

Doerr, Annabelle; Fitzenberger, Bernd

Working Paper

Konzeptionelle Lehren aus der ersten Evaluationsrunde der Branchenmindestlöhne in Deutschland

ZEW Discussion Papers, No. 15-073 [rev.]

Provided in Cooperation with:

ZEW - Leibniz Centre for European Economic Research

Suggested Citation: Doerr, Annabelle; Fitzenberger, Bernd (2015) : Konzeptionelle Lehren aus der ersten Evaluationsrunde der Branchenmindestlöhne in Deutschland, ZEW Discussion Papers, No. 15-073 [rev.], Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim

This Version is available at:

<https://hdl.handle.net/10419/141455>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

Discussion Paper No. 15-073

**Konzeptionelle Lehren aus der
ersten Evaluationsrunde der
Branchenmindestlöhne in Deutschland**

Annabelle Doerr and Bernd Fitzenberger

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 15-073

**Konzeptionelle Lehren aus der
ersten Evaluationsrunde der
Branchenmindestlöhne in Deutschland**

Annabelle Doerr and Bernd Fitzenberger

Erste Version: Oktober 2015

Diese Version: Dezember 2015

Download this ZEW Discussion Paper from our ftp server:

<http://ftp.zew.de/pub/zew-docs/dp/dp15073.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von
neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung
der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other
economists in order to encourage discussion and suggestions for revisions. The authors are solely
responsible for the contents which do not necessarily represent the opinion of the ZEW.

Konzeptionelle Lehren aus der ersten Evaluationsrunde der Branchenmindestlöhne in Deutschland*

Annabelle Doerr

Walter Eucken Institut, WWZ Universität Basel

Bernd Fitzenberger

Humboldt-Universität zu Berlin, IFS, CESifo, IZA, ROA, ZEW

Erste Version: Oktober 2015

Diese Version: Dezember 2015

Zusammenfassung

Dieser Aufsatz liefert eine kritische Zusammenfassung und Beschreibung der vom 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 stärkere Evidenzbasierung der 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 hätten die Variation der geschätzten Effekte und die Eignung einer Kontrollgruppe stärker diskutiert und eine Abschätzung der Verzerrung in den geschätzten Effekten vorgenommen werden sollen.

Stichworte: Mindestlöhne, Differenz - von - Differenzen (DvD), Evaluationsstudien

*Dieser Aufsatz 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. Wir danken den TeilnehmerInnen des ZEW-Workshops 'Verbesserung und Erweiterung der Methoden zur Mindestlohnevaluation' in 2013 und der IZA-Konferenz 'A Minimum Wage for Germany? Learning from European Experience' in 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.

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 ökonometrischen 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 Beispiel für einer 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).

Tabelle 1 liefert einen Überblick über die in der ersten Evaluationsrunde betrachteten Branchen und der jeweiligen Evaluationsstudien. 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 Basis der im Rahmen der Evaluationen erhobenen Befragungsdaten analysiert werden.

Allen Studien in Tabelle 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 Regelung betroffenen Branchen, weil

Tabelle 1: Übersicht der Evaluationsstudien zu Branchenmindestlöhnen

Branchen- typ ^a	Branche	Evaluationsstudie ^b	Forschungseinrichtungen der AutorInnen der Studien ^c
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

Hinweise: ^a Branchentyp ist eine in diesem Aufsatz vorgenommene Einteilung der acht Branchen in drei Typen, die im Text in Abschnitt 3 erläutert wird, ^b siehe Literaturliste; ^c 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.

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 Kausalanalysen in der ersten Evaluationsrunde fast ausschließlich auf Differenz-von-Differenzen (DvD) Ansätzen. Diese erlauben 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 ökonometrischen 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).

Der Aufsatz geht bewusst nicht auf die umfangreichen Forschungsarbeiten ein, die im Nachgang zu den Evaluationsstudien entstanden sind (bspw. 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). Diese Studien stellen einen äußerst begrüßenswerten Mehrwert der Evaluationsstudien dar. Es ist sinnvoll, dass sich die Autoren der Evaluationsstudien erfolgreich einer wissenschaftlichen Qualitätskontrolle stellen.

Im Folgenden werden der in den ökonometrischen Kausalanalysen der Evaluationsstudien dominierende DvD-Ansatz und die zugrundeliegenden Annahmen erklärt. Anschließend folgt eine Systematisierung und Bewertung der vorliegenden Ergebnisse. 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 unsere 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.¹ 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.

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

¹Dieser Abschnitt kann keinen Überblick über die Methoden zur ökonometrischen Schätzung von Kausaleffekten liefern. Hierzu sei bspw. auf die Überblicksartikel von Imbens und Wooldridge (2013), Lechner (2010), Heckman et al. (1999) oder das Lehrbuch Bauer et al. (2010) sowie die dort zitierte Literatur verwiesen. Der Übersichtsartikel Fitzenberger und Hujer (2002) diskutiert die ökonometrische Schätzung von Kausaleffekten im Kontext der Analyse der Wirkungen der Aktiven Arbeitsmarktpolitik in Deutschland.

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. 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.²

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 sie weisen eine eher unterdurchschnittliche Beschäftigungsentwicklung auf. Für das Beispiel folgt, 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 DvD-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

²Eine konditionale DvD-Schätzung kann bspw. mit folgender Regression umgesetzt werden

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

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 ε_{jt} der Fehlerterm. Der Koeffizient des Interaktionseffektes zwischen Zeit und Treatmentgruppe, θ^{DiD} , liefert die DvD-Schätzung, da er den Unterschied in der zeitlichen Veränderung zwischen Treatment- und Kontrollgruppe beschreibt.

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 (antizipieren), 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, da sich damals absehbar die Baukonjunktur deutlich und nachhaltig abschwächte, vor allem in Ostdeutschland, und gleichzeitig mit einer weiteren Zunahme an ausländischen Beschäftigten gerechnet wurde (König und Möller, 2008, Abschnitt 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 Möglichkeit, 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). Als Alternative diskutiert Lechner (2010) den von Athey and Imbens (2006) vorgeschlagenen Changes-in-Changes Ansatz als skalierungsunabhängiges, 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. Umsetzungsschwierigkeiten treten bei diskret verteilten Zielgrößen auf.

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 Spillovereffekte 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 daraus, 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, Lohnkostenerhöhungen 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 Spillovereffekten 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. Daher steigen auch Löhne oberhalb des Mindestlohns.

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 (bspw. dominiert im Regelfall der Substitutionseffekt bei einer Kontrollgruppe aus der gleichen Branche im Regelfall den Skaleneffekt, siehe Fitzenberger, 2009) eine Einschätzung der Richtung der Verzerrung der geschätzten Kausaleffekte vorgenommen werden kann.

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. Tabelle 1). Die Gruppierung berücksichtigt vorhandene Strukturähnlichkeiten zwischen den einzelnen Branchen. Die Tabellen 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 Branchenkon-

Tabelle 2: Übersicht der Studien zum Branchentyp 1 – Teil 1

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

Quelle: Möller et al (2011), Aretz et al (2011), eigene Zusammenstellung.

junktur wieder ausgesetzt. Die sich hieraus ergebende Variation in den Mindestlohnregelungen, konnten die Evaluationsstudien zur empirischen Messung der Mindestlohneffekte nutzen.

Eine systematische Übersicht der Studien für den Branchentyp 1 finden sich in Tabelle 2 und 3. Alle vier Studien verfolgen ähnliche Kontrollgruppenansätze. Zum einen wird eine passende Kontrollbranche verwendet, die nicht einem allgemeinverbindlichen Mindestlohn unterliegt.

Tabelle 3: Übersicht der Studien zum Branchentyp 1 – Teil 2

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

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

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 sowohl für West- als auch für Ostdeutschland vor- und nachgelagerte Branchen sowie eine baunahe und bauferne Vergleichsbranche ausgewählt. Es ist anzumerken, dass Spillovereffekte auch bei vor- oder nachgelagerten Branchen über Input-Output-Beziehungen auftreten können, wenn auch in einem geringerem 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 sind. 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 sowie 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) verwendet die vorhandene Variation in den Daten, um geeignete Kontrollgruppen zu finden. Bei dieser Methode basiert die Auswahl geeigneter Kontrollgruppen nicht auf sachlogischen Überlegungen, sondern die Daten bestimmen auf Basis eines geeigneten statistischen Kriteriums die ähnlichsten Kontrollgruppen durch eine entsprechende Gewichtung. 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. Erhöhung mindestens zum Mindestlohn entlohnt wurden, verwenden die Studien zum Elektro-, Maler- und Lackiererhandwerk nur Personen, die vor Einführung des Mindestlohns mindestens diesen und maximal 15% darüber verdienten. Diese Vorgehensweise ist vor dem Hintergrund möglicher Spillovereffekte 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 Spillovereffekte 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. Desweiteren können Substitutions- und Skaleneffekte in der Arbeitsnachfrage eine Rolle spielen.

Auf Betriebsebene werden Betriebe als Kontrollbetriebe verwendet, die über dem gültigen

Mindestlohniveau entlohnend 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 erst möglich macht. Für das Bauhauptgewerbe und für das Dachdeckerhandwerk wurden erfreulicherweise 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 auch hier naheliegend gewesen, Placebo-Tests durchzuführen.

In der Evaluation zum Bauhauptgewerbe wird zusätzlich eine Variation des DvD-Ansatzes verwendet, der die Effekte als Funktion der Eingriffsintensität modelliert, was prinzipiell sinnvoll ist. Der sogenannte inkrementelle DvD-Schätzer (Dolton et al, 2010, 2012) kontrolliert für die strukturellen Besonderheiten zwischen Betroffenen und Nichtbetroffenen, indem die Eingriffsintensität vor der Mindestlohneinführung in die Schätzung aufgenommen wird. Dieser Koeffizient wird für die tatsächliche Eingriffsintensität *nach* der Mindestlohneinführung verwendet, um für die zeitvariable Selektion, der von den Mindestlohnregelungen Betroffenen zu kontrollieren. Der inkrementelle Effekt der tatsächlichen Eingriffsintensität nach der Mindestlohneinführung (Interaktionseffekt mit einer Jahresdummyvariablen) ist dann der geschätzte Effekt der Mindestlohnregelungen relativ zu dem Zeitraum vor der Mindestlohneinführung. Er stellt somit einen kumulativen Effekt dar.³

³Formal basiert der in der Evaluationsstudie für das Bauhauptgewerbe umgesetzte inkrementelle DvD-Schätzer auf folgender Regression (siehe auch Fußnote 2)

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

wobei Y_t Jahresdummies, J_j Dummies für Beobachtung j (Betrieb, Region) und P_{jt} die Eingriffsintensität des Mindestlohns (bspw. der Anteil der Beschäftigten in einer Region/einem Betrieb mit einem Lohn unterhalb des

Es ist möglich, dass die Eingriffsintensität selbst im Zeitverlauf von den Mindestlohnregelungen in der Vergangenheit beeinflusst wird. Eine solche Endogenität der Eingriffsintensität kann die Selektionskorrektur je nach Zielgröße auf Basis des vor der Mindestlohneinführung geschätzten Zusammenhangs verzerren. Weiterhin betreffen die kumuliert geschätzten Effekte die Wirkungen unterschiedlicher Mindestlohnregelungen für eine sich im Zeitverlauf verändernde Treatmentgruppe. Dies stellt, wie auch Dolton et al (2010, 2012) konstatieren, eine kausale Interpretation der inkrementellen Effekte relativ zu der Zeit vor Einführung des Mindestlohns in Frage. Eine Modifikation (Beschränkung) des inkrementellen DvD-Schätzers für die Veränderung zwischen zwei Zeitperioden auf Basis der Dummy für die Beobachtungseinheiten in der ersten Periode (oder stattdessen der Eingriffsintensität in der zweiten Periode) als zeitinvariant unterstellte Selektion und einer Interaktion der Eingriffsintensität mit einer Dummyvariable für die zweite Periode erschien uns daher als Robustheitscheck analog zu einer Standard-DvD-Schätzung sinnvoll. Wenn sich die geschätzten Mindestlohneffekte 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 Tabelle 2 und 3 zusammengefasst. Zur Gewährleistung der internen Validität der Ergebnisse ist die Durchführung umfangreicher Placebo-Tests prinzipiell zu befürworten.

Die Studie zum Mindestlohn im Bauhauptgewerbe unternimmt 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. Tabelle 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 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 liefern positive Effekte auf das Lohnwachstum in beiden Teilen Deutschlands für die Mindestlohneinführung 1997.

Die Untersuchungen zur Einführung des Mindestlohns II und weiterer Erhöhungen liefern keine robusten Ergebnisse (Möller et al, 2011, vgl. S.168 ff.). Grundsätzlich sollte die Auswahl

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 $\theta_0 P_{jt}$ wird auf Basis der Beobachtungen vor Einführung des Mindestlohns bis 1996 geschätzt (offen ist wie P_{jt} für die Jahre 1993 bis 1996 bestimmt werden kann - in Abgrenzung von J_j) und dieser Effekt kontrolliert ab 1997 für 'strukturelle' Besonderheiten je nach Betroffenheit durch den Mindestlohn, die jedoch nicht per se durch den Mindestlohn verursacht sind. θ_t^{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 θ_t^{DiD} zwischen zwei benachbarten Zeitpunkten zu interpretieren.

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. Es besteht sozusagen die Gefahr eines statistischen Overfittings. Deshalb schlagen wir in diesem Zusammenhang die Verwendung eines statistischen Regularisierungsverfahrens vor, das ein quantitatives Prüfverfahrens zur Bestimmung der Kontrollgruppe umfasst.

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 eben dieser 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 Lackiererhandwerk. Zudem liefern Robustheitsanalysen und Placebo-Tests Hinweise dafür, dass die Annahme paralleler Trends in diesen Studien häufig verletzt ist (Kirchmann et al, 2011b, vgl. S.198 ff., S.211 ff.) und (Kirchmann et al, 2011a, vgl. S.201 ff., S.211 ff., S.215 ff.). Zur Erhöhung der internen Validität der Ergebnisse erachten wir weitergehende Analysen (ggf. mit synthetischen Kontrollgruppen) als notwendig.

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 Abschnitt 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.194 ff.).

In den Evaluationen zum Bauhauptgewerbe und Dachdeckerhandwerk ergeben sich posi-

tive Effekte auf das Lohnwachstum. Die Analyse des Dachdeckerhandwerks kann sehr präzise Stundenlohninformationen benutzen, so dass eine Abgrenzungsproblematik zwischen den Vergleichsgruppen geringer einzuschätzen ist. In beiden Studien finden sich Hinweise für das Auftreten von Spillovereffekten. Diese werden in der Studie zum Dachdeckerhandwerk sehr ausführlich diskutiert. Im oberen Bereich der Lohnverteilung kommt es zu negativen Lohneffekten. Insgesamt geht die Lohnspreizung in Ostdeutschland stark zurück (Aretz et al, 2011, vgl. S.307 ff).

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 externe Validität der Evaluationsstudien für den Branchentyp 1 ist folglich beschränkt und es stellt sich die Frage nach der Politikendogenität der Einführung der Mindestlöhne.

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 Unterschieden 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.

Tabelle 4: Übersicht der Studien zum Branchentyp 2

	<i>Gebäudereinigung (IAQ)</i>	<i>Wäschereidienstleistungen im Objektkundengeschäft (IAQ)</i>
Mindestlohnregelung	Einführung im Jahr 2007 Erhöhung im Jahr 2008 Allgemeinverbindlichkeit von Tarifverträgen seit 1970er Jahren. Betrifft nur gewerblich Beschäftigte, d.h. nicht alle Beschäftigte, die Reinigungstätigkeiten durchführen.	Einführung im Jahr 2009 Jährliche Erhöhungen bis 2012
Daten	Eigene Betriebsbefragung Expertenbefragung Administrative Individualdaten Mikrozensus	Qualitative Interviews Amtliche Statistik Administrative Individualdaten Daten der Bundesagentur für Arbeit (BA-Daten)
Kontrollgruppen	Kontrollgruppen außerhalb der Branche 1. Köche und Kellner, 2. Berufe des Wach- und Sicherheitsgewerbe, 3. Innenreiniger in anderen Branchen, Kontrollgruppen innerhalb der Branche keine	Kontrollgruppen außerhalb der Branche keine Kausalanalyse durchgeführt Kontrollgruppen innerhalb der Branche 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 Wirtschaftszweigklassifikation (WKZ) und die Definition des Geltungsbereichs des Mindestlohns.
Ergebnisse	Beschäftigung (Regionenebene) Tariflohnänderung 2004: keine Effekte feststellbar Beschäftigung (Beschäftigtenebene) Mindestlohneinführung Vergleich KG1: keine Effekte Vergleich KG2/KG3: negative Effekte Mindestloohnerhöhung Vergleich KG1/KG2: positive Effekte Mindestloohnerhöhung Vergleich KG3: keine Effekte	Beschäftigung keine Kausalanalyse durchgeführt Löhne keine Kausalanalyse durchgeführt

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

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, standardisierte 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 Tabelle 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 Mindestlohnpflicht bestand. Hinzu kommt, dass ebenso ex post in administrativen Daten die genaue Abgrenzung der Branche, für die die Mindestlohnpflicht 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) durchgeführt werden können. 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.

Die Studie zur Gebäudereinigung Bosch et al (2011b) konnte trotz der beschränkten Datenverfügbarkeit einen Kontrollgruppenansatz umsetzen. 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 Mindestlöhne 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. Tabellen 27 und 28). Die Einführung der Mindestlöhne führt zu signifikant negativen Effekten im Vergleich zu den erstgenannten Kontrollgruppen. Die Effekte der Mindestloohnerhöhung im Jahr 2008 erweisen sich hingegen als positiv (Bosch et al, 2011b, vgl. Tabelle 29).

Die interne Validität der Ergebnisse für die Gebäudereinigungsbranche ist kritisch zu bewerten. Zur Überprüfung der Plausibilität der Annahme paralleler Trends werden zeitliche Entwick-

lungen der relevanten Zielgrößen für die Treatment- und Kontrollgruppe dargestellt (Bosch et al, 2011b, vgl. S.250 ff). 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 Abschnitt 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 sind. Die Anwendung von Placebo-tests kann dazu beitragen die interne Validität der Untersuchung zu verbessern. Eine Verwendung synthetischer Kontrollgruppen aus mehreren anderen Branchen 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 Mindestlohns zum Zeitpunkt der Erstellung der Gutachten extrem 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 Vermei-

Tabelle 5: Übersicht der Studien zum Branchentyp 3

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

Quelle: Kirchmann et al (2011c), Egel et al (2011), eigene Zusammenstellung.

dung von Verzerrungen durch die Berechnung aus anderen Datensätzen. Desweiteren 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 auch, durch Multiproduktunternehmen gekennzeichnet ist, wie beispielsweise Dienstleistungen in der Objektbetreuung oder Hausmeistertätigkeiten. Auch 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. Tabelle 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 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. Tabelle 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 des-

halb 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 entlohnnten. 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 entlohnnten. Für die Abfallwirtschaft werden mögliche Spillovereffekte 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. Tabelle 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, so dass 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 Grundgesamtheit 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 ökonometrische 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 ökonometrischen 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

im Jahr 2011 in Auftrag gab. Der Fokus in diesem Aufsatz liegt auf der ökonometrische Wirkungsanalyse. Trotz aller Detailkritik in diesem Aufsatz 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 vorgelegten Evaluationsstudien können zum Teil erstmalig tatsächliche Daten vor und nach Einführung von Mindestlöhnen verwenden und bei allen Unterschieden sind sich die Gutachten darin einig, dass ein Kontrollgruppenansatz zur Ermittlung kausaler Wirkungen notwendig ist. Die kritischen Ausführungen in diesem Aufsatz illustrieren exemplarisch wichtige Herausforderungen auf dem *steinigen Weg* zu einer evidenzbasierten Wirtschaftspolitik. Wir beschränken uns auf die Evaluationsstudien, in der Form wie sie vom Ministerium rezipiert und veröffentlicht wurden, um hieran exemplarisch eine Diskussion der Qualität der Studien anzustoßen, wie sie in den politischen Prozess einfließen. Die Evaluationsstudien sind unter einem starken Zeitdruck entstanden. In einigen Fällen lagen nicht einmal administrative Daten für den Zeitraum nach Einführung der Mindestlöhne vor und deshalb wurden teilweise Befragungsdaten erhoben. Diese Schwierigkeiten relativieren teilweise die kritischen Einschätzungen in diesem Aufsatz.

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 notwendig, um zukünftige Evaluationen zu verbessern und ein kontinuierliches Monitoring der Wirkungen des Mindestlohns zu ermöglichen, wie es durch die Low-Pay-Commission in Großbritannien erfolgt und wie es auch von Seiten der Mindestlohnkommission in Deutschland in Zukunft erfolgen sollte.

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, die Studien entwickeln ihre Kontrollgruppendesigns in unterschiedlicher Weise. 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 Kontrolleinheiten 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 erfordert. Die Studien verwenden zu Recht Placebo-Tests, um die Plausibilität der Annahme paralleler Trends zwischen Treatment- und Kontrollgruppen zu stützen. Allerdings werden die Fälle der Ablehnung der Placebo-Tests meist nicht inhaltlich weiter 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, das sogenannte Changes-in-Changes (CiC) Verfahren (Athey and Imbens, 2006) wäre für stetige Zielgrößen eine sinnvolle Ergänzung zu DvD-Ansätzen, da es die Skalierungsabhängigkeit der identifizierenden Annahmen des DvD-Ansatzes umgeht. Es ist möglich heterogene Mindestlohnwirkungen entlang der Verteilung unterschiedlich skaliert Zielgrößen zu ermitteln. 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. Abschließend ist anzumerken, dass sowohl der synthetische Kontrollgruppenansatz als auch das CiC-Verfahren - ebenso wie der DvD-Ansatz - auf der SUTVA-Annahme beruhen, die Spillovereffekte ausschließt.

Viertens, in einigen Studien wurde eine mögliche Verletzung 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 Erkennt-

nisse 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 expliziter in den Gutachten untersucht werden können. Angesichts von Spillovereffekten wäre auch das Zusammenspiel der Ergebnisse für verschiedene Zielgrößen vor dem Hintergrund empirisch plausibler Verhaltensreaktionen (Arbeitsnachfrageelastizitäten) zu interpretieren. Darüberhinaus könnten synthetische Kontrollgruppen einen besseren Vergleichsmaßstab liefern und deren Zusammensetzung könnte durch Placebo-Tests überprüft werden. 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, aufgrund der Spezifika der einzelnen Branchen und der unterschiedlichen Höhe der Mindestlöhne ist die externe Validität der Ergebnisse begrenzt. Für zukünftige Evaluationen ist es sinnvoll in noch stärkerem Umfang ein einheitliches 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 vergleichende Diskussion der Ergebnisse über verschiedene Branchen statt.

Sechstens, bei der Schätzung der Effekte von Branchenmindestlöhnen ist dem möglichen Problem der Politikendogenität Rechnung zu tragen. Diese besteht darin, dass ein Mindestlohn in einer Branche nur dann eingeführt wird, wenn die Akteure (Tarifparteien etc.) in einer Branche dies anstreben.

Siebtens, obwohl reine Zeitreihenevidenz keine Kausalanalyse ersetzen kann, so liefert sie doch 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 Tabelle 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 Dachdeckergewerbe (Aretz et al, 2011). In dieser Bran-

che 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.

Literatur

- Abadie A, Diamond A, Hainmueller J (2010) Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association* 115:493–505
- Aretz B, Arntz M, Gottschalk S, Gregory T, Niefert M, Christian R (2011) ZEW-Studie zu den Wirkungen eines Mindestlohns in der Dachdeckerwirtschaft. Bundesministerium für Arbeit und Soziales, Mannheim
- Athey S, Imbens GW (2006) Identification and inference in nonlinear difference-in-differences models. *Econometrica* 74(2):431–497
- Bauer TK, Fertig M, Schmidt CM (2009) Empirische Wirtschaftsforschung. Berlin/Heidelberg
- Boockmann B, Buch CM, Schnitzer M (2014) Evidenzbasierte Wirtschaftspolitik in Deutschland: Defizite und Potentiale. *Perspektiven der Wirtschaftspolitik* 15(4):307–323
- Bosch G, Hieming B, Mesaros L, Weinkopf C (2011a) Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Wäschereidienstleistungen im Objektkundengeschäft. Bundesministerium für Arbeit und Soziales
- Bosch G, Kalina T, Kern C, Neuffer S, Schwarzkopf M, Weinkopf C (2011b) Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Gebäudereinigung. Bundesministerium für Arbeit und Soziales
- Dolton P, Rosazza-Bondibene C, Wadsworth J (2010) The UK national minimum wage in retrospect. *Fiscal studies* 31(4):509–534
- Dolton P, Rosazza-Bondibene C, Wadsworth J (2012) Employment, inequality and the UK national minimum wage over the medium-term. *Oxford Bulletin of Economics and Statistics* 74(1):78–106

- Egeln J, Gottschalk S, Gürtzgen N, Niefert M, Rammer C, Sprietsma M (2011) ZEW-Studie zu den Wirkungen eines Mindestlohns in der Abfallwirtschaft. Bundesministerium für Arbeit und Soziales
- Fitzenberger B (2009) Anmerkungen zur Mindestlohn-debatte: Elastizitäten, Strukturparameter und Topfschlagen. *Zeitschrift für Arbeitsmarktforschung* 42(1):85–92
- Fitzenberger B, Hujer R (2002) Stand und Perspektiven der Evaluation der aktiven Arbeitsmarktpolitik in Deutschland. *Perspektiven der Wirtschaftspolitik* 3(2):139–158
- Franz W (2013) *Arbeitsmarktökonomik*, 8. Auflage. Springer Berlin Heidelberg
- Heckman JJ, LaLonde RJ, Smith JA (1999) The economics and econometrics of active labor market programs. In: Ashenfelter O, Card D (eds), *Handbook of labor economics* 3:1865–2097
- ifo Schnelldienst (2008) Mindestlohn: Für und Wider - Aufruf der Präsidenten und Direktoren der Wirtschaftsforschungsinstitute und Beiträge weiterer Wirtschaftswissenschaftler. Ausgabe 6, ifo Institut München, <https://www.wcesifo-group.de/DocDL/SD-6-08pdf> [Heruntergeladen am 10 Sept 15]
- Imbens GM, Wooldridge JM (2008) Recent developments in the econometrics of program evaluation. Discussion Paper No w14251, National Bureau of Economic Research
- 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 (2011a) Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Elektrohandwerk. Bundesministerium für Arbeit und Soziales
- 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 (2011b) Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Maler- und Lackiererhandwerk. Bundesministerium für Arbeit und Soziales
- 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 (2011c) Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Pflege. Bundesministerium für Arbeit und Soziales
- König M, Möller J (2008) Mindestlohneffekte des Entsendegesetzes? Eine Mikrodatenanalyse für die deutsche Bauwirtschaft. *Zeitschrift für Arbeitsmarktforschung* 41(2-3):327–346

- Lechner M (2010) The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics* 4(3):165–224
- Manning A (2013) Minimum wages: a view from the UK. *Perspektiven der Wirtschaftspolitik* 14(1-2):57–66
- 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 (2011) Evaluation bestehender gesetzlicher Mindestlohnregelungen Branche: Bauhauptgewerbe. Bundesministerium für Arbeit und Soziales