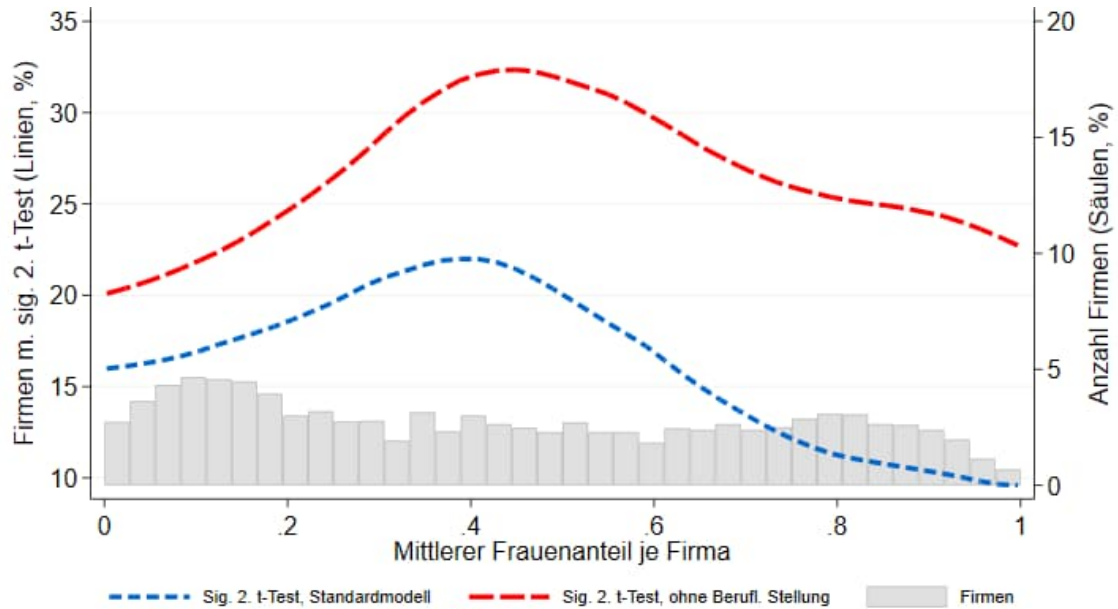
	(1)	(2)	(1)-(2)
	Öffentlich	Privat	
	Mittelwert (Std.abw.)	Mittelwert (Std.abw.)	Mittlere Differenz
Geschätzter Geschlechterkoeffizient (Standardmodell)	-0.058	-0.049	-0.008***
	(0.002)	(0.003)	
Signifikanzrate 2. Test	17.379	15.186	2.193*
	(0.704)	(0.961)	
Medianlohn	6904.894	6334.822	570.072***
	(41.136)	(43.227)	
Bezahlt Zulagen	0.612	0.650	-0.039**
	(0.009)	(0.013)	
Frauenanteil	0.453	0.411	0.042***
	(0.005)	(0.008)	
50-249 Mitarbeitende	0.658	0.655	0.002
	(0.009)	(0.013)	
250-999 Mitarbeitende	0.271	0.258	0.013
	(0.008)	(0.012)	
1000+ Mitarbeitende	0.071	0.087	-0.016*
	(0.005)	(0.007)	
Privatfirma	0.914	0.875	0.039***
	(0.005)	(0.009)	
Mittleres Alter Belegschaft	42.601	42.353	0.248**
	(0.071)	(0.103)	
Mittlere Bildungsjahre	12.487	11.841	0.646***
	(0.031)	(0.046)	
N	2918	1430	

Anmerkungen: Die Anzahl der Sterne in der letzten Spalte zeigt das Ergebnis eines paarweisen t-Tests, der prüft ob die mittlere Differenz statistisch von Null unterscheidbar ist. *p<0.1; **p<0.05; ***p<0.01

Quelle: LSE 2020, eigene Berechnungen.

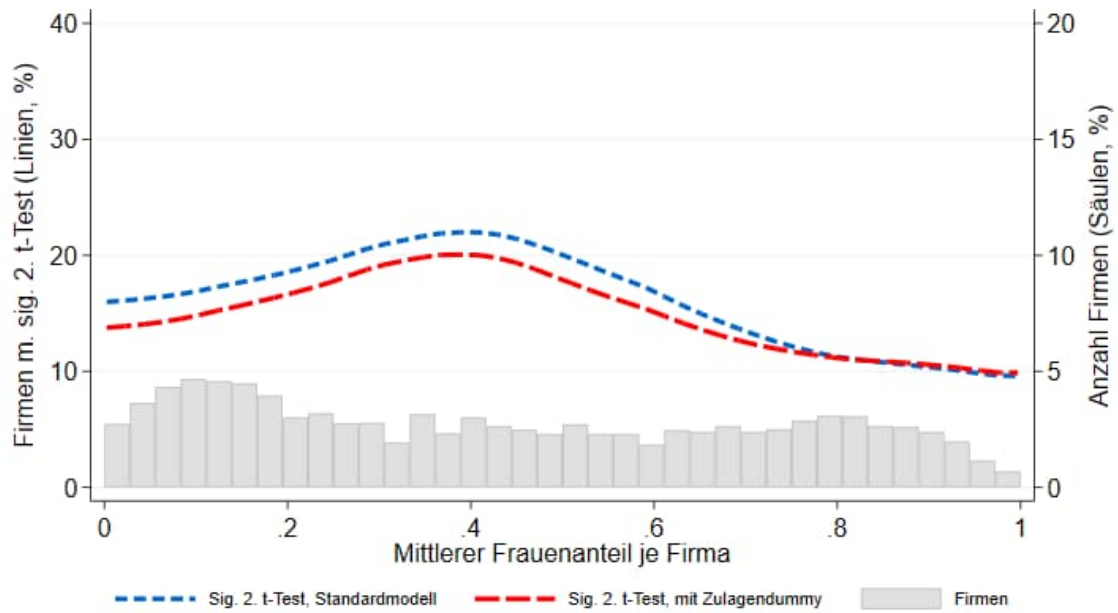
A.2 Unterschiede nach Frauenanteil der Firma

Abbildung 11: Signifikanzrate nach Frauenanteil, ohne berufliche Stellung, alle Firmen



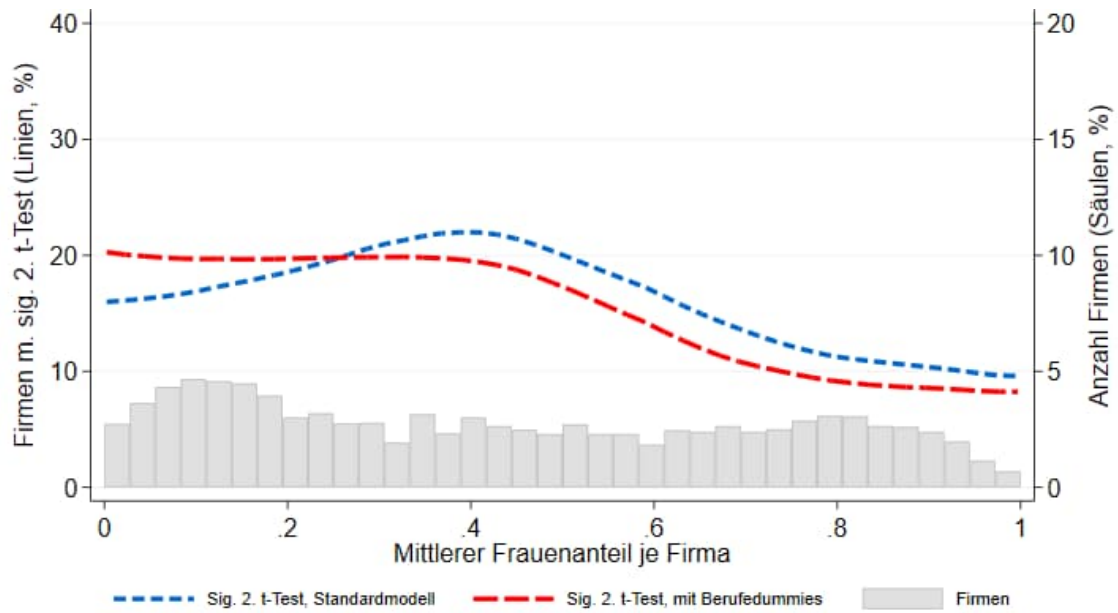
Anmerkungen: Die Kurven geben an, welcher Anteil von Firmen für den gegebenen Frauenanteil und die Modellspezifikation den Lohngleichheitstest nicht bestehen würde, also im t-Test der zweiten Stufe ein signifikantes Ergebnis hätte (linke Achse). Zu Grunde liegen Regressionsmodelle mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die Säulen entsprechen der Anzahl der Firmen. Quelle: Lohnstrukturhebung 2020, eigene Berechnungen.

Abbildung 12: Signifikanzrate nach Frauenanteil, Zulagen, alle Firmen



Anmerkungen: Die Kurven geben an, welcher Anteil von Firmen für den gegebenen Frauenanteil und die Modellspezifikation den Lohngleichheitstest nicht bestehen würde, also im t-Test der zweiten Stufe ein signifikantes Ergebnis hätte (linke Achse). Zu Grunde liegen Regressionsmodelle mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die Säulen entsprechen der Anzahl der Firmen. Quelle: Lohnstrukturhebung 2020, eigene Berechnungen.

Abbildung 13: Signifikanzrate nach Frauenanteil, Berufskategorien, alle Firmen



Anmerkungen: Die Kurven geben an, welcher Anteil von Firmen für den gegebenen Frauenanteil und die Modellspezifikation den Lohngleichheitstest nicht bestehen würde, also im t-Test der zweiten Stufe ein signifikantes Ergebnis hätte (linke Achse). Zu Grunde liegen Regressionsmodelle mit dem Logib-Standardmodell, die für alle Firmen mit mindestens 50 Mitarbeitenden separat geschätzt wurden. Die Säulen entsprechen der Anzahl der Firmen. Quelle: Lohnstrukturhebung 2020, eigene Berechnungen.

