

# Konzeptionelle Lehren aus der ersten Evaluationsrunde der Branchenmindestlöhne in Deutschland

Bernd Fitzenberger<sup>1</sup> · Annabelle Doerr<sup>2</sup>

Angenommen: 19. August 2016 / Online publiziert: 19. September 2016  
© The Author(s) 2016. This article is available at SpringerLink with Open Access.

**Zusammenfassung** Dieser Aufsatz liefert eine kritische Zusammenfassung und Beschreibung der vom Bundesministerium für Arbeit und Soziales (BMAS) in Auftrag gegebenen Evaluationsstudien zu den Branchenmindestlöhnen in Deutschland, wobei der Schwerpunkt auf den ökonometrischen Kausalanalysen liegt. Die Evaluation der Branchenmindestlöhne ist ein wichtiges und erfolgreiches Beispiel für eine evidenzbasierte Wirtschaftspolitik. Allen Evaluationsstudien gemein ist die Schwierigkeit der Identifikation der tatsächlich vom Mindestlohn betroffenen

Beschäftigten in den jeweiligen Branchen. Die identifizierenden Annahmen und die Umsetzung des verwendeten Differenz-von-Differenzen-Ansatzes (DvD-Ansatz) sowie die Wahl der Kontrollgruppen werden kritisch diskutiert und es werden alternative Ansätze aufgezeigt. In vielen Fällen hätte die Variation der geschätzten Effekte und die Eignung einer Kontrollgruppe stärker diskutiert und eine Abschätzung der Verzerrung in den geschätzten Effekten vorgenommen werden sollen.

Mindestlöhne · Differenz-von-Differenzen (DvD) · Evaluationsstudien

---

Ein Teil dieses Aufsatzes basiert auf den von den Autoren geschriebenen Teilen des Gutachtens „Vorbereitende Forschung für die zweite Evaluationsrunde Mindestlöhne – Verbesserung und Erweiterung der Evaluationsmethoden“, das das ZEW Mannheim zusammen mit den Autoren im Jahr 2013 im Auftrag des Bundesministeriums für Arbeit und Soziales erstellt hat (siehe ZEW, 2013). Wir danken zwei anonymen Gutachtern sowie dem betreuenden Herausgeber dieser Zeitschrift für hilfreiche Kommentare, die zu einer Verbesserung des Aufsatzes geführt haben. Weiterhin danken wir den TeilnehmerInnen des ZEW-Workshops „Verbesserung und Erweiterung der Methoden zur Mindestlohn-evaluation“ im Jahr 2013 und der IZA-Konferenz „A Minimum Wage for Germany? Learning from European Experience“ im Jahr 2014 sowie Melanie Arntz und zahlreichen weiteren KollegInnen, mit denen wir über die Evaluation der Wirkungen von Mindestlöhnen sprechen konnten, für hilfreiche Kommentare. Für diesen Beitrag tragen wir alleine die Verantwortung.

---

✉ Bernd Fitzenberger  
bernd.fitzenberger@wiwi.hu-berlin.de

Annabelle Doerr  
doerr@eucken.de

<sup>1</sup> Humboldt-Universität zu Berlin, Spandauer  
Straße 1, 10099 Berlin, Deutschland

<sup>2</sup> Walter Eucken Institut, Goethestraße 10, 79100 Freiburg,  
Deutschland

## Conceptual lessons from the evaluation studies on sectoral minimum wages in Germany

**Abstract** This paper provides a summary and a critical assessment of the evaluation studies on sectoral minimum wages in Germany, which were commissioned by the Federal Ministry of Labour and Social Affairs (BMAS). The evaluation of sectoral minimum wages is an important and successful example for the move towards evidence-based economic policy. All evaluation studies share the difficulty to identify the employees in the sectors which were actually covered by the sector specific minimum wage. We provide a critical discussion of the identifying assumptions, of the implementation of the difference-in-differences estimator, and of the choice of control groups. We discuss some alternative methodological approaches. We think that it would have been useful to analyze further the heterogeneity of the effect estimates and the choice of appropriate control groups. Furthermore, it would have been useful to assess the likely bias of the estimated effects.

## 1 Einleitung

... I do think the studies that attempt to exploit actual variation in minimum wages are considerably better than those simulations that purport to estimate the impact of minimum wages on employment without any actual data on minimum wages.

Manning (2013, S. 64)

Seit dem Inkrafttreten des Arbeitnehmerentsendegesetzes (AEntG) im Jahre 1996 können tarifvertragliche Bestimmungen per Allgemeinverbindlichkeitserklärung oder Rechtsverordnung auf alle Beschäftigten einer Branche ausgeweitet werden. Somit wurde eine rechtliche Grundlage zur Einführung von Branchenmindestlöhnen geschaffen. Während in der frühen Phase nur das Bauhaupt- und zwei Baunebengewerbe von dem Gesetz betroffen waren, wurden bis zum Jahre 2010 insgesamt 11 Branchen in das AEntG aufgenommen. Die Auswirkung des Mindestlohns auf die Beschäftigung, den Arbeitnehmerschutz und den Wettbewerb wurde im Auftrag des Bundesministeriums für Arbeit und Soziales (BMAS) und vor dem Hintergrund des §24 AEntG durch verschiedene Forschungseinrichtungen für acht Branchen untersucht. Die Ergebnisse lagen im November 2011 vor. Dieser Aufsatz liefert eine kritische Zusammenfassung und Beschreibung dieser Evaluationen der Branchenmindestlöhne, wobei der Schwerpunkt auf den ökonomischen Kausalanalysen liegt. Angesichts der Einführung des gesetzlichen Mindestlohns in Deutschland im Jahr 2015, soll dieser Aufsatz einen Beitrag zur Diskussion einer methodisch angemessenen Evaluation von Mindestlöhnen liefern.

Trotz aller in diesem Beitrag diskutierten Detailkritik an den Evaluationsstudien zu den Branchenmindestlöhnen ist festzuhalten, dass diese ein wichtiges und erfolgreiches Bei-

spiel für eine stärkere Evidenzbasierung der Wirtschaftspolitik in Deutschland darstellen, wie sie in der aktuellen Debatte von vielen Wirtschaftswissenschaftlern angemahnt wird (vgl. Boockmann et al. 2014; sowie zur Einordnung in die Mindestlohndebatte Manning 2013; und Fitzenberger 2009).

Einen Überblick über die in der ersten Evaluationsrunde betrachteten Branchen und der jeweiligen Evaluationsstudien befindet sich in Tab. 1. Die Ausgangslage der betrachteten Branchen war sehr heterogen sowohl im Hinblick auf die bestehenden Marktverhältnisse als auch bezogen auf den Zeitpunkt und die konkrete Ausgestaltung der Mindestlohneinführung. Ebenso heterogen war die für Evaluationszwecke zur Verfügung stehende Datenbasis. Die Studien zu den Branchen aus dem Bauhaupt- und Baunebengewerbe (Branchentyp 1) basieren auf umfangreichen administrativen Datensätzen. Dagegen konnten die Evaluationsstudien für Branchen mit einer kurzen Laufzeit des Mindestlohns (Branchen des Typs 2 oder 3) teilweise nur auf im Rahmen der Evaluationen erhobene Befragungsdaten zurückgreifen.

Allen Studien in Tab. 1 gemein ist die Schwierigkeit der Identifikation der tatsächlich vom Mindestlohn betroffenen Beschäftigten. Für einige Branchen ergaben sich Ungenauigkeiten bei der Abgrenzung der durch die tarifvertraglichen Regelungen betroffenen Branchen, weil die in den Sekundärdaten vorliegende Wirtschaftszweigklassifikation vielfach vom Geltungsbereich der Mindestlöhne abweicht. Zudem ist für die genaue Abgrenzung der betroffenen Beschäftigten eine präzise Messung der Stundenlöhne (ohne Sonderzahlungen) notwendig. Die meisten bisher verwendeten Sekundärdaten erfassen den Stundenlohn jedoch gar nicht oder nur unpräzise.

Trotz der heterogenen Ausgangslage im Hinblick auf die Mindestlohnregelungen in den einzelnen Branchen (Niveau, Einführung und Veränderungen der Mindestlöhne) und aufgrund begrenzter Datenverfügbarkeit basieren die Kausal-

**Tab. 1** Übersicht der Evaluationsstudien zu Branchenmindestlöhnen

Branchentyp <sup>a</sup>	Branche	Evaluationsstudie <sup>b</sup>	Forschungseinrichtungen der AutorInnen der Studien <sup>c</sup>
1	Bauhauptgewerbe	Möller et al. (2011)	IAB, RWI, ISG
1	Dachdeckerhandwerk	Aretz et al. (2011)	ZEW
1	Elektrohandwerk	Kirchmann et al. (2011a)	IAW
1	Maler- und Lackiererhandwerk	Kirchmann et al. (2011b)	IAW
2	Wäschereidienstleistungen im Objektkundengeschäft	Bosch et al. (2011b)	IAQ
2	Gebäudereinigung	Bosch et al. (2011a)	IAQ
3	Pflegebranche	Kirchmann et al. (2011c)	IAW
3	Abfallwirtschaft	Egeln et al. (2011)	ZEW

<sup>a</sup> Branchentyp ist eine in diesem Aufsatz vorgenommene Einteilung der acht Branchen in drei Typen, die im Text in Abschn. 3 erläutert wird,

<sup>b</sup> siehe Literaturliste; <sup>c</sup> IAB Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg; IAQ Institut Arbeit und Qualifikation; Duisburg-Essen; IAW Institut für angewandte Wirtschaftsforschung, Tübingen; ISG Institut für Sozialforschung und Gesellschaftspolitik, Köln; RWI Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Essen; ZEW Zentrum für Europäische Wirtschaftsforschung, Mannheim.

analysen in der ersten Evaluationsrunde fast ausschließlich auf Differenz-von-Differenzen (DvD) Ansätzen. Die Anwendung dieser Ansätze erlaubt unter bestimmten Annahmen die Schätzung der kausalen Wirkung eines Mindestlohns auf die zu untersuchenden Zielgrößen. Im Rahmen dieser Studie werden die Ergebnisse der ersten Evaluationsrunde im Hinblick auf die interne und externe Validität der ökonomischen Schätzungen der Mindestlohnwirkungen systematisch zusammengefasst und kritisch beurteilt, um daraus konzeptionelle Lehren zu ziehen. Im Sinne einer Evidenzbasierung der Wirtschaftspolitik halten wir eine Diskussion der Qualität von Studien nach deren Erstellung für notwendig – auch um zukünftige Evaluationen zu verbessern. Dies ist im Sinne einer *lernenden Evaluation* sinnvoll, die zu einem kontinuierlichen Monitoring der Wirkungen der Mindestlöhne fortgeführt werden kann, wie es exemplarisch von der Low-Pay-Commission in Großbritannien durchgeführt wird (Manning 2013).

Die in unserem Aufsatz diskutierten Themen und Schlussfolgerungen sind nicht nur für die Analyse der Branchenmindestlöhne relevant. Auch im Falle des gesetzlichen Mindestlohns wird es Unterschiede in der Betroffenheit zwischen Branchen, Regionen oder Firmen geben. Spillover-Effekte werden aufgrund der insgesamt höheren Betroffenheit eine noch größere Rolle spielen und die Datenlage wird vergleichbar sein. Zukünftige Evaluationen werden also mit ähnlichen Herausforderungen konfrontiert sein, wie bei der Wirkungsanalyse der Branchenmindestlöhne. Es ist deshalb zu erwarten, dass die hier diskutierten Ansätze zur Anwendung kommen. Gleichwohl ist jede Politikmaßnahme spezifisch und es gibt kein einheitliches Analysemodell, das in gleicher Weise auf alle Maßnahmen angewendet werden kann. Ein spezifisches Ziel unseres Aufsatzes ist die Diskussion von Kriterien für die Beurteilung der Eignung eines Analysemodells. In diesem Sinne liefert unser Aufsatz einen Beitrag für die Analyse zukünftiger Fragestellungen.

Der Aufsatz geht bewusst nicht im Detail auf die umfangreichen Forschungsarbeiten ein, die im Nachgang zu den Evaluationsstudien entstanden sind (im Sonderband dieser Zeitschrift über „Minimum wages in Germany“ im Jahr 2012 und im Sonderband des German Economic Review über „Economic Effects of Minimum Wages in Germany“ im Jahr 2013 sowie exemplarisch die Studien Felder 2014, Frings et al. 2013, oder Rattenhuber 2014). Diese Studien stellen einen äußerst begrüßenswerten Mehrwert der Evaluationsstudien dar, aber sie nehmen im Regelfall weder eine kritische Einschätzung der in den Gutachten verwendeten methodischen Ansätze vor noch greifen sie die in diesem Aufsatz gemachten methodischen Vorschläge

auf (die noch nicht veröffentlichte Arbeit von Felder 2014, ist hier eine Ausnahme).<sup>1</sup>

Im Folgenden werden der in den ökonomischen Kausalanalysen der Evaluationsstudien dominierende DvD-Ansatz und die zugrundeliegenden Annahmen erklärt. Anschließend folgen eine Systematisierung und Bewertung der vorliegenden Evaluationsergebnisse. Die Evaluationsstudien haben eine große Zahl von Ergebnisvariablen untersucht. Wir konzentrieren uns in diesem Beitrag ausschließlich auf eine konzeptionelle Einordnung der Ergebnisse für Beschäftigung und Löhne. Im Sinne einer übersichtlichen Darstellung werden die untersuchten Branchen in drei Obergruppen zusammengefasst. Diese Gruppierung berücksichtigt vorhandene Strukturähnlichkeiten zwischen den einzelnen Branchen. Abschließend fassen wir die Lehren aus der ersten Evaluationsrunde zusammen.

## 2 Methodische Konzepte der ersten Evaluationsrunde

Die Schätzung des kausalen Effekts des Mindestlohns auf die relevanten Zielgrößen wurde in den Evaluationsstudien, sofern es die Datenlage zuzulassen schien, mit Hilfe eines DvD-Ansatzes durchgeführt.<sup>2</sup> Um den kausalen Effekt einer Maßnahme/Treatment (hier: Einführung oder Veränderung von Mindestlöhnen) zu bestimmen, wird wie bei einem klassischen Experiment eine Gruppe von Betroffenen mit einer Gruppe von Nichtbetroffenen vor und nach Einführung der Maßnahme hinsichtlich einer vorher definierten Zielgröße (Ergebnisvariable) verglichen. Der Vorteil des DvD-Ansatzes besteht darin, dass Zeittrends oder strukturelle Einflüsse, die beide Vergleichsgruppen gleichermaßen betreffen, durch Differenzenbildung herausgerechnet werden.

<sup>1</sup> Es ist natürlich sinnvoll, dass sich die Autoren der Evaluationsstudien erfolgreich einer wissenschaftlichen Qualitätskontrolle stellen. Allerdings werden nicht alle wirtschaftspolitischen Evaluationsstudien in wissenschaftlichen Zeitschriften publiziert, selbst wenn sie wissenschaftlich fundiert sind. Auch, um eine selektive Hervorhebung veröffentlichter Studien zu vermeiden, fokussieren wir uns auf die durch das BMAS veröffentlichten Studien, die direkt in die Politikberatung eingeflossen sind.

<sup>2</sup> Dieser Abschnitt kann keinen Überblick über die Methoden zur ökonomischen Schätzung von Kausaleffekten liefern. Hierzu sei bspw. auf die Überblicksartikel von Imbens und Wooldridge (2009), Lechner (2010), Heckman et al. (1999) oder das Lehrbuch Bauer et al. (2009), sowie die dort zitierte Literatur verwiesen. Der Übersichtsartikel Fitzenberger und Hujer (2002) diskutiert die ökonomische Schätzung von Kausaleffekten im Kontext der Analyse der Wirkungen der aktiven Arbeitsmarktpolitik in Deutschland. Einen Überblick zu den Methoden und Ergebnissen der Mindestlohnforschung im internationalen Kontext geben Card und Krueger (1995) sowie Neumark and Wascher (2008).

Die Umsetzung des Verfahrens verlangt eine präzise Abgrenzung der vom Mindestlohn betroffenen Beobachtungseinheiten (Treatmentgruppe) und geeigneter nicht betroffener Kontrolleinheiten (Kontrollgruppe). Die Differenz zwischen der Veränderung der Zielgröße (z. B. Lohn oder Beschäftigung) für die Treatmentgruppe und der Veränderung für die Kontrollgruppe schätzt den kausalen Effekt der Maßnahme. Der Identifikation des kausalen Maßnahmeeffektes liegt die Annahme zugrunde, dass die beobachtete Veränderung für die Kontrollgruppe der kontrafaktischen Veränderung für die Treatmentgruppe entspricht, wenn die Maßnahme nicht durchgeführt worden wäre. Die sogenannte *Annahme paralleler Trends* impliziert eine Parallelität in der Entwicklung der Zielgröße zwischen Treatment- und Kontrollgruppe, wenn das Treatment nicht stattgefunden hätte. Ist die Annahme plausibel, entsteht erst durch das Treatment ein Unterschied in der Entwicklung der Zielgröße. Wenn die zeitliche Veränderung zusätzlich von Eigenschaften der Beobachtungseinheiten abhängt, dann kann ein konditionaler DvD-Schätzer zur Anwendung kommen. Hierfür werden zunächst Treatment- und Kontrollgruppe in den relevanten beobachteten Kontrollvariablen vergleichbar gemacht. Dies erfolgt zum Beispiel indem im Rahmen einer Regression die Kontrollvariablen als erklärende Variablen berücksichtigt werden oder indem durch Matching eine Kontrollgruppe bestimmt wird, die der Treatmentgruppe im Durchschnitt in den Kontrollvariablen entspricht.<sup>3</sup>

Politikmaßnahmen wie die Mindestlohneinführung werden in diesem Ansatz als natürliches Experiment aufgefasst. Da eine Datenbasis zur kontrafaktischen Situation, d. h. einer Situation ohne Mindestlohn nicht existiert, kann die Annahme paralleler Trends nicht formal getestet werden. Werden zwei Branchen miteinander verglichen, kann eine Verletzung der Annahme aus einem makroökonomischen Schock resultieren, der die Branchen in unterschiedlicher Weise trifft. Eine Verletzung der Annahme bei Analysen auf Betriebsebene kann exemplarisch im Hinblick auf das Innovationsverhalten diskutiert werden. Wenn innovationsfreudige Betriebe im stärkeren Umfang Beschäftigte durch Maschinen ersetzen und gleichzeitig höhere Löhne zahlen, dann sind innovationsfreudigere Betriebe weniger vom Mindestlohn betroffen und weisen eine unterdurchschnittliche Beschäftigungsentwicklung auf. Für das Beispiel folgt,

<sup>3</sup> Eine konditionale DvD-Schätzung kann bspw. mit folgender Regression umgesetzt werden

$$E_{jt} = \gamma_0 + \gamma_1 Y_{1,t} + \theta_0 D_j + \theta^{\text{DvD}} Y_{1,t} D_j + \delta X_{jt} + \varepsilon_{jt},$$

wobei  $t = 0, 1$  Zeitperiode (vor/nach Mindestlohneinführung/-erhöhung),  $E_{jt}$  Zielgröße für Beobachtungseinheit  $j$  in Periode  $t$ ,  $Y_{1,t}$  Zeitdummy für Periode 1,  $D_j$  Dummy für Treatmentgruppe,  $X_{jt}$  Kontrollvariablen und  $\varepsilon_{jt}$  der Fehlerterm. Der Koeffizient des Interaktionseffektes zwischen Zeit und Treatmentgruppe,  $\theta^{\text{DvD}}$ , liefert die DvD-Schätzung, da er den Unterschied in der zeitlichen Veränderung zwischen Treatment- und Kontrollgruppe beschreibt.

dass ein DvD-Ansatz die Beschäftigungseffekte des Mindestlohns zu positiv einschätzt, falls den Unterschieden im Innovationsverhalten nicht Rechnung getragen wird.

Die Plausibilität der Annahme paralleler Trends wird häufig mit Hilfe von sogenannten Placebo-Tests überprüft. Bei Placebo-Tests werden fiktive Treatments simuliert, die zu einem anderen Zeitpunkt als das echte Treatment stattfinden. Das Schätzergebnis, das unter oben vorgestellten Annahmen den kausalen Treatmenteffekt misst, wird typischerweise für einen Zeitraum vor der tatsächlichen Maßnahme (typischerweise vor der Mindestlohneinführung) berechnet. Ergeben sich dabei signifikante Effekte, obwohl das Treatment noch gar nicht stattgefunden hat, gibt dies einen Hinweis auf eine Verletzung der Annahme paralleler Trends. Da der geschätzte Effekt nicht auf das Treatment zurückgeführt werden kann, muss es andere Unterschiede zwischen der Treatment- und der Kontrollgruppe geben, die eine unterschiedliche Entwicklung in der Zielgröße verursachen. Allerdings sind in diesem Zusammenhang mögliche Antizipationseffekte zu berücksichtigen, die dazu führen können, dass bereits vorab signifikante Auswirkungen des Treatments gemessen werden können. So ist es möglich, dass die Betroffenen bereits im Vorfeld auf ein bekannt gewordenes Treatment reagieren und Anpassungen in der Zielgröße vornehmen.

Ein fundamentaler Kritikpunkt an der Annahme eines natürlichen Experimentes ergibt sich aus der Möglichkeit der Politikendogenität der branchenspezifischen Mindestlohnregelungen in Deutschland. Es ist denkbar, dass die Tarifparteien in einer Branche gerade deshalb eine Mindestlohnregelung erwirken, weil sie für diese Branche eine bestimmte Entwicklung erwarten, die von der Entwicklung in den Kontrollgruppen abweicht, auch wenn in der Vergangenheit parallele Trends beobachtet wurden. Dieser Aspekt könnte bspw. zum Zeitpunkt der Einführung des Branchenmindestlohns im Bauhauptgewerbe 1997 eine Rolle gespielt haben. Damals war absehbar, dass sich die Baukonjunktur vor allem in Ostdeutschland deutlich und nachhaltig abschwächen würde. Gleichzeitig wurde mit einer weiteren Zunahme an ausländischen Beschäftigten gerechnet (König und Möller 2008, Abschn. 2). Das mögliche Vorliegen von Politikendogenität stellt die Anwendung eines DvD-Ansatzes basierend auf einer Kontrollgruppe aus anderen Branchen in Frage, wenn für die Vergleichsbranchen nicht die gleiche negative Beschäftigungsentwicklung für die Zukunft erwartet wurde. In diesem Fall würde ein DvD-Ansatz den Beschäftigungseffekt des Mindestlohns zu negativ einschätzen.

Der Skalierung der Zielgröße kommt im Rahmen des DvD-Ansatzes eine besondere Bedeutung zu, da die Annahme paralleler Trends nicht gleichzeitig für eine lineare und eine nicht-lineare Skalierung der Zielgröße gelten kann. Bei nicht-linearen Transformationen besteht die Möglich-

keit, dass die Annahme paralleler Trends nicht gleichzeitig für eine Variable vor und nach der Skalierung zutrifft. Somit hängt die Plausibilität der identifizierenden Annahme davon ab, die „richtige Skalierung“ gefunden zu haben (Lechner 2010).

Lechner (2010) diskutiert den von Athey und Imbens (2006) vorgeschlagenen Changes-in-Changes (CiC) Ansatz als alternatives, skalierungsunabhängiges und nichtparametrisches Verfahren. Dabei wird die kontrafaktische Verteilung der Zielgröße für die Treatmentgruppe nach Einführung des Mindestlohns auf Basis der Verteilung für die Kontrollgruppe zu diesem Zeitpunkt und der Differenz der Verteilungen vor Einführung des Mindestlohns geschätzt. Allerdings treten bei diesem Verfahren Umsetzungsschwierigkeiten bei diskret verteilten Zielgrößen auf (z. B. Beschäftigungsvariablen). Lechner (2010, Abschn. 3.2.5) vermutet, dass Schwierigkeiten bei der Verwendung von Kontrollvariablen („Fluch der Dimensionen“) erklären, warum der CiC-Ansatz bis heute nur selten angewandt wurde (er benennt eine frühere Version der Arbeit von Havnes und Mogstad 2015, als einzige Ausnahme neben der ursprünglichen Studie von Athey und Imbens 2006).<sup>4</sup>

Eine weitere zentrale Annahme des DvD-Ansatzes ist die sogenannte *SUTVA-Annahme* (Stable Unit Treatment Value Assumption). Der Treatmentstatus einer Beobachtungseinheit darf keinen Einfluss auf die Zielgrößen der Kontrollgruppen haben. Zu Verletzungen der Annahme kommt es, wenn Spillover-Effekte in Löhnen oder Substitutions- oder Skaleneffekte in der Arbeitsnachfrage auftreten (Franz 2013).

Substitutionseffekte treten auf, wenn relativ teurer gewordene Arbeitskräfte durch Beschäftigte aus der Kontrollgruppe ersetzt werden und somit die Gesamtbeschäftigung in der Kontrollgruppe steigt. Skaleneffekte ergeben sich dadurch, dass die durch die Mindestlohnregelungen ausgelöste Erhöhung der Produktionskosten dazu führt, dass die produzierte Menge und damit die eingesetzte Beschäftigung zurückgehen. Der negative Skaleneffekt ist umso stärker, je weniger die Unternehmen dazu in der Lage sind, Lohn-

kostenerhöhungen auf die Absatzpreise zu überwälzen. Die Überwälzungsspielräume sind umso geringer, je preisreagibler die Nachfrage nach den produzierten Gütern ist. Substitutions- und Skaleneffekte in der Beschäftigung können wiederum indirekt eine Reaktion der Löhne für die betroffenen Beschäftigten durch Veränderung in den Knappheitsrelationen am Arbeitsmarkt auslösen.

Unter Spillover-Effekten im engeren Sinn, die nicht aus der Arbeitsnachfrage erklärt werden können, sind direkte Lohneffekte für vom Mindestlohn nicht direkt betroffene Beschäftigte aufzufassen. Diese können zum Beispiel auftreten, wenn eine betriebliche Lohndifferenzierung aus Motivationsgründen für die Beschäftigten auch nach der Mindestlohneinführung aufrechterhalten wird und daher die Löhne oberhalb des Mindestlohns steigen.

Eine Verletzung der identifizierenden Annahmen führt zu einer Verzerrung der Schätzergebnisse, die dann nicht mehr kausal interpretiert werden können. Gleichwohl argumentieren wir, dass unter plausiblen Annahmen eine Einschätzung der Richtung der Verzerrung der geschätzten Kausaleffekte vorgenommen werden kann (bspw. dominiert im Regelfall der Substitutionseffekt bei einer Kontrollgruppe aus der gleichen Branche den Skaleneffekt, siehe Fitzenberger 2009).

Im Folgenden werden die acht Studien aus der ersten Evaluationsrunde der Mindestlöhne hinsichtlich der genannten methodischen Aspekte beleuchtet und die Ergebnisse vor diesem Hintergrund kritisch bewertet.

### 3 Systematisierung, Darstellung und Bewertung der Evaluationsergebnisse

Für eine übersichtliche Darstellung der Bewertung werden die im Rahmen der Evaluation untersuchten Branchen in drei Branchentypen zusammengefasst (vgl. Tab. 1). Die Gruppierung berücksichtigt vorhandene Strukturähnlichkeiten zwischen den einzelnen Branchen. Die Tab. 2–5 geben einen systematischen Überblick der methodischen Ansätze und Ergebnisse.

#### 3.1 Ergebnisse für Branchentyp 1

##### 3.1.1 Kontrollgruppen

Zum Branchentyp 1 gehören vier Branchen, die traditionelle Handwerke mit Facharbeiterdominanz, hohen Qualitätsstandards und starker Regulierung umfassen. Darunter fallen das Bauhauptgewerbe, das Dachdeckerhandwerk, das Malerei- und Lackiererhandwerk sowie das Elektrohandwerk. Mindestlöhne wurden in diesen Branchen schon ab 1997 durch das AEntG eingeführt, zwischenzeitlich stark modifiziert und zeitweise in Abhängigkeit von der Bran-

<sup>4</sup> Aktuelle Arbeiten von Melly und Santangelo (2015) zeigen, dass es für stetige Zielgrößen mit Hilfe von (bedingten) Quantilsregressionen unter bestimmten Annahmen (z. B. bedingte Rankinvarianz) möglich ist, die bedingte kontrafaktische Verteilung dieser Variablen zu schätzen. Aus diesen bedingten Schätzungen lassen sich die Schätzungen der unbedingten kontrafaktischen Verteilungen ableiten. Die verwendeten funktionalen Formannahmen sind dabei nicht stärker als beim DvD-Ansatz. Die Studie von Havnes und Mogstad (2015, Abschn. 5.3.4) umfasst eine erfolgreiche und gut publizierte Anwendung der CiC-Methode als Robustheitscheck – allerdings ohne auf Kovariate zu konditionieren. Die Studie verdeutlicht, dass auch nichtkonditionale oder nichtparametrische DvD-Quantilsregressionen verwendet werden können. Mit Hilfe der geschätzten Quantileffekte lassen sich die durchschnittlichen Effekte bestimmen und in einer Robustheitsanalyse mit den DvD-Schätzungen für Mittelwerte vergleichen.

chenkonjunktur wieder ausgesetzt. Die sich hieraus ergebende Variation in den Mindestlohnregelungen, konnten die Evaluationsstudien zur Identifikation der Mindestlohneffekte nutzen.

Eine systematische Übersicht der Studien für den Branchentyp 1 befindet sich in Tab. 2 und 3. Alle vier Studien verfolgen ähnliche Kontrollgruppenansätze. Zum einen wird eine passende Kontrollbranche verwendet, die nicht einem allgemeinverbindlichen Mindestlohn unterliegt. Zum anderen werden Beschäftigte und Betriebe aus der gleichen Branche verwendet, die nicht oder nur gering durch den Mindestlohn betroffen sind.

Die Studie zum Bauhauptgewerbe nutzt zur Auswahl geeigneter Kontrollbranchen ein quantitatives Prüfverfahren. Ausschlaggebend für die Auswahl der Kontrollbranchen ist eine starke Ähnlichkeit in der Lohn- und Beschäftigungsentwicklung zwischen dem Bauhauptgewerbe und den potenziellen Vergleichsbranchen. Dabei werden jeweils für West- und Ostdeutschland vor- und nachgelagerte Branchen sowie eine baunahe und bauferne Vergleichsbranche ausgewählt. Es ist anzumerken, dass Spillover-Effekte auch bei vor- oder nachgelagerten Branchen über Input-Output-Beziehungen auftreten können, wenn auch in einem geringeren Umfang als innerhalb der direkt von der Mindestlohnregelung betroffenen Branche. Ob die ausgewählten Kontrollbranchen geeignet sind, den kausalen Effekt der Politikmaßnahme zu identifizieren hängt u. a. davon ab, ob die Annahme paralleler Trends für die jeweiligen Kontrollbranchen plausibel ist. Im Rahmen der Studie werden zahlreiche Placebo-Tests durchgeführt, um die Plausibilität der Annahme zu rechtfertigen.

Im Rahmen der Evaluation des Mindestlohns im Dachdeckerhandwerk werden drei verschiedene Kontrollbranchen ausgewählt und hinsichtlich einer Reihe von Kennzahlen zur Beschreibung der konjunkturellen Situation und des Wettbewerbs auf dem Arbeitsmarkt verglichen. Letztendlich zeigen sich zwischen dem Dachdeckerhandwerk und der Installationsbranche die meisten Ähnlichkeiten. Ein quantitatives Prüfverfahren zur Kontrollgruppenauswahl wird nicht angewandt. Die Studien zu Mindestlöhnen im Elektrohandwerk und im Maler- und Lackiererhandwerk nutzen mehrere potenzielle Kontrollgruppen. Diese werden nicht auf Basis eines quantitativen Prüfverfahrens ermittelt, sondern ihre Auswahl wird sachlogisch begründet.

Somit ist die Studie zum Bauhauptgewerbe die einzige Studie, die ein quantitatives Prüfverfahren zur Auswahl einer oder mehrerer geeigneter Kontrollgruppen nutzt. Der in der internationalen Forschungsliteratur viel beachtete synthetische Kontrollgruppenansatz (Abadie et al. 2010; 2015; Dube and Zipperer 2015) verwendet die vorhandene Variation in den Daten, um geeignete Kontrollgruppen zu finden. Bei dieser Methode basiert die Auswahl geeigneter Kontrollgruppen nicht auf sachlogischen Überlegungen, son-

dern die Daten bestimmen auf Basis eines geeigneten statistischen Kriteriums die ähnlichsten Kontrollgruppen durch eine entsprechende Gewichtung. Gegenüber der diskretionären Auswahl der Kontrollbranchen in den Gutachten zu den Branchenmindestlöhnen hat der synthetische Kontrollgruppenansatz den Vorteil, dass mehrere Kontrollbranchen mit entsprechender Gewichtung gepoolt werden können, um einen besseren Vergleichsmaßstab (eine bessere synthetische Kontrollgruppe) zu ermitteln.<sup>5</sup> Das im Bauhauptgewerbe angewandte Prüfverfahren gewichtet hingegen alle betrachteten Kontrollgruppen gleich.

Alle dem Branchentyp 1 zugehörigen Studien verwenden auch Personen innerhalb der jeweiligen Branche als Vergleichsgruppe. Während diese in den Studien zum Dachdeckerhandwerk und zum Bauhauptgewerbe aus allen Personen besteht, die bereits vor der Mindestlohneinführung bzw. Mindestlohnerhöhung mindestens zum (neuen) Mindestlohn entlohnt wurden, verwenden die Studien zum Elektro-, Maler- und Lackiererhandwerk nur Personen, die

<sup>5</sup> Es werden statistische Verfahren zur Maximierung der Prognosegüte für die Wahl der Gewichtungsparameter genutzt (analog zu Kreuzvalidierungsverfahren zur Bandweitenwahl bei nichtparametrischen Schätzungen). Dabei besteht die Gefahr eines statistischen Overfittings, wenn für das Verfahren zu viel Flexibilität zugelassen wird, um die Vergleichbarkeit vor Einführung des Mindestlohns zu erhöhen (vgl. die kritischen Ausführungen in Kaul et al. (2016) im Hinblick auf die Prädiktion der Ergebnisvariablen vor dem Treatment). Aus statistischer Sicht ist es wichtig, für die Überprüfung der Prognosegüte ein hinreichend großes Validierungssample von Beobachtungen für die Mindestlohnbranche vor der Einführung oder Veränderung des Mindestlohns verwenden zu können, um die Prognosegüte zu optimieren. Die Studie von Felder (2014) zur Schätzung der Mindestlohnwirkungen im Dachdeckerhandwerk zeigt exemplarisch die Schwierigkeiten in der Umsetzung der Methode synthetischer Kontrollgruppen auf, wenn die Anzahl der Kontrollbranchen und die Anzahl der Zeitperioden vor der Mindestlohneinführung nicht sehr groß ist. Die von Felder (2014) identifizierte optimale Gewichtung der Kontrollbranchen weicht stark von den in der BMAS Studie verwendeten Gewichten ab. Die Studie findet für unterschiedliche Kontrollbranchen unterschiedliche Schätzergebnisse, was nicht per se gegen die Methode spricht. Weiterhin legen Placebo-Schätzungen nahe, dass die statistische Streuung der geschätzten Treatmenteffekte sehr groß ist. Die Problematik hängt mit der geringen Stichprobengröße zusammen. Möglicherweise könnte in der Studie von Felder (2014) die Auswahl der synthetischen Kontrollgruppen durch weitere Überprüfung der Prognosegüte und durch weitere Placebo-Tests robuster werden und im Ergebnis eindeutiger Ergebnisse liefern. Das ist an dieser Stelle spekulativ, da sich die Arbeit laut Auskunft der Autorin in Überarbeitung befindet. Die Auswahl der Kontrollbranche ist eine wichtige Dimension der Schätzung des Vergleichsmaßstabs. Während einfache Standardanwendungen des synthetischen Kontrollgruppenansatzes (Abadie et al. 2015; Dube and Zipperer 2015; Felder 2014) gewichtete Vergleichsgruppen für Aggregate (Branchen, Staaten) schätzen, untersuchen die Mindestlohngutachten Unterschiede in der Betroffenheit der Firmen oder Regionen in einer Branche, um die Mindestlohneffekte zu identifizieren. Felder (2014) beschränkt sich darauf eine synthetische Vergleichsbranche für das gesamte Dachdeckerhandwerk in Ostdeutschland bzw. Westdeutschland zu ermitteln, um dann aggregierte Ergebnisse zu vergleichen.

vor Einführung des Mindestlohns mindestens diesen und maximal 15 % darüber verdienten. Diese Vorgehensweise ist vor dem Hintergrund möglicher Spillover-Effekte in diesem Bereich der Lohnverteilung zu bewerten. Die gewählte Lohngrenze scheint ad hoc festgelegt worden zu sein. Hier stellt sich die Frage, ob und in welcher Weise die Ergebnisse von der Höhe der Lohngrenze abhängen.

Zudem stellt sich das Abgrenzungsproblem a fortiori, da keine Stundenlöhne beobachtet werden. In diesem eng definierten Bereich der Lohnverteilung besteht bei einer unpräzisen Beobachtung der Stundenlöhne die Gefahr einer fälschlichen Zuweisung in die Treatment- oder Kontrollgruppe. Wenn Spillover-Effekte auftreten, ist die zweite zentrale Annahme des DvD-Ansatzes verletzt, da auch die Kontrollgruppe indirekt durch die Mindestlohneinführung betroffen ist. Dies ist beispielsweise denkbar, wenn aus Motivationsgründen die betriebliche Lohnspreizung nach Einführung des Mindestlohns aufrechterhalten wird und deshalb auch über dem Mindestlohn liegende Löhne steigen. Ebenso können Substitutions- und Skaleneffekte in der Arbeitsnachfrage eine Rolle spielen.

Auf Betriebsebene werden Betriebe als Kontrollbetriebe verwendet, die über dem gültigen Mindestlohniveau entlohnen bzw. bei denen kein mindestlohnbedingter Anstieg der Lohnkosten dokumentiert wurde. Die Evaluationsstudien zu Mindestlöhnen in der Elektrobranche und im Maler- und Lackiererhandwerk verwenden zusätzlich Beschäftigte aus Betrieben dieser Branchen, die schon vor der Mindestlohneinführung tarifgebunden waren. Neben möglichen Substitutionseffekten besteht im Rahmen dieses Ansatzes auch das Risiko der Endogenität der Tarifbindung, d.h. Betriebe, die schon vor Einführung des Mindestlohns tarifgebunden waren, können im Hinblick auf die betrachteten Zielgrößen eine Selektion von Betrieben darstellen.

### 3.1.2 Methoden und Daten

Alle Studien für den Branchentyp 1 können auf eine solide und umfangreiche Datenbasis zurückgreifen, die eine aussagekräftige Wirkungsanalyse möglich macht. Für das Bauhauptgewerbe und für das Dachdeckerhandwerk wurden im Rahmen der Evaluation neue administrative Daten zur Verfügung gestellt, die eine Berechnung der Stundenlöhne ermöglichen.

Grundsätzlich besteht die Möglichkeit den DvD-Ansatz zu einem DvDvD-Ansatz zu erweitern, wenn eine weitere Kontrollgruppe zur Verfügung steht. Die Studie zu Mindestlöhnen im Dachdeckerhandwerk nutzt den erweiterten DvD-Ansatz, um branchenspezifische Effekte auszuschließen und die Robustheit der Ergebnisse zu überprüfen. Dieser Ansatz unterstellt eine Parallelität in den Veränderungen der Zeittrends, eine Annahme deren Plausibilität schwieriger als für einfache DvD-Verfahren durch Placebo-Tests zu

überprüfen ist. Gleichwohl wäre es hier naheliegend gewesen Placebo-Tests durchzuführen.

In der Evaluation zum Bauhauptgewerbe wird zusätzlich eine Variation des DvD-Ansatzes verwendet, die die Mindestlohneffekte als Funktion der Eingriffsintensität modelliert. Der sogenannte inkrementelle DvD-Schätzer (Dolton et al. 2010; 2012) nutzt als identifizierende Variation die z.B. regionalen Unterschiede in der Veränderung der Eingriffsintensitäten über die Zeit. Um für strukturelle Unterschiede zwischen Betroffenen und Nichtbetroffenen (z.B. Regionen) zu kontrollieren, wird die aktuelle Eingriffsintensität  $P_{jt}$  vor und nach der Mindestlohneinführung in die Schätzung aufgenommen. Der stufenweise (inkrementelle) Effekt der tatsächlichen Eingriffsintensität nach der Mindestlohneinführung wird als Interaktionseffekt von  $P_{jt}$  mit einer Jahresdummy geschätzt. Dieser kumulative Effekt wird als Effekt der Mindestlohnregelung relativ zu dem Zeitraum vor der Mindestlohneinführung interpretiert (siehe folgende Fußnote für eine formale Darstellung).<sup>6</sup> Der geschätzte Koeffizient des Haupteffekts von  $P_{jt}$  kontrolliert für strukturelle Unterschiede zwischen Betroffenen und Nichtbetroffenen.

Die Ausnutzung der regionale Variation der Eingriffsintensitäten nach Kontrolle für strukturelle Unterschiede ist grundsätzlich sinnvoll. Allerdings sind im Rahmen dieses Ansatzes folgende Probleme zu beachten: Erstens besteht vermutlich ein Endogenitätsproblem der Veränderung der Eingriffsintensität über die Zeit, d.h. die Veränderung in der Eingriffsintensität im Zeitverlauf wird von den Mindestlohnregelungen in der Vergangenheit beeinflusst. Zweitens reflektieren die kumuliert geschätzten Effekte die Wirkungen unterschiedlicher Mindestlohnregelungen (z.B. Einführung und spätere Erhöhung) für eine sich im Zeitverlauf verändernde Treatmentgruppe. Dies stellt, wie auch Dolton et al. (2010), (2012) konstatieren, eine kausale Interpreta-

<sup>6</sup> Formal basiert der in der Evaluationsstudie für das Bauhauptgewerbe umgesetzte inkrementelle DvD-Schätzer auf folgender Regression (siehe auch Fußnote 3)

$$E_{jt} = J_j + \sum_{t=93}^{08} \gamma_t Y_t + \theta_0 P_{jt} + \sum_{t=97}^{08} \theta_t^{\text{DID}} Y_t P_{jt} + \delta X_{jt} + \varepsilon_{jt},$$

wobei  $J_j$  Dummies für Beobachtung  $j$  (z.B. Betrieb oder Region),  $Y_t$  Jahresdummies und  $P_{jt}$  die Eingriffsintensität des Mindestlohns (bspw. der Anteil der Beschäftigten in einer Region/einem Betrieb mit einem Lohn unterhalb des Mindestlohns oder am Mindestlohn) darstellen.  $P_{jt}$  misst den Effekt der Einführung bzw. Veränderung des Mindestlohns, d.h. die Wirkung des Mindestlohns geht über die Eingriffsintensität. Der sogenannte durchlaufende Effekt der Eingriffsintensität  $\theta_0 P_{jt}$  soll für strukturelle Unterschiede zwischen den Betroffenen und Nichtbetroffenen kontrollieren, die auch schon vor Mindestlohneinführung vorhanden waren.  $\theta_t^{\text{DID}}$  beschreibt den kumulierten Mindestlohneffekt relativ zur Zeit vor Einführung der Mindestlohns. Dolton et al. (2010, 2012) sprechen sich dafür aus, nur die inkrementellen Veränderungen von  $\theta_t^{\text{DID}}$  zwischen zwei benachbarten Zeitpunkten zu interpretieren.

**Tab. 2** Übersicht der Studien zum Branchentyp 1 – Teil 1

	Bauhauptgewerbe (IAB, RWI, ISG)	Dachdeckerhandwerk (ZEW)
Mindestlohnregelung	Mindestlohn I (ML I): Einführung 1997 Mindestlohn II (ML II): Einführung 2003, Abschaffung 2009 (Ost) Regelmäßige Erhöhungen, starke Erhöhung 1999	Einführung 1997 Erneuerungen zwischen 2000 und 2010 Zeitweise Perioden ohne Mindestlohnregelungen
Daten	Administrative Linked-Employer-Employee Daten Bau Daten aus inländischem und ausländischem Meldeverfahren zur Urlaubskasse (SOKA Bau) Eigene Beschäftigtenbefragung ELVIRA 2011 (Baustatistische Online-Datenbank) Mikrozensus	Administrative Individualdaten Amtliche Statistik CATI-Unternehmensbefragung (CATI) Daten der Bundesagentur für Arbeit (BA-Daten) Daten der Lohnausgleichskasse (LAK) Mannheimer Unternehmenspanel (MUP) Qualitative Interviews mit Dachdeckerbetrieben
Kontrollgruppen	<b>Kontrollgruppen außerhalb der Branche (KG)</b> 1. baunah (Maler- und Lackiererhandwerk), 2. vorgelagert (Herstellung von Kalksandstein), 3. nachgelagert (West: Holzmöbel, Ost: Möbeltischlerei), 4. baufern (West: Pappe/Papier, Ost: Güterbeförderung Kfz) <b>Kontrollgruppen innerhalb der Branche</b> Beschäftigte des Bauhauptgewerbes, die bereits vor Einführung/Erhöhung des Mindestlohns mindestens auf diesem Niveau entlohnt wurden. Betriebe aus dem Bauhauptgewerbe, die über Mindestlohn entlohnen.	<b>Kontrollgruppen außerhalb der Branche</b> 1. Installationsgewerbe (KG1) <b>Kontrollgruppen innerhalb der Branche</b> Beschäftigte der Dachdeckerbranche, die bereits vor Einführung/Erhöhung des Mindestlohns mind. auf diesem Niveau entlohnt wurden. Dachdeckerbetriebe ohne mindestlohnbedingten Anstieg der Lohnkosten.
Eingriffsintensität	<b>Beschäftigte</b> Mindestlohn I: 3 % (West) und 17 % (Ost) Mindestlohn II: 11 % (West) und 30 % (Ost) <b>Betriebe</b> Anteil betroffener Betriebe (1997 ML I): 27 % (West) und 71 % (Ost) Anteil betroffener Betriebe (2003 ML I + II): 67 % (West) und 89 % (Ost)	<b>Beschäftigte</b> Einführung: 3 % (West) und 12 % (Ost) Bis 2008 (bezogen auf nächste Erhöhung): 10 % (West) und 60 % (Ost) <b>Betriebe</b> Anteil betroffener Betriebe: 30 % (West) und 90 % (Ost)
Ergebnisse	<b>Beschäftigung (Regionenebene)</b> Vergleich mit allen Kontrollbranchen: Keine Effekte <b>Beschäftigung (Beschäftigtenebene) Einführung</b> Weiterbeschäftigungswahrscheinlichkeit: Negative Effekte (Ost) Positive Effekte (West) <b>Beschäftigung (Betriebsebene) Einführung</b> Einstellungswahrscheinlichkeit: Negative Effekte (Ost) Keine Effekte (West)	<b>Beschäftigung (Beschäftigtenebene)</b> Weiterbeschäftigungswahrscheinlichkeit: Vergleich KG1: Negative Effekte (West/Ost) Vergleich KG innerhalb: Negative Effekte (West/Ost) <b>Beschäftigung (Betriebsebene)</b> Vergleich KG1: Negative Effekte (West/Ost) Vergleich KG innerhalb: Negative Effekte (West/Ost)



**Tab. 2** Übersicht der Studien zum Branchentyp 1 – Teil 1 (Fortsetzung)

Bauhauptgewerbe (IAB, RWI, ISG)	Dachdeckerhandwerk (ZEW)
<b>Löhne (Beschäftigtenebene), Einführung</b>	<b>Löhne (Beschäftigtenebene)</b>
Vergleich KG innerhalb:	Vergleich KG innerhalb:
Positive Effekte (West/Ost)	Negative Effekte (West/Ost)
Vergleich KG 1: Positive Effekte (West/Ost)	Vergleich KG1:
Vergleich KG 2: Keine Effekte (West), positive Effekte (Ost)	Positive Effekte in unteren Dezilen (Ost)
Vergleich KG 3: Positive Effekte (West/Ost)	Positive Effekte in vielen Dezilen (West)
Vergleich KG 4: Negative Effekte (West/Ost)	<b>Löhne (Betriebsebene)</b>
<b>Löhne (Betriebsebene), Einführung</b>	Vergleich KG innerhalb: Keine Effekte
innerhalb BHG: Positive Effekte (West/Ost)	

Quelle: Möller et al. (2011), Aretz et al. (2011), eigene Zusammenstellung. Anmerkung: Im Bauhauptgewerbe wird zwischen zwei Mindestlohnstufen unterschieden. Der Mindestlohn I gilt für einfache Bau- und Montagetätigkeiten, die keine Qualifikation voraussetzen. Der Mindestlohn II wird für fachlich begrenzte Tätigkeiten bezahlt.

tion der inkrementellen Effekte über die Mindestlohneinführung hinaus in Frage. Wir schlagen deshalb vor eine modifizierte Version des inkrementellen DvD-Schätzers als Robustheitscheck durchzuführen. Dabei sollte die Veränderungen zwischen jeweils zwei Zeitperioden analog zu einer Standard-DvD Schätzung betrachtet werden. Wenn sich die geschätzten Mindestlohneffekte aus dem inkrementellen Design und dem Robustheitscheck signifikant unterscheiden, dann ist es notwendig die Plausibilität der impliziten Annahmen über die Selektivität der Treatmentgruppe expliziter zu diskutieren, um eine angemessene Schätzung bestimmen zu können.

### 3.1.3 Ergebnisse

Die Ergebnisse der Wirkungsanalysen für die Zielgrößen Beschäftigung und Lohn sind in Tab. 2 und 3 zusammengefasst. Zur Gewährleistung der internen Validität der Ergebnisse ist die Durchführung umfangreicher Placebo-Tests prinzipiell zu befürworten. Die Gefahr eines statistischen Overfittings sollte dabei diskutiert werden.

Die Studie zum Mindestlohn im Bauhauptgewerbe präsentiert zahlreiche Placebo-Tests. Für die Zielgröße Beschäftigung scheint die Annahme paralleler Trends für alle gewählten Kontrollgruppen plausibel. Das Schätzverfahren liefert keine signifikanten Beschäftigungseffekte in West- und Ostdeutschland (Möller et al. 2011, vgl. Tab. 6.4). Es ist auffällig, dass die Placebo-Tests vor allem dann gelingen (sprich parallele Trends werden nicht verworfen), wenn auch nach der Mindestlohneinführung keine signifikanten Effekte gemessen werden. Die Placebo-Experimente werden auch für die Zielgröße Lohnwachstum durchgeführt. Diese gelingen für die ausgewählten Kontrollgruppen außerhalb des Bauhauptgewerbes nie und für Kontrollgruppen innerhalb der Branche nur für Westdeutschland. Die

Ergebnisse zeigen positive Effekte auf das Lohnwachstum in beiden Teilen Deutschlands für die Mindestlohneinführung im Jahre 1997.

Die Untersuchungen zur Einführung des Mindestlohns II und weiterer Erhöhungen liefern keine robusten Ergebnisse (Möller et al. 2011, vgl. S. 168ff.). Grundsätzlich sollte die Auswahl geeigneter Kontrollgruppen nicht allein durch das Gelingen von Placebo-Tests motiviert sein, da dies auch von der Skalierung der Zielgröße abhängen kann. Des Weiteren steigt mit der Anzahl an durchgeführten Tests auch die Wahrscheinlichkeit dafür, dass einer der Placebo-Tests gelingt, obwohl die Kontrollgruppe nicht geeignet sein könnte. Dann besteht die Gefahr eines statistischen Overfittings. Wir schlagen in diesem Zusammenhang vor, das quantitative Prüfverfahren in Analogie zu Kreuzvalidierungsverfahren zu gestalten. Die Güte des Modells sollte auf Basis von Prognosen außerhalb der Schätzstichprobe in Verbindung mit einer gleichzeitigen Bestrafung von einer zu hohen Flexibilität innerhalb der Schätzstichprobe berechnet werden. Wenn eine stärkere Flexibilität des Schätzansatzes bei der Wahl der Kontrollgruppe mit einem höheren Prognosefehler außerhalb der Schätzstichprobe verbunden ist, dann sollte diese Flexibilität im statistischen Sinne stärker bestraft werden (analog zur Konstruktion des angepassten Bestimmtheitsmaß für eine lineare Regression).

Während die Studie zum Mindestlohn im Dachdeckerhandwerk nur eine Vergleichsbranche verwendet, nutzen die anderen Studien mehrere Kontrollbranchen. Der Vergleich zu nur einer Kontrollbranche beschränkt die Interpretation der Ergebnisse auf einen Vergleich mit ebendieser einen Branche. Die Anwendung mehrerer Kontrollbranchen kann zu dem Problem führen, dass sich die Ergebnisse je nach Wahl der Kontrollgruppe unterscheiden. Dies zeigt sich in den Studien zum Elektrohandwerk und Maler- und Lackierhandwerk. Zudem liefern Robustheitsanalysen und Pla-

**Tab. 3** Übersicht der Studien zum Branchentyp 1 – Teil 2

	Elektrohandwerk (IAW)	Maler- und Lackiererhandwerk (IAW)
Mindestlohnregelung	Einführung 1997 Abschaffung 2003 Wiedereinführung 2007	Einführung 2003 Erhöhung und Ausweitung des Mindestlohns 2009
Daten	Administrative Individualdaten Amtliche Statistik Betriebshistorikpanel Eigene Fragebogenumfrage Expertenbefragung Qualitative Interviews	Administrative Individualdaten Amtliche Statistik Betriebshistorikpanel Eigene Fragebogenumfrage Mikrozensus Qualitative Interviews
Kontrollgruppen	<b>Kontrollgruppen außerhalb der Branche</b> 1A. Verschiedene Branchen mit Mindestlohn, 1B. Verschiedene Branchen ohne Mindestlohn, 2. Elektrohandwerker in anderen Branchen, <b>Kontrollgruppen innerhalb der Branche</b> 3. Beschäftigte des Elektrohandwerks, die bereits vor Mindestlohneinführung mindestens auf und max. 15 % über Mindestlohnniveau entlohnt wurden. 4. Beschäftigte in Betrieben, die vor Einführung des Mindestlohns tarifgebunden waren.	<b>Kontrollgruppen außerhalb der Branche</b> 1A. Verschiedene Branchen mit Mindestlohn, 1B. Verschiedene Branchen ohne Mindestlohn, 2. Maler und Lackierer in anderen Branchen, <b>Kontrollgruppen innerhalb der Branche</b> 3. Beschäftigte des Maler- und Lackiererhandwerks, die bereits vor Mindestlohneinführung mindestens auf und max. 15 % über Mindestlohnniveau entlohnt wurden. 4. Beschäftigte in Betrieben, die vor Einführung des Mindestlohns tarifgebunden waren.
Eingriffsintensität	<b>Beschäftigte</b> Einführung 1997: in Ostdeutschland stärker als im Westen Abschaffung 2003 und Wiedereinführung 2007: zeigt kaum Effekte <b>Betriebe</b> Anteil Betriebe (1997): 30 % (West) und 90 % (Ost)	<b>Beschäftigte</b> Einführung 2003: Intensität (West) gering, nur Ungelernte und Leiharbeiter Einführung 2003: Intensität (Ost) 50 % Gelernte <b>Betriebe</b> Erhöhung 2009: 50 % der Betriebe
Ergebnisse	<b>Beschäftigung (Beschäftigtenebene)</b> Vergleich mit KG1 und KG2: Keine Effekte Vergleich KG innerhalb: Keine Effekte <b>Löhne (Beschäftigtenebene)</b> Einführung KG1B und KG3: Positiver Effekt ist im Osten stärker Abschaffung: Keine Effekte Wiedereinführung 2007 KG1 und KG2: Positive Effekte Vergleich KG3: Negative Effekte <b>Löhne (Betriebsebene)</b> Wiedereinführung 2007: Keine Effekte	<b>Beschäftigung (Beschäftigtenebene)</b> Vergleich mit KG1 und KG2: Keine Effekte (West/Ost) Vergleich KG innerhalb: Keine Effekte <b>Löhne (Beschäftigtenebene)</b> Vergleich alle KG: Keine Effekte (West), positive Effekte (Ost) <b>Löhne (Betriebsebene)</b> Vergleich KG1: Keine Effekte (West), positive Effekte (Ost)

Quelle: Kirchmann et al. (2011a), Kirchmann et al. (2011b), eigene Zusammenstellung.

cebo-Tests Hinweise dafür, dass die Annahme paralleler Trends in diesen Studien häufig verletzt ist (Kirchmann et al. 2011b, vgl. S. 198ff., S. 211ff.) und (Kirchmann et al. 2011a, vgl. S. 201ff., S. 211ff., S. 215ff.). Zur Erhöhung der internen Validität der Ergebnisse erachten wir weitergehende Analysen (ggf. mit synthetischen Kontrollgruppen) als notwendig. Die Auswahl der Kontrollbranche ist allerdings nur eine, aber natürliche wichtige Dimension in der Schätzung des Vergleichsmaßstabs. Während einfa-

che Standardanwendungen des synthetischen Kontrollgruppenansatzes (Abadie et al. 2015; Dube and Zipperer 2015; Felder 2014) gewichtete Vergleichsgruppen für Aggregate (Branchen, Staaten) schätzen, untersuchen die Mindestlohngutachten ebenfalls Unterschiede in der Betroffenheit der Firmen oder Regionen in der Branche, um die Mindest-

lohneffekte zu schätzen. Der synthetische Kontrollgruppenansatz ist diesbezüglich zu erweitern.<sup>7</sup>

Die Ergebnisse des inkrementellen DvD-Ansatzes in der Studie zum Bauhauptgewerbe liefern einen negativen permanenten Unterschied zwischen den Vergleichsgruppen. Die Ergebnisse für Interaktionen der Betroffenheit und Indikatorvariablen für die einzelnen Jahre zwischen 1999 bis 2001 (in diesen Jahren gab es einen Mindestlohn) sind positiv und signifikant. Laut diesen Ergebnissen führte der Mindestlohn in diesen Jahren zu positiven Beschäftigungseffekten relativ zur Periode vor Einführung des Mindestlohns. Im Hinblick auf das Lohnwachstum zeigen sich in frühen Jahren positive Effekte in West- und Ostdeutschland. Danach schwanken die Ergebnisse vor allem in Ostdeutschland (Möller et al. 2011, vgl. S. 247.). Aufgrund der in Abschn. 3.1.2 geäußerten Vorbehalte erscheinen die Ergebnisse des inkrementellen DvD-Schätzers ab dem zweiten Jahr nach Einführung des Mindestlohns als schwer interpretierbar.

Die Ergebnisse, die auf Basis eines Vergleichs mit einer Kontrollgruppe innerhalb der Branche geschätzt wurden, müssen vor dem Hintergrund möglicher Spillover- und Substitutionseffekte interpretiert werden. Die Belastbarkeit der Ergebnisse hängt maßgeblich von der Datenqualität ab, da eine unpräzise Abgrenzung der Vergleichsgruppen zu verzerrten Ergebnissen führt. Für das Elektrohandwerk ergeben sich bei einem Vergleich mit nicht betroffenen Beschäftigten innerhalb der Branche negative Beschäftigungseffekte der Mindestlohneinführung 1997 für Westdeutschland sowie der Wiedereinführung 2007 in Ostdeutschland (Kirchmann et al. 2011a, vgl. S. 199). Auch die Effekte auf die Löhne sind negativ. Für das Maler- und Lackiererhandwerk zeigen sich keine signifikanten Effekte für beide Zielgrößen (Kirchmann et al. 2011b, vgl. S. 194ff).

In den Evaluationen zum Bauhauptgewerbe und Dachdeckerhandwerk ergeben sich positive Effekte auf das Lohnwachstum. Die Analyse des Dachdeckerhandwerks kann sehr präzise Stundenlohninformationen benutzen, so dass die Abgrenzung der Vergleichsgruppe einfacher ist.

<sup>7</sup> Dies birgt die Chance für eine verbesserte Auswahl einer synthetischen Kontrollgruppe für jede Firma oder Region in der Treatmentgruppe, siehe auch Fußnote 5. Dies erfordert somit eine mehrdimensionale Auswahl der potenziellen Kontrollbeobachtungen (Branche  $\times$  Region bzw. Branche  $\times$  Firma). Die Analogie zu Matchingverfahren (Fitzenberger and Hujer 2002; Imbens and Wooldridge 2009) liegt auf der Hand. Dube und Zipperer (2015) beschränken sich darauf, synthetische Vergleichsregionen (Staaten) für die betrachteten Fälle der Mindestloohnerhöhungen zu ermitteln. Die zweite Dimension liegt in der Analyse verschiedener Mindestloohnerhöhungen, da einzelne Staaten zu unterschiedlichen Zeitpunkten eine Erhöhung des Mindestlohns vornehmen. Somit beschränkt sich die Analyse nicht auf *einen aggregierten Treatmentfall*. In diesem Punkt stimmen wir nicht mit Felder (2014) überein, die argumentiert, dass die Methode nur für aggregierte Ergebnisse verwendbar ist.

In beiden Studien finden sich Hinweise für das Auftreten von Spillover-Effekten. Diese werden in der Studie zum Dachdeckerhandwerk sehr ausführlich diskutiert. Im oberen Bereich der Lohnverteilung kommt es zu negativen Lohn-effekten. Insgesamt geht die Lohnspreizung in Ostdeutschland stark zurück (Aretz et al. 2011, vgl. S. 307ff).

Der Beschäftigungseffekt fällt bei einem Vergleich mit Kontrollpersonen innerhalb des Dachdeckerhandwerks negativer aus als bei einem Vergleich mit dem Installationsgewerbe. Eine Verzerrung der Ergebnisse ist hier möglich, wenn es zu indirekten positiven Auswirkungen auf die nichtbetroffenen Kontrollpersonen kommt. Es ist festzuhalten, dass eine weitergehende, substantielle Analyse dieser Aspekte sinnvoll wäre. Wenn Schätzungen der Substitutions- und Skaleneffekte in der Arbeitsnachfrage vorliegen und gleichzeitig die Spillovereffekte in den Löhnen erfasst werden, könnte die Verzerrung in den DvD-Schätzungen quantifiziert werden. Unter der Annahme, dass der Substitutionseffekt den Skaleneffekt dominiert (Fitzenberger 2009), wäre eine Aussage über die Richtung der Verzerrung möglich.

Die Branchen aus dem Bauhaupt- und Baunebengewerbe weisen mehrere Besonderheiten auf (Produktionsbedingungen, Wettbewerb, Beschäftigtenstruktur usw.) und die starke Zunahme ausländischer Arbeitnehmer in diesen Branchen war Anlass für die Einführung von Mindestlöhnen Ende der 1990er Jahre. Gleichzeitig agierten die Baubranchen unter der besonderen konjunkturellen Lage nach der Wiedervereinigung. Zum einen erfolgte Ende der 1990er Jahre ein vorhersehbarer Schrumpfungsprozess der Branche nach Ende des wiedervereinigungsbedingten Booms und zum anderen war die Zunahme der Beschäftigung ausländischer Arbeitnehmer Auslöser der Einführung von Mindestlöhnen. Die Mindestlohnregelungen wirken zudem auf die seinerzeit vor allem im Baugewerbe weit verbreitete Arbeitnehmerentsendung, die jedoch in den Gutachten nicht erfasst werden konnte (Meier and Munz 2008). Folglich ist die externe Validität der Evaluationsstudien für den Branchentyp 1 aufgrund der besonderen Situation der Baubranchen beschränkt. Zudem stellt sich angesichts des Vorhersehbarkeit des Schrumpfungsprozesses insbesondere für die Baubranchen die Frage nach einer möglichen Politikendogenität der Mindestlohnregelungen.

Zusammenfassend lässt sich für Branchentyp 1 festhalten, dass die vorgelegten Studien sorgfältige Analysen auf einer breiten Datenbasis und mit einem weiten Methodenspektrum umgesetzt haben. Allerdings fallen die Ergebnisse je nach Wahl der Kontrollgruppe unterschiedlich aus. Die Studien für das Bauhauptgewerbe und für das Dachdeckerhandwerk konnten innovative Datensätze nutzen, diese mit unterschiedlichen methodischen Ansätzen untersuchen und die Ergebnisse dahingehend diskutieren. Für das Bauhauptgewerbe und für das Dachdeckerhandwerk sind trotz Unter-

**Tab. 4** Übersicht der Studien zum Branchentyp 2

	Gebäudereinigung (IAQ)	Wäschereidienstleistungen im Objektkundengeschäft (IAQ)
Mindestlohnregelung	Einführung 2007 Erhöhung 2008 Allgemeinverbindlichkeit von Tarifverträgen seit 1970er Jahren. Betrifft nur gewerblich Beschäftigte, d. h. nicht alle Beschäftigte, die Reinigungstätigkeiten durchführen.	Einführung 2009 Jährliche Erhöhungen bis 2012
Daten	Administrative Individualdaten Eigene Betriebsbefragung Expertenbefragung Mikrozensus	Administrative Individualdaten Amtliche Statistik Daten der Bundesagentur für Arbeit (BA-Daten) Qualitative Interviews
Kontrollgruppen	<b>Kontrollgruppen außerhalb der Branche</b> 1. Köche und Kellner, 2. Berufe des Wach- und Sicherheitsgewerbes, 3. Innenreiniger in anderen Branchen, <b>Kontrollgruppen innerhalb der Branche</b> keine	<b>Kontrollgruppen außerhalb der Branche</b> Keine Kausalanalyse durchgeführt <b>Kontrollgruppen innerhalb der Branche</b> Keine Kausalanalyse durchgeführt
Eingriffsintensität	Direkte Wirkung des Mindestlohns wird als gering eingeschätzt, da die Tariflöhne seit den 1970er Jahren allgemeinverbindlich sind.	Nicht ermittelbar, da Identifikation des Geltungsbereichs und somit der Betroffenheit nicht möglich.  Abgrenzungsproblematik durch die Wirtschaftszweikklassifikation (WKZ) und die Definition des Geltungsbereichs des Mindestlohns.
Ergebnisse	<b>Beschäftigung (Regionenebene)</b> Tariflohnänderung 2004: Keine Effekte <b>Beschäftigung (Beschäftigtenebene)</b> Mindestlohneinführung Vergleich KG1: Keine Effekte Vergleich KG2/KG3: Negative Effekte Mindestlohnerhöhung Vergleich KG1/KG2: Positive Effekte Mindestlohnerhöhung Vergleich KG3: Keine Effekte	<b>Beschäftigung</b> Keine Kausalanalyse durchgeführt <b>Löhne</b> Keine Kausalanalyse durchgeführt

Quelle: Bosch et al. (2011b); Bosch et al. (2011a); eigene Zusammenstellung.

schieden je nach Kontrollgruppe negative Beschäftigungseffekte, vor allem bei hoher Eingriffsintensität, zu konstatieren, die allerdings angesichts teilweise sehr hoher Eingriffsintensität nicht sehr stark ausfallen. Demgegenüber finden sich keine negativen Beschäftigungseffekte im Maler- und Lackiererhandwerk und im Elektrohandwerk. Meist, aber nicht in allen Fällen finden sich positive Lohneffekte und in mehreren Fällen wirkt sich der Mindestlohn augenscheinlich auch auf die Lohnstruktur oberhalb des Mindestlohnes aus.

### 3.2 Ergebnisse für Branchentyp 2

Zum zweiten Branchentyp gehören die Wäschereidienstleistungen im Objektkundengeschäft und die Gebäudereinigung. Hierbei handelt es sich um zwei arbeitsintensive Branchen, in denen die Tätigkeiten vor allem einfache, stan-

dardisierte Dienstleistungen mit starkem körperlichem Einsatz umfassen. Charakteristisch für beide Branchen ist ein hoher Anteil an weiblichen Beschäftigten. Die ersten Mindestlöhne wurden ab dem Jahr 2007 eingeführt.

In Tab. 4 werden die wichtigsten Punkte der Studien systematisch zusammengefasst. Im Vergleich zu den vier Studien aus dem Bauhaupt- und Baunebengewerbe leiden die Studien zum Branchentyp 2 unter starken Problemen der Datenverfügbarkeit. Dies liegt an der relativ kurzen Zeitperiode zwischen der Einführung der Mindestlöhne und dem Zeitpunkt der Evaluation. Außerdem bestehen aufgrund der begrenzten Verfügbarkeit von Informationen zu den Wirtschaftsklassen erhebliche Abgrenzungsprobleme der vom Mindestlohn betroffenen Betriebe und Personen und damit ebenfalls erhebliche Abgrenzungsprobleme bei der Konstruktion von Vergleichsgruppen.

In der Studie zum Mindestlohn im Wäschereigewerbe wurde aufgrund der unzureichenden Datenlage keine ökonometrische Wirkungsanalyse durchgeführt. Die Ergebnisse der Studie basieren daher ausschließlich auf Aussagen aus Expertengesprächen und Befragungen (Bosch et al. 2011a). Die Ergebnisse verdeutlichen, dass zum Zeitpunkt der Evaluation eine hohe Unsicherheit auf Seiten der Wäschereibetriebe über die Mindestlohnspflicht bestand. Hinzu kommt, dass ebenso ex post in administrativen Daten die genaue Abgrenzung der Branche, für die die Mindestlohnspflicht gilt, schwierig ist. Gleichwohl hätte nach unserer Einschätzung eine kausale Wirkungsanalyse, die die unvollständige Betroffenheit von Betrieben als Treatment verwendet, in Kombination mit modernen Kontrollgruppenansätzen (synthetische Kontrollgruppen unter Berücksichtigung der Heterogenität innerhalb der Branche, siehe Fußnoten 5 und 7) durchgeführt werden können, die unter Umständen belastbare Ergebnisse geliefert hätte. Der Evaluationsstudie (Bosch et al. 2011a) ist jedoch zugutezuhalten, dass eine ökonometrische Wirkungsanalyse im Falle des Mindestlohns im Wäschereigewerbe besondere Schwierigkeiten aufweist, sowohl im Hinblick auf den Zeitdruck, unter dem die Studien zu erstellen waren, als auch im Hinblick auf die besonderen Abgrenzungsprobleme der Betroffenheit durch den Mindestlohn.

In der Studie zur Gebäudereinigung (Bosch et al. 2011b) wurde trotz der beschränkten Datenverfügbarkeit ein Kontrollgruppenansatz umgesetzt. Ausgangspunkt für die Bildung der Vergleichsgruppen sind die regional unterschiedlichen Tariflohnänderungen im April 2004 sowie die Einführung der Mindestlöhne im Juli 2007 und die anschließende Erhöhung der tariflich vereinbarten Mindestlöhne ab März 2008. Zur Bildung von Kontrollgruppen außerhalb der Gebäudereinigungsbranche wurden Kunden- und Beschäftigungsstrukturen verglichen. In dieser Hinsicht weisen die branchenbezogenen Kontrollgruppen aus dem Wachgewerbe und aus der getränkegeprägten Gastronomie die größte Ähnlichkeit mit der Gebäudereinigungsbranche auf. In der Beschäftigungsstruktur gibt es vor allem bei den Arbeitszeiten und dem Geschlechteranteil Unterschiede zum Wachgewerbe. Die Kundenstruktur ist in beiden Branchen sehr ähnlich, wobei in der getränkegeprägten Gastronomie deutliche saisonale und klimabedingte Schwankungen auftreten. Eine weitere Kontrollgruppe besteht aus Reinigungskräften, die in anderen Branchen beschäftigt sind und somit vergleichbare Tätigkeiten ausüben, aber keinen Mindestlohnanspruch besitzen. Die Heterogenität der Branchen, in denen Reinigungskräfte außerhalb des Gebäudereinigerhandwerks beschäftigt sind, ist jedoch nur begrenzt kontrollierbar.

Die Ergebnisse zeigen ein uneinheitliches Bild der Tariflohnänderungen im Jahre 2004 und der Aufnahme in das AEntG und der damit verbundenen Wirksamkeit der Min-

destlöhne im Juli 2007. Für die Zielgröße Beschäftigung zeigt sich keine eindeutige Mindestlohnwirkung. Die Effekte werden auf Regionalebene bestimmt und sind je nach Region positiv oder negativ und meistens insignifikant (Bosch et al. 2011b, vgl. Tab. 27 und 28). Die Einführung der Mindestlöhne führt zu signifikant negativen Effekten im Vergleich zu den erstgenannten Kontrollgruppen. Die Effekte der Mindestlohnerhöhung im Jahr 2008 erweisen sich hingegen als positiv (Bosch et al. 2011b, vgl. Tab. 29).

Die interne Validität der Ergebnisse für die Gebäudereinigungsbranche ist kritisch zu bewerten. Zur Überprüfung der Plausibilität der Annahme paralleler Trends werden zeitliche Entwicklungen der relevanten Zielgrößen für die Treatment- und Kontrollgruppe dargestellt (Bosch et al. 2011b, vgl. S. 250ff). Diese branchenspezifischen Entwicklungspfade stellen die Gültigkeit der identifizierenden Annahme infrage und sprechen somit gegen die Eignung der Kontrollgruppen. Insbesondere für die Beschäftigungsentwicklung scheint diese wichtige Annahme wenig plausibel. Es werden keine Placebo-Tests zur weiteren Plausibilitätsprüfung durchgeführt.

Die geschätzten negativen Effekte der Mindestlohneinführung im Jahre 2007 und die positiven Effekte der Erhöhung im darauffolgenden Jahr könnten auch als Hinweis auf eine Endogenität des Treatments (zur Politikendogenität siehe Abschn. 2) angesehen werden. Es ist durchaus denkbar, dass die Einführung der Mindestlöhne und die spätere Erhöhung im Hinblick auf zukünftige Beschäftigungseffekte von den Tarifparteien ausgehandelt worden sind. Diesem Problem wird durch den gewählten Evaluationsansatz nicht Rechnung getragen.

Die Wirkungsanalyse zum Mindestlohn in der Gebäudereinigung zeigt, wie wichtig eine ausführliche Analyse und Diskussion über die Eignung der gewählten Kontrollgruppen im Hinblick auf die erforderlichen Annahmen des Schätzmodells ist. Die Anwendung von Placebotests kann dazu beitragen die interne Validität der Untersuchung zu verbessern. Eine Verwendung synthetischer Kontrollgruppen aus mehreren anderen Branchen (unter Berücksichtigung der Heterogenität innerhalb der Branche, siehe Fußnoten 5 und 7) oder einer Kontrollgruppe innerhalb der Gebäudereinigungsbranche wäre sinnvoll.

Zusammenfassend lassen sich für Branchentyp 2 folgende Punkte festhalten. Für die Wäschereidienstleistungen im Objektkundengeschäft wurde leider nicht versucht, eine ökonometrische Wirkungsanalyse des Mindestlohns durchzuführen. In der ökonometrischen Wirkungsanalyse für die Gebäudereinigung wurde die Eignung des gewählten Evaluationsansatzes und der gewählten Kontrollgruppen nicht hinreichend kritisch diskutiert. Daher ist es schwierig die fehlende Robustheit der erzielten Ergebnisse angemessen einzuordnen. Beiden Gutachten ist zugutezuhalten, dass die Datenlage für eine ökonometrische Evaluation des Min-

**Tab. 5** Übersicht der Studien zum Branchentyp 3

	Pflegebranche (IAW)	Abfallwirtschaft (ZEW)
Mindestlohnregelung	Einführung 2010 Erhöhung 2012	Einführung 2010
Daten	Amtliche Statistik Befragungsdaten Mikrozensus Qualitative Interviews	Amtliche Statistik CATI-Unternehmensbefragung Mannheimer Unternehmenspanel (MUP) Qualitative Interviews
Kontrollgruppen	<b>Kontrollgruppen außerhalb der Branche</b> Keine verwendet <b>Kontrollgruppen innerhalb der Branche</b> Einrichtungen, deren Beschäftigte nach kollektivrechtlichen Verträgen mind. in der Höhe des Mindestlohns entlohnt werden.	<b>Kontrollgruppen außerhalb der Branche</b> Betriebe des Garten- und Landschaftsbaus <b>Kontrollgruppen innerhalb der Branche</b> Betriebe der Abfallwirtschaft, die keine Lohnerhöhung bei den Mitarbeitern in Folge des Mindestlohns vornehmen mussten.
Eingriffsintensität	<b>Beschäftigte:</b> West: 4,48 % ambulant, 5,83 % stationär Ost: 6,17 % ambulant, 4,31 % stationär <b>Betriebe:</b> West: Mindestlohn von 8,50€ entspricht bei 66,7 % (69,4 %) der stationären (ambulanten) Einrichtungen dem Medianlohn. Ost: Mindestlohn von 7,50€ entspricht bei 81,1 % (73,2 %) der stationären (ambulanten) Einrichtungen dem Medianlohn.	<b>Beschäftigte:</b> Durchschnittliche Branchenbetroffenheit 6 % <b>Unternehmensebene:</b> Anteil an Unternehmen mit mind. einem Arbeitnehmer mit einem Stundenlohn unterhalb von 8€ vor Mindestlohneinführung: 15,1 % (West), 41,5 % (Ost)
Ergebnisse	<b>Beschäftigung (Betriebsebene)</b> Vergleich mit KG: Keine Effekte (außer West negativ für ambulante Dienste) <b>Löhne (Betriebsebene)</b> Vergleich mit KG: Keine Effekte (außer positiv für Pflegehilfskräfte)	<b>Beschäftigung (Betriebsebene)</b> Vergleich mit KG: Keine Effekte <b>Löhne (Betriebsebene)</b> Keine Kausalanalyse durchgeführt

Quelle: Kirchmann et al. (2011c); Egelin et al. (2011); eigene Zusammenstellung.

destlohns zum Zeitpunkt der Erstellung der Gutachten sehr schwierig war.

### 3.3 Ergebnisse für Branchentyp 3

Der dritte Branchentyp umfasst die Branchen Abfallwirtschaft und Pflege. Beide Branchen sind durch ein Nebeneinander öffentlicher und privater Anbieter gekennzeichnet. Die Qualitätssicherung der Leistungserstellung hat in diesen Branchen einen hohen Stellenwert. Mindestlöhne wurden in diesen Branchen erst im Jahr 2010 eingeführt.

In beiden Studien waren amtliche Daten nicht oder nur beschränkt verfügbar. Für die Pflegebranche lagen keinerlei amtliche Daten vor, die den Zeitraum nach der Mindestlohneinführung abdeckten, so dass die Autoren eine eigene Befragung in Auftrag gaben. Hierbei wurde durch ein Befragungsunternehmen eine Online-Befragung durchgeführt, die den Zeitraum kurz vor der Mindestlohneinführung und den Zeitraum kurz danach abdeckt. Der Vorteil dieser Datengewinnung liegt in der präzisen Abfrage der Stundenlöhne und der damit verbundenen Vermeidung von Ver-

zerrungen durch die Berechnung aus anderen Datensätzen. Des Weiteren ist eine präzise Abgrenzung der vom Mindestlohn betroffenen Einrichtungen und Personen möglich, da die Befragung nicht nur Fragen zur Lohnstruktur sondern auch zu den Einrichtungen und dem Arbeitsumfeld umfasst. Auch die Studie zum Mindestlohn in der Abfallwirtschaft ergänzt die beschränkte Datenbasis durch eigens für die Evaluation erhobene Befragungsdaten.

Kontrollgruppen aus anderen Vergleichsbranchen werden nur in der Studie zur Abfallwirtschaft verwendet. Als Kontrollgruppe außerhalb der Branche dient der Garten- und Landschaftsbau, weil diese Branche, wie die Abfallwirtschaft durch Multiproduktunternehmen gekennzeichnet ist, wie beispielsweise Dienstleistungen in der Objektbetreuung oder Hausmeister Tätigkeiten. In dieser Branche gibt es eine spezifische Wettbewerbssituation, die durch ein Nebeneinander von öffentlichen und privaten Anbietern sowie durch geringe Qualifikationsanforderungen an die Beschäftigten gekennzeichnet ist. In einem umfassenden Branchenbild wird die Annahme paralleler Trends untersucht. Mögliche Verletzungen der Annahme zeigen sich vor allem in

der Sektorengröße, der Umsatzentwicklung sowie der Unternehmensanzahl und deren Größe. Aus diesen Unterschieden heraus wird die Kontrollgruppe in der Evaluation ausschließlich für den Teilbereich der Abfallbeseitigung verwendet. Somit untersucht die Wirkungsanalyse den Effekt der Mindestlohneinführung auf Betriebe im Teilbereich der Abfallbeseitigung im Vergleich zu Betrieben aus dem Landschafts- und Gartenbau. Es zeigen sich keine signifikanten Beschäftigungseffekte (Egeln et al. 2011, vgl. Tab. A2).

Die externe Validität dieses Ergebnisses ist beschränkt, da es nur für den Teilbereich der Abfallbeseitigung gilt. Gerade in diesem Teilbereich der Abfallwirtschaft ist die Nachfrage vergleichsweise konstant und die Möglichkeiten Beschäftigung abzubauen oder diese zu ersetzen sind gering. Die Verwendung der weiteren Teilmärkte der Abfallwirtschaft wäre eventuell möglich gewesen, wenn die Eignung der Kontrollgruppe als Vergleichsbranche für diese Teilmärkte durch Placebo-Tests überprüft worden wäre. Außerdem hätte das Analysepotenzial unter Anwendung eines synthetischen Kontrollgruppendesigns (unter Berücksichtigung der Heterogenität innerhalb der Branche, siehe Fußnoten 5 und 7) erhöht werden können.

In beiden Studien werden Kontrollgruppen innerhalb der untersuchten Branche verwendet. Die Studie zum Mindestlohn in der Pflege wählt die Gruppe der tarifgebundenen Einrichtungen als Kontrollgruppe, da diese nicht vom Mindestlohn betroffen sind, aber dieselbe Entwicklung zu erwarten hatten, wie die vom Mindestlohn betroffene Gruppe. Substitutionseffekte im Hinblick auf andere Branchen erscheinen vernachlässigbar, weil die Pflege branchenspezifische Tätigkeiten erfordert. Arbeitsnachfrageinduzierte Substitutionseffekte zwischen Treatment- und Kontrollgruppe auf Betriebsebene sind per Definition ausgeschlossen, da die Tarifbindung eine Eigenschaft der Einrichtung und nicht der Arbeitnehmer ist. Die Arbeitnehmer aus der Kontrollgruppe (also aus Betrieben mit Tarifbindung) sind somit nicht qualifizierter oder für einen Arbeitgeber interessanter als jene in der Treatmentgruppe. Allerdings besteht die Möglichkeit der Substitution durch die Nachfrage nach Pflegeleistungen in der Kontrollgruppe, da deren Lohnkosten nicht ansteigen.

Da die Evaluation der Pflegebranche die Effekte der Mindestlohneinführung auf Betriebsebene untersucht, erfolgt die Abgrenzung der Vergleichsgruppen einzig auf Basis des Kriteriums der Tarifbindung. Die Zuordnung zur Treatment- und Kontrollgruppe ist nicht eindeutig, da eine Tarifbindung auch freiwillig angewendet worden sein kann, indem sie beispielsweise in einem individuellen Arbeitsvertrag vereinbart wurde. Es besteht weiterhin eine mögliche Endogenität der Tarifbindung selbst. Die Lohnhöhe und die freiwillige Tarifbindung können von der individuellen Leistungsfähigkeit des Betriebes abhängen. Das Branchenbild zeigt, dass die Tarifbindung vom Anbietertyp anhängig ist.

Öffentlich-rechtliche sowie freigemeinnützige und kirchliche Anbieter entlohnen weitaus häufiger nach Tarif als private Träger. Die Vergleichbarkeit mit privaten Anbietern, die in stärkerem Wettbewerb stehen und die keine Tarifbindung aufweisen, ist daher fraglich. Die Interpretation der Ergebnisse sollte deshalb vor dem Hintergrund einer möglichen Endogenität der Tarifbindung erfolgen.

Insgesamt zeigen sich nur wenig signifikante Ergebnisse für den Pflegesektor. Eine Ausnahme sind negative Beschäftigungseffekte in Westdeutschland bei ambulanten Diensten. Die Effekte auf die Stundenlöhne sind im Osten stärker als im Westen und nur für Pflegehilfskräfte signifikant (Kirchmann et al. 2011c, vgl. Tab. 10.17–10.19). Die Placebo-Tests können nicht mit demselben Datensatz durchgeführt werden, da die Panelbeobachtungen nur die zwei Perioden umfassen, die für die Wirkungsanalyse verwendet werden. In der Studie werden deshalb Lohninformationen aus administrativen Individualdaten genutzt. Es werden Placebo-Tests auf Beschäftigtenebene durchgeführt, die nur bedingt mit den Ergebnissen der Wirkungsanalyse auf Betriebsebene vergleichbar sind. Eine Durchführung solcher Robustheitsanalysen ist grundsätzlich zu befürworten. In diesem Fall tragen sie jedoch nur wenig zur Bewertung der internen Validität der Ergebnisse bei.

Die Evaluation zur Abfallwirtschaft definiert Betriebe als Treatmentgruppe, die vor Mindestlohneinführung mindestens einen Arbeitnehmer unterhalb des Mindestlohns entlohnten. Als Kontrollgruppe werden die Betriebe gewählt, die keinem Mitarbeiter in Folge des Mindestlohns eine Lohnerhöhung zahlen mussten, also alle Betriebe, die ihre Mitarbeiter ausnahmslos oberhalb des Mindestlohns entlohnten. Für die Abfallwirtschaft werden mögliche Spillover-Effekte als wenig relevant angesehen, da das Beschäftigungswachstum in der Kontrollgruppe niedriger ausfällt als in der Treatmentgruppe. Substitutionseffekte sind zwischen regulär Beschäftigten und Teilzeitbeschäftigten sowie Beschäftigten des zweiten Arbeitsmarktes (1-Euro-Jobber) zu erwarten. Diese werden in der Studie deskriptiv untersucht. Der Anteil von 1-Euro Jobbern ist sehr gering. Der Anteil von Teilzeitbeschäftigten ist in Unternehmen, die den Mindestlohn umgehen, höher (Egeln et al. 2011, vgl. Tab. 33). Aufgrund der Datenlage konnten diese Effekte nicht weitergehend untersucht werden, sollten aber in zukünftigen Analysen stärker berücksichtigt werden.

Durch die Verwendung von Befragungsdaten können beide Studien präzise Informationen über Stundenlöhne nutzen, sodass keine Verzerrungen durch die Berechnung aus anderen Datensätzen entstehen. Im Hinblick auf die externe Validität der Ergebnisse ist noch zu erwähnen, dass die verwendeten Daten nicht repräsentativ für die Gesamtheit sind. Ein Abgleich mit amtlichen Daten, das Durchführen von vergleichbaren Robustheitsanalysen und die Ermittlung von Effekten auf Individualebene wären

ratsam, wenn in zukünftigen Evaluationen die interne und externe Validität erhöht werden sollen.

Zusammenfassend lässt sich für Branchentyp 3 positiv festhalten, dass beide Studien versucht haben auf Basis von Befragungsdaten eine belastbare ökonomische Evaluation des Mindestlohns, trotz des kurzen Zeitraums zwischen der Einführung des Mindestlohns und der Erstellung der Gutachten, durchzuführen. Es wurden nahezu keine signifikanten Beschäftigungseffekte in beiden Branchen und kaum Lohneffekte in der Pflegebranche gefunden. Die externe Validität der erzielten ökonomischen Ergebnisse ist als kritisch einzuschätzen. Weitergehende Untersuchungen wären im Hinblick auf die Teilzeitbeschäftigung in der Pflegebranche und auf die Abgrenzung zwischen privaten und öffentlichen Anbietern geboten.

#### 4 Zusammenfassung und Ausblick

Dieser Aufsatz liefert einen Überblick und eine kritische Bewertung der acht Evaluationsstudien der branchenspezifischen Mindestlöhne, die das Bundesministerium für Arbeit und Soziales (BMAS) im Jahr 2011 in Auftrag gab. Der Fokus in diesem Aufsatz liegt auf der ökonomischen Wirkungsanalyse. Trotz aller Detailkritik ist festzuhalten, dass die Evaluation der Branchenmindestlöhne im Jahr 2011 ein wichtiges und erfolgreiches Beispiel für einer stärkere Evidenzbasierung der Wirtschaftspolitik in Deutschland darstellt. Dies wird insbesondere durch einen Vergleich mit dem Stand der Diskussion zum Ende des letzten Dekade deutlich (vgl. Fitzenberger 2009 und die Beiträge im Sonderheft des ifo Schnelldiensts 2008). Manning (2013) sieht Parallelen im Stand der Diskussion in Großbritannien vor Beginn der Arbeit der Low-Pay-Commission und dem Stand der Diskussion in Deutschland vor einigen Jahren. Angesichts der Einführung des gesetzlichen Mindestlohns in Deutschland im Jahr 2015 kommt der Evaluation der Branchenmindestlöhne eine besondere Bedeutung zu.

Die kritischen Ausführungen in diesem Aufsatz illustrieren exemplarisch wichtige Herausforderungen auf dem *steirigen Weg* zu einer evidenzbasierten Wirtschaftspolitik. Wir beschränken uns auf die Evaluationsstudien, in der Form wie sie vom BMAS rezipiert und veröffentlicht wurden, um hieran exemplarisch eine Diskussion der Qualität der Studien, wie sie in den politischen Prozess einfließen, anzustoßen. Die Evaluationsstudien sind unter einem starken Zeitdruck entstanden und den Gutachten wurden starke Vorgaben im Hinblick auf die Herangehensweise gemacht (im Rahmen der jeweiligen Leistungsbeschreibung sowie zahlreicher weiterer Absprachen mit dem BMAS im Rahmen des Projektverlaufs). In einigen Fällen lagen nicht einmal administrative Daten für den Zeitraum nach Einführung der Mindestlöhne vor. Aus diesem Grund wurden zusätzlich

Befragungsdaten erhoben. Diese Schwierigkeiten relativieren teilweise die kritischen Einschätzungen in diesem Aufsatz. Generell sollten bei anspruchsvollen Ausschreibungen dieser Art Quantität und Qualität der in Auftrag gegebenen Forschungsarbeiten in einem angemessenen Verhältnis stehen, damit Forschungsmethoden und Daten dem Stand der Forschung entsprechen.

Im Sinne einer Evidenzbasierung der Wirtschaftspolitik halten wir grundsätzlich eine kritische Diskussion der Qualität von Studien nach deren Erstellung für notwendig – auch um zukünftige Evaluationen zu verbessern – und stellen fest, dass dies viel zu selten passiert. In Sinne einer *lernenden Evaluation* ist eine kritische Diskussion der Evaluationsergebnisse wichtig. Dies ermöglicht auch ein kontinuierliches Monitoring der Wirkungen des Mindestlohns, wie es durch die Low-Pay-Commission in Großbritannien erfolgt und wie es auch von Seiten der Mindestlohnkommission in Deutschland in Zukunft erfolgen sollte. An dieser Stelle ist auf folgende Einschränkung hinzuweisen: Obwohl unser Aufsatz punktuell Implikationen für den politischen Prozess erwähnt, umfasst er keine umfassende Diskussion der politischen Gestaltung der Evaluation der Mindestlohnwirkungen. Obgleich wünschenswert, würde dies den Rahmen dieses Aufsatzes sprengen.

Welche konzeptionelle Lehren sind aus der ersten Evaluationsrunde zu ziehen? Erstens, aus der Systematisierung und Bewertung der Studien geht hervor, dass sich trotz ähnlicher methodischer Ansätze (fast immer werden DvD-Schätzer verwendet) ein sehr vielschichtiges Bild aus den acht Gutachten ergibt und dass die Qualität der Ergebnisse stark mit der Verfügbarkeit von administrativen Arbeitnehmer-Arbeitgeber-Daten zusammenhängt. Trotz der teilweisen Verfügbarkeit solcher Daten besteht in vielen Bereichen ein Abgrenzungsproblem im Hinblick auf die von den Mindestlohnregelungen betroffenen Beschäftigten. Fehlende Stundenlohninformationen und eine nicht eindeutig bestimmbare Branchenzugehörigkeit in den Daten sind Gründe dafür. Schließlich stellt sich auch die Frage, ob Betriebe die Branchenzugehörigkeit manipulieren können, um allfälligen Mindestlohnregelungen zu entgehen. In einigen wenigen Baubranchen konnten administrative Arbeitnehmer-Arbeitgeber-Daten bereitgestellt werden, die bisher der Wissenschaft nicht zur Verfügung standen und auf deren Basis die Stundenlöhne berechnet werden konnten. Dies ist ein Mehrwert einer *lernenden Evaluation*.

Zweitens ist festzuhalten, dass die Studien das Kontrollgruppendesign in unterschiedlicher Weise entwickelten. Während die Studie zum Bauhauptgewerbe ein quantitatives Prüfverfahren zur Auswahl der geeigneten Kontrollgruppen nutzt, wird diese in anderen Studien sachlogisch begründet. Weiterhin verwendet die Studie zur Pflegebranche nur eine interne Kontrollgruppe. Keine der Studien bezog sich auf den von Abadie et al. (2010) vorgeschlagenen



synthetischen Kontrollgruppenansatz. Dieser Ansatz weist den Vorteil auf, dass zur Identifizierung einer geeigneten Kontrollgruppe die Variation über alle möglichen Kontrollgruppen genutzt wird, anstatt wie zum Beispiel in der Studie zum Bauhauptgewerbe, alle möglichen Kontrollgruppen gleich zu gewichten. Allerdings wäre die Umsetzung auf die spezifische Fragestellung anzupassen, die u. a. eine angemessene Modellierung der Dynamik in den Ergebnisvariablen und der Heterogenität innerhalb der Branchen erfordert. Die Verwendung des synthetischen Kontrollgruppenansatzes zur Schätzung der Wirkungen von Mindestlöhnen verspricht neue Erkenntnisse (Dube and Zipperer 2015). Die Anwendbarkeit für die Analyse der Branchenmindestlöhne in Deutschland ist zukünftig genauer zu untersuchen (Felder 2014; Kaul et al. 2016). Die hier betrachteten Evaluationsstudien zu den Branchenmindestlöhnen verwenden zu Recht Placebo-Tests, um die Plausibilität der Annahme paralleler Trends zwischen Treatment- und Kontrollgruppen zu stützen. In Fällen der Ablehnung der Placebo-Tests werden diese meist nicht inhaltlich interpretiert und zur Verbesserung des Analyseansatzes eingesetzt. Für das Gutachten zur Gebäudereinigungsbranche ist kritisch anzumerken, dass einzelne DvD-Ergebnisse trotz Verletzung der Placebo-Tests inhaltlich interpretiert werden. Der inkrementelle DvD-Ansatz, der für das Bauhauptgewerbe zur Anwendung kommt, weist Interpretationsprobleme im Hinblick auf die Dynamik der Mindestlohnwirkungen auf, die in diesem Gutachten nicht hinreichend diskutiert werden.

Drittens kann das Changes-in-Changes (CiC) Verfahren (Athey und Imbens 2006) für stetige Zielgrößen eine sinnvolle Ergänzung zu DvD-Ansätzen darstellen, da es die Skalierungsabhängigkeit der identifizierenden Annahmen umgeht. Es ist möglich heterogene Mindestlohnwirkungen entlang der Verteilung unterschiedlich skalierte Zielgrößen zu ermitteln (vgl. Havnes und Mogstad 2015, für eine erfolgreiche Anwendung der Methode). Allerdings setzt dieses Verfahren voraus, dass die Veränderungen in der Verteilung der Zielgrößen in der Kontrollgruppe die kontrafaktischen Veränderungen entlang der Verteilung für die Treatmentgruppe ohne Treatment abbilden. Diese Variante der Annahme paralleler Trends ist im Einzelfall kritisch zu beleuchten. Für stetige Ergebnisvariablen kann das CiC-Verfahren mit Hilfe von (bedingten) Quantilsregressionen umgesetzt werden, um den Einfluss von Kovariaten zu berücksichtigen (vgl. Melly und Santangelo 2015). Umsetzungsprobleme bestehen bei diskreten Ergebnisvariablen. Es liegen nach unserem Kenntnisstand noch keine Erfahrungswerte für die Analyse von Mindestlohnwirkungen vor, so dass offen ist, ob die Methode belastbare Ergebnisse für die Analyse der Mindestlohnwirkungen liefern kann – nach unserer Einschätzung wäre es jedoch einen Versuch wert. Für alle vorgeschlagenen Ansätze gilt, wie auch für den

DvD-Ansatz, dass die Plausibilität und mögliche Verletzungen der SUTVA-Annahme diskutiert werden müssen.

Unsere vierte Anmerkung bezieht sich auf die SUTVA-Annahme. In einigen Studien wurden mögliche Verletzungen der identifizierenden Annahme des DvD-Ansatzes und das mögliche Auftreten von Spillover-Effekten diskutiert. Im Regelfall wurde allerdings die Robustheit von Ergebnissen gegenüber der Wahl der Kontrollgruppe als erstrebenswertes und einzig inhaltlich interpretierbares Ergebnis angesehen. Umgekehrt können sich jedoch aus der Variation der Ergebnisse je nach Wahl der Kontrollgruppe neue Erkenntnisse ergeben, sprich diese Unterschiede können Anhaltspunkte für die ökonomische Wirkungsweise der Mindestlöhne liefern. Wenn unterschiedliche Kontrollgruppen in unterschiedlicher Weise von Spillover-Effekten betroffen sind, spiegelt sich dies in einer Variation der DvD-Schätzungen je nach Wahl der Kontrollgruppe wieder und hätte explizit in den Gutachten dahingehend untersucht werden können. Angesichts von Spillover-Effekten ist das Zusammenspiel der Ergebnisse für verschiedene Zielgrößen vor dem Hintergrund empirisch plausibler Verhaltensreaktionen (Arbeitsnachfrageelastizitäten) zu interpretieren. Darüber hinaus könnten synthetische Kontrollgruppen einen besseren Vergleichsmaßstab liefern und deren Zusammensetzung könnte durch Placebo-Tests überprüft werden. Wenn allerdings die Placebo-Tests zur Modellwahl (Wahl der Kontrollgruppe) verwendet werden, besteht die Gefahr eines statistischen Overfittings. Dem sollte durch eine weitere Überprüfung der Prognosegüte des Modells ausserhalb der Schätzstichprobe entgegengewirkt werden. Dazu sind hinreichend große Datensätze notwendig. Schließlich lassen sich Spillover-Effekte im Rahmen von strukturellen Schätzungen der Arbeitsnachfrage, des Arbeitsangebots, der Güternachfrage oder eines monopsonistischen Arbeitsmarktes explizit modellieren. Die strukturellen Modelle, die typischerweise auf stärkeren Modellierungsannahmen als die Schätzungen der Kausaleffekte beruhen, sollten so spezifiziert und kalibriert werden, dass sie die ohne weitergehende Annahmen an die Modellstruktur geschätzten Kausaleffekte replizieren.

Fünftens ist die externe Validität der Ergebnisse aufgrund der Spezifika der einzelnen Branchen und der unterschiedlichen Höhe der Mindestlöhne begrenzt. Es ist wichtig, für jede Branche einen angemessenen Evaluationsansatz zu wählen, der sich nach Branche unterscheiden kann. Zusätzlich ist es sinnvoll über die Branchen hinweg Schätzungen mit einem einheitlichen Forschungsdesign anzuwenden. Dies kann die externe Validität der einzelnen Ergebnisse erhöhen, da eventuelle Unterschiede in den Ergebnissen bei Anwendung des gleichen Forschungsdesigns auf unterschiedliche Branchen erklärt werden müssten. Es ist zwar so, dass drei Institute (IAQ, IAW, ZEW) an mehreren Gutachten beteiligt waren, aber es fand keine vergleichenden

de Diskussion der Ergebnisse über verschiedene Branchen statt. Bspw. wäre es interessant zu wissen, ob man auch für andere Branchen negative Beschäftigungseffekte wie für das Dachdeckerhandwerk findet, wenn man für die anderen Branchen den gleichen Evaluationsansatz wie im Gutachten für das Dachdeckerhandwerk anwendet.

Sechstens ist dem bei der Schätzung der Effekte von Branchenmindestlöhnen möglichen Problem der Politikendogenität Rechnung zu tragen. Politikendogenität tritt auf, wenn ein Mindestlohn in einer Branche nur dann eingeführt wird, wenn dies von Akteurer (Tarifparteien etc.) in einer Branche angestrebt wird.

Siebtens möchten wir festhalten, dass eine reine Zeitreihenevidenz keine Kausalanalyse ersetzen kann. Dennoch liefert sie eine wichtige deskriptive Evidenz. Beispielsweise stellt Manning (2013, S. 61–62) für Großbritannien graphisch die zeitliche Entwicklung der Lohnungleichheit und der Höhe des Mindestlohns relativ zum Median dar. Manning (2013) interpretiert dies als Hinweis darauf, dass ein Mindestlohn die Ungleichheit reduziert. Die in den Evaluationsstudien geschätzten Kausaleffekte sollten stärker im Lichte der deskriptiven Zeitreihenevidenz zu den Mindestlöhnen diskutiert werden, als dies bisher der Fall war.

Lassen sich belastbare inhaltliche Lehren aus der ersten Evaluationsrunde ziehen? Die verschiedenen Evaluationsstudien liefern ein äußerst vielschichtiges Bild (siehe Tab. 2–5). Generell ist festzuhalten, dass mögliche negative Beschäftigungseffekte von Mindestlöhnen geringer ausfallen, als dies von Kritikern befürchtet wurde (ifo Schnelldienst 2008). Ein eindrückliches Beispiel liefert der Mindestlohn im Dachdeckerhandwerk (Aretz et al. 2011). In dieser Branche liegt der Mindestlohn in Ostdeutschland über dem Median des Lohnes und es finden sich in eher geringem Umfang negative Beschäftigungseffekte. Gleichzeitig sind jedoch deutliche Anpassungsreaktionen in dieser Branche zu konstatieren. Zum einen ging die Dynamik der Beschäftigung zurück und zum anderen reduzierten sich die Lohnunterschiede im oberen Bereich der Lohnverteilung. Ein offener Punkt in den Gutachten ist die kausale Schätzung der durch den Mindestlohn ausgelösten Preisüberwälzung. Die vielschichtigen Anpassungs- und Ausweichreaktionen auf Firmenebene, die sich in verschiedenen Evaluationsstudien andeuten, und deren Bewertung sollten Themen zukünftiger Evaluationsstudien zu Mindestlöhnen in Deutschland werden.

**Open Access** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

## Literatur

- Abadie, A., Diamond, A., Hainmueller, J.: Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *J Am Stat Assoc* **115**, 493–505 (2010)
- Abadie, A., Diamond, A., Hainmueller, J.: Comparative politics and the synthetic control method. *Am J Pol Sci* **59**(2), 495–510 (2015)
- Aretz, B., Arntz, M., Gottschalk, S., Gregory, T., Niefert, M., Christian, R.: ZEW-Studie zu den Wirkungen eines Mindestlohns in der Dachdeckerwirtschaft. Bundesministerium für Arbeit und Soziales, Mannheim (2011)
- Athey, S., Imbens, G.W.: Identification and inference in nonlinear difference-in-differences models. *Econometrica* **74**(2), 431–497 (2006)
- Bauer, T.K., Fertig, M., Schmidt, C.M.: Empirische Wirtschaftsforschung. Springer-Verlag, Berlin/Heidelberg (2009)
- Boockmann, B., Buch, C.M., Schnitzer, M.: Evidenzbasierte Wirtschaftspolitik in Deutschland: Defizite und Potentiale. *Perspekt Wirtschaftspolitik* **15**(4), 307–323 (2014)
- Bosch, G., Hieming, B., Mesaros, L., Weinkopf, C.: Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Wäschereidienstleistungen im Objektkundengeschäft. Bundesministerium für Arbeit und Soziales. [http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-waeschereidienstleistungen.pdf?\\_\\_blob=publicationFile](http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-waeschereidienstleistungen.pdf?__blob=publicationFile) (2011a). Zugegriffen: 30. Juni 2016
- Bosch, G., Kalina, T., Kern, C., Neuffer, S., Schwarzkopf, M., Weinkopf, C.: Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Gebäudereinigung. Bundesministerium für Arbeit und Soziales. [http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-gebaedereinigung.pdf?\\_\\_blob=publicationFile](http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-gebaedereinigung.pdf?__blob=publicationFile) (2011b). Zugegriffen: 30. Juni 2016
- Card, D., Krueger, A.B.: Myth and measurement: The new economics of the minimum wage. Princeton University Press, Princeton, NJ (1995)
- Dolton, P., Rosazza-Bondibene, C., Wadsworth, J.: The UK national minimum wage in retrospect. *Fisc Stud* **31**(4), 509–534 (2010)
- Dolton, P., Rosazza-Bondibene, C., Wadsworth, J.: Employment, inequality and the UK national minimum wage over the medium-term. *Oxf Bull Econ Stat* **74**(1), 78–106 (2012)
- Dube, A., Zipperer, B.: Pooling multiple case studies using synthetic controls: an application to minimum wage policies. IZA Discussion Paper, Bd. 8944. (2015)
- Egeln, J., Gottschalk, S., Gürtzgen, N., Niefert, M., Rammer, C., Sprietsma, M.: ZEW-Studie zu den Wirkungen eines Mindestlohns in der Abfallwirtschaft. Bundesministerium für Arbeit und Soziales. [http://ftp.zew.de/pub/zew-docs/gutachten/Abfallendbericht\\_inkl\\_Anlagen.pdf](http://ftp.zew.de/pub/zew-docs/gutachten/Abfallendbericht_inkl_Anlagen.pdf) (2011). Zugegriffen: 30. Juni 2016
- Felder, R.: The minimum wage in the german roofing sector – an evaluation with the synthetic control method. University of Mannheim, Mannheim (2014). Unveröffentlichte Masterarbeit
- Fitzenberger, B.: Anmerkungen zur Mindestlohndebatte: Elastizitäten, Strukturparameter und Topfschlagen. *Zeitschrift für Arbeitsmarktforschung* **42**(1), 85–92 (2009)
- Fitzenberger, B., Hujer, R.: Stand und Perspektiven der Evaluation der aktiven Arbeitsmarktpolitik in Deutschland. *Perspekt Wirtschaftspolitik* **3**(2), 139–158 (2002)
- Franz, W.: Arbeitsmarktökonomik, 8. Aufl. Springer, Berlin, Heidelberg (2013)
- Frings, H., Berge, P. vom, Paloyo, A.R.: High-impact minimum wages and heterogeneous regions. ZBW-Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg (2013). Unveröffentlichte Arbeit
- Havnes, T., Mogstad, M.: Is universal child care leveling the playing field? *J Public Econ* **127**, 100–114 (2015)

- Heckman, J.J., LaLonde, R.J., Smith, J.A.: The economics and econometrics of active labor market programs. In: Ashenfelter, O., Card, D. (Hrsg.) *Handbook of labor economics*, Bd. 3, S. 1865–2097. (1999)
- ifo Schnelldienst: Mindestlohn: Für und Wider – Aufruf der Präsidenten und Direktoren der Wirtschaftsforschungsinstitute und Beiträge weiterer Wirtschaftswissenschaftler (2008). Ausgabe 6, ifo Institut München. <https://www.cesifo-group.de/DocDL/SD-6-08.pdf>, Zugegriffen: 10 Sept 15
- Imbens, G.M., Wooldridge, J.M.: Recent developments in the econometrics of program evaluation. *J Econ Lit* **47**(1), 5–86 (2009)
- Kaul, A., Klößner, S., Pfeifer, G., Schieler, M.: Synthetic control methods: never use all pre-intervention outcomes as economic predictors (2016). Unveröffentlichte Arbeit. [http://www.oekonometrie.uni-saarland.de/papers/SCM\\_Predictors.pdf](http://www.oekonometrie.uni-saarland.de/papers/SCM_Predictors.pdf), Zugegriffen: 30 Mai 2016
- Kirchmann, A., Koch, A., Krumm, R., Klee, G., Rosemann, M., Strotmann, H., Kleimann, R., Boockmann, B., Späth, J., Verbeek, H., Sappl, R., Harsch, K., Klempt, C., Neumann, M., Rattenhuber, P.: Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Elektrohandwerk. Bundesministerium für Arbeit und Soziales. [http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-elektrohandwerk.%20pdf?\\_\\_blob=publicationFile](http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-elektrohandwerk.%20pdf?__blob=publicationFile) (2011a). Zugegriffen: 30. Juni 2016
- Kirchmann, A., Koch, A., Krumm, R., Klee, G., Rosemann, M., Strotmann, H., Kleimann, R., Boockmann, B., Späth, J., Verbeek, H., Sappl, R., Harsch, K., Klempt, C., Neumann, M., Rattenhuber, P.: Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Maler- und Lackiererhandwerk. Bundesministerium für Arbeit und Soziales. [http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-maler-und-lackierhandwerk.pdf?\\_\\_blob=publicationFile](http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-maler-und-lackierhandwerk.pdf?__blob=publicationFile) (2011b). Zugegriffen: 30. Juni 2016
- Kirchmann, A., Koch, A., Krumm, R., Klee, G., Strotmann, H., Kleimann, R., Boockmann, B., Späth, J., Verbeek, H., Sappl, R., Harsch, K., Klempt, C., Neumann, M., Rattenhuber, P.: Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Pflege. Bundesministerium für Arbeit und Soziales. [http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-pflegebranche.pdf?\\_\\_blob=publicationFile](http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Meldungen/evaluation-mindestlohn-pflegebranche.pdf?__blob=publicationFile) (2011c). Zugegriffen: 30. Juni 2016
- König, M., Möller, J.: Mindestlohneffekte des Entsendegesetzes? Eine Mikrodatenanalyse für die deutsche Bauwirtschaft. *Zeitschrift für ArbeitsmarktForschung* **41**(2-3), 327–346 (2008)
- Lechner, M.: The estimation of causal effects by difference-in-difference methods. *Found Trends® Econom* **4**(3), 165–224 (2010)
- Manning, A.: Minimum wages: a view from the UK. *Perspekt Wirtschaftspolitik* **14**(1-2), 57–66 (2013)
- Meier, V., Munz, S.: Beschäftigungseffekte von Mindestlöhnen unter Vernachlässigung der Hauptbetroffenen: Kommentar zu König und Möller. *Ifo Schnelld* **61**(15), 30–32 (2008)
- Melly, B., Santangelo, G.: The changes-in-changes model with covariates. Universität Bern, Bern (2015). Unveröffentlichte Arbeit
- Möller, J., Bender, S., König, M., Vom Berge, P., Umkehrer, M., Wolter, S., Schaffner, S., Bachmann, R., Frings, H., Janßen-Timmen, R., Paloyo, A., Tamm, M., Fertig, M., Apel, H.: Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Bauhauptgewerbe. Bundesministerium für Arbeit und Soziales. <http://doku.iab.de/graeuepap/2011/evaluation-mindestlohn-bauhauptgewerbe.pdf> (2011). Zugegriffen: 30. Juni 2016
- Neumark, D., Wascher, W.L.: *Minimum wages*. MIT Press, Cambridge, Massachusetts (2008)
- Rattenhuber, P.: Building the minimum wage: the distributional impact of Germany's first sectoral minimum wage on wages and hours across different wage bargaining regimes. *Empir Econ* **46**(4), 1429–1446 (2014)
- ZEW [Aretz B, Arntz M, Bonin H, Butschek A, Dörr A, Fitzenberger B, Gregory T, Gürtzgen N, Stichnoth H, Walter T]: Vorbereitende Forschung für die zweite Evaluationsrunde Mindestlöhne: Verbesserung und Erweiterung der Evaluationsmethoden (2013). Projektbericht. [ftp://ftp.zew.de/pub/zew-docs/gutachten/Mindestlohn\\_Maerz2014.pdf](ftp://ftp.zew.de/pub/zew-docs/gutachten/Mindestlohn_Maerz2014.pdf), Zugegriffen: 30 Mai 2016

**Bernd Fitzenberger** hat seit April 2015 den Lehrstuhl für Ökonometrie an der Humboldt-Universität zu Berlin inne. Davor war er Professor in Dresden, Mannheim, Frankfurt und Freiburg. An der Stanford University promovierte er 1993 mit einem Ph.D. in Economics. 1998 habilitierte er sich an der Universität Konstanz. Seine Hauptforschungsgebiete sind Dynamik der Lohnstruktur und Beschäftigungsentwicklung, Gewerkschaften (Tarifbindung, Organisationsgrad, Lohnstrukturen und Beschäftigung), Evaluation von Maßnahmen der Arbeitsmarkt- und Bildungspolitik, berufliche Bildung und Übergang von Schule zu Beruf, Beschäftigung von Müttern, Evaluationsmethoden und die Methode der Quantilsregressionen.

**Annabelle Doerr** ist seit Oktober 2015 Forschungsreferentin am Walter Eucken Institut in Freiburg und Post-Doc im Bereich Arbeitsmarktökonomie an der Wirtschaftswissenschaftlichen Fakultät der Universität Basel. Im September 2015 schloss sie ihre Promotion an der Albert-Ludwigs Universität in Freiburg ab. Während ihrer Promotion war sie von 2010-2014 Stipendiatin des Graduiertenprogramms des Instituts für Arbeitsmarkt- und Berufsforschung (IAB) in Nürnberg. Ihre Forschungsschwerpunkte liegen in der empirischen Arbeitsmarktforschung insbesondere der Evaluation von Programmen der aktiven Arbeitsmarktpolitik, der Mindestlohnforschung und der Messung von Schwarzarbeit.