

Elterngeldreform



Bericht von Mathias Huebener, Daniel Kühnle und C. Katharina Spieß

Einführung des Elterngeldes hat Ungleichheit
in kindlicher Entwicklung nicht erhöht 519

Interview mit C. Katharina Spieß

»Positive Effekte des Elterngeldes werden
mit Blick auf kindliche Entwicklung nicht geschmälert« 530

Bericht von Tim Kaiser und Lukas Menkhoff

Schlechte Evaluierung rentiert sich kaum:
Lehren aus dem Bereich der finanziellen Bildung 531

Bericht von Alexander S. Kritikos und Anselm Mattes

In Deutschland sinkt die Zahl
der Betriebsgründungen weiterhin,
nicht aber in Berlin 539

Am aktuellen Rand Kommentar von Claus Michelsen

Die Mietpreisbremse ist keine Investitionsbremse 548



DIW Berlin – Deutsches Institut
für Wirtschaftsforschung e.V.
Mohrenstraße 58, 10117 Berlin
T +49 30 897 89 -0
F +49 30 897 89 -200
84. Jahrgang
27. Juni 2017

Herausgeberinnen und Herausgeber

Prof. Dr. Tomaso Duso
Dr. Ferdinand Fichtner
Prof. Marcel Fratzscher, Ph.D.
Prof. Dr. Peter Haan
Prof. Dr. Claudia Kemfert
Prof. Dr. Lukas Menkhoff
Prof. Johanna Mollerstrom, Ph.D.
Prof. Karsten Neuhoff, Ph.D.
Prof. Dr. Jürgen Schupp
Prof. Dr. C. Katharina Spieß
Prof. Dr. Gert G. Wagner

Chefredaktion

Dr. Gritje Hartmann
Dr. Wolf-Peter Schill

Redaktion

Renate Bogdanovic
Dr. Franziska Bremus
Prof. Dr. Christian Dreger
Sebastian Kollmann
Markus Reiniger
Mathilde Richter
Miranda Siegel
Dr. Alexander Zerrahn

Lektorat

Dr. Johannes Geyer
Felix Weinhardt, Ph.D.
Alexander Eickelpasch

Vertrieb

DIW Berlin Leserservice
Postfach 74
77649 Offenburg
leserservice@diw.de
Tel. (01806) 14 00 50 25
20 Cent pro Anruf
ISSN 0012-1304
ISSN 1860-8787 (Online)

Gestaltung

Edenspiekermann

Satz

eScriptum GmbH & Co KG, Berlin

Druck

USE gGmbH, Berlin

Nachdruck und sonstige Verbreitung –
auch auszugsweise – nur mit Quellen-
angabe und unter Zusendung eines
Belegexemplars an die Serviceabteilung
Kommunikation des DIW Berlin
(kundenservice@diw.de) zulässig.

Gedruckt auf 100 % Recyclingpapier.



Der DIW Wochenbericht wirft einen unabhängigen Blick auf die Wirtschaftsentwicklung in Deutschland und der Welt. Er richtet sich an die Medien sowie an Führungskräfte in Politik, Wirtschaft und Gesellschaft. Wenn Sie sich für ein Abonnement interessieren, können Sie zwischen den folgenden Optionen wählen:

Standard-Abo: 179,90 Euro im Jahr (inkl. MwSt. und Versand).

Studenten-Abo: 49,90 Euro.

Probe-Abo: 14,90 Euro für sechs Hefte.

Bestellungen richten Sie bitte an leserservice@diw.de oder den DIW Berlin Leserservice, Postfach 74, 77649 Offenburg; Tel. (01806) 14 00 50 25, 20 Cent/Anruf aus dem dt. Festnetz, 60 Cent maximal/Anruf aus dem Mobilnetz. Abbestellungen von Abonnements spätestens sechs Wochen vor Laufzeitende

NEWSLETTER DES DIW BERLIN



Der DIW Newsletter liefert Ihnen wöchentlich auf Ihre Interessen zugeschnittene Informationen zu Forschungsergebnissen, Publikationen, Nachrichten und Veranstaltungen des Instituts: Wählen Sie bei der Anmeldung die Themen und Formate aus, die Sie interessieren. Ihre Auswahl können Sie jederzeit ändern, oder den Newsletter abbestellen. Nutzen Sie hierfür bitte den entsprechenden Link am Ende des Newsletters.

>> Hier Newsletter des DIW Berlin abonnieren: www.diw.de/newsletter

RÜCKBLLENDE: IM WOCHENBERICHT VOR 40 JAHREN

Zur längerfristigen Entwicklung der Wohnungsnachfrage

Die Entwicklung der Wohnungsbautätigkeit wird derzeit unter konjunkturellen wie strukturpolitischen Gesichtspunkten diskutiert. Aus konjunkturellen Gründen erschien es zunächst notwendig, einem weiteren Rückgang der Baunachfrage entgegenzuwirken, der nach dem Auslaufen wichtiger Wohnungsbau-Förderungsprogramme zu erwarten gewesen wäre. So hat das Bundeskabinett die Fortsetzung des Regionalprogramms bis 1978 beschlossen, das nach dem ursprünglichen Finanzplan bereits in diesem Jahr auslaufen sollte. Entscheidungen über Umfang und Struktur der Wohnungsbauförderung auf längere Frist stehen indes noch nicht.

Die Zahl der jährlich fertiggestellten Wohnungen ging nach 1973 stark zurück. 1976 wurden erstmals seit 1950 weniger als 400 000 Wohnungen gebaut. Die Baugenehmigungen, die den Fertigstellungen um ein bis zwei Jahre vorauslaufen, liegen seit 1975 weit unter diesem Niveau und dürften in diesem Jahr kaum die Zahl von 350 000 Wohnungen überschreiten.

aus dem Wochenbericht Nr. 26 vom 30. Juni 1977

Einführung des Elterngeldes hat Ungleichheit in kindlicher Entwicklung nicht erhöht

Von Mathias Huebener, Daniel Kühnle und C. Katharina Spieß

Die Reform des Elterngeldes hat nicht dazu geführt, dass Ungleichheiten in der kindlichen Entwicklung zunehmen. Das zeigt die vorliegende Studie des DIW Berlin, die als eine der ersten anhand der Ergebnisse von Schuleingangsuntersuchungen der Frage nachgeht, welche Effekte die Einführung des Elterngeldes auf sprachliche und motorische Fähigkeiten, die sozio-emotionale Stabilität und den schulischen Förderbedarf hatte. Als das Elterngeld im Jahr 2007 das Bundeserziehungsgeld ersetzte, mutmaßten Kritikerinnen und Kritiker der Reform, dass sich die Ungleichheit erhöhen könnte. Hintergrund ist, dass im Gegensatz zum Erziehungsgeld auch viele Familien im mittleren und oberen Einkommensbereich nicht unerhebliche staatliche Leistungen erhalten. Somit können nun auch besser gebildete Mütter häufiger und länger vom Job pausieren und damit grundsätzlich mehr Zeit mit ihren Kindern verbringen. Dennoch zeigen sich sowohl für Kinder bildungsnaher als auch für Kinder weniger gebildeter Eltern trotz sehr präziser Schätzergebnisse keine statistisch signifikanten Effekte der Elterngeldreform auf die erfassten Entwicklungsmaße. Somit werden die anderen Effekte des Elterngeldes – etwa eine höhere wirtschaftliche Stabilität von Familien und die insgesamt gestiegene Erwerbstätigkeit von Müttern – nicht geschmälert.

Die Elterngeldreform war eine der bedeutendsten familienpolitischen Reformen der vergangenen Jahrzehnte. Zum 1. Januar 2007 wurde das Bundeserziehungsgeld durch eine Lohnersatzleistung für denjenigen Elternanteil abgelöst, der in den ersten Lebensmonaten eines Kindes seine Erwerbsarbeit unterbricht oder reduziert, um das eigene Kind zu betreuen. Die Wirkungen dieser Reform wurden vielfältig untersucht¹, insbesondere die Auswirkungen auf die wirtschaftliche Situation der Familien und auch das Erwerbsverhalten der Mütter und Väter. Darüber hinaus standen die Folgen auf die Geburtenrate im Fokus. Vereinzelt wurden auch andere Aspekte wie das elterliche Wohlbefinden oder die elterliche Arbeitsteilung bei der Haus- und Familienarbeit analysiert. Zu den Wirkungen des Elterngeldes auf die kindliche Entwicklung liegen bislang allerdings so gut wie keine Befunde vor. Der vorliegende Bericht adressiert diese Forschungslücke auf Basis von Daten aus Schuleingangsuntersuchungen.²

Große Unterschiede in der Wirkung auf Einkommen und Erwerbstätigkeit

Um mögliche Effekte des Elterngeldes auf die kindliche Entwicklung zu untersuchen, ist es zentral, zunächst die Wirkungen des Elterngeldes auf Eltern zu betrachten. Dabei muss berücksichtigt werden, dass sich mit dem Elterngeld einhergehende Veränderungen unterschiedlich auf Eltern auswirken.

Vor der Reform konnten Eltern das Bundeserziehungsgeld beantragen. Dabei handelte es sich um eine bedarfs-

¹ Für eine zusammenfassende Darstellung insbesondere der ökonomischen Wirkungsforschung zum Elterngeld vgl. Mathias Huebener et al. (2016): 10 Jahre Elterngeld: Eine wichtige familienpolitische Maßnahme. DIW Wochenbericht Nr. 49, 1159–1166 (online verfügbar, abgerufen am 16. Juni 2017. Dies gilt auch für alle anderen Online-Quellen dieses Berichts, sofern nicht anders vermerkt).

² Für eine ausführliche Beschreibung der Analysen vgl. Mathias Huebener, Daniel Kühnle und C. Katharina Spieß (2017): Paid Parental Leave and Child Development: Evidence from the 2007 German Parental Benefit Reform and Administrative Data. DIW Discussion Paper Nr. 1651 (online verfügbar).

Tabelle 1

Bezug von Erziehungsgeld (vor 2007) bzw. Elterngeld (ab 2007)

	Alle	Schulabschluss der Mutter	
		Haupt-/Realschule	(Fach-)Abitur
Jährliches Nettohaushaltseinkommen vor der Geburt in Euro	31 712	27 268	37 531
<i>Für Geburten vor dem 1. Januar 2007 gelten die Regelungen zum Erziehungsgeld</i>			
Bezug für bis zu sechs Monate, Anteil in Prozent	77,3	84,1	71,1
Bezug für sechs bis zwölf Monate, Anteil in Prozent	47,1	53,0	39,8
Bezug für mehr als zwölf Monate, Anteil in Prozent	39,9	45,3	33,0
Anzahl der Beobachtungen	311	173	138
<i>Für Geburten ab dem 1. Januar 2007 gelten die Regelungen zum Elterngeld</i>			
Bezug, Anteil in Prozent	fast 100	fast 100	fast 100
Monatliches Elterngeld von Müttern in Euro	634	563	771
Anteil der Väter in Elternzeit in Prozent	12,8	9,3	20,9
Monatliches Elterngeld von Vätern in Euro	1 061	864	1 190
Anzahl der Beobachtungen	197	124	73

Anmerkung: Ausgeschlossen sind Angaben von Beamtinnen und Beamten sowie selbständigen Müttern.

Quelle: Eigene Berechnungen basierend auf SOEPv30 für Kinder geboren zwischen 2005 und 2008.

© DIW Berlin 2017

geprüfte Transferleistung, die – abhängig vom Haushaltseinkommen – in einer Höhe von bis zu 300 Euro im Monat für bis zu zwei Jahre gezahlt wurde. Haushalte, die ein bestimmtes Einkommen überschritten, waren nicht anspruchsberechtigt. Die Einkommenshöchstgrenze sank mit dem Alter des Kindes und nach sechs beziehungsweise zwölf Monaten reduzierten sich die Leistungen weiter.³ Berechnungen auf Basis des Sozio-oekonomischen Panels (SOEP)⁴, einem für Deutschland repräsentativen Haushalts- und Personenpanel, zeigen, dass in den ersten sechs Monaten nach der Geburt eines Kindes ungefähr 77 Prozent der Eltern Erziehungsgeld erhal-

3 Der Bezug von Erziehungsgeld für die ersten sechs Monate nach der Geburt des Kindes war nur möglich, wenn das pauschalierte Nettoeinkommen eines Paares unter 30 000 Euro lag (23 000 Euro bei Alleinerziehenden). Ab dem siebten Lebensmonat betrug diese Grenze 16 500 Euro für ein Paar (13 500 Euro für Alleinerziehende). Überstieg das pauschalierte Nettoeinkommen diese niedrigeren Einkommensgrenzen, wurde das Erziehungsgeld nicht ganz entzogen, sondern gemindert. Der Regelbetrag verringerte sich um 5,2 Prozent von jenem Teil des Einkommens, das die Grenzen überstieg. Das geminderte Erziehungsgeld gab es nur für Familien mit pauschalitem Nettoeinkommen bis 22 086 Euro (Paare) beziehungsweise bis 19 086 Euro (Alleinerziehende), siehe Huebener et al. (2016), a. a. O.

4 Vgl. Gert G. Wagner, Joachim R. Frick und Jürgen Schupp (2007): The German Socio-Economic Panel Study (SOEP): Scope, evolution and enhancements. Schmollers Jahrbuch 127(1), 139-169.

ten hatten, während es im Zeitraum von sechs bis zwölf Monaten nach der Geburt nur noch 47 Prozent waren (Tabelle 1). Nach dem ersten Jahr sank der Anteil derjenigen, die Erziehungsgeld bezogen, auf etwa 40 Prozent.

Für Kinder, die ab dem 1. Januar 2007 geboren wurden, hat das Elterngeld das Erziehungsgeld abgelöst. Das Elterngeld beträgt als Lohnersatzleistung 67 Prozent des vorherigen Nettoerwerbseinkommens⁵, maximal 1 800 Euro im Monat für ein Jahr. Ein Mindestbetrag von 300 Euro im Monat wird unabhängig vom vorherigen Erwerbseinkommen gezahlt.⁶ Geht der zweite Elternteil ebenfalls in Elternzeit, kann die Familie insgesamt bis zu 14 Monate Elterngeld beziehen.⁷ Folglich hat die Elterngeldreform im Vergleich zum Erziehungsgeld die Höhe der Geldleistung und die Bezugsdauer verändert. Daten des Statistischen Bundesamtes zufolge beziehen etwa 96 Prozent der Eltern Elterngeld⁸, wobei Mütter durchschnittlich 634 Euro erhalten und Väter etwa 1 061 Euro (Tabelle 1).

Regelungen zur Elternzeit blieben durch die Einführung des Elterngeldes weitgehend unverändert: Sie kann weiterhin drei Jahre nach Geburt des Kindes in Anspruch genommen werden. Regelungen zu einer Teilzeiterwerbstätigkeit bis zu 30 Stunden pro Woche blieben ebenfalls unverändert. Bei einer Teilzeiterwerbstätigkeit reduziert sich das Elterngeld.

Insgesamt unterscheiden sich die Wirkungen der Reform stark nach dem Einkommen vor der Geburt des Kindes: Mütter mit einem höheren Erwerbseinkommen und einem höheren Haushaltseinkommen erhielten vor der Reform entweder nur sechs Monate Erziehungsgeld oder keine Leistungen. Nach der Reform erhalten sie nun Elterngeld für das gesamte erste Lebensjahr des Kindes, wenn sie ihre Erwerbstätigkeit entsprechend unterbrechen beziehungsweise reduzieren. Im zweiten Jahr nach der Geburt des Kindes hat sich für sie nur insofern etwas geändert, als sich der Elterngeldbezug um zwei Monate erhöht, wenn auch der Partner zur Betreu-

5 Im Detail stellt sich die Regelung wie folgt dar: Bei (Netto-)Erwerbseinkommen zwischen 1 000 und 1 200 Euro ersetzt das Elterngeld das nach der Geburt wegfallende Einkommen zu 67 Prozent. Für Geringverdienerinnen und Geringverdiener mit einem Einkommen unter 1 000 Euro vor der Geburt des Kindes steigt die Ersatzrate schrittweise auf bis zu 100 Prozent. Für Nettoeinkommen ab 1 200 Euro vor der Geburt des Kindes sinkt die Ersatzrate des Elterngeldes von 67 auf 65 Prozent, vgl. auch Huebener et al. (2016), a. a. O.

6 Beziehen Familien neben Elterngeld auch noch Leistungen nach dem Sozialgesetzbuch (SGB II), wird seit dem Jahr 2011 der Mindestbetrag von 300 Euro des Elterngeldes als Einkommen angerechnet.

7 Auch Alleinerziehende können bis zu 14 Monate Elterngeld beziehen. Typischerweise beziehen Mütter über einen deutlich längeren Zeitraum Elterngeld als Väter.

8 Vgl. Statistisches Bundesamt (2008): Öffentliche Sozialleistungen. Statistik zum Elterngeld. Elterngeld für Geburten 2007. Anträge von Januar 2007 bis Juni 2008 (online verfügbar).

ung des Kindes seine Erwerbsarbeit unterbricht. Mütter im unteren Einkommensbereich hingegen hatten vor der Elterngeldreform bis zu zwei Jahre Anspruch auf Erziehungsgeld. Mit dem Elterngeld können sie beziehungsweise ihr Haushalt eine entsprechende Leistung für maximal 14 Monate beziehen.

Betrachtet man verschiedene Bildungsgruppen, erhalten Mütter mit einem höheren Bildungsabschluss (Fachabitur oder allgemeine Hochschulreife) im Durchschnitt ein höheres Elterngeld als Mütter mit einem niedrigeren Bildungsniveau (Haupt- oder Realschulabschluss, Tabelle 1). Da das Elterngeld in Form einer Lohnersatzleistung ausbezahlt wird, ist dies zu erwarten, denn Mütter mit niedrigerer Bildung erzielen in der Regel ein geringeres Einkommen. So war der Anteil derjenigen Eltern, die Erziehungsgeld bezogen, unter den Müttern mit Abitur deutlich geringer als unter Müttern mit einer niedrigen oder mittleren Schulbildung: 40 Prozent der Mütter mit Abitur bezogen länger als sechs Monate Elterngeld, während es bei Müttern mit einem niedrigeren Schulabschluss 53 Prozent waren. Im zweiten Lebensjahr eines Kindes erhielten 33 Prozent der Mütter mit Abitur und 45 Prozent der Mütter mit geringerer Schulbildung Erziehungsgeld. Seit Einführung des Elterngeldes sind nahezu alle Mütter bezugsberechtigt. Mütter mit einer höheren Schulbildung bekommen durchschnittlich 771 Euro im Monat, andere Mütter erhalten 563 Euro. Hinzu kommt, dass die Partnermonate eher genommen werden, wenn beide Partner über einen hohen Bildungsabschluss verfügen.⁹ Die resultierenden Umverteilungswirkungen haben der Elterngeldreform von verschiedenen Seiten Kritik eingebracht: Angesichts ohnehin schon bestehender sozialer Ungleichheiten zwischen Kindern wurde bemängelt, dass einkommensstärkere Haushalte in höherem Maße profitieren als einkommensschwächere.¹⁰ Somit könnten bestehende Ungleichheiten zunehmen.

Mit der Elterngeldreform gehen auch unterschiedliche Wirkungen auf die mütterliche Erwerbstätigkeit einher. Mehrere bereits vorliegende Studien zeigen, dass bedingt durch das Elterngeld mehr Mütter mit hohem Einkommen als zuvor ihre Erwerbstätigkeit im ersten Jahr nach der Geburt des Kindes reduzieren. Demgegenüber erhöhen Mütter mit eher geringem Einkommen ihre Erwerbstätigkeit im zweiten Lebensjahr des Kindes, was mit der insgesamt kürzeren Bezugsdauer des Elterngeldes zusammenhängt.¹¹

⁹ Vgl. Heike Trappe (2013): Väterzeit – das Elterngeld als Beschleuniger von Gleichstellung? Zeitschrift für Familienforschung, 25(2), 238–264.

¹⁰ Vgl. zum Beispiel Annette Henninger et al. (2008): Geschlechtergleichheit oder „exklusive Emanzipation“? Ungleichheitssoziologische Implikationen der aktuellen familienpolitischen Reformen. Berliner Journal für Soziologie 18 (1), 99–128.

¹¹ Vgl. zum Beispiel Katharina Wrohlich et al. (2012): Elterngeld Monitor. DIW Politikberatung kompakt Nr. 61 (online verfügbar); Jochen Kluge und

Mögliche Wirkungen der Elterngeld-einführung auf kindliche Entwicklung

Welche Effekte kann die Einführung des Elterngeldes auf die kindliche Entwicklung haben? Und wie könnte sich die Reform auf Entwicklungsunterschiede zwischen Kindern mit unterschiedlichem sozioökonomischem Hintergrund auswirken? Geht man davon aus, dass ein verändertes Erwerbsverhalten einer Mutter den Umfang der Zeit beeinflusst, die sie mit ihrem Kind oder ihren Kindern verbringt, und nimmt man zusätzlich an, dass mehr Zeit für die Kinder einen positiven Einfluss auf die kindliche Entwicklung hat, lässt sich folgender Zusammenhang vermuten: Die Entwicklung von Kindern bildungsnaher Mütter könnte sich verbessern, da sie aufgrund der reduzierten Erwerbstätigkeit im ersten Jahr nach der Geburt grundsätzlich mehr Zeit mit ihren Kindern verbringen können. Die Entwicklung von Kindern bildungsferner Mütter könnte hingegen negativ beeinflusst werden, da diese Mütter infolge der Elterngeldreform im zweiten Lebensjahr weniger Zeit mit ihren Kindern verbringen können. Hinzu kommt, dass insbesondere Paare, bei denen beide ein höheres Bildungsniveau haben, die Partnermonate nutzen und damit beide Elternteile mehr Zeit mit ihren Kindern verbringen können. Auch dies würde Bildungsungleichheiten zwischen Kindern potentiell erhöhen. So könnten Entwicklungsungleichheiten, die beispielsweise auf den Bildungshintergrund der Eltern zurückgehen, durch das Elterngeld verstärkt werden.

Allerdings sind entsprechende Effekte von einer Vielzahl weiterer Faktoren abhängig. So gibt es beispielsweise empirische Hinweise darauf, dass der Zusammenhang zwischen der mütterlichen Erwerbstätigkeit und der kindlichen Entwicklung vom Erwerbsumfang abhängig ist.¹² Insbesondere wenn die Mutter relativ bald nach der Geburt in Vollzeit erwerbstätig ist, kann dies negative Effekte auf die kindliche Entwicklung haben. Geht die Mutter hingegen einer Teilzeitarbeit nach, sind die Effekte geringer – sofern diese überhaupt nachweisbar sind. Außerdem zeigen mehrere Studien, dass der Zusammenhang zwischen einem Mehr an elterlicher

Marcus Tamm (2013): Parental leave regulations, mothers' labor force attachment and fathers' childcare involvement: evidence from a natural experiment. *Journal of Population Economics* 26(3), 983–1005; Jochen Kluge und Sebastian Schmitz (2014): Mittelfristige Effekte der Elterngeldreform in Ost- und Westdeutschland. *Vierteljahreshefte zur Wirtschaftsforschung* 83(1), 163–181 (online verfügbar); Annette Bergemann und Regina T. Riphahn (2015): Maternal employment effects of paid parental leave. IZA Discussion Paper 9073; sowie Johannes Geyer et al. (2015): The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. *Labour Economics* 36, 84–98.

¹² Vgl. zum Beispiel Jeanne Brooks-Gunn, Wen-Jui Han und Jane Waldfogel (2010): First-year maternal employment and child development in the first seven years. *Monographs of the Society for Research in Child Development* 75, 1–147.

Kasten 1

Datengrundlage

Die Analysen zum Einfluss der Elterngeldreform auf die Entwicklung von Kindern basieren auf Daten von Schuleingangsuntersuchungen des Landes Schleswig-Holstein.¹ Schuleingangsuntersuchung dienen dazu, Entwicklungsverzögerungen festzustellen und die Schulfähigkeit zu beurteilen. Sie werden in der Regel kurz vor der geplanten Einschulung von erfahrenen Expertinnen und Experten des *Kinder- und Jugendärztlichen Dienstes der Gesundheitsämter* durchgeführt. Die einzelnen Untersuchungsbereiche sind standardisiert und basieren auf wissenschaftlich fundierten Verfahren.²

Auf Basis dieser Daten ist es möglich, die kindliche Entwicklung über unterschiedliche Indikatoren abzubilden. Die vorliegenden Analysen konzentrieren sich auf vier Indikatoren (Tabelle 1):

- *Sprachliche Fähigkeiten:* Diese werden unter anderem auf Basis des kindlichen Wortschatzes, der Satzlänge, Grammatik und Satzstellung sowie auf Basis der Fähigkeit, vorgegebene Sätze und Pseudowörter nachzusprechen, beurteilt.
- *Motorische Fähigkeiten:* Sie werden unter anderem darüber eingeschätzt, wie häufig ein Kind innerhalb einer vorgegebenen Zeit seitlich mit beiden Beinen hin- und herspringen kann, wie lange es auf einem Bein stehen kann und wie häufig das Kind auf einem Bein hüpfen kann.
- *Verhaltensauffälligkeiten:* Diese Beurteilung erfolgt auf Basis der aktuellen Symptomatik, von Vorbefunden, Beobachtungen der Schulärztinnen und -ärzte während der Untersuchung und teilweise unter Einbezug eines standardisierten Elternfragebogens zur Erfassung der Verhaltensstärken und -auffälligkeiten des Kindes.³
- *Schulärztliche Empfehlung zur Einschulung ohne Förderbedarf:* Dieses Maß basiert auf den vorangegangenen Untersuchungen und fasst den Gesamteindruck des Ex-

perten oder der Expertin des *Kinder- und Jugendärztlichen Dienstes der Gesundheitsämter* zum Entwicklungsstand des Kindes zusammen.

Insgesamt führen die Einschätzungen in Schleswig-Holstein nicht dazu, dass Einschulungen aufgeschoben werden. Vielmehr helfen die Untersuchungen, den spezifischen Förderbedarf zu identifizieren. Aufgeschobene Einschulungen sind nur in besonderen Ausnahmefällen möglich, etwa wenn gesundheitliche Gründe gegen eine Einschulung sprechen.

Neben der kindlichen Entwicklung werden während der Schuleingangsuntersuchung auch Informationen zum familiären Hintergrund des Kindes erfragt, die Eltern freiwillig beantworten können. Dazu zählen beispielsweise Angaben zur Familiensituation, dem Bildungsstand der Eltern oder dem Migrationshintergrund.

Der Datensatz, auf dem die Analysen beruhen, umfasst Kinder, die zwischen Juli 2005 und Juni 2008 geboren wurden. In die Analysen der durchschnittlichen Reformeffekte gehen 44 997 Kinder ein, für die Informationen zu allen vier betrachteten Entwicklungsmaßen vorliegen. Die Analysen, die die Reformeffekte separat nach dem Bildungshintergrund der Mutter untersuchen, basieren auf 36 748 Beobachtungen, da die freiwilligen Angaben zum elterlichen Bildungshintergrund teilweise fehlen.⁴

In Schleswig-Holstein leben etwa 3,6 Prozent der gesamten Bevölkerung in Deutschland. Da der Datensatz die Grundgesamtheit der Kinder Schleswig-Holsteins abdeckt, sind die Ergebnisse für dieses Bundesland und die entsprechenden Schülerkohorten repräsentativ. Um zu bestimmen, inwiefern die Analysen repräsentativ für Deutschland sind, werden sozioökonomische und soziodemografische Merkmale der Bevölkerung Schleswig-Holsteins mit den übrigen (west-)deutschen Bundesländern verglichen (Tabelle 2). Mit wenigen Ausnahmen unterscheidet sich die Population Schleswig-Holsteins im Mittel kaum von den anderen (west-)deutschen Bundesländern. Unterschiede gibt es lediglich dahingehend, dass in Schleswig-Holstein im Durchschnitt ein geringerer Anteil der Einwohnerinnen und Einwohner einen Migrationshintergrund hat, ein geringerer Anteil in städtischen Regionen lebt und dass die Verfügbarkeit von Plätzen in Kindertageseinrichtungen für Kinder über drei Jahren geringer ist, allerdings nicht die Plätze für Kinder unter drei Jahren.

1 Die Autorin und die Autoren dieses Berichts danken dem Ministerium für Soziales, Gesundheit, Wissenschaft und Gleichstellung des Landes Schleswig-Holstein für die Bereitstellung der Daten. Für die Analysen und die Interpretation der Ergebnisse tragen die Autorin und die Autoren des vorliegenden Berichts die alleinige Verantwortung. Schuleingangsuntersuchungen sind für Kinder in fast allen Bundesländern gesetzlich vorgeschrieben. Der Zugang zu diesen Daten zu Forschungszwecken ist jedoch nur eingeschränkt möglich.

2 Für eine detaillierte Darstellung der verwendeten Daten und Untersuchungsgrößen vgl. Huebener, Kuehnle und Spiess (2017), a. a. O.

3 Der Elternfragebogen basiert auf dem wissenschaftlich validierten „Strength and Difficulties Questionnaire“, vgl. SDQ (2017): Information for researchers and professionals about the Strengths & Difficulties Questionnaire (online verfügbar).

4 Ob Eltern in der Schuleingangsuntersuchung Angaben zu ihrem Bildungshintergrund machen, ist statistisch unabhängig von der Einführung der Elterngeldreform. Daher ist davon auszugehen, dass fehlende Angaben die geschätzten Effekte nicht verzerren.

Tabelle 1

Entwicklungsmaße von Schulanfängerinnen und Schulanfängern

Anteile in Prozent

	Alle Kinder	Nach Schulabschluss der Mutter	
		Haupt-/Realschule	(Fach-)Abitur
Ausreichende Sprachfähigkeiten	71,5	67,1	77,7
Ausreichende sozial-emotionale Stabilität	81,0	77,6	85,1
Ausreichende motorische Fähigkeiten	82,5	80,3	86,0
Schulfähig ohne Förderbedarf	84,0	80,9	91,1
Anzahl der Beobachtungen	44 997	22 492	14 256

Quelle: Eigene Berechnungen basierend auf Daten der Schuleingangsuntersuchungen des Landes Schleswig-Holstein für Kinder, die zwischen Juli 2005 und Juni 2008 geboren wurden.

© DIW Berlin 2017

Tabelle 2

Vergleich sozio-ökonomischer und demografischer Merkmale der Bevölkerung Schleswig-Holsteins mit der Bevölkerung in anderen Bundesländern

Anteile in Prozent, sofern nicht anders vermerkt

	Schleswig-Holstein	Westdeutschland	Deutschland
Durchschnittsalter (in Jahren)	44,6	43,9	44,2
Frauen	52	52	52
Unverheiratet	38	39	39
Verheiratete Personen	47	47	47
Geschiedene Personen	7	6	6
Haushaltsgröße (Personenzahl)	2,67	2,74	2,68
Kinder im Haushalt (Anzahl)	0,94	0,98	0,93
In Deutschland geboren	89	85	86
Erwerbstätig	45	47	46
Arbeitsuchend	3	3	4
Nicht erwerbstätig	51	50	50
Frauenenerwerbsquote	80	80	81
Nutzungsquote Kindertageseinrichtung, Kinder unter drei Jahren ¹	12	12	15
Nutzungsquote Kindertageseinrichtung, Kinder zwischen drei bis unter sechs Jahren ¹	84	91	91
Väter, die Elterngeld bezogen haben ²	18	20	21
Höchster Bildungsstand (ISCED-97)³			
ISCED 1,2,3c	25	27	25
ISCED 3a	4	5	5
ISCED 3b	47	42	43
ISCED 4	6	6	6
ISCED 5 oder höher	18	20	20
Individuelles Einkommen (in Euro, netto pro Monat)			
weniger als 1 100	49	50	50
1 100 bis 2 300	40	39	40
2 300 bis 3 600	7	7	6
3 600 bis 5 000	3	3	3
5 000 bis 18 000	1	1	1
Gemeindegröße (Einwohnerinnen und Einwohner)			
weniger als 2 000	19	5	6
2 000 bis 5 000	11	9	9
5 000 bis 10 000	12	12	11
10 000 bis 50 000	33	35	33
50 000 bis 100 000	8	10	9
mehr als 100 000	16	31	32
Anzahl der Beobachtungen	17 718	368 654	452 732

Anmerkung: Die Angaben für Westdeutschland und Gesamtdeutschland schließen Schleswig-Holstein aus.

1 Angaben für 2008.

2 Basierend auf Statistisches Bundesamt (2010): Elterngeld für Geburten 2008 nach Kreisen, Wiesbaden. Es handelt sich um Väter, deren Kinder 2008 geboren wurden.

3 ISCED steht für die International Standard Classification of Education.

Quelle: Eigene Berechnungen basierend auf dem Deutschen Mikrozensus 2009.

© DIW Berlin 2017

Zeit und der kindlichen Entwicklung nicht eindeutig ist. Unter anderem mit zunehmendem Alter des Kindes ist die gemeinsame Zeit anders zu bewerten.¹³ Ferner hängen Effekte auf die kindliche Entwicklung auch sehr stark davon ab, welche Betreuungsalternativen genutzt werden und wie deren Qualität ist.¹⁴ Insgesamt ist es letztlich also eine empirische Frage, wie das Elterngeld auf die kindliche Entwicklung wirkt und inwiefern dadurch Entwicklungsunterschiede zwischen Kindern mit unterschiedlichem Bildungshintergrund der Eltern verändert werden.

Internationale Studien weisen auf geringe Auswirkungen hin

Die Frage, inwiefern sich die Einführung des Elterngeldes auf die kindliche Entwicklung auswirkt, wurde bislang kaum untersucht. Allerdings liegen einige Studien für andere Länder oder auch frühere Reformen in Deutschland vor, die ähnliche Veränderungen im Hinblick auf unterschiedliche kindliche Entwicklungsmaße und die Gesundheit von Kindern analysieren.¹⁵ Dabei sind Studien, die die *Einführung* entsprechender Leistungen untersuchen, von solchen zu unterscheiden, die *Ausweitungen* bereits eingeführter Elterngeld- und Elternzeitregelungen betrachten. Letztgenannte kommen mehrheitlich zu dem Ergebnis, dass es kaum Effekte auf die kindliche Entwicklung gibt (gemessen an schulischen Leistungen, der kognitiven Entwicklung von Kindern und dem späteren Arbeitsmarkterfolg). Dies gilt insbesondere für Studien, die die Verlängerung von Elternzeitregelungen im ersten Lebensjahr des Kindes betrachten.¹⁶ Keine dieser bisher vorliegenden, in Fachzeitschriften veröffentlichten, Studien analysiert allerdings eine Reform, die – wie das Elterngeld in Deutschland – die Bezugsdauer verkürzt, dafür aber die Geldleistung im ersten Lebensjahr erhöht. Einzig eine auf SOEP-Daten basierende Studie befasst sich mit den kurzfristigen Wirkungen der deutschen Elterngeldreform auf die kindliche Entwicklung in den ersten drei Lebensjahren.¹⁷

13 Vgl. zum Beispiel Ariel Kalil, Rebecca Ryan und Michael Corey (2012): Diverging destinies: Maternal education and the developmental gradient in time with children. *Demography* 49 (4), 1361–1383.

14 Vgl. zum Beispiel Brooks-Gunn, Han und Waldfogel. (2010), a. a. O.

15 Für eine Zusammenfassung der bisherigen Studien vgl. Mathias Huebener (2016): Parental leave policies and child development: A review of empirical findings. DIW Roundup Nr. 102 (online verfügbar).

16 Die wenigen Studien, die die Einführung entsprechender Regelungen betrachten, können positive Effekte auf die kindliche Gesundheit, den späteren schulischen Erfolg der Kinder und deren Löhne im Erwachsenenalter messen. Sofern Veränderungen untersucht werden, die das zweite Lebensjahr betreffen, zeigen sich heterogene Effekte: Teilweise sind für bestimmte Gruppen positive, teilweise aber auch negative Effekte, auf die Entwicklung von Kindern beziehungsweise deren späteren Arbeitsmarkterfolg festzustellen.

17 Vgl. Katrin Huber (2015): Moving to an earnings-related parental leave system – do heterogeneous effects on parents make some children worse off? SOEPpapers on Multidisciplinary Panel Data Research Nr. 791 (online verfügbar). Für eine sozialwissenschaftliche Studie außerhalb der Ökonomie, die sich mit der

Vermutlich auch aufgrund der geringen Stichprobengröße kommt die Analyse jedoch nicht zu eindeutigen Befunden.¹⁸

Auswirkungen auf kindliche Entwicklung werden auf Basis von Schuleingangsdaten untersucht

Die vorliegende Analyse betrachtet die Auswirkungen des Elterngeldes auf kindliche Entwicklungsmaße, die kurz vor der Einschulung eines Kindes im Rahmen von Schuleingangsuntersuchungen erfasst wurden. Die Kinder, die für die Studie berücksichtigt wurden, sind daher im Durchschnitt sechs Jahre alt. Dieser Zeitpunkt ist besonders bedeutend, da der Entwicklungsstand des Kindes vor Schulbeginn maßgeblich mit dem schulischen Erfolg und damit auch der späteren Integration auf dem Arbeitsmarkt sowie vielen weiteren Faktoren wie der Gesundheit zusammenhängt. Die in die Analyse einbezogenen Kinder wurden zwischen Juli 2005 und Juni 2008 geboren.

Die Untersuchung basiert auf Schuleingangsdaten des Landes Schleswig-Holstein.¹⁹ Die Ergebnisse sind für dieses Bundesland und die entsprechenden Schülerkohorten repräsentativ. Da sich die Bevölkerung Schleswig-Holsteins mit Blick auf Merkmale wie das Bildungsniveau oder das Haushaltseinkommen im Mittel nicht wesentlich von der Bevölkerung anderer Bundesländer unterscheidet, ist davon auszugehen, dass die Ergebnisse Bedeutung über Schleswig-Holstein hinaus haben (Tabelle 2 in Kasten 1).

Die Informationen der Schuleingangsuntersuchung zum Entwicklungsstand der Kinder beruhen auf Einschätzungen von Expertinnen und Experten des *Kinder- und Jugendärztlichen Dienstes der Gesundheitsämter*. Bei diesen Untersuchungen wird der Entwicklungsstand des Kindes erfasst und festgestellt, ob ein schulischer Förderbedarf vorliegt. Auf Basis der Daten dieser Untersuchungen ist es möglich, die kindliche Entwicklung über unterschiedliche Indikatoren abzubilden. In diesem Bericht konzentriert sich die Analyse auf vier Indikatoren, die in der familien- und bildungsökonomischen

Wirkung des Elterngeldes auf das kindliche Wohlergehen befasst, vgl. Axel Schölmerich et al. (2013): Endbericht des Moduls Wohlergehen von Kindern, erstellt im Auftrag der Geschäftsstelle Gesamtevaluation der ehe- und familienbezogenen Leistungen in Deutschland, Kapitel 6.6. Diese Studie stellt keine kurzfristigen Effekte auf Kinder fest.

18 Nur sehr wenige Studien untersuchen explizit, wie sich ein Anstieg in der väterlichen Beteiligung an der Elternzeit auf die kindliche Entwicklung auswirkt. Für eine der wenigen Ausnahmen vgl. Cools, S., Fiva, J. H., & Kirkebøen, L. J. (2015). Causal effects of paternity leave on children and parents. *Scandinavian Journal of Economics*, 117(3), 801–828.

19 Obwohl Schuleingangsuntersuchungen in fast allen Bundesländern durchgeführt werden, ist Schleswig-Holstein eines der wenigen Bundesländer, das den Zugang zu diesen Daten zu Forschungszwecken ermöglicht hat.

Tabelle 2

Effekte der Elterngeldeinführung auf Entwicklungsmaße von Schulanfängerinnen und Schulanfängern

Entwicklungsmaß des Kindes	Mittelwert (in Prozent)	Effekte auf Entwicklungsmaße (in Prozentpunkten)			
		Effekt auf alle Kinder	differenziert nach Schulabschluss der Mutter		Effekt auf Entwicklungsungleichheiten (Differenz aus Spalten 4 und 5)
			Haupt-/Realschule	(Fach-)Abitur	
	(1)	(2)	(3)	(4)	(5)
Ausreichende Sprachfähigkeiten	71,5	-0,74 (0,91) [0,1327]	-1,81 (1,34) [0,1286]	-0,92 (1,57) [0,1007]	0,89 (2,06)
Ausreichende sozial-emotionale Stabilität	81,0	-0,35 (0,81) [0,0951]	-1,27 (1,20) [0,1042]	1,12 (1,35) [0,0777]	2,38 (1,81)
Ausreichende motorische Fähigkeiten	82,5	-0,67 (0,77) [0,0802]	-0,51 (1,15) [0,0842]	-1,27 (1,27) [0,0765]	-0,76 (1,72)
Schulfähig ohne Förderbedarf	84,0	0,48 (0,73) [0,1110]	0,32 (1,10) [0,1139]	0,79 (1,05) [0,0660]	0,47 (1,52)
Anzahl der Beobachtungen	44 997	44 997	22 492	14 256	36 748

Anmerkung: Die Schätzmethode wird in Kasten 2 beschrieben. Jeder Koeffizient der Spalten 2 bis 4 beruht auf einer eigenen Schätzung. Der Standardfehler ist in runden Klammern angegeben, das Bestimmtheitsmaß (R²) in eckigen Klammern. Spalte 1 gibt den Stichprobenmittelwert aus Tabelle 2 in Kasten 1 wieder und bezieht sich auf den Anteil an Kindern, bei denen ein ausreichender Entwicklungsstand diagnostiziert wurde. Die Spalten 2 bis 4 geben die Veränderung in Prozentpunkten an, die die Elterngeldreform auf die Diagnose eines ausreichenden Entwicklungsstandes hatte.

Quelle: Eigene Berechnungen basierend auf Daten der Schuleingangsuntersuchungen des Landes Schleswig-Holstein für Kinder, die zwischen Juli 2005 und Juni 2008 geboren wurden.

Forschung besonders relevant sind und die grundsätzlich von elterlichen Aktivitäten beeinflusst werden können: erstens sprachliche Fähigkeiten, zweitens motorische Fähigkeiten, drittens Verhaltensauffälligkeiten und viertens eine zusammenfassende schulärztliche Empfehlung zur Einschulung ohne Förderbedarf.

Im Durchschnitt der analysierten Daten weisen etwa 72 Prozent der Kinder keine Sprachauffälligkeiten auf. Bei etwa 83 Prozent der Kinder sind keine motorischen Auffälligkeiten zu beobachten und bei 81 Prozent keine Verhaltensauffälligkeiten. Insgesamt wird 84 Prozent der Kinder eine Empfehlung zur Einschulung ohne Förderbedarf ausgesprochen (Tabelle 2, Spalte 1).

Auf Basis dieser Daten wird die Wirkung des Elterngeldes auf die vier genannten Entwicklungsmaße untersucht. Anhand einer quasi-experimentellen Methode wird der kausale Effekt des Elterngeldes auf die kindliche Entwicklung von Schulanfängerinnen und Schulanfängern geschätzt (Kasten 2). Insgesamt basieren die multivariaten Regressionsschätzungen auf einem

Datensatz, der Informationen zu fast 45 000 Kindern umfasst.²⁰

Elterngeld beeinflusst die vier untersuchten Entwicklungsmaße nicht

Das Elterngeld hat im Durchschnitt – zunächst unabhängig vom Schulabschluss der Mutter – auf keines der vier betrachteten kindlichen Entwicklungsmaße einen statistisch signifikanten Einfluss (Tabelle 2, Spalte 2). Hinzu kommt, dass die geschätzten Effekte nahe Null liegen. Dies könnte daraus resultieren, dass manche Kinder von der Einführung des Elterngeldes profitieren, andere hingegen benachteiligt werden und sich insgesamt beide Effekte gegenseitig aufheben. Differenziert man die Analyse daher nach dem Bildungshintergrund der Mutter, bleibt der Befund jedoch bestehen: keine sta-

²⁰ Zur Verwendung dieses Datensatzes siehe auch eine Studie von Christina Felfe und Rafael Lalive (2014): Does Early Child Care Help or Hurt Children's Development? IZA Discussion Paper 8484 (online verfügbar). Diese Studie untersucht unter Verwendung des gleichen Datensatzes die Effekte des Ausbaus von Kindertageseinrichtungen auf die kindliche Entwicklung.

Kasten 2

Methode

Um die ursächlichen Effekte der Einführung des Elterngeldes auf die kindliche Entwicklung zu untersuchen, werden Ansätze des Regressions-Diskontinuitäten-Ansatzes mit denen des Differenz-von-Differenzen-Ansatzes kombiniert. Die verwendete Methode schätzt den Gesamteffekt der Einführung des Elterngeldes im Vergleich zur Situation, in der nur anspruchsberechtigte Eltern Erziehungsgeld beziehen konnten.¹

Die Grundidee dieses Ansatzes ist, die Entwicklungsindikatoren von Kindern zu vergleichen, von denen einige zufällig das Elterngeld bekommen haben und andere per Zufall das Erziehungsgeld. Durch die institutionelle Reform zum 1. Januar 2007 wurde diese Idee als „natürliches Experiment“ umgesetzt, da der Geburtstermin des Kindes entscheidet, welcher Gruppe die Eltern angehören: Da die Einführung des Elterngeldes erst am 29.09.2006 vom Deutschen Bundestag beschlossen wurde, waren von der Reform betroffene Mütter zum Zeitpunkt der Bekanntgabe bereits schwanger. Damit ergibt sich die Möglichkeit, dass Eltern „quasi“ zufällig unter die Elterngeldregelung statt die Erziehungsgeldregelung fallen. Erwartungsgemäß gibt es keine statistischen Unterschiede in Bezug auf die Charakteristika der Kinder und des familiären Hintergrundes, die mit der Einführung des Elterngeldes zusammenhängen.²

¹ Für Details zur Methodik sowie zu den umfangreichen Robustheitsprüfungen siehe Huebener, Kuehne und Spiess (2017), a.a.O.

Für die Schätzung des kausalen Reformeffekts können daher Entwicklungsmaße von Kindern verglichen werden, die kurz vor und kurz nach der Einführung der Elterngeldreform geboren wurden. Um die Präzision der statistischen Schätzung zu erhöhen, werden Kinder miteinander verglichen, die innerhalb von sechs Monaten vor beziehungsweise nach der Elterngeldreform geboren wurden.

Aufgrund dieses erweiterten Geburtszeitfensters, das für den Mittelwertvergleich herangezogen wird, könnte es zu Mittelwertverzerrungen kommen. Diese könnten beispielsweise dadurch entstehen, dass Kinder, die zwischen Juli und Dezember geboren werden, zum Zeitpunkt der Schuleingangsuntersuchung schon älter sind. Um mögliche Mittelwertunterschiede durch solche verzerrenden Einflüsse, die unabhängig von der Elterngeldreform

² Empirische Evidenz deutet darauf hin, dass es zu leichten Geburtsverschiebungen in unmittelbarer Nähe des Reformdatums (1. Januar 2007) kam. Vgl. aktuell Hendrik Jürges (2017): Financial incentives, timing of births, and infant health: a closer look into the delivery room. The European Journal of Health Economics, 18(2), 195–208. In den Hauptanalysen des vorliegenden Berichts werden deshalb Kinder, die in unmittelbarer Nähe des Reformdatums geboren wurden, ausgeschlossen.

tistisch signifikanten Effekte des Elterngeldes mit Blick auf sprachliche und motorische Fähigkeiten, Verhalten und Förderbedarf der Kinder (Tabelle 2, Spalten 3 und 4). Weder die Entwicklungsmaße der Kinder von Müttern mit höherer Bildung, noch der Kinder von Müttern mit einem niedrigeren Bildungsniveau werden durch das Elterngeld beeinflusst. Die Unterschiede zwischen den Kindern unterschiedlicher Bildungsgruppen verändern sich durch das Elterngeld also nicht (Tabelle 2, Spalte 5).

Um sicherzustellen, dass diese geringen Effekte nicht das Resultat der spezifischen Messung sind, wurden weitere Analysen durchgeführt. Zunächst wurde betrachtet, inwieweit die verwendeten Entwicklungsmaße der Kinder mit anderen Faktoren wie dem Alter zum Untersuchungszeitpunkt, dem Geschlecht des Kindes und dem Bildungsstand der Mutter korrelieren.²¹ Beispielsweise verbessert sich mit dem Alter des Kindes die Wahrchein-

²¹ Diese Analysen beruhen auf multivariaten Regressionen, die zahlreiche Faktoren gleichzeitig berücksichtigen, wodurch die ermittelten Unterschiede

lichkeit, keine Sprachauffälligkeiten aufzuweisen, und zwar um 6,2 Prozentpunkte pro Lebensjahr. Vergleicht man diesbezügliche Entwicklungsunterschiede zwischen Geschlechtern, haben Mädchen eine um etwa 7,2 Prozentpunkte höhere Wahrscheinlichkeit, mit Blick auf sprachliche Fähigkeiten nicht auffällig zu sein. Bemerkenswert ist auch der Unterschied nach dem Bildungshintergrund der Mutter: Kinder, deren Mutter kein Abitur hat, werden mit einer um 9,5 Prozentpunkte höheren Wahrscheinlichkeit als sprachauffällig eingeschätzt als Kinder von Müttern mit Abitur. Im Hinblick auf die anderen Entwicklungsmaße lassen sich ähnliche Unterschiede festmachen (Abbildung).

Diesen Entwicklungsunterschieden in Bezug auf das Alter der Kinder, ihrem Geschlecht und dem mütterlichen Bildungshintergrund werden die geschätzten Effekte der Elterngeldreform gegenübergestellt (Abbildung). Dabei

eine *ceteris paribus*-Interpretation erhalten. Für Details dieser Analysen siehe Huebener, Kuehne und Spiess (2017), a.a.O.

auftreten, zu verhindern, werden Unterschiede, die auch für Kinder anderer Geburtsjahrgänge bestehen, herausgerechnet.

Mithilfe multivariater Regressionen können außerdem andere Faktoren in die Analyse eingeschlossen werden, die Einfluss auf kindliche Entwicklungsmaße haben, aber von der Elterngeldreform unabhängig sind. Damit kann die Präzision der statistischen Schätzung erhöht werden. Solche Faktoren sind zum Beispiel generelle Unterschiede zwischen Kindern bestimmter Geburtsmonate, Geburtskohorten und Landkreise beziehungsweise kreisfreier Städte. Weiterhin werden der Migrationshintergrund und das Alter zum Zeitpunkt der Schuleingangsuntersuchung berücksichtigt. Die Methode rechnet auch Einflüsse heraus, die im Zeitverlauf einen zunehmenden Einfluss auf die kindliche Entwicklung haben können, aber unabhängig von der Einführung des Elterngeldes sind, wie generelle Veränderungen in der gesamtwirtschaftlichen Situation oder der Ausbau des Angebots an Plätzen in der Kindertagesbetreuung. Auch hierdurch erhöht sich die Genauigkeit der Schätzung.

Die Hauptannahme dieser Methode basiert darauf, dass Entwicklungsunterschiede zwischen Kindern, die sechs Monate vor und nach der Elterngeldreform geboren wurden, die gleichen gewesen wären wie die von Kindern, die im gleichen Monat ein Jahr zuvor und danach auf die Welt kamen, wenn die Elterngeldreform nicht eingeführt worden wäre. Umfangreiche Plausi-

bilitätschecks geben keinen Hinweis auf eine Verletzung dieser Annahme.

Während das methodische Vorgehen die Verzerrung der Effekte durch zahlreiche andere Faktoren eines einfachen Vorher-Nachher-Vergleichs berücksichtigen kann, haben die Ergebnisse eine besonders hohe (interne) Validität für Kinder, die kurz nach der Einführung des Elterngeldes geboren wurden. Reaktionsanpassungen von Eltern später geborener Kinder über die Zeit (beispielsweise verursacht durch Veränderungen sozialer Normen in der Inanspruchnahme des Elterngeldes durch Väter) oder Effekte aufgrund von Veränderungen im Fertilitätsverhalten³ auf die Kinderentwicklung werden durch diese Methode nicht erfasst.

Weiterhin ist die externe Validität der Ergebnisse von der Repräsentativität des institutionellen, wirtschaftlichen und gesellschaftlichen Kontexts geprägt, in dem die Analyse durchgeführt wurde.⁴

3 So beeinflusst die Zahl der Kinder beziehungsweise der Zeitpunkt der Geburten das Alter der Eltern und in der Folge das Haushaltseinkommen bei Geburt der Kinder.

4 Zur Vergleichbarkeit in Bezug auf sozioökonomische und soziodemografische Merkmale Schleswig-Holsteins mit (West-)Deutschland insgesamt siehe Kasten 1 des vorliegenden Berichts.

wird besonders deutlich, wie klein die Effekte der Elterngeldreform auf die kindliche Entwicklung ausfallen und dass sie kaum von Null verschieden sind.

In anderen, in diesem Bericht nicht weiter dargestellten, Berechnungen wurde außerdem geprüft, ob sich Entwicklungsunterschiede bezüglich des Bildungshintergrundes der Väter zeigen.²² Auch dabei ergaben sich keine statistisch und ökonomisch signifikanten Effekte auf die kindliche Entwicklung oder die zwischen den beiden Bildungsgruppen bestehenden Entwicklungsunterschiede.

Bis zu dieser Stelle wurde jeweils die Gesamtschätzung zu sprachlichen und motorischen Entwicklungsauffälligkeiten sowie der sozio-emotionalen Stabilität

aus den Schuleingangsuntersuchungen herangezogen. Die Ergebnisse sagen also aus, dass die Elterngeldreform mit Blick auf Gesamtschätzungen wie „sprachlich auffällig“ oder „sprachlich unauffällig“ keine Effekte hat. Dies schließt aber nicht aus, dass beispielsweise manche Kinder ohne Sprachauffälligkeiten durch die Einführung des Elterngeldes zum Zeitpunkt der Einschulung dennoch weiter entwickelt sind, als sie es ohne die Elterngeldreform wären. Eine zusätzliche Schätzung mit den differenzierten Testergebnissen, auf denen die Gesamtschätzung beruht²³, zeigt jedoch, dass auch dabei keine statistisch signifikanten Effekte des Elterngeldes feststellbar sind – vielmehr weisen die Ergebnisse wiederum auf einen sehr präzise geschätzten Nulleffekt der Reform hin. Eine umfangreiche Anzahl weiterer Tests mit Blick auf die Methode und die berücksich-

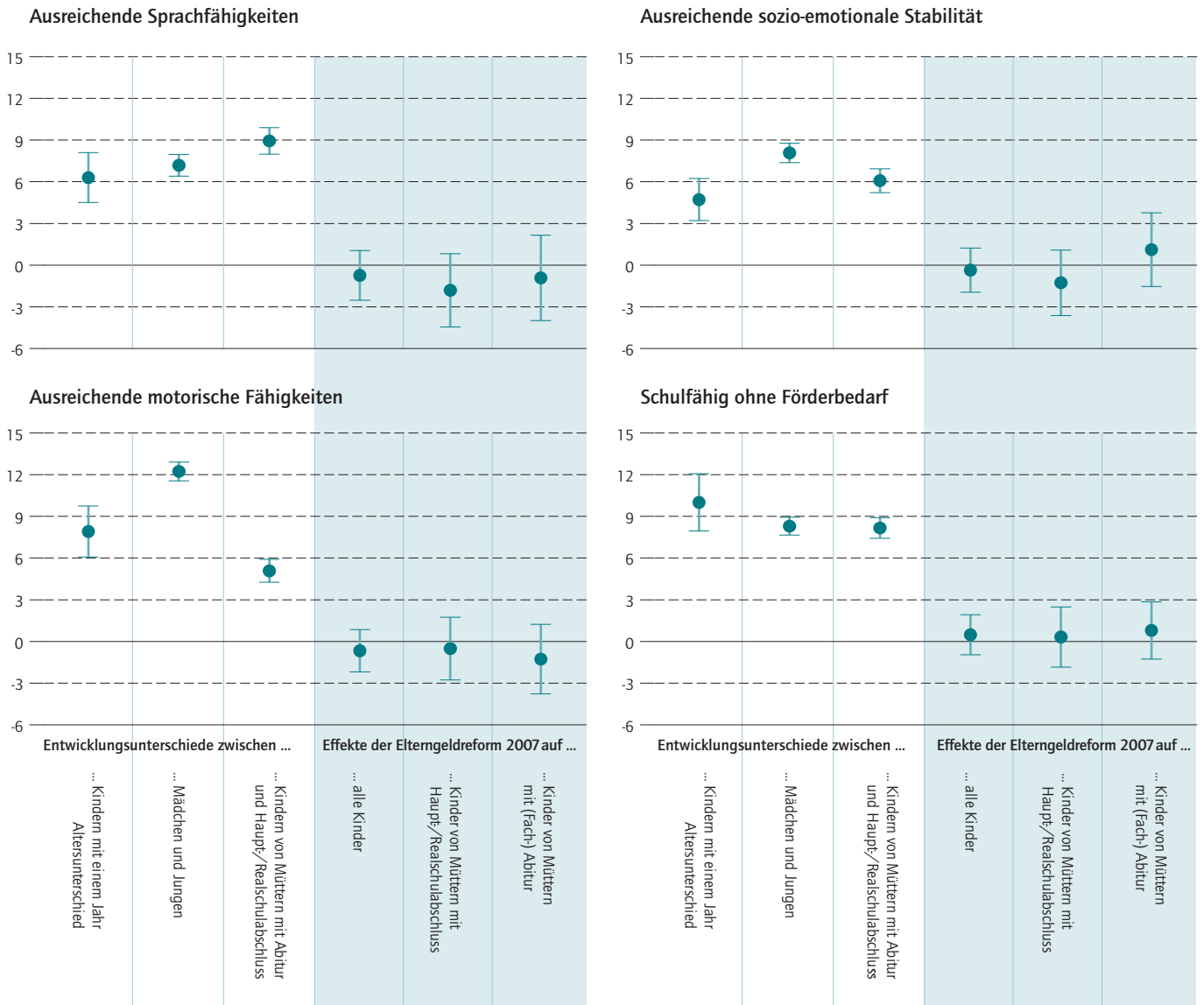
22 Allerdings ist anzumerken, dass sich die Analysen auf den Zeitraum der Elterngeldeinführung beziehen und in diesem Zeitraum der Anteil von Vätern, die Elterngeld bezogen haben, noch nicht so hoch war wie heute. Vgl. dazu zum Beispiel Huebener et al. (2016), a. a. O.

23 Diese Analysen beruhen auf einer kleineren Zahl von Beobachtungen, da die Informationen nicht für alle Landkreise und kreisfreien Städte vorliegen.

Abbildung

Entwicklungsunterschiede von Kindern nach unterschiedlichen Merkmalen im Vergleich zu den Effekten des Elterngeldes

In Prozentpunkten



Anmerkung: Die Abbildung stellt Entwicklungsunterschiede in kindlichen Entwicklungsmaßen in Bezug auf das Alter der Kinder, ihrem Geschlecht und dem mütterlichen Bildungshintergrund dar. Diesen werden zum Vergleich die geschätzten Effekte der Elterngeldreform gegenübergestellt. Die vertikalen, graugrünen Bänder stellen das Konfidenzintervall (95 Prozent) der statistischen Schätzergebnisse dar.

Quelle: Eigene Berechnungen basierend auf Daten der Schuleingangsuntersuchungen des Landes Schleswig-Holstein für Kinder, die zwischen Juli 2005 und Juni 2008 geboren wurden.

© DIW Berlin 2017

Die Effekte der Elterngeldreform sind minimal im Vergleich zu Entwicklungsunterschieden nach anderen Merkmalen.

tigten Daten bestätigen die Robustheit der ausgewiesenen Nulleffekte.²⁴

24 Neben der Unterscheidung der Effekte nach mütterlichem und väterlichem Bildungshintergrund wurden die Analysen auch separat nach dem höchsten Haushaltsbildungsabschluss durchgeführt. Außerdem wurde nach Unterschieden in den Effekte nach Geschlecht und in Interaktion mit den elterlichen Bildungsabschlüssen geschaut, da Jungen im Vergleich zu Mädchen typischerweise sensibler auf Veränderungen im Hinblick auf ihre Entwicklung reagieren.

Diese Ergebnisse sind vor dem Hintergrund anderer Studien zu ähnlichen Reformen in internationalen Kontexten weniger überraschend. So konnten vergleichbare

Auch hier konnten keine signifikanten robusten Effekte der Elterngeldreform festgestellt werden. Für Einzelheiten zu diesen Analysen siehe Huebener, Kuehne und Spiess (2017), a. a. O.

Studien, abgesehen von wenigen Ausnahmen, ebenfalls keine Effekte auf kindliche Entwicklungsmaße finden.

Fazit

Die Elterngeldreform des Jahres 2007 hatte weitreichende Effekte auf die wirtschaftliche Stabilität von Familien und das Erwerbsverhalten von Müttern und Vätern. Dennoch beeinflusst sie, wie die vorliegende Studie zeigt, nicht die untersuchten kindlichen Entwicklungsmaße. Trotz sehr präzise geschätzter Ergebnisse, die zahlreichen Robustheitstests standhalten, zeigen sich keine statistisch signifikanten Effekte auf die sprachlichen und motorischen Fähigkeiten von Schulanfängerinnen und Schulanfängern, deren sozio-emotionale Stabilität und den Förderbedarf – auch nicht mit Blick auf Entwicklungsunterschiede zwischen Kindern mit unterschiedlichem elterlichem Bildungshintergrund.

Dieser Befund kann vielfältige Ursachen haben. So waren Mütter mit einem höheren Bildungsabschluss, die infolge der Elterngeldreform tendenziell ihre Erwerbstätigkeit im ersten Lebensjahr häufiger und länger reduzieren als zuvor, vor der Elterngeldreform meist nicht vollzeit-, sondern teilzeiterwerbstätig. Dies hat auf die kindliche Entwicklung – wenn überhaupt – geringere Effekte als eine Vollzeiterwerbstätigkeit. Hinzu kommt, dass das Elterngeld vor allem mit Blick auf die zweite Hälfte des ersten Lebensjahres des Kindes dafür sorgt, dass besser gebildete Mütter in geringerem Maße erwerbstätig sind als zuvor. Mit Blick auf die ersten sechs Lebensmonate eines Kindes, die sehr entscheidend für die kindliche Entwicklung sind, gab es hingegen kaum Änderungen: Vor und nach der Elterngeldreform waren und sind die Mütter in diesem Zeitraum meist nicht erwerbstätig.

Mütter mit einem niedrigeren Bildungsabschluss gehen seit der Elterngeldreform tendenziell häufiger im zweiten Lebensjahr einer Erwerbstätigkeit nach als zuvor. Allerdings arbeiten die Mütter dann häufig in Teilzeit. Entscheidend ist zudem, wie und wo die Kinder betreut werden, wenn die Mutter erwerbstätig ist. Die für diese Studie verwendeten Daten geben dazu im Detail jedoch

keine Auskunft. Offensichtlich führt die alternative Betreuung – etwa in einer Kindertageseinrichtung oder bei Großeltern – aber nicht zu messbaren Veränderungen in den erfassten kindlichen Entwicklungsmaßen bei Schulanfängerinnen und Schulanfängern.

Dies ist insofern bemerkenswert, als andere Analysen auf Basis gleicher Daten zeigen können, dass beispielsweise der Ausbau der Kindertagesbetreuung einzelne Entwicklungsmaße mitunter signifikant verändert hat. Für die vorliegende Studie wurden also Entwicklungsmaße betrachtet, die durchaus durch familienpolitische Veränderungen beeinflussbar sind.²⁵

Bei der Bewertung der Ergebnisse darf man allerdings nicht außer Acht lassen, dass hier vier Entwicklungsmaße für Schulanfängerinnen und Schulanfänger betrachtet wurden. Streng genommen kann nicht ausgeschlossen werden, dass Effekte im Hinblick auf andere Entwicklungsmaße wie die mathematisch-naturwissenschaftlichen Fähigkeiten auftreten oder sich längerfristige Effekte einstellen. Die vorliegenden Ergebnisse sowie die anderer internationaler Studien lassen dies aber eher unwahrscheinlich erscheinen.

Zudem muss bedacht werden, dass in dieser Studie die Wirkungen des Elterngeldes in seinen Regelungen der Jahre 2007 und 2008 untersucht wurden. Zwischenzeitlich wurden Anpassungen vorgenommen, für die nicht mit Sicherheit ausgeschlossen werden kann, dass sie zu anderen Ergebnissen führen würden. So hat das *ElterngeldPlus* zum 1. Juli 2015 flexiblere Möglichkeiten geschaffen, während des Elterngeldbezugs in Teilzeit erwerbstätig zu sein, ohne sich dabei gegenüber anderen Paaren viel schlechter zu stellen. Allerdings erscheint es eher unwahrscheinlich, dass diese relativ geringen Veränderungen substantiell andere Ergebnisse hervorbringen würden als die hier untersuchte Elterngeldreform, die eine vergleichsweise große Veränderung darstellte.

²⁵ Vgl. Felfe und Lalive (2014), a. a. O.

Mathias Huebener ist Wissenschaftlicher Mitarbeiter der Abteilung Bildung und Familie am DIW Berlin | mhuebener@diw.de

Daniel Kühnle ist Wissenschaftlicher Mitarbeiter des Lehrstuhls für Statistik und empirische Wirtschaftsforschung an der Universität Erlangen-Nürnberg | daniel.kuehnle@wiso.uni-erlangen.de

C. Katharina Spieß ist Leiterin der Abteilung Bildung und Familie am DIW Berlin und Professorin für Bildungs- und Familienökonomie an der Freien Universität Berlin | kspiess@diw.de

JEL: J13, J18, J22, J24

Keywords: parental leave benefit, child development, skill formation, school readiness, motor skills, language skills, socio-emotional stability, socio-economic differences, inequality



Prof. Dr. C. Katharina Spieß, Leiterin der Abteilung Bildung und Familie am DIW Berlin

INTERVIEW MIT C. KATHARINA SPIESS

»Positive Effekte des Elterngeldes werden mit Blick auf kindliche Entwicklung nicht geschmälert«

1. Frau Spieß, die Auswirkungen des 2007 eingeführten Elterngeldes sind schon oft untersucht worden. Was ist neu an Ihrem Ansatz? Die Auswirkungen des Elterngeldes wurden vielfach untersucht im Hinblick auf die wirtschaftliche Stabilität von Familien sowie auf die Erwerbstätigkeit von Müttern und beispielsweise auch im Hinblick auf deren Zufriedenheit. Was wir in unserem neuen Forschungsprojekt untersucht haben, sind die Auswirkungen des Elterngeldes auf die kindliche Entwicklung, also die Auswirkungen auf die Kinder selbst.
2. Wie haben Sie die Effekte auf die kindliche Entwicklung analysiert? Wir haben die Auswirkungen des Elterngeldes auf die kindliche Entwicklung anhand von Entwicklungsmaßen von Schulanfängern untersucht. Wir hatten die gesamten Schuleingangsdaten mehrerer Geburtsjahrgänge in Schleswig-Holstein, an deren Beispiel wir die Untersuchung durchgeführt haben. Wir haben die Entwicklung von Kindern anhand von vier Entwicklungsmaßen abbilden können: Erstens die Sprachfähigkeiten von Kindern, zweitens ihre motorische Entwicklung, drittens ihre sozioemotionale Entwicklung und viertens, ob eine Schulfähigkeit ohne Förderbedarf vorliegt.
3. Zu welchem Ergebnis sind Sie gekommen? Wir konnten feststellen, dass die Elterngeldreform keine Effekte auf die Entwicklung der Kinder hatte. Wir waren selbst zunächst überrascht und haben deshalb untersucht, ob es Unterschiede zwischen unterschiedlichen Gruppen gibt. Dabei haben wir nach dem Bildungshintergrund der Eltern differenziert und uns angeschaut, ob sich die Entwicklungsunterschiede zwischen Kindern, deren Eltern eine höhere Bildung hatten, und Eltern mit geringerer Bildung unterscheiden. Auch da konnten wir aber feststellen, dass es keine Unterschiede gibt, dass also die Ungleichheiten in der kindlichen Entwicklung durch das Elterngeld nicht zugenommen haben.
4. Es wurde ja vielfach bemängelt, dass einkommensstärkere Haushalte stärker vom Elterngeld profitieren als einkommensschwächere Haushalte. Wie ist Ihr Ergebnis vor diesem Hintergrund zu erklären? Es könnte erwartet werden, dass einkommensstärkere Haushalte durch die

Anreize der Reform nun auch mehr Zeit mit den Kindern verbringen als einkommensschwächere Haushalte. Da das Einkommen stark mit der Bildung korreliert, hätten wir erwartet, dass die Ungleichheit zwischen Kindern höher und niedriger gebildeter Mütter tatsächlich zunimmt, was wir aber nicht bestätigen können. Wir haben unterschiedliche Erklärungsansätze. Wir wissen aus anderen Studien, dass einkommensstärkere Eltern auch vor der Einführung des Elterngeldes in der Regel einer Teilzeiterwerbstätigkeit nachgegangen sind und auch in den ersten sechs Lebensmonaten des Kindes, die für die Entwicklung sehr wichtig sind, teilweise auch Eltern ohne Elterngeld nicht erwerbstätig waren. Die Veränderungen in dieser sehr wichtigen Phase von Kindern waren demnach gar nicht so groß, wie man erwarten würde. Bei einkommensschwächeren Eltern ist es so, dass sie im zweiten Lebensjahr des Kindes ihre Erwerbstätigkeit erhöhen und damit grundsätzlich weniger Zeit mit dem Kind verbringen können. Wir vermuten aber, dass hier die grundsätzlich geringere elterliche Zeit keine negativen Auswirkungen auf die Kinder hat, weil diese vielleicht in einer guten Kindertageseinrichtung untergebracht sind beziehungsweise die Qualität der Zeit mit den Eltern gar nicht abgenommen hat.

5. Halten Sie die aktuelle Regelung des Elterngeldes für richtig, oder gibt es Anpassungsbedarf? Das Elterngeld, so wie es eingeführt worden ist, hat sehr viele positive Aspekte im Hinblick auf die wirtschaftliche Stabilität von Familien sowie im Hinblick auf eine höhere Erwerbstätigkeit von Müttern, gerade im unteren Einkommensbereich, und ist somit zum Beispiel sehr wichtig für die Bekämpfung von Altersarmut. Wir sehen, dass die kindliche Entwicklung nicht unter der aktuellen Regelung leidet beziehungsweise können wir keine negativen Effekte messen, sowohl im Hinblick auf die allgemeine Entwicklung als auch im Hinblick auf Ungleichheit. Von daher beurteilen wir das Elterngeld als eine erfolgreiche familienpolitische Maßnahme.

Das Gespräch führte Erich Wittenberg.



Das vollständige Interview zum Anhören finden Sie auf www.diw.de/interview

Schlechte Evaluierung rentiert sich kaum: Lehren aus dem Bereich der finanziellen Bildung

Von Tim Kaiser und Lukas Menkhoff

Die Verbesserung finanzieller Bildung ist inzwischen weltweit ein etabliertes Ziel der Wirtschaftspolitik, das über vielfältige Initiativen erreicht werden soll. Es verfügen aber nur wenige Wirkungsevaluierungen über den wissenschaftlich wünschenswerten Stand, um die Leistungen dieser Initiativen sicher bewerten zu können. Dieser Bericht erläutert die verschiedenen praktizierten Formen der Evaluierung und zeigt, dass „schlechte“ Evaluierung, zum Beispiel in Form von bloßen Vorher-Nachher-Vergleichen, zu verzerrten Bewertungen führt und typischerweise das Ergebnis beschönigt. Diese Schwächen bedeuten, dass die Träger finanzieller Bildungsmaßnahmen und die wirtschaftspolitisch Verantwortlichen die Wirkung ihrer Aktion eventuell überschätzen und die wahren Probleme nicht gut erkennen. Von methodisch schlechten Evaluierungen sollte abgesehen werden, um stattdessen fachgerechte Evaluierungsmethoden zu definieren und einzusetzen.

Die unzureichende finanzielle Bildung vieler Menschen ist ein identifiziertes und detailliert beschriebenes Problem.¹ Die OECD hat bereits im Jahr 2005 dazu aufgefordert, finanzielle Bildung in Schulen zu verankern² und die Erfassung finanzieller Kompetenz im Rahmen der international vergleichenden Schulleistungsstudie PISA wurde erstmals im Jahr 2012 durchgeführt. Zahlreiche Initiativen und Maßnahmen sollen zudem die Finanzbildung verbessern: So werden vielerorts am Arbeitsplatz Trainings zur Altersvorsorge angeboten, auch existieren individuelle Trainingsangebote zu Teilaspekten wie Sparen oder Schulden.³ Ziel dieser Angebote ist es, die Kenntnis über Finanzprodukte und Zusammenhänge zu verbessern und auf das Finanzverhalten der einzelnen Akteure einzuwirken, um ihnen adäquate Konsum-, Kredit- oder Anlageentscheidungen zu ermöglichen.

Diesen verschiedenen Maßnahmen konnten bisher aber keine eindeutig erwiesenen Erfolge zugeschrieben werden.⁴ So wurde ein „Mangel jeglicher zwingenden Evidenz einer positiven Wirkung [von finanzieller Bildung]“ konstatiert.⁵ Offensichtlich herrscht eine große Diskrepanz zwischen dem großen Aufwand für die Durchführung der Maßnahmen und dem gerin-

1 Annamaria Lusardi und Olivia S. Mitchell (2014): The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, 52(1), 5–44; Leora Klapper, Annamaria Lusardi und Peter van Oudheusden (2015): Financial literacy around the world: Insights from the Standard and Poor's Rating Services global financial literacy survey (online verfügbar, abgerufen am 19. Juni 2017. Dies gilt, insofern nicht anders vermerkt, für alle Online Quellen in diesem Bericht).

2 Vgl. OECD (2005): Recommendation on Principles and Good Practices for Financial Education, OECD Publishing.

3 Für eine detailliertere Auflistung möglicher Maßnahmen siehe Tim Kaiser und Lukas Menkhoff (2017): Does financial education impact financial literacy and financial behavior, and, if so, when? *World Bank Economic Review* (forthcoming) und DIW Discussion Paper 1562 (revised).

4 Justine S. Hastings, Brigitte C. Madrian und William L. Skimmyhorn (2013): Financial literacy, financial education, and economic outcomes. *Annual Review of Economics*, 5, 347–373.

5 Daniel Fernandes, John G. Lynch Jr. und Richard G. Netemeyer (2014): Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8), 1861–1883.

gen Wissen über ihre Wirksamkeit. Studien kommen sogar zu gegensätzlichen Urteilen über die Effektivität finanzieller Bildung, teilweise weil unterschiedliche Maßstäbe an den Erfolg angelegt werden. Diese Situation ist für Träger finanzieller Bildungsmaßnahmen und die wirtschaftspolitisch Verantwortlichen unbefriedigend.

Dieser Wochenbericht widmet sich der Beantwortung von drei Fragen: Worauf kommt es bei einer guten Evaluierung von Bildungsmaßnahmen an? Wie unterscheiden sich hier „schlechte“ und „gute“ Evaluierung? Kann man überhaupt eine Aussage darüber treffen, ob finanzielle Bildung eine erwünschte Wirkung hat? Anhand eines Samples von Studien zur Wirkung finanzieller Bildungsmaßnahmen wird gezeigt, dass falsche Evaluierung zu „geschönten“ Ergebnissen führt. Wissenschaftlich überzeugende Evaluierung stellt weitaus höhere Anforderungen an Wirksamkeit.

Die Praxis der Wirkungsevaluierung ist sehr heterogen

Wirtschaftspolitische Maßnahmen werden nicht immer evaluiert: Wenn die Zusammenhänge gut bekannt sind, spart man sich lieber den Aufwand. Bei neuartigen Maßnahmen ist eine begleitende Wirkungsevaluierung dennoch sinnvoll, weil man zum einen wissen möchte, ob die Maßnahmen überhaupt zur Zielerreichung beitragen und zum anderen, wie man die Effektivität eventuell verbessern kann. Im Bereich der finanziellen Bildung, in dem seit mehreren Jahrzehnten verschiedene Träger aktiv sind, wird seit langem immer wieder evaluiert, häufig jedoch nicht nach den derzeit wissenschaftlich wünschenswerten Standards.

Einfache Ex post-Befragungen

Eine einfache und preiswerte Form der Evaluierung ist die Ex-post-Befragung der TeilnehmerInnen. Eine solche Methode schreibt sämtliche Änderungen den Ergebnissen der Intervention zu. Nach einem einwöchigen Training zur Finanzbildung werden beispielsweise die TeilnehmerInnen zum Abschluss nach ihren Erfahrungen gefragt. Die Fragen zielen darauf ab, zu erfahren, ob es Lernerfolge gab, worin diese bestehen, ob es zu geplanten Verhaltensänderungen kommt oder was man verbessern könnte. Das Hauptproblem solch einer Befragung ist offensichtlich, dass es sich um qualitative Selbsteinschätzungen handelt, die zudem noch unter dem unmittelbaren Eindruck des Trainings stehen, sodass die Wirkung tendenziell überschätzt wird. So gewonnene Erkenntnisse können hilfreich sein, sollten aber nicht mit einer Evidenz für eine kausale Wirkung der Maßnahme verwechselt werden.

Ex-ante- und Ex-post-Befragungen

Etwas aufwendiger, aber potentiell informativer sind zweimalige Befragungen, also einmal vor einem Training (ex ante) und ein zweites Mal danach (ex post). Aus der Befragung zu Beginn ergibt sich eine sogenannte Baseline-Information, die den Referenzpunkt für gemessene Veränderungen bildet. Allerdings bleibt unklar, ob während des Trainings – gerade wenn es sich über längere Perioden hinzieht – andere Dinge eintreten, die unabhängig von der Intervention einen Einfluss auf die gemessenen Werte haben können. Vielleicht haben die Medien Finanzthemen stark aufgegriffen, weil es zufällig – parallel zur Bildungsmaßnahme – starke Verwerfungen im Finanzbereich gab. Dann hätten sich die TeilnehmerInnen eventuell stark mit Finanzthemen beschäftigt und etwas gelernt, aber nur teilweise wegen der Intervention. Ein Kontrafaktum wird nicht beobachtet.

Evaluierung mit Vergleichsgruppe

Deshalb werden Evaluierungen bevorzugt mit einer Vergleichsgruppe durchgeführt (vgl. Kasten 1). Um im genannten Beispiel zu bleiben, würden neben den TeilnehmerInnen der Bildungsmaßnahme zur gleichen Zeit und in vergleichbarer Form auch andere Personen befragt, die kein Training erhalten. Man kann im Nachhinein die Veränderung bei der Zielgruppe mit Training mit einer möglichen Veränderung bei der Kontrollgruppe vergleichen. Die Selektion der TeilnehmerInnen ist hierbei die größte Herausforderung. Die Praxis zeigt, dass Menschen, die an Trainings teilnehmen, besondere Züge aufweisen: Sie haben beispielsweise ein besonders großes Interesse am Thema, sind überhaupt neugierig und aufgeschlossen für Neues. Es verwundert dann wenig, dass die Trainings mit solchen Personen gute Effekte hervorbringen, das ist aber nicht verallgemeinerungsfähig. Dem kann man partiell entgegen wirken, indem man für beobachtbare Charakteristika der Personen kontrolliert und gegebenenfalls korrigiert (sogenannte Matching-Verfahren).

Randomisierte Evaluierung

Ein probateres Mittel, um eine solche Kontrollgruppe zu identifizieren, und somit ein valides Kontrafaktum zu generieren, ist die Randomisierung: Hierbei entscheidet allein der Zufall, ob ein Individuum in der Treatment- oder der Kontrollgruppe landet und nicht etwa das Individuum selbst.⁶ Beispielsweise wird eine Gruppe per Zufallsentscheidung, wie den Wurf eines Würfels, in zwei Untergruppen aufgeteilt. Dadurch kann – bei einer großen Zahl von Individuen – davon ausgegangen werden, dass beide

⁶ Vgl. Julian C. Jamison (2016): The entry of randomized assignment into the social sciences. Working Paper (online verfügbar).

Kasten 1

Anforderung der Messung kausaler Wirkung

Wirkungsevaluierung meint die Identifikation von kausalen Effekten einer Intervention auf eine oder mehrere Ergebnisvariable(n). Während die Identifikation von kausalen Wirkungszusammenhängen in den Naturwissenschaften in der Regel durch die Isolierung von Umwelteinflüssen im Labor ohne Gefahr von Fehlschlüssen (hohe interne Validität) zu realisieren ist, gestaltet sich Wirkungsevaluierung in den Sozialwissenschaften häufig schwieriger: Menschliches Handeln ist selten unter Laborbedingungen zu beobachten, und wenn dies der Fall ist (zum Beispiel bei computergestützten Labor-Experimenten zur Überprüfung von ökonomischen Theorien), dann stellt sich die Frage nach der Generalisierbarkeit der ermittelten Ergebnisse (externe Validität) auf Situationen außerhalb des Labors.

Wirkungsevaluierung von Politikmaßnahmen, wie zum Beispiel einer Bildungsmaßnahme im Bereich finanzieller Bildung, bedarf daher der Beobachtung von menschlichem Handeln im natürlichen Raum (Feld), ist dennoch an einer möglichst genauen Identifikation von kausalen Wirkungszusammenhängen (hohe interne Validität) interessiert. Um die Unterschiede in der Qualität und Glaubwürdigkeit der verschiedenen Forschungs-Designs zu verstehen, hilft es, die allgemeine Frage nach der kausalen Wirkung einer Maßnahme zu formalisieren und zwei Kernkonzepte der Wirkungsevaluierung (kausale Inferenz und Kontrafaktum) näher zu erläutern:¹

¹ Vgl. Paul J. Gertler et al. (2016): *Impact evaluation in practice*. World Bank Publications.

Unter Wirkung (Δ) wird die Differenz zwischen einer Ergebnisvariable einer Person mit Intervention ($Y|I=1$) und derselben Ergebnisvariable derselben Person ohne Intervention ($Y|I=0$) verstanden: $\Delta = (Y|I=1) - (Y|I=0)$. $I=1$ könnte hierbei eine Maßnahme zur Stärkung finanzieller Bildung bedeuten; das interessierende Ergebnis Y wäre beispielsweise die Summe der Ersparnisse in Euro oder der Anteil von richtig beantworteten Fragen in einem Leistungstest zum finanziellen Wissen. Man würde also gerne die Ergebnisvariable (zum Beispiel die Ersparnisse) zum selben Zeitpunkt für dieselbe Person mit und ohne Intervention beobachten. Es ist offensichtlich, dass diese Art von Beobachtung nicht möglich ist, da zum Zeitpunkt der (Nicht-) Intervention die Entscheidung für einen Zustand getroffen wurde, der die gleichzeitige Beobachtung des alternativen Zustandes unmöglich macht.

Wissenschaftliche Wirkungsevaluierung versucht daher ein alternatives, valides *Kontrafaktum* zu identifizieren, um den kausalen Effekt der Intervention möglichst präzise zu schätzen. Hierzu wird die Wirkung auf Ebene von Gruppen gemessen: Existiert eine Gruppe mit Intervention (Treatment-/Interventions-Gruppe) und eine Gruppe ohne Intervention (Kontroll-/Vergleichs-Gruppe) so kann man die Mittelwerte für die Ergebnisvariable beider Gruppen miteinander vergleichen – wenn interne Validität gegeben ist, spiegelt die Mittelwertdifferenz den kausalen Effekt wieder. So betrachtet besteht ein Kernproblem der Wirkungsevaluierung darin, eine valide Kontrollgruppe zu finden.²

² Vgl. Paul J. Gertler et al. (2016), a.a.O., 52ff.

Gruppen im Mittel vollkommen identisch sind und zwar sowohl in Bezug auf beobachtbare Merkmale, aber auch – und dies ist wichtig – in Bezug auf nicht beobachtbare Merkmale. Der einzige systematische Unterschied zwischen den beiden Gruppen ist dann die experimentelle Intervention (zum Beispiel die Bildungsmaßnahme). Dies ist das ideale sozialwissenschaftliche Experiment und die sauberste Methode, kausale Effekte zu identifizieren. Ihre Bedeutung hat jüngst stark zugenommen, allerdings existieren auch Grenzen (vgl. Kasten 2).

Schlechte Evaluierung finanzieller Bildungsmaßnahmen überschätzt deren Effektivität

Von den vier beschriebenen Formen der Evaluierung wird im Folgenden nur noch auf die zwei oder drei bes-

ten Formen eingegangen – Ex-Post-Befragungen werden also außen vor gelassen. Konkret basiert dieser Bericht auf einer Analyse von 143 Arbeiten.⁷ Diese Studien wurden in drei Gruppen klassifiziert, je nach wissenschaftlicher Qualität der Evaluierungsmethode. Gruppe 1 bilden hierbei Studien, die Vorher-Nachher-Vergleiche von nur einer Gruppe verwenden. Gruppe 2 sind Studien, die auf Quasi-Experimenten oder natürlichen Experimenten beruhen (beobachtetes Kontrafaktum). Gruppe 3 besteht aus Studien, die randomisierte Evaluierungen nutzen.

⁷ Vgl. Tim Kaiser und Lukas Menkhoff (2017), a. a. O. In dieser Meta-Analyse arbeiten wir mit einem Datensatz aus 126 Studien. Die 17 weiteren Studien wurden im Rahmen einer systematischen Literaturrecherche identifiziert und wegen Nicht-Erfüllung der Inklusionskriterien hinsichtlich der methodischen Qualität nicht in die Empirische Analyse eingeschlossen.

Kasten 2

Verbreitung und Grenzen randomisierter Evaluierung

Randomisierte Evaluierungen haben sich erst seit den neunziger Jahren in der Ökonomie durchgesetzt. Im Bereich der finanziellen Bildung gab es in den letzten zehn Jahren ein deutlicher Anstieg von experimentellen Evaluierungen (Abbildung). Pro Kalenderjahr sind alle aussagekräftigen Evaluierungsstudien ermittelt worden und davon der Anteil an randomisierten Studien, sogenannten RCTs (Randomized Controlled Trials).

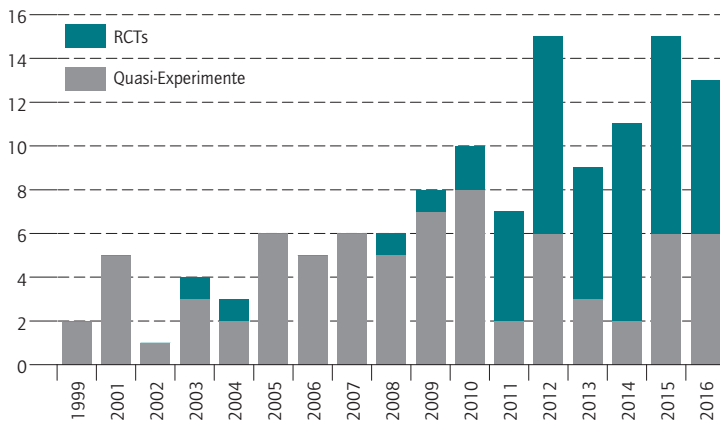
RCTs gelten als besonders aussagekräftig im Hinblick auf kausale Aussagen. Allerdings sind die Anforderungen an RCTs hoch, was den Mitteleinsatz anbelangt, und sie sind spezifisch, was die Fragestellung anbelangt. So gibt es neben Interventionen, die lediglich Opportunitätskosten verursachen (wie Bildungsmaßnahmen) auch Interventionen (beispielsweise in der Medizin), die aus ethischen Gründen nicht am Menschen durchführbar sind. Daneben kann es Ereignisse geben, die nicht in Experimenten herbeigeführt werden können, wie Naturkatastrophen, Finanzkrisen oder große Innovationen. Manche Fragen erfordern für ihre Analyse auch lange Zeiträume, für die sich RCTs nicht eignen, unter anderem weil man die Umfeldbedingungen mit den Jahren immer schlechter kontrollieren kann.

Eine randomisierte Evaluierung ist vergleichsweise aufwendig und leider nicht immer durchführbar. Eine wichtige Bedingung ist dabei, dass die zu vergleichenden Gruppen (mit und ohne Intervention) ähnlich sind, was für eine räumliche Nähe spricht, aber gleichzeitig darf es keine Übertragung (Spillover-Effekte) von den Trainierten zu den Untrainierten geben. Weiterhin bestehen häufiger ethische oder praktische Vorbehalte, manche Personen in den Genuss einer Intervention kommen zu lassen und andere nicht. Manchmal wird die Intervention mit der anfangs unberücksichtigten (Kontroll-) Gruppe nachgeholt; dennoch erweist sich diese empfundene Ungleichbehandlung in der Praxis als ein Nachteil einer randomisierten Evaluierung.

Trotz all dieser Einschränkungen lässt sich aber feststellen, dass RCTs sehr geeignet sind, um kurz- und mittelfristige Effekte von Bildungsmaßnahmen zu evaluieren, was häufig der relevante Zeithorizont ist.

Abbildung

Anzahl der Studien zur Wirkung finanzieller Bildung nach Forschungsdesign pro Jahr



Quelle: Eigene Berechnungen.

© DIW Berlin 2017

Randomized Controlled Trials machen seit fünf Jahren die Mehrheit der Studien aus.

Um Vergleichbarkeit hinsichtlich der Wirksamkeit finanzieller Bildungsmaßnahmen herzustellen, werden für alle Studien sogenannte Effektstärken⁸ berechnet, also ein standardisiertes Maß, das etwas über die Wirksamkeit der Bildungsmaßnahme im Hinblick auf den angestrebten Erfolg aussagt (Abbildung 1).

Die weniger rigorosen Gruppe-1-Studien weisen demnach zwei Eigenschaften auf: Erstens liegt ihre durchschnittliche Effektstärke deutlich über derjenigen der Gruppen 2 und 3. Breite Konfidenzintervalle bei den

Ergebnissen weisen zweitens auf eine erhebliche Unsicherheit in Bezug auf die geschätzten Effekte hin. Offensichtlich liefern schwächere Evaluierungsmethoden tendenziell ein nach oben verzerrtes und zugleich unpräziseres Bild. Deshalb werden weitere Fragen nur noch anhand der 126 Studien der Gruppe 2 und 3 erläutert.

Selbst bei rigorosen Studien beeinflussen die Forschungsmethoden die Ergebnisse stark

Für die Studien der Gruppen 2 und 3 wird die Wirkung, die eine Variation der Messung auf die geschätzte Effek-

⁸ Mark W. Lipsey und David B. Wilson (2001). Practical meta-analysis. Sage, Thousand Oaks, CA.

tivität der Bildungsmaßnahme hat, untersucht. Es geht hier also ausschließlich um die Beeinflussung des Messergebnisses durch die Art der Messung. Hierzu werden vier mögliche Einflüsse betrachtet:

1. Studiendesign: Berücksichtigt wird, ob die Studie von ihrem Design her dem höchsten Anspruch gerecht wird und als Randomized Controlled Trial (RCT) angelegt ist. Da RCTs am stärksten positiven Effekten durch Selbstselektion entgegenwirken (siehe Kasten 2), wird hier ein negatives Vorzeichen des Koeffizienten erwartet.

2. Erfassung der Zielgruppe: Bei einer Bildungsintervention wird vorab definiert, wer zur Zielgruppe gehört. Bei der Messung der Effektivität hat man dann – sofern das Studiendesign diese Unterscheidung zulässt – zwei Möglichkeiten, die Zielgruppe zu definieren. Man misst entweder die Effektivität für diejenigen, die wirklich an der Maßnahme teilgenommen haben, das ist der sogenannte TOT-Effekt (*treatment on the treated*), oder man misst die Wirkung bei der ursprünglich angestrebten Zielgruppe, der sogenannte ITT-Effekt (*intended to treat*). Da typischerweise nicht alle Mitglieder einer Zielgruppe erreicht werden, weil manche nicht kooperieren, liegt der TOT-Effekt systematisch höher. Dieser höhere Effekt wäre weniger wichtig, wenn er zufällig auftreten würde. Tatsächlich jedoch zeigt die Praxis, dass gerade diejenigen durch Bildungsmaßnahmen schwer zu erreichen sind, die es am nötigsten hätten und deshalb ausdrücklich zur Zielgruppe gehören.

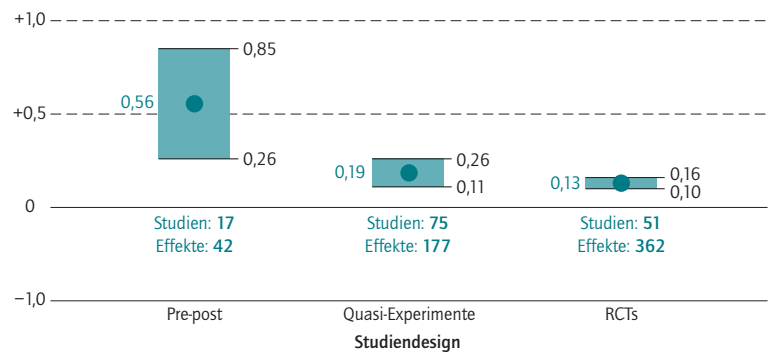
3. Messabstand: Typischerweise lässt die Erinnerung an Gelerntes mit der Zeit nach. Folglich wirkt sich der zeitliche Abstand zwischen einer Bildungsintervention und der Messung des Erfolgs nachteilig auf die gemessene Effektivität aus.

4. Studiengröße: Studien mit mehr TeilnehmerInnen weisen tendenziell kleinere Standardfehler für jede Schätzung aus. Insofern führen größere Studien zu mehr Verlässlichkeit in der Aussage.

In dem untersuchten Sample aus 126 Studien erklären die kodierten Variablen zum Studiendesign rund 16 Prozent der Varianz in den Effektstärken (Tabelle 1). Im Einzelnen wird ersichtlich, dass RCTs tendenziell kleinere Effektstärken berichten (negatives Vorzeichen), dass die Operationalisierung als TOT-Effekt mit höheren geschätzten Effektstärken einhergeht, dass ein größerer Abstand von Intervention und Messung höhere Effekte erwarten lässt (jedoch nicht mit einem besonders starken Einfluss), und dass Studien mit größeren Standardfehlern systematisch zu einer Verzerrung zugunsten von zu optimistisch eingeschätzter Wirksamkeit gelangen.

Abbildung 1

Empirischer Unterschied in den berichteten Effektstärken (Hedges' g) nach Studiendesign von Maßnahmen finanzieller Bildung



Quelle: Eigene Berechnungen.

© DIW Berlin 2017

Weniger rigorose Studiendesigns führen zu nach oben verzerrten und mit hoher Unsicherheit geschätzten Ergebnissen.

Tabelle 1

Einflüsse des Studiendesigns auf die geschätzten Effektstärken von Maßnahmen finanzieller Bildung

	(1)
	Abhängige Variable: Effektstärke
RCT	-0,073* (0,041)
TOT	0,092* (0,049)
Delay	-0,000*** (0,000)
SE	1,132** (0,552)
Konstante	0,090** (0,035)
Adj. R2	0,160
n (Studien)	126
n (Effektstärken)	539

Anmerkung: Koeffizienten sind Ergebnisse einer OLS Regression. Abhängige Variable ist Effektstärke einer Intervention zur finanziellen Bildung auf finanzielles Wissen oder finanzielles Verhalten. Standardfehler (geclustert auf Studienebene) in Klammern unter den Koeffizienten.

* p<0.1, ** p<0.05, *** p<0.01

Quelle: Eigene Berechnungen.

© DIW Berlin 2017

Tabelle 2

Zusammenfassung der durchschnittlichen Wirksamkeit finanzieller Bildung in der Literatur

Abhängige Variable:	Finanzielles Wissen		Finanzielles Verhalten	
	Alle Studien	Nur RCTs	Alle Studien	Nur RCTs
Effektstärke				
(1) OLS	0,263*** (0,041)	0,209*** (0,033)	0,086*** (0,012)	0,082*** (0,014)
(2) WLS (1/SE)	0,191*** (0,025)	0,175*** (0,024)	0,026** (0,011)	0,067*** (0,013)
(3) WLS (1/SE2)	0,135*** (0,019)	0,155*** (0,015)	0,002* (0,001)	0,051*** (0,010)
(4) Metareg	0,297*** (0,042)	0,234*** (0,039)	0,079*** (0,009)	0,075*** (0,013)
(5) Robumeta	0,287*** (0,037)	0,237*** (0,039)	0,064*** (0,008)	0,078*** (0,012)
n(Studien)	67	33	90	40
n(Effektstärken)	190	135	349	227

Anmerkung: Abhängige Variable ist die Effektstärke (Hedges' g).

Die Tabelle zeigt Ergebnisse verschiedener Meta-Analyse-Modelle. Ausgewiesen sind die (gewichteten) mittleren Effektstärken pro Outcome-Typ und Studiendesign. Standardfehler sind in Klammern.

* p<0.1, ** p<0.05, *** p<0.01

Zeile (1) zeigt die Ergebnisse einer ungewichteten OLS-Regression. Zeile (2) zeigt Ergebnisse einer unrestricted weighted least squares regression mit dem inversen Standardfehler als Regressionsgewicht, vgl. Tom D. Stanley und Hristos Doucouliagos (2015). Neither fixed nor random: weighted least squares meta-analysis. *Statistics in Medicine* 34(13), 2115–2127. Zeile (3) zeigt eine identische Schätzung mit der inversen Varianz als Gewicht. Zeile (4) zeigt Ergebnisse einer Random-Effects-Metaregression mit korrigierten Standardfehlern nach Knapp und Hartung, vgl. Rebecca DerSimonian und Nan Laird (1986). *Meta-analysis in clinical trials. Controlled clinical trials*, 7(3), 177–188; Guido Knapp und Joachim Hartung (2003). *Improved tests for a random effects meta-regression with a single covariate. Statistics in Medicine*, 22(17), 2693–2710. Diese Schätzung arbeitet mit einer (synthetischen) Effektstärke pro Studie. Zeile (5) zeigt Ergebnisse einer Robust variance meta-regression with dependent effect size estimates, vgl. Larry V. Hedges, Elizabeth Tipton und Matthew C. Johnson. (2010). *Research Synthesis Methods*. 1(1), 39–65.

© DIW Berlin 2017

Der signifikante Koeffizient des Standardfehlers der Effektstärke (SE) gibt zudem einen Hinweis auf einen vorhandenen Publikationsbias: Wenn Effektstärken systematisch mit dem dazugehörigen Standardfehler im Zusammenhang stehen, bedeutet dies, dass Studien mit insignifikanten Ergebnissen seltener publiziert werden. Dies bezeichnet man als *file drawer problem*, weil ForscherInnen „schwache“ Forschungsergebnisse „in eine Schublade verbannen“. ⁹ Dieses Publikationsverhalten

⁹ Vgl. Tom D. Stanley (2001): Wheat from chaff: Meta-Analysis as quantitative literature review. *Journal of Economic Perspectives* 15(3), 131–150.

verzerrt eine Analyse der durchschnittlichen Wirksamkeit tendenziell nach oben, da die insignifikanten Ergebnisse nicht im Datensatz vorhanden sind. Eine andere Erklärung für diesen positiven Zusammenhang können allerdings auch ex-ante Power-Berechnungen sein, die die Samplegröße schon vor Beginn der Studie an die erwarteten Effektstärken anpassen.

Festhalten lässt sich, dass die gemessene Wirkung einer Maßnahme größer sein wird, wenn die Evaluierung nicht als RCT erfolgt, wenn der TOT-Effekt angegeben wird und/oder wenn der Messabstand möglichst geringgehalten wird. Schließlich sind Aussagen aus Studien mit geringerer Varianz (zugleich eher größere Studien) tendenziell verlässlicher.

Finanzbildung ist wirksam, finden auch rigorose Studien

Wie bereits erwähnt, sind die Urteile zur Wirksamkeit recht unterschiedlich. Neben den qualitativen (narrativen) Überblicksarbeiten, die auf subjektiver Auswahl und Interpretation der Literatur beruhen, gibt es quantitative Meta-Analysen. Diese haben den Anspruch, alle verfügbaren Forschungsarbeiten in einer einheitlichen und nachvollziehbaren Weise auszuwerten. Damit sind sie einerseits rigide, andererseits weniger subjektiv. Leider liegt für die Analyse der Wirksamkeit von Maßnahmen finanzieller Bildung bisher nur eine umfangreiche Meta-Analyse vor. ¹⁰

Diese Arbeit wirkte wie ein Paukenschlag in der einschlägigen Literatur. Während bis dahin die Vorstellung vorherrschte, Bildungsinterventionen würden mehr oder weniger stark wirken, stellten die Autoren in den Raum, es gäbe keine nachweisbare Wirksamkeit. Im Durchschnitt aller von ihnen betrachteten Studien kommen sie auf eine Effektstärke von etwa 0,07. Effektstärken unterhalb von 0,2 werden als klein bezeichnet (und oberhalb von 0,8 als groß). Aber noch wesentlicher ist die Feststellung, dass selbst dieser geringe Wert noch zu hoch ist, wenn man methodisch sauber misst. Die Effektstärke von allen damals erfassten RCTs, als bester Messung, liegt nur noch bei 0,02 und ist damit klar insignifikant und ökonomisch bedeutungslos.

Eine Gruppe von AutorInnen der Weltbank hat daraufhin mit einer eigenen Meta-Analyse reagiert. ¹¹ Auch sie weisen neben Erfolgen auf Probleme hin und darauf, dass die erwünschten Wirkungen, beispielsweise ein besse-

¹⁰ Daniel Fernandes, John G. Lynch Jr. und Richard G. Netemeyer (2014), a. a. O.

¹¹ Margret Miller et al. (2015): Can you help someone become financially capable? A meta-analysis of the literature. *World Bank Research Observer*, 30(2), 220–246.

res Verständnis finanzieller Produkte oder ein rationales Spar- oder Anlageverhalten, nicht immer eintreten. Gerade beim Kreditverhalten scheinen Bildungsmaßnahmen oft wirkungslos zu bleiben: Menschen nehmen immer noch zu viele Schulden auf und können Verbindlichkeiten nicht bedienen. Der Nachteil dieser Studie ist ihre geringe Abdeckung: Es wurden nur 19 Primärstudien berücksichtigt, so dass vertiefende Aussagen kaum generalisiert werden können.

Maßnahmen wirksam sowohl auf Finanzwissen als auch auf Finanzverhalten

Neue, eigene Untersuchungen kommen zu zwei wesentlichen Ergebnissen: Maßnahmen finanzieller Bildung sind meist wirksam, auch bei Verwendung rigoroser Evaluierungsmethoden, und die Streubreite der Wirkungen ist hoch.

Interventionen der finanziellen Bildung beeinflussen sowohl das Wissen als auch das Verhalten positiv. Dieser Befund gilt auch dann, wenn höchste methodische Ansprüche an die Evaluierungen gestellt werden, das heißt, das Sample auf randomisierte Experimente beschränkt wird. Der mittlere Effekt von Interventionen auf das finanzielle Wissen im Sample aller 67 berücksichtigten Studien der Gruppe 2 und 3 liegt je nach Schätzmethode in der Meta-Analyse zwischen 0,135 und 0,297 (Tabelle 2). Betrachtet man nur die Ergebnisse von Studien der Gruppe 3 (RCTs) liegt der ermittelte Effekt zwischen 0,155 und 0,234. Damit wird deutlich, dass Maßnahmen zur Stärkung des finanziellen Wissens ähnlich wirksam sind wie Bildungsinterventionen im Bereich der Naturwissenschaften an Hochschulen.¹²

Die Effektstärke auf tatsächliches Finanzverhalten ist zwar deutlich geringer (zwischen 0,002 und 0,086 für alle Studien und zwischen 0,051 und 0,082 für RCTs), jedoch existiert eine signifikant positive, messbare Evidenz für einen kausalen Effekt der Bildungsmaßnahmen in sämtlichen durchgeführten Meta-Analysen. Zudem sind diese Effektstärken durchaus vergleichbar mit Interventionen aus anderen Domänen wie den Gesundheitswissenschaften, in denen es darum geht, Menschen zu einer Änderung ihres Verhaltens anzuregen.¹³

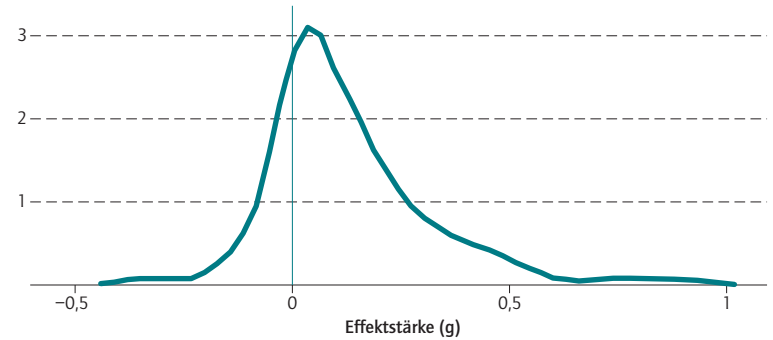
¹² Vgl. Scott Freeman et al. (2014): Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415.

¹³ David B. Portnoy et al. (2008): Computer-delivered interventions for health promotion and behavioral risk reduction: A meta-analysis of 75 randomized controlled trials, 1988-2007. *Preventive Medicine*, 47(1), 3-16.

Abbildung 2

Verteilung (Dichteschätzung) der kodierten Effektstärken bei den untersuchten Studien zur Wirkung finanzieller Bildung

Für $g < 1$



Quelle: Eigene Berechnungen.

© DIW Berlin 2017

Die empirischen Effektstärken in der Literatur zur finanziellen Bildung sind äußerst heterogen.

Ergebnisse sind sehr heterogen

Es muss allerdings konstatiert werden, dass eine recht breite Streuung in den geschätzten Effektstärken besteht. Dies wird an die Verteilung aller kodierten (ungewichteten) Effektstärken für das eingeführte Sample von 126 Studien deutlich (Abbildung 2). Neben den oben diskutierten Erklärungsfaktoren des Studiendesigns gibt es weitere Determinanten, auf die hier nicht eingegangen wird, und einen erheblichen unerklärten „Rest“.¹⁴ In jedem Fall verdeutlicht die große Spannweite an ermittelten Wirkungen, dass man grundsätzlich von erfolgreichen Maßnahmen lernen kann; aber damit dies systematisch möglich ist, bedarf es einer hinreichend großen Zahl sorgfältig evaluierter Bildungsmaßnahmen.

Fazit: Rigorose Evaluierung notwendig

Finanzielle Bildung hat in den letzten Jahren erhöhte Aufmerksamkeit bekommen, sowohl von privaten Trägern als auch von Schulbehörden und der Wirtschaftspolitik. Es wird in diesem Bereich viel investiert, aber gründliche Evaluierungen der getroffenen Maßnahmen sind rar. Dabei kann man nur aus methodisch überzeugenden Studien verlässlich lernen. Einfache Methoden wie Vorher-Nachher-Befragungen haben zwar den Vorteil, schnell und preiswert zu sein, weisen aber zwei

¹⁴ Vgl. Tim Kaiser und Lukas Menkhoff (2017), a. a. O.

große Nachteile auf: Sie überschätzen die Wirksamkeit der Bildungsmaßnahmen systematisch und sind wenig präzise, so dass man kaum etwas über die Erfolgsbedingungen von Bildungsinitiativen lernen kann.

Will man wissen, welche Art finanzieller Bildung in welcher Form nachhaltig wirkt, sollte man Gelegenheiten zum Beispiel bei der Einführung neuer Curricula nut-

zen, um systematisch und verlässlich daraus für spätere Maßnahmen zu lernen.

Bildungsmaßnahmen wirken sich stärker auf das Finanzwissen als auf das Finanzverhalten aus. Fachgerechte Evaluierung würde es möglicherweise erlauben, diese Diskrepanz zu adressieren und gezielter auf die Verbesserung des Finanzverhaltens hinzuwirken.

Tim Kaiser ist wissenschaftlicher Mitarbeiter der Abteilung Weltwirtschaft am DIW Berlin | tkaiser@diw.de

Lukas Menkhoff ist Leiter der Abteilung Weltwirtschaft am DIW Berlin | lmekhoff@diw.de

JEL: D 14, I 21

Keywords: Financial education, financial literacy, financial behavior, meta-analysis, meta-regression, impact evaluation

In Deutschland sinkt die Zahl der Betriebsgründungen weiterhin, nicht aber in Berlin

Von Alexander S. Kritikos und Anselm Mattes

Berlin war auch im Jahr 2016 die deutsche Gründungshauptstadt. Pro Kopf wurden deutlich mehr Betriebe sowohl im Vergleich zum Bundesdurchschnitt als auch im Vergleich zu anderen Großstädten wie beispielsweise Hamburg eröffnet. Außerdem nimmt im Bund die Gesamtzahl der Gründungen seit dem Jahr 2011 weiterhin kontinuierlich ab, während in Berlin dieser Abwärtstrend gestoppt wurde: Im Jahr 2016 starteten ähnlich viele Menschen einen eigenen Betrieb wie im Jahr 2015. Schlusslicht im Vergleich der zwölf größten deutschen Städte ist Stuttgart mit der niedrigsten Gründungsneigung. Nach wie vor fällt in Berlin vor allem die hohe Bedeutung von Betriebseröffnungen durch ausländische Bürgerinnen und Bürger auf, die fast 50 Prozent aller Gründungen ausmachen. Bei Betriebsgründungen mit Innovationspotential haben München und Köln die Nase vorn. In diesem Segment der forschungs- und wissensintensiven Gründungen verliert Berlin ähnlich wie viele andere Großstädte an Dynamik.

Unternehmensgründungen beeinflussen auf vielfältige Art und Weise die nationale und regionale Wirtschaft, etwa durch die Umsetzung neuer Produktideen oder durch die Intensivierung des Wettbewerbs und die Schaffung von Arbeitsplätzen. Insgesamt können Gründungen einen wichtigen Beitrag zum wirtschaftlichen Strukturwandel und zum Produktivitätswachstum leisten.

Vor diesem Hintergrund untersucht der Beitrag die Entwicklung des Gründungsgeschehens im Bund und in Berlin für das Jahr 2016 im Vergleich zu den Vorjahren. Darüber hinaus wird die Entwicklung in den zwölf größten deutschen Städten für das Jahr 2015 in die Perspektive genommen.¹ Die Analysen basieren auf Daten der Gewerbeanzeigenstatistik (Kasten 1). Dabei wird im weiteren Verlauf zwischen Gründungen von *echten Betrieben* und Kleingewerbegründungen zum Vollerwerb und zum Nebenerwerb unterschieden. Darüber hinaus legt der Bericht einen Fokus auf Gründungen durch Ausländerinnen und Ausländer sowie auf Gründungen mit Innovationspotential im Bereich der forschungsintensiven Industrien und wissensintensiven Dienstleistungen, die als besonders wichtig für die zukünftige Entwicklung einer Region oder eines Landes gelten.

Abwärtstrend in Berlin vorerst gestoppt, im Bund setzt er sich fort

Im Jahr 2016 wurden in Berlin knapp 37 000 Betriebe eröffnet, nahezu gleich viele wie im Jahr 2015 (Tabelle 1). Damit ist der kontinuierliche Abwärtstrend der letzten fünf Jahre – im Jahr 2011 zählte man in Berlin noch knapp 42 000 Gründungen – vorerst gestoppt. Im Bund

¹ Die Studie schließt an frühere Untersuchungen des DIW Berlin zum Gründungsgeschehen an. Siehe zuletzt Alexander S. Kritikos (2016): Berlin: Hauptstadt der Gründungen, aber (noch) nicht der schnell wachsenden Unternehmen. DIW Wochenbericht Nr. 29, 637–644 (online verfügbar, abgerufen am 20. Juni 2017). Dies gilt auch für alle anderen Online-Quellen dieses Berichts, sofern nicht anders vermerkt). Der vorliegende Bericht basiert auf einer Studie der DIW Econ GmbH im Auftrag der BBB Bürgschaftsbank zu Berlin-Brandenburg: „Gründerindex Ausgabe 2017“, Berlin 2017 (online verfügbar). Darin sind weitere Details zur Datenbasis und den verwendeten Methoden sowie weitere Ergebnisse enthalten.

Kasten 1

Datengrundlage Gewerbeanzeigenstatistik

Grundlage der Erhebung ist die *Gewerbeanzeigenstatistik*, die Informationen über die obligatorischen Anmeldungen bei den Gewerbeldeämtern liefert. Gegenüber anderen Statistiken zu Betriebsgründungen hat die Gewerbeanzeigenstatistik den Vorteil einer Vollerhebung. Sie liefert auch verlässliche Daten für kleinere regionale Einheiten und erlaubt damit einen aussagekräftigen Vergleich der Gründungsneigung auf der Ebene einzelner Städte.¹

Die Gewerbeanzeigenstatistik unterscheidet drei verschiedene Gründungsformen: (1) Als sogenannte *echte Betriebe* werden alle Gründungen durch juristische Personen und Personengesellschaften erfasst. Bei der Gründung eines *echten Betriebs* durch eine natürliche Person muss dieser entweder im Handelsregister eingetragen sein, eine Handwerkskarte besitzen oder mindestens einen Arbeitnehmer oder eine Arbeitnehmerin beschäftigen. (2) Kleingewerbegründungen zum Vollerwerb sind alle Einzelgründer beziehungsweise Einzelgründerinnen, die ihren beruflichen Mittelpunkt in das neue Gewerbe legen. (3) Dem gegenüber haben Gründer und Gründerinnen von Kleingewerbebetrieben zum Nebenerwerb (noch) nicht ihren beruflichen Mittelpunkt in der selbstständigen Tätigkeit. Nicht betrachtet werden Übernahmen und Übergaben sowie die Umwandlung

¹ Für eine umfassende Darstellung der Vor- und Nachteile der Gewerbeanzeigenstatistik im Vergleich zu anderen Datenquellen zur Erfassung des Gründungsgeschehens, wie dem Mikrozensus, der Statistik der sozialversicherungspflichtig Beschäftigten der Bundesagentur für Arbeit, dem Mannheimer Unternehmerpanel des Zentrums für Europäische Wirtschaftsforschung, dem Global Entrepreneurship Monitor (GEM), sowie dem KfW-Gründungsmonitor, siehe Michael Fritsch, Alexander S. Kritikos und Alina Rusakova (2012): *Selbstständigkeit in Deutschland: Der Trend zeigt seit langem nach oben*. DIW Wochenbericht Nr. 4/2012 (online abrufbar).

von Unternehmen. Zweigniederlassungen und unselbstständige Zweigstellen bleiben bei den Analysen ebenso außer Acht. Nicht erfasst werden Berufe, die nach der Gewerbeordnung nicht meldepflichtig sind (freie Berufe, Versicherungen sowie die Urproduktion wie Land- und Forstwirtschaft oder Bergbau).

Die Gewerbeanzeigenstatistik enthält auch Informationen darüber, in welcher Branche ein neu gegründeter Betrieb aktiv ist. Die Brancheneinteilung erfolgt anhand der amtlichen Klassifikation der Wirtschaftszweige des Jahres 2008 (WZ 2008).² Der Schwerpunkt der Gründungstätigkeit liegt im Dienstleistungsbereich, insbesondere bei den *sonstigen wirtschaftlichen Dienstleistungen*, den *freiberuflichen, technischen und wissenschaftlichen Dienstleistungen* sowie den *sonstigen Dienstleistungen*. Zu den freiberuflichen, technischen und wissenschaftlichen Dienstleistungen zählen die Rechts- und Steuerberatung, die Wirtschaftsprüfung, das Auftragsmanagement und die PR- sowie Unternehmensberatung, Architektur- und Ingenieurbüros, naturwissenschaftliche Analyselabors, Forschungs- und Entwicklungsdienstleistungen, Werbeagenturen und Marktforschungsinstitute, Designstudios, Fotografische Dienstleistungen, Übersetzer/Dolmetscher und das Veterinärwesen. Zu den sonstigen wirtschaftlichen Dienstleistungen gehören unter anderem die Vermietung von beweglichen Sachen (inklusive Videotheken), Leasing-Agenturen, Arbeitskräftevermittlung und Arbeitskräfteverleih, Reisebüros und Reiseveranstalter, Wach- und Sicherheitsdienste, Gebäude-Dienstleistungen, Messe-, Ausstellungs- und Kongressveranstalter sowie Call-Center. Die sonstigen Dienst-

² Siehe die Klassifikation der Wirtschaftszweige des Statistischen Bundesamtes (online abrufbar).

gingen die Gründungen von etwas über 533 000 im Jahr 2015 auf 518 000 eröffnete Gewerbe zurück; ein Rückgang von knapp drei Prozent. Im Jahr 2011 gab es im Bund noch 635 000 gewerbliche Gründungen (Tabelle 1).

Unterscheidet man nach Kleingewerbegründungen und *echten Betriebsgründungen*, zeigt sich ein differenzierteres Bild. In Berlin gibt es seit dem Jahr 2013 einen Aufwärtstrend bei der Gründung von *echten Betrieben*, der auch im Jahr 2016 weiter anhält; man zählte im Jahr 2016 knapp 7 000 *echte Gründungen*, 2013 waren es noch knapp 6 100. Im Bund stiegen sie im Jahr 2016 auf knapp 90 000, nachdem es im Jahr 2015 noch rund 86 000 gewesen waren. Damit konnte auf Bundesebene zumindest in diesem wichtigen Bereich der Gründung von *echten Betrieben* der Abwärtstrend gestoppt werden.

Insgesamt fällt auf, dass der Rückgang der Gewerbe Gründungen im Bund vor allem auf den Rückgang von Kleingewerbegründungen im Vollerwerb zurückzuführen ist.

Die meisten Gründungen gibt es im Dienstleistungsbereich

In Berlin wie im Bund nimmt die Gründung von Betrieben im Bereich der Dienstleistungen breiten Raum ein. Knapp 36 Prozent aller Gründungen finden in Berlin allein in drei Teilbereichen des Dienstleistungssektors statt: den *sonstigen wirtschaftlichen Dienstleistungen*, den *freiberuflichen, technischen und wissenschaftlichen Dienstleistungen* sowie den *sonstigen Dienstleistungen* (zur Erläuterung der Branchengliederung siehe Kasten 1). Im Bund war es mit knapp 35 Prozent ein ähnlich hoher Anteil

leistungen umfassen die Tätigkeiten von Interessenvertretungen (wie Wirtschafts- und Arbeitgeberverbände, Berufsorganisationen, kirchliche Vereinigungen und politische Parteien), die Reparatur von Datenverarbeitungsgeräten und Gebrauchsgütern und eine Vielzahl von persönlichen und anderen Dienstleistungen (wie Solarien, Wäschereien, Friseur- und Kosmetiksalons sowie Bestattungsinstitute).

Weiterhin erlaubt die Gewerbeanzeigenstatistik auch eine Unterscheidung nach Gründungen mit und ohne Innovationspotential. Zu diesem Zweck wird das Gründungsgeschehen auf Basis der Kategorisierung forschungsintensiver Industrien und wissensintensiver Dienstleistungen der Expertenkommission Forschung und Innovation (EFI) analysiert, die ebenfalls auf der WZ 2008 aufbaut.

Forschungs- und wissensintensive Branchen

Forschungsintensive Industrien sind:

- Verarbeitendes Gewerbe
 - Herstellung von chemischen Erzeugnissen (WZ 20)
 - Herstellung von pharmazeutischen Erzeugnissen (WZ 21)
 - Herstellung von Datenverarbeitungsgeräten, elektronischen und optischen Erzeugnissen (WZ 26)
 - Herstellung von elektrischen Ausrüstungen (WZ 27)
 - Maschinenbau (WZ 28)
 - Herstellung von Kraftwagen und Kraftwagenteilen (WZ 29)
 - Sonstiger Fahrzeugbau (WZ 30)

Wissensintensive Dienstleistungen sind:

- Information und Kommunikation
 - Verlagswesen (WZ 58)
 - Herstellung, Verleih und Vertrieb von Filmen und Fernsehprogrammen; Kinos; Tonstudios und Verlegen von Musik (WZ 59)
 - Rundfunkveranstalter (WZ 60)
 - Telekommunikation (WZ 61)
 - Erbringung von Dienstleistungen der Informationstechnologie (WZ 62)
 - Informationsdienstleistungen (WZ 63)
- Erbringung von Finanz- und Versicherungsdienstleistungen
 - Erbringung von Finanzdienstleistungen (WZ 64)
 - Versicherungen, Rückversicherungen und Pensionskassen ohne Sozialversicherungen (WZ 65)
 - Mit Finanz- und Versicherungsdienstleistungen verbundene Tätigkeiten (WZ 66)
- Erbringung von freiberuflichen und wissenschaftlichen Dienstleistungen
 - Rechts- und Steuerberatung, Wirtschaftsprüfung (WZ 69)
 - Verwaltung und Führung von Unternehmen und Betrieben; Unternehmensberatung (WZ 70)
 - Architektur- und Ingenieurbüros; technische, physikalische und chemische Untersuchung (WZ 71)
 - Forschung und Entwicklung (WZ 72)
 - Werbung und Marktforschung (WZ 73)
 - Sonstige freiberufliche, wissenschaftliche und technische Tätigkeiten (WZ 74)
 - Veterinärwesen (WZ 75)
- Sonstige wissensintensive Dienstleistungen
 - Gesundheitswesen (WZ 86)
 - Kreative, künstlerische und unterhaltende Tätigkeiten (WZ 90)
 - Bibliotheken, Archive, Museen, botanische und zoologische Gärten (WZ 91)

(Tabelle 2). Weiterhin spielt die Bauwirtschaft eine bedeutende Rolle (mit 23 Prozent in Berlin und 14 Prozent im Bund). Der viel höhere Anteil des Bausektors in Berlin wird von der hohen Zahl der Einzelgründungen getrieben. Betrachtet man nur die Gründungen der *echten Betriebe* in diesem Sektor, so reduziert sich der Anteil in Berlin auf elf Prozent, während es im Bund nach wie vor 14 Prozent sind.

Noch ist der Saldo der Gründungen und Stilllegungen positiv

Für ein umfassendes Bild des Gründungsgeschehens stellen wir den Gewerbeanmeldungen die zugehörigen Gewerbeabmeldungen gegenüber (Tabelle 3). Der Saldo zwischen Neugründungen von Betrieben und Stilllegun-

gen ist sowohl in Berlin als auch in Gesamtdeutschland positiv, das heißt es gibt einen Nettozuwachs im Unternehmensbestand. Allerdings wird der Nettozuwachs in Berlin in den letzten Jahren immer kleiner; auch im Bund gingen die Salden bis 2014 deutlich zurück. Dies liegt vor allem an dem Segment des Kleingewerbes zum Vollerwerb, bei dem es in Berlin im Jahr 2016 quasi zu einem Nullwachstum kam und im Bund – wie schon in den letzten Jahren – sogar zu einer Reduktion des Unternehmensbestandes; das heißt es wurden mehr Betriebe geschlossen als neu eröffnet. Sehr positiv verläuft dagegen die Entwicklung bei den *echten Betrieben*. In Berlin kamen im Jahr netto fast 2 000 und im Bund rund 15.700 *echte Betriebe* neu dazu. In diesem Bereich ist in Berlin schon länger ein positiver Trend zu beobachten, im Bund dagegen verlief die Entwicklung in den

Tabelle 1

Gewerbegründungen Berlin

	2011	2012	2013	2014	2015	2016
Echte Betriebsgründungen	6 405	6 470	6 098	6 286	6 511	6 971
Kleingewerbegründungen	35 516	35 228	34 194	32 524	30 207	29 766
darunter Vollerwerb	25 049	23 591	22 490	21 397	19 334	17 930
darunter Nebenerwerb	10 467	11 637	11 704	11 127	10 873	11 836
Gewerbegründungen insgesamt	41 921	41 698	40 292	38 810	36 718	36 737

Gewerbegründungen Bund

	2011	2012	2013	2014	2015	2016
Echte Betriebsgründungen	100 225	92 017	88 294	86 263	86 377	89 786
Kleingewerbegründungen	534 707	484 548	486 921	461 678	447 120	428 272
darunter Vollerwerb	293 018	243 351	238 039	210 562	198 417	178 357
darunter Nebenerwerb	241 689	241 197	248 882	251 116	248 703	249 915
Gewerbegründungen insgesamt	634 932	576 565	575 215	547 941	533 497	518 058

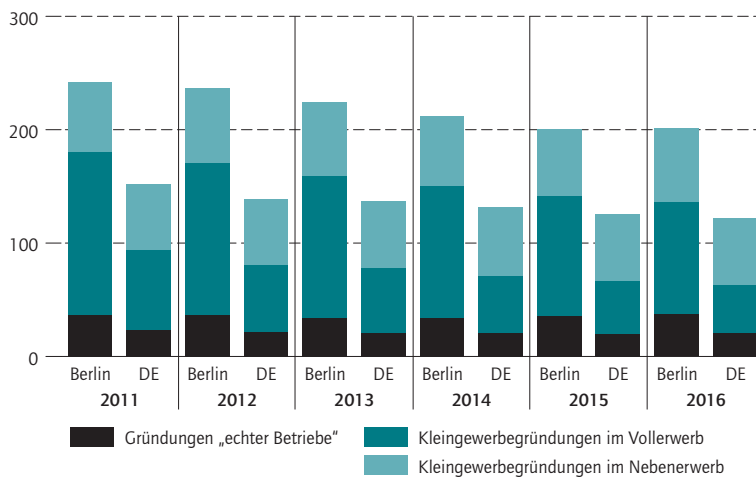
Quelle: Gewerbeanzeigenstatistik.

© DIW Berlin 2017

Abbildung 1

Gründungsneigung im Bund und in Berlin

Pro 10 000 Erwerbspersonen



Quellen: Gewerbeanzeigenstatistik; Mikrozensus; eigene Berechnungen.

© DIW Berlin 2017

Die Gründungsneigung ist in Berlin seit Jahren höher als im Bund.

letzten Jahren uneinheitlich, am aktuellen Rand kam es aber von 2015 auf 2016 zu einem sprunghaften Anstieg von 50 Prozent.

Berlin bleibt Gründungshauptstadt

Für einen Vergleich des Gründungsgeschehens zwischen verschiedenen Regionen oder Städten eignet sich die *Gründungsneigung*, also die Zahl der Gründungen pro 10 000 Erwerbspersonen (Kasten 2), besser als die absoluten Gründungszahlen. Im Jahr 2016 kam Berlin auf einen Wert von 201 Gründungen je 10 000 Erwerbspersonen (Abbildung 1). Dies entspricht fast exakt dem Vorjahreswert. Im Bund sind die Werte mit 123 Gründungen demgegenüber deutlich niedriger.²

Auch bei den *echten Betrieben* liegt Berlin deutlich vor dem Bund. Im Jahr 2016 wurden in Berlin 38 solcher Gründungen je 10 000 Erwerbspersonen registriert, dem steht ein Wert von 21 im Bund gegenüber. Bei den Kleingewerbegründungen zum Vollerwerb sind es mit 98 zu 42 neuen Gewerben sogar mehr als doppelt so viele.

Unter den zwölf größten deutschen Städten nimmt Berlin bei der Gründungsneigung nach wie vor den ersten Platz ein (Abbildung 2).³ Berlin war im Jahr 2015 mit 182 Gewerbegründungen je 10 000 erwerbsaktiven Personen deutlich gründungsfreudiger als die Städte in der Vergleichsgruppe: Leipzig (133), Hamburg (130) und München (126) liegen deutlich dahinter. Stuttgart (75) und Bremen (82) weisen die niedrigsten Werte bei der Gründungsneigung aus.

Die hohe Gründungsneigung in Berlin wird von Ausländern und Ausländerinnen getragen

Migrantinnen und Migranten gelten in vielen Ländern als gründungsfreudiger als einheimische Beschäftigte. Das gilt auch für Deutschland⁴ und kann unterschiedliche Gründe haben, nicht nur schlechtere Chancen am Arbeitsmarkt, sondern auch unterschiedliche Gründungsmotive oder eine andere Form von Kreativität und Innovationsbereitschaft. Vor dem Hintergrund der seit mehreren Jahren rückläufigen Gründungsbereitschaft gilt es, die höhere Gründungsneigung dieser Bevölkerungsgruppe gesondert zu analysieren.

Unterscheidet man entsprechend nach Nationalitäten, zeigt sich, dass die Gründungsneigung bei Deutschen bundesweit nur bei fünf Gründungen pro 1 000 Einwohn-

² Auch im Stadtstaatenvergleich liegt Berlin im Jahr 2016 in der Gründungsneigung weiterhin vor Hamburg und mit noch viel größerem Abstand vor Bremen. Siehe auch IfM Bonn (2017): Gewerbliche Existenzgründungsintensität 2006 bis 2016 in Deutschland nach Bundesländern (online verfügbar).

³ Der Städtevergleich beruht auf einer leicht veränderten Bezugsgröße, nämlich der Zahl der Erwerbsaktiven. Für Details siehe Kasten 2.

⁴ Siehe etwa René Leicht und Stefan Berwing (2016): „Gründungspotenziale von Menschen mit ausländischen Wurzeln: Entwicklungen, Erfolgsfaktoren, Hemmnisse.“ Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. IfM Bonn.

Tabelle 2

Anzahl der Gründungen, ausgewählte Branchen (2016)

	Absolut				Relativ (Prozent)			
	Gründungen insgesamt		Echte Betriebe		Gründungen insgesamt		Echte Betriebe	
	Berlin	Bund	Berlin	Bund	Berlin	Bund	Berlin	Bund
Baugewerbe	8 453	71 073	760	12 688	23,0	13,7	10,9	14,1
sonstige wirtschaftliche Dienstleistungen	4 890	67 589	722	7 523	13,3	13,0	10,4	8,4
Freiberufliche, technische und wissenschaftliche Dienstleistungen	4 153	58 692	972	11 511	11,3	11,3	13,9	12,8
Einzelhandel	3 573	77 121	774	10 253	9,7	14,9	11,1	11,4
sonstige Dienstleistungen	4 129	53 870	469	4 117	11,2	10,4	6,7	4,6
Gastgewerbe	2 168	31 046	849	9 508	5,9	6,0	12,2	10,6

Quellen: Gewerbeanzeigenstatistik; eigene Berechnungen.

© DIW Berlin 2017

Tabelle 3

Salden der Gewerbemeldungen in Berlin und Deutschland, 2011 bis 2016

	2011		2012		2013		2014		2015		2016	
	Berlin	Bund	Berlin	Bund	Berlin	Bund	Berlin	Bund	Berlin	Bund	Berlin	Bund
„echte Betriebe“	1 633	19 508	1 642	12 673	1 478	10 273	1 622	9 970	1 620	10 445	1 977	15 713
Kleingewerbe	12 683	80 591	9 549	34 898	9 650	44 870	7 698	16 153	6 957	20 399	5 395	11 722
darunter Vollerwerb	7 336	-8 924	3 884	-48 742	4 120	-38 082	2 507	-61 726	1 780	-53 971	266	-56 201
darunter Nebenerwerb	5 347	89 515	5 665	83 640	5 530	82 952	5 191	77 879	5 177	74 370	5 129	67 923
Saldo insgesamt	14 316	100 099	11 191	47 571	11 128	55 143	9 320	26 123	8 577	30 844	7 372	27 435

Quellen: Gewerbeanzeigenstatistik; eigene Berechnungen.

© DIW Berlin 2017

ner oder Einwohnerinnen liegt, während Personen ohne deutsche Staatsangehörigkeit deutlich häufiger gründen, im Bundesdurchschnitt kommen sie auf knapp 16 Gründungen.⁵ Die Unterschiede zwischen den Städten sind jedoch groß: Während Berlin mit über 30 Gründungen pro 1 000 ausländische Personen einen sehr hohen Wert aufweist (Tabelle 3 und Abbildung 3), markiert Stuttgart mit knapp zehn Gründungen pro 1 000 ausländischen Personen die untere Grenze.

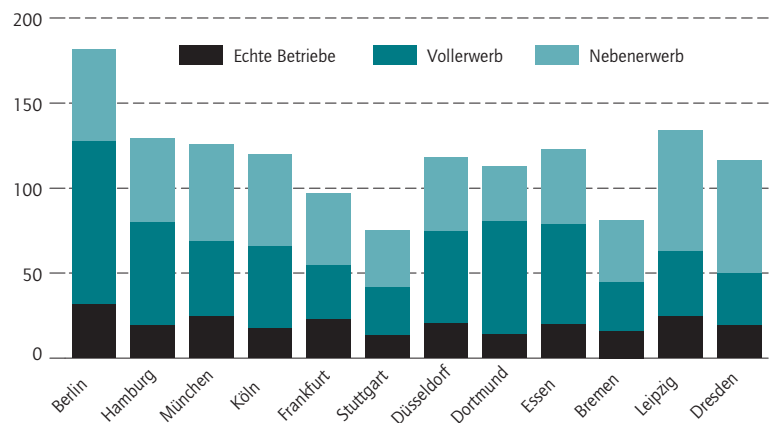
Dagegen sind Unterschiede bei der Gründungsneigung der Deutschen deutlich weniger ausgeprägt: Berlin liegt mit einem Wert von knapp sieben Gründungen pro 1 000 Deutsche im Mittelfeld der betrachteten großen Städte; hier sind Düsseldorf, Köln, Frankfurt und Leipzig vorne. Eine sogar im Vergleich zum Bund unterdurchschnittli-

che Gründungsneigung von Bürgerinnen und Bürgern mit deutscher Staatsangehörigkeit findet sich dagegen in Stuttgart und Dortmund.

Abbildung 2

Gründungsneigung in den zwölf größten deutschen Städten

Insgesamt pro 10 000 Erwerbsaktive, 2015



Quellen: Forschungsdatenzentrum der amtlichen Statistik der Länder; Bundesagentur für Arbeit, Arbeitskreis Erwerbstätigenrechnung; eigene Berechnungen.

© DIW Berlin 2017

Berlin ist deutlich gründungsfreudiger als andere Großstädte.

Kasten 2

Messung der Gründungsneigung

Um das Gründungsgeschehen zwischen verschiedenen Ländern, Regionen und Städten zu vergleichen, wird in der Regel nicht auf die absolute Anzahl an Betriebsgründungen abgestellt. Stattdessen wird die so genannte Gründungsneigung herangezogen, die unterschiedliche Größenverhältnisse berücksichtigt. Diese wird üblicherweise als Anzahl der Unternehmensgründungen pro 10 000 Erwerbspersonen definiert. Als Erwerbspersonen werden alle Personen erfasst, die eine auf Erwerb gerichtete Tätigkeit ausüben (das heißt die Erwerbstätigen) oder suchen (das heißt die Erwerbslosen).

In diesem Bericht wird die Gründungsneigung für Berlin und den Bund entsprechend diesem Konzept berechnet. Allerdings wird für den Vergleich des Gründungsgeschehens zwischen den zwölf größten Städten Deutschlands sowie für die Analyse der Gründungsneigung nach Staatsbürgerschaft von diesem Konzept leicht abgewichen:

Die untersuchten Städte sind Zentren von Ballungsräumen und verfügen über eine wirtschaftliche Anziehungskraft, die über ihr eigentliches Stadtgebiet hinausgeht. Daher verfügen diese Städte über ein deutlich größeres Arbeitskräftepotential und damit auch über ein größeres Gründungspotential, als Erwerbsperso-

nen innerhalb der Stadtgrenzen leben. Um diesem Umstand zumindest näherungsweise Rechnung zu tragen, führen wir die Zahlen der Arbeitskräfte am Arbeitsort sowie die Arbeitslosen am Wohnort zu den *Erwerbsaktiven* als neuer Normierungsgröße zusammen. Weiterhin ist zu beachten, dass die letzten verfügbaren Werte für den Städtevergleich sich auf das Jahr 2015 beziehen. Daher weichen die im Städtevergleich genannten Werte von den Werten für Berlin im Jahr 2016 leicht ab.

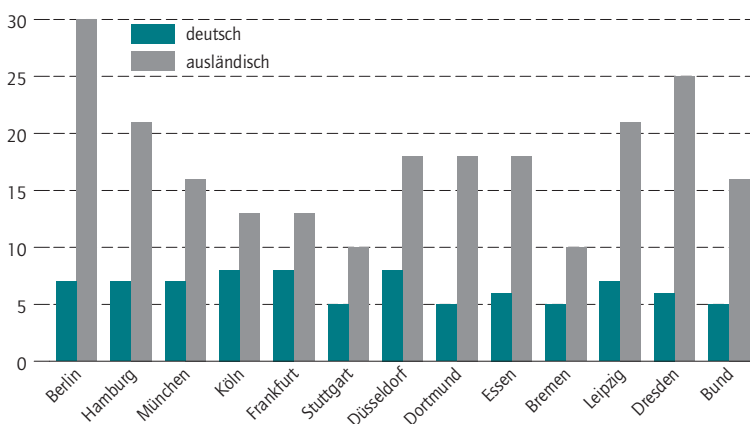
Bei der Analyse der Gründungsneigung von Personen mit verschiedenen Staatsbürgerschaften wird ebenfalls nicht die Zahl der Erwerbspersonen als Bezugsgröße verwendet, weil Daten zur Erwerbstätigkeit nach Staatsangehörigkeit in dieser Gliederungstiefe nicht zur Verfügung stehen. Stattdessen wird die Anzahl der Einwohner und Einwohnerinnen mit entsprechender Nationalität verwendet.

Die Wahl der verschiedenen Bezugsgrößen ist – in den praktischen Grenzen der Datenverfügbarkeit – davon geleitet, die optimale Datengrundlage für die jeweils spezifische Fragestellung zu nutzen, auch wenn dies den Vergleich zwischen den verschiedenen Analyseteilen etwas erschwert.

Abbildung 3

Gründungsneigung nach Nationalität

Pro 1 000 Einwohner, 2015



Quellen: Gewerbeanzeigenstatistik Forschungszentrum der Länder; Ausländerstatistik Statistisches Bundesamt; eigene Berechnungen.

© DIW Berlin 2017

Ohne die ausländischen Gründungen läge Berlin nur im Mittelfeld.

Während in der Gesamtschau Stuttgart somit das Schlusslicht der zwölf Großstädte bildet (da sowohl Deutsche wie auch Nichtdeutsche dort die niedrigste Gründungsneigung haben, Abbildung 3), ist die hohe Gründungsneigung in Berlin ein in erster Linie durch die ausländischen Einwohner und Einwohnerinnen getragener Effekt. Insgesamt wurden im Jahr 2015 über 17 000 Betriebe von Personen ohne deutsche Staatsbürgerschaft gegründet; dies entspricht einem Anteil von 47 Prozent. Die Mehrheit dieser Gründungen stammt von Bürgerinnen und Bürgern aus Polen mit über 6 000 neuen Betrieben, gefolgt von Rumänien und Bulgarien (Tabelle 3).

Innovative Gründungen: Berlin mit München, Köln und Hamburg in der Spitzengruppe

Für das zukünftige Wachstum und die Wettbewerbsfähigkeit eines Landes oder einer Region sind Gründungen in den technologischen, innovativen und kreativen Branchen von besonderer Bedeutung. Betrachtet man – auf Basis der für die Expertenkommission Forschung und Innovation (EFI) geführten Liste forschungsintensiver Industrien und wissensintensiver Dienstleistungen – Gründungen mit erhöhten Innovationspotenzial

Tabelle 4

Gewerbeanmeldungen im Jahr 2015

Gründungen insgesamt nach Nationalität

	Deutsche	Ausländer	darunter:					
			Polen	Rumänen	Türken	Bulgaren	Italiener	Kroaten
Berlin	19 435	17 247	6 058	2 027	1 391	1 465	564	292
Hamburg	10 596	5 805	2 056	707	447	610	77	194
München	7 266	6 403	1 219	724	414	291	221	852
Köln	6 487	2 728	861	170	502	111	166	44
Frankfurt	4 025	2 658	328	219	348	114	165	216
Stuttgart	2 425	1 468	162	138	181	52	80	238
Düsseldorf	3 675	2 648	526	313	221	130	96	77
Dortmund	2 241	1 753	616	197	257	119	38	33
Essen	2 920	1 527	447	139	244	55	44	45
Bremen	2 184	899	200	62	168	87	13	20
Leipzig	3 845	835	38	111	44	28	40	29
Dresden	3 142	837	71	190	17	77	18	88
Bund	390 041	143 058	33 622	20 408	16 114	9 442	6 389	6 337

Quelle: Gewerbeanzeigenstatistik, Forschungsdatenzentrum der Länder.

© DIW Berlin 2017

Tabelle 5

Anzahl der Betriebsgründungen mit Innovationspotential

Gründungen im Bereich der forschungsintensiven Industrien und wissensintensiven Dienstleistungen

	forschungsintensive Industrie		wissensintensive Dienstleistungen		Summe (forschungsintensive Industrie + wissensintensive Dienstleistungen)	
	2014	2015	2014	2015	2014	2015
Berlin	228	210	6 267	6 013	6 495	6 223
Hamburg	26	34	3 933	3 796	3 959	3 830
München	54	38	3 873	3 774	3 927	3 812
Köln	20	14	2 690	2 669	2 710	2 683
Frankfurt	4	6	1 840	1 877	1 844	1 883
Stuttgart	20	26	896	935	916	961
Düsseldorf	16	12	1 220	1 240	1 236	1 252
Dortmund	6	6	518	509	524	515
Essen	17	7	956	745	973	752
Bremen	66	31	575	485	641	516
Leipzig	8	9	972	953	980	962
Dresden	8	8	761	719	769	727
Bund	2 714	2 730	101 170	99 354	103 884	102 084

Quelle: Gewerbeanzeigenstatistik, Forschungsdatenzentrum der Länder.

© DIW Berlin 2017

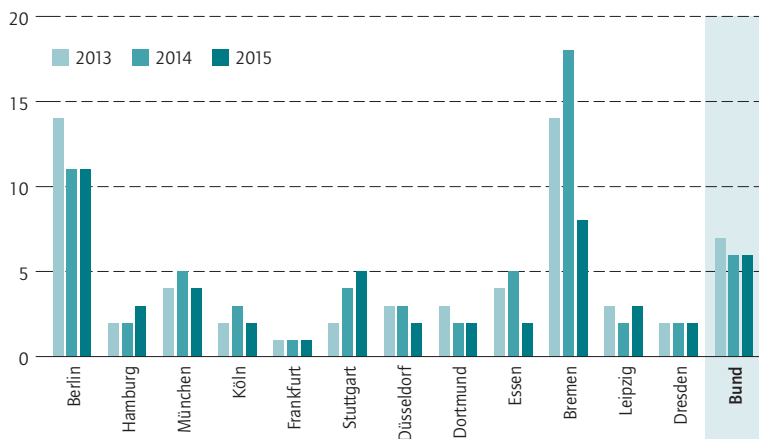
(Kasten 1), so zeigt sich das Gründungsgeschehen im gesamten Bundesgebiet relativ stabil. Zwar sanken die Gründungen von wissensintensiven Dienstleistungen im Jahr 2015 im Vergleich zu 2014 leicht von knapp über auf knapp unter 100 000 Gründungen, während die Zahl der Gründungen in den forschungsintensiven Industrien im Vergleich zwischen den beiden Jahren nahezu unverändert blieben. Im Städtevergleich weisen in diesem Bereich München und Köln die höchste Grün-

dungsneigung auf, Berlin liegt auf dem dritten Platz und in etwa gleichauf mit Hamburg sowie knapp vor Leipzig. Zum Vergleich: Im Jahr 2015 wurden in Berlin 6 223 Gründungen mit Innovationspotential erfasst (Tabelle 4), also knapp 17 Prozent aller Gründungen oder rund 340 Gründungen auf 100 000 erwerbsaktive Personen (Abbildungen 4 und 5), in München insgesamt 3 812 Gründungen, immerhin 28 Prozent aller Gründungen oder rund 378 Gründungen auf 100 000 erwerbsaktive Personen.

Abbildung 4

Forschungsintensive Industrien in den zwölf Vergleichsstädten

Gründungen je 100 000 erwerbsaktive Personen



Quellen: Forschungsdatenzentrum der Länder; Bundesagentur für Arbeit; Arbeitskreis Erwerbstätigenrechnung des Bundes und der Länder; eigene Berechnungen.

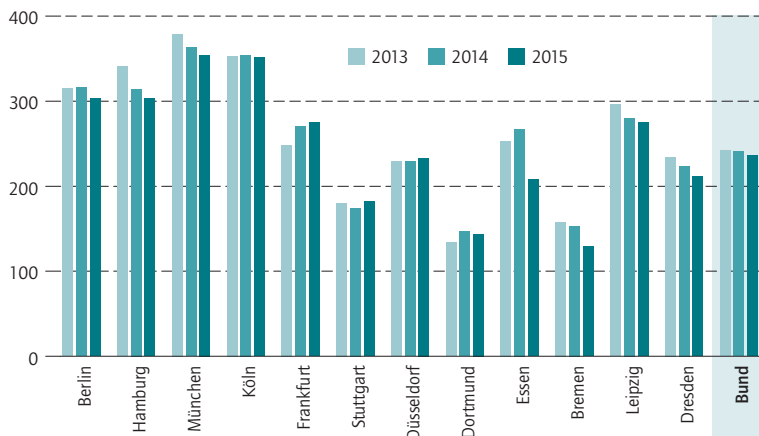
© DIW Berlin 2017

Berlin und Bremen liegen als einzige der Vergleichsstädte über dem Bundesdurchschnitt.

Abbildung 5

Wissensintensive Dienstleistungen in den zwölf Vergleichsstädten

Gründungen je 100 000 erwerbsaktive Personen



Quellen: Forschungsdatenzentrum der Länder; Bundesagentur für Arbeit; Arbeitskreis Erwerbstätigenrechnung des Bundes und der Länder; eigene Berechnungen.

© DIW Berlin 2017

München, Hamburg, Berlin und Köln führen den Städtevergleich an.

In allen Städten dominieren die Gründungen im Bereich der wissensintensiven Dienstleistungen. Das trifft auch auf Berlin zu mit 6 013 Gründungen gegenüber 210 neuen Betrieben bei den forschungsintensiven Industrien. Die Anzahl von forschungsintensiven Gründungen im verarbeitenden Gewerbe ist – trotz ihrer zuletzt rückläufigen Tendenz in den letzten drei Jahren – in Berlin

dennoch auffallend hoch. Das dürfte unter anderem auf die Nähe zu Hochschulen (sogenannte Clustereffekte) zurückgehen, etwa in verschiedenen „Hotspots“⁶ wie Berlin-Adlershof, wo sich besonders viele industrielle High-Tech-Gründungen ansiedeln.⁷

Insgesamt fällt aber auch auf, dass in nahezu allen Städten (mit Ausnahme von Frankfurt am Main) die Bereitschaft für innovative Gründungen stagniert und vereinzelt sogar deutlich rückläufig ist.

Fazit

Während in Deutschland das Gründungsgeschehen im sechsten Jahr rückläufig blieb, wurde in Berlin der Abwärtstrend im Jahr 2016 gestoppt. Berlin bleibt auch weiterhin Gründungshauptstadt, getragen von seinen ausländischen Bürgerinnen und Bürgern, die in dieser Stadt im Gegensatz zu den Deutschen eine besonders hohe Gründungsaktivität aufweisen. Dass ausländische Mitbürgerinnen und Mitbürger häufiger Betriebe gründen, gilt auch für die gesamte Bundesrepublik und für alle betrachteten Großstädte. In Berlin ist dieses Phänomen jedoch besonders ausgeprägt. Bei den Gründungen in den forschungsintensiven Industrien und wissensintensiven Dienstleistungen nimmt Berlin – bezogen auf seine Größe – Platz drei ein, hinter München und Köln.

Fokussiert man auf die jüngste Entwicklung bei den Gründungen mit Wachstumspotential, das sind zum einen Gründungen *echter Betriebe* und zum anderen Gründungen von Betrieben in Sektoren mit Innovationspotential, findet sich ein eher uneinheitliches Gründungsgeschehen. Zum einen steigt die Zahl der *echten Gründungen* jüngst wieder an. Zum anderen stagniert oder sinkt im Bund und in vielen Großstädten, darunter auch in Berlin, die Bereitschaft, sich mit innovativen und damit wohl auch riskanteren Gründungen an den Markt zu wagen. Zwar ist der Rückgang noch nicht besorgniserregend. Dennoch ist es an der Politik, sich vermehrt darüber Gedanken zu machen, Gründungen mit Innovationspotential durch bessere institutionelle und finanzielle Rahmenbedingungen zu unterstützen. Eine wichtige Maßnahme in diese Richtung wäre die Vereinfachung der bürokratischen Prozeduren zur Eröffnung eines Betriebs, die nach wie vor viele zeitintensive Schritte umfassen. Entsprechend befindet sich Deutschland hier seit vielen Jahren im internationalen Vergleich weit hinten und nimmt laut Weltbank

⁶ Julian Kahn (2017): Innovationserhebung Berlin 2016, Technologiestiftung, Berlin (online abrufbar).

⁷ Siehe DIW Econ (2015): Hightech für Berlin, die regionalwirtschaftliche Bedeutung der Technologiezentren in Adlershof, Berlin.

index *Ease of Doing Business* derzeit den Platz 114 ein.⁸ So besteht die Anmeldung einer GmbH in Deutschland aus neun Prozeduren, die 10,5 Tage dauern, in Neuseeland ist es dagegen nur eine einzige Prozedur, die in einem Tag zu erledigen ist.

8 Siehe World Bank (2017): *Ease of Doing Business* (online verfügbar). Dort werden auch die konkreten Prozeduren für die Eröffnung etwa einer GmbH in einzelnen Schritten inklusive Kosten und zeitlicher Beanspruchung beschrieben.

Alexander S. Kritik ist Forschungsdirektor im Vorstandsbereich des DIW Berlin und Professor für Volkswirtschaftslehre an der Universität Potsdam | akritikos@diw.de

JEL: D22, L26,

Keywords: Startups, Entrepreneurship, Deutschland, Berlin, Hamburg, Munich

Es sei abschließend aber auch betont, dass die Gründung eines – unter Umständen innovativen – Betriebs erst der Anfang für eine zukünftig positive Entwicklung in einer Region ist. Wirkung entfalten solche Gründungen erst dann, wenn sie sich zu schnell wachsenden Unternehmen weiter entwickeln. Ob und wo das in Deutschland der Fall ist, kann in der vorgelegten Untersuchung nicht geklärt werden. Hierfür ist weitere Forschungsarbeit nötig, für die zunächst geeignete Daten identifiziert werden müssen.

Anselm Mattes ist Manager der DIW Econ GmbH | amattes@diw-econ.de



Dr. Claus Michelsen, Wissenschaftlicher Mitarbeiter in den Abteilungen Konjunkturpolitik und Klimapolitik am DIW Berlin
Der Beitrag gibt die Meinung des Autors wieder.

Die Mietpreisbremse ist keine Investitionsbremse

Die Mietpreisbremse ist kaum eingeführt, da schaffen die ersten beiden Bundesländer sie wieder ab. Die Erfahrungen zeigten, so heißt es im Koalitionsvertrag von CDU, FDP und Grünen in Schleswig-Holstein, „dass der angestrebte Effekt, nämlich die Sicherung stabiler Mietpreise, nicht eingetreten ist“. Die künftige schwarz-gelbe Koalition in Nordrhein-Westfalen meint sogar, die Mietpreisbremse hätte in ihrer Wirkungslosigkeit die privaten Investitionen in den Wohnungsbau gebremst.

Es ist erstaunlich, mit welcher Sicherheit derartige Urteile gefällt werden. Zwar gibt es mittlerweile erste Studien zur Wirkung der Mietpreisbremse. Diese besagen einstimmig, dass ihr Einfluss auf Mieten gering sei. Zunächst sagt das vor allem etwas über die wirkungslose Ausgestaltung aus statt über die Maßnahme an sich. Nichts ist außerdem bekannt über die Auswirkungen auf die private Investitionstätigkeit. Daten sind in diesem Zusammenhang schlicht Mangelware.

Es wäre einfach darauf hinzuweisen, wie widersprüchlich der NRW-Koalitionsvertrag ist: Die Mietpreisbremse soll wirkungslos gewesen sein, gleichzeitig aber die Investitionen in den Wohnungsbau gebremst haben. Wenn überhaupt dürfte sie jedoch eine eher positive Wirkung auf die Investitionstätigkeit entfalten. So beschränkt die Mietpreisbremse die Mieten für Gebäude mit Baufertigstellung bis Oktober 2014. Alle danach erstmalig vermieteten Neubauten und umfassend modernisierten Wohnungen sind ausgenommen.

Der ökonomischen Theorie nach wird die Nachfrage nach den regulierungsbedingt „günstigen“ Mieten im Bestandssegment künstlich angefacht. Mehr Menschen wollen innerhalb der Städte Wohnraum finden, als dies ohne Mietpreisbremse der Fall gewesen wäre. Die zusätzliche Nachfrage treibt die Mieten im Neubausegment zusätzlich an, die Rentabilität der Investitionen dort steigt und dürfte weitere Bauherren auf den Markt locken. Nicht umsonst sind es die Baugenehmigungen von Mehrfamilienhäusern, die derzeit die Baukonjunktur stützen. Strittig ist, was mit dem Bestand regulierter Wohnungen passiert – hier wird befürchtet, dass weniger in deren Instand-

haltung und damit deren Qualität investiert wird. Langfristig müssten sogar Gebäude abgerissen werden. Diese so aus dem regulierten Markt gedrängten Wohnungen würden der Theorie zufolge aber wieder durch unregulierte Neubauten ersetzt. Kurzum: Es ist aus theoretischer Sicht unwahrscheinlich, dass die Investitionstätigkeit reduziert würde – wahrscheinlicher ist, dass gerade wegen der Mietpreisbremse mehr Wohnraum neu gebaut und so das Angebot langfristig sogar über die Menge eines unregulierten Marktes hinaus steigen würde.

In der Tat löst die Mietpreisbremse nicht das Wohnungsproblem für Geringverdienerinnen und Geringverdiener. Diese dürften vor allem damit konfrontiert sein, dass Wohlhabendere den Zuschlag für eine Wohnung erhalten. Dieses Problem würde aber weder das Wohngeld, noch eine Preisgrenze lösen. Marktzugang könnten nur Wohnberechtigungsscheine und ein entsprechender Ausbau des sozialen Wohnungsbaus sicherstellen.

Die Zahl der Wohnungen mit Sozialbindung ist in den letzten Jahren kontinuierlich gesunken – eine Steigerung der Bautätigkeit in diesem Segment wäre sinnvoll, um Geringverdienenden den Marktzugang zu erhalten. Die neuen Landesregierungen täten gut daran, die schon angestrebten Bemühungen für mehr kommunalen und sozialen Wohnungsbau, Nachverdichtung und Baulanderschließung zu unterstützen – die bundespolitischen Weichenstellungen sind in diesem Zusammenhang bereits erfolgt.

Zwar ist vor allem auf der Angebotsseite das Problem der derzeitigen Wohnungspolitik zu suchen. Dennoch kann die Mietpreisbremse temporär für Entlastung sorgen und die Spitze der Wohnungsmarktknappheit untertunneln. Notwendig wäre es, die offensichtlichen Schwächen der Regulierung zu beheben. Diese sind weitgehend bekannt und liegen vor allem darin, dass Verstöße kaum sanktioniert und die zulässigen Miethöhen nicht transparent nachvollzogen werden können. Das ließe sich leicht ändern.

Dieser Kommentar ist in einer längeren Version am 26. Juni 2017 bei ZEIT Online erschienen.