



Bundesinstitut  
für Bevölkerungsforschung

DIW BERLIN

Stiftung  
Ravensburger  
Verlag

# Anhang

Vertiefende Analysen und Materialien



## Oma und Opa gefragt?

Veränderungen in der Enkelbetreuung -  
Wohlbefinden von Eltern - Wohlergehen von Kindern

BIB.BEVÖLKERUNGS.STUDIEN 1 | 2022

## **Oma und Opa gefragt?**

Veränderungen in der Enkelbetreuung -  
Wohlbefinden von Eltern - Wohlergehen von Kindern

Bundesinstitut für Bevölkerungsforschung  
Friedrich-Ebert-Allee 4  
65185 Wiesbaden  
Telefon: 0611-752235  
E-Mail: [post@bib.bund.de](mailto:post@bib.bund.de)  
De-Mail: [kontakt@bib-bund.de-mail.de](mailto:kontakt@bib-bund.de-mail.de)  
Internet: [www.bib.bund.de](http://www.bib.bund.de)

### **Autorinnen und Autoren:**

Mara Barschkett; Ludovica Gambaro, Ph.D.; Clara Schäper,  
Univ.-Prof. Dr. C. Katharina Spieß; Elena Ziege. Dank an Josefine Koebe

### **Layout, Satz und Redigierung:**

Christian Fiedler, Sybille Steinmetz, Ulrike Brunner, Britta Müller

### **Titelbild:**

Adobe Stock | NDABCREATIVITY

### **Fotos:**

Anja Koehler (Seite 6)

Vervielfältigung und Verbreitung, auch auszugsweise,  
mit Quellenangabe gestattet

© Bundesinstitut für Bevölkerungsforschung (BiB) 2022

urn:nbn:de:bib-var-2022-034

DOI: <https://doi.org/10.12765/bro-2022-01>

ISSN 2702-7651 (Print)

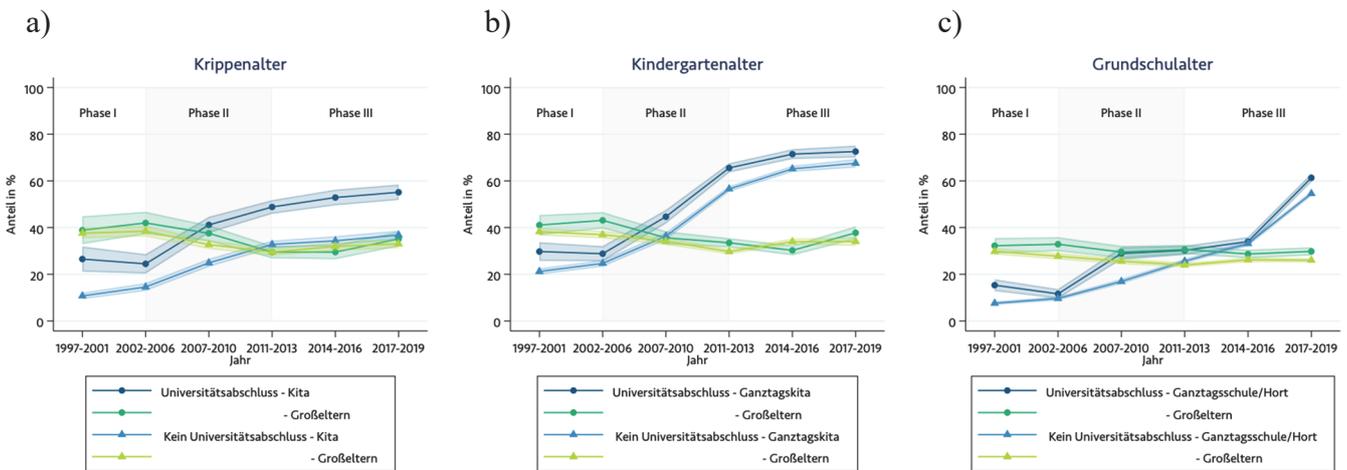
ISSN 2702-766X (Online)

# A. Zusätzliche Analysen

## A1. Differenziertere Analysen

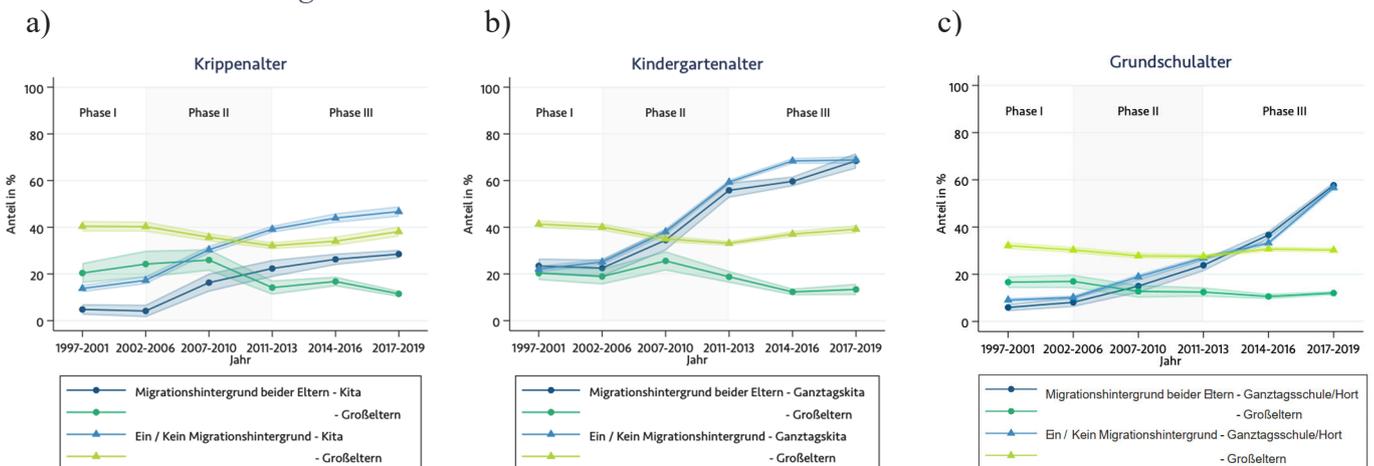
Im Folgenden zeigen wir zusätzliche Analysen der Entwicklung der Großelternbetreuung und der Kindertagesbetreuung, sowie der Entwicklung der Großelternbetreuung in Kombination mit anderen Betreuungsformen. Abbildungen A1 und A2 werden jeweils in Kapitel 4.1.4 beschrieben. Abbildungen A3-A5 ergänzen die Analysen in Kapitel 4.1.5.

Abbildung A 1: Entwicklung der Kita-Nutzung und der Großelternbetreuung nach Bildungshintergrund der Mutter und Alter des Kindes



Anmerkung: Die helleren Balken, die die Linien umrahmen zeigen 95 Prozent-Konfidenzintervalle. In dunkelgrün (dunkelblau) sehen wir den Anteil der Kinder in Großelternbetreuung (formeller Betreuung), deren Mutter einen Universitätsabschluss hat. In hellgrün (hellblau) sehen wir den Anteil der Kinder in Großelternbetreuung (formeller Betreuung), deren Mutter keinen Universitätsabschluss hat. Quelle: Eigene Berechnungen auf Basis von SOEP (1997-2019).

Abbildung A 2: Entwicklung der Kita-Nutzung und der Großelternbetreuung nach Migrationshintergrund der Eltern und Alter des Kindes



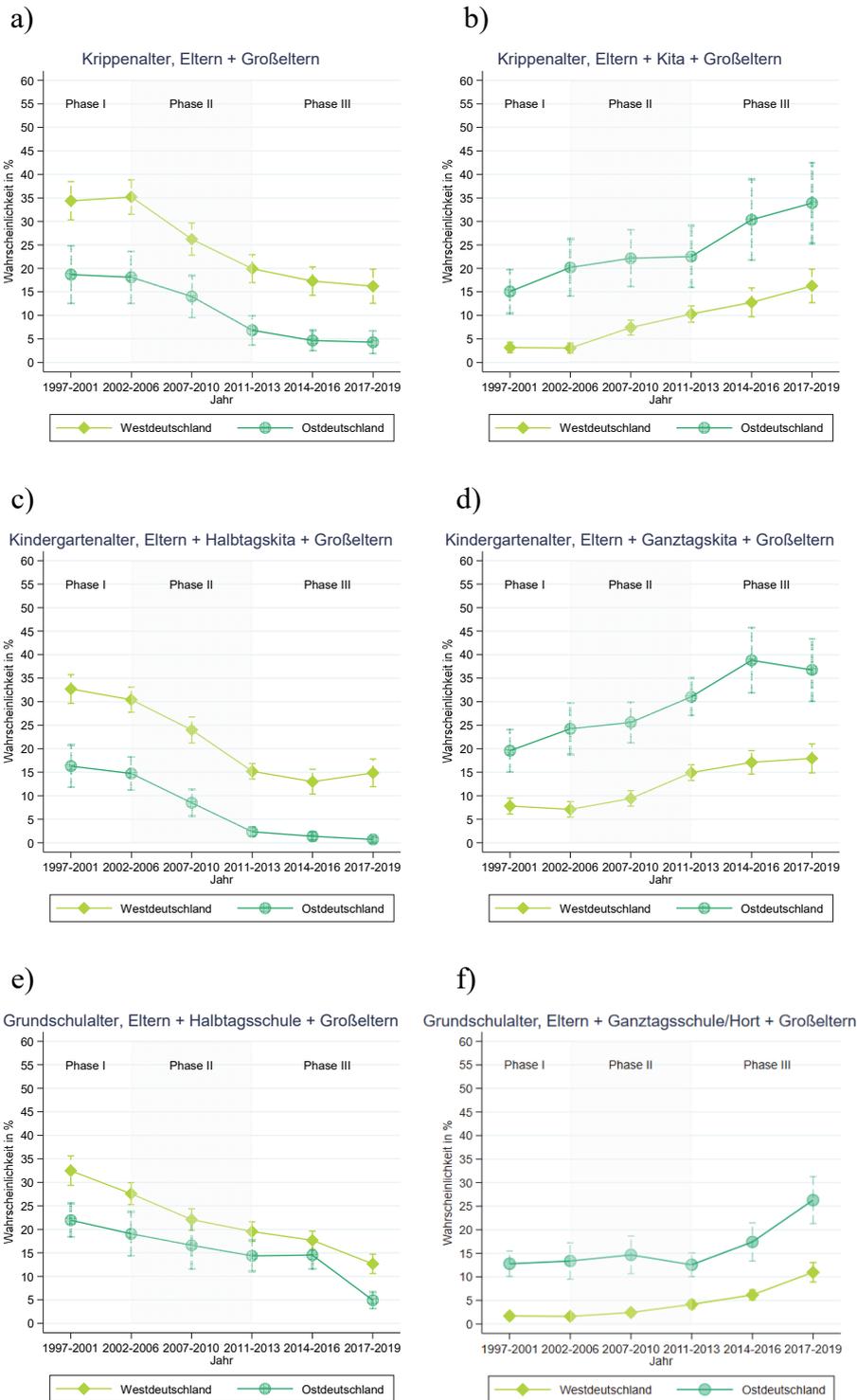
Anmerkung: Die helleren Balken, die die Linien umrahmen zeigen 95 Prozent-Konfidenzintervalle. In dunkelgrün (dunkelblau) sehen wir den Anteil der Kinder in Großelternbetreuung (formeller Betreuung), deren Eltern beide einen direkten Migrationshintergrund haben. In hellgrün (hellblau) sehen wir den Anteil der Kinder in Großelternbetreuung (formeller Betreuung), die ein oder kein Elternteil mit Migrationshintergrund haben. Quelle: Eigene Berechnungen auf Basis von SOEP (1997-2019).

**Differenziertere Analyse der Entwicklung der Großelternbetreuung in Kombination mit anderen Betreuungsformen unter Berücksichtigung weiterer Faktoren.** Zur differenzierten Untersuchung der Trends in den beschriebenen Betreuungskombinationen wurden – in Analogie zum Vorgehen in Kapitel 4.1.4, Logit-Modelle geschätzt, die es erlauben weitere Faktoren, welche Betreuungskombinationen beeinflussen, zu berücksichtigen. So wurde das Alter des betreuten Kindes, die Anzahl der Kinder unter 16 Jahren im Haushalt, sowie das Alter der Mutter in den Modellen mit aufgenommen. Außerdem wurden in die Modelle jeweils Interaktionsterme zwischen regionaler Zugehörigkeit und den Zeitintervallen eingefügt. Fernerhin können wir mit den nachfolgenden Analysen auch in diesem Kontext untersuchen, inwiefern es sich bei den bereits beschriebenen Veränderungen um solche handelt, die auch im statistischen Sinne signifikante Veränderungen sind.

Abbildung A3 stellt die geschätzten Nutzungswahrscheinlichkeiten von Betreuungskombinationen dar. Wir vergleichen dabei, ob sich für diejenigen, die die Großeltern als Betreuungspersonen hinzuziehen, Betreuungskombinationen über die Zeit verändert haben. Konkret, schätzen wir die Wahrscheinlichkeiten, dass die Großelternbetreuung nur mit der Betreuung durch Eltern (bei Kindern im Krippenalter), die Großeltern- und Elternbetreuung mit einer Kita- Halbtagsbetreuung oder Halbtagschule (bei Kindern im Kindergarten- bzw. Grundschulalter) kombiniert wird (linke Seite der Abbildungen A3-A5). Dies vergleichen wir mit den geschätzten Wahrscheinlichkeiten, dass die Großeltern, wenn sie neben den Eltern an einer regelmäßigen Betreuung beteiligt sind, diese bei Kindern im Krippenalter mit einer Kita-Betreuung kombinieren und bei älteren Kindern, diese mit einer ganztägigen Kita-Betreuung oder dem schulischen Ganztags vereinbaren (rechte Seite der Abbildungen A3-A5).

Abbildung A3a zeigt, dass **Kinder im Krippenalter** in *Phase III* mit einer signifikant geringeren Wahrscheinlichkeit ausschließlich von den Eltern und Großeltern betreut werden als in *Phase I*, und zwar in West-, sowie in Ostdeutschland. Abbildung A3b zeigt, dass die Wahrscheinlichkeit, dass ein Kind zusätzlich zu den Eltern regelmäßig in einer Kombination aus Kita und Großelternbetreuung betreut wird sowohl in West- als auch in Ostdeutschland im statistischen Sinne signifikant und in einem ähnlichen Umfang angestiegen ist. Betrachten wir Abbildungen A3c bis f, zeigt sich, dass diese Trends sich auch für die **beiden älteren Altersgruppen fortsetzen**. Wir können also festhalten, dass die beobachteten Verschiebungen hin zu einer vermehrten Betreuung von Kindern durch Großeltern in Kombinationen mit einer ganztägigen Kita-Betreuung und dem schulischen Ganztags im statistischen Sinne zunehmen.

Abbildung A 3: Entwicklung der Nutzungswahrscheinlichkeiten von Betreuungskombinationen, nur Kinder, die von den Großeltern betreut werden, nach Region und Alter des Kindes



Anmerkung: Logistische Regression, gewichtet. Die gestrichelten Linien zeigen die 95 Prozent-Konfidenzintervalle.  
Quelle: Eigene Berechnungen auf Basis von SOEP (1997-2019).

Eine weitere Frage, die wir in diesem Kapitel auf der Basis von Schätzmodellen beantworten wollen, ist, ob sich die dargestellten Veränderungen in den Nutzungswahrscheinlichkeiten nach dem Bildungshintergrund der Mutter und dem Migrationshintergrund beider Eltern unterscheiden. Kombinieren Familien, bei denen die Mutter einen höheren Bildungshintergrund hat und sie damit auch eher erwerbstätig ist bzw. ein höheres Erwerbsvolumen aufweist, die großelterliche Kinderbetreuung signifikant häufiger mit Kitas oder dem schulischen Ganztags? Nutzen andere Familien die großelterliche Kinderbetreuung weiterhin, ohne sie mit einer Kita-Betreuung oder dem schulischen Ganztags zu kombinieren?

**Veränderungen nach Bildungshintergrund der Mutter.** Abbildung A4 zeigt die Unterschiede in den Nutzungswahrscheinlichkeiten über die Zeit jeweils für die drei Altersgruppen und nach dem Bildungshintergrund der Mutter.<sup>1</sup>

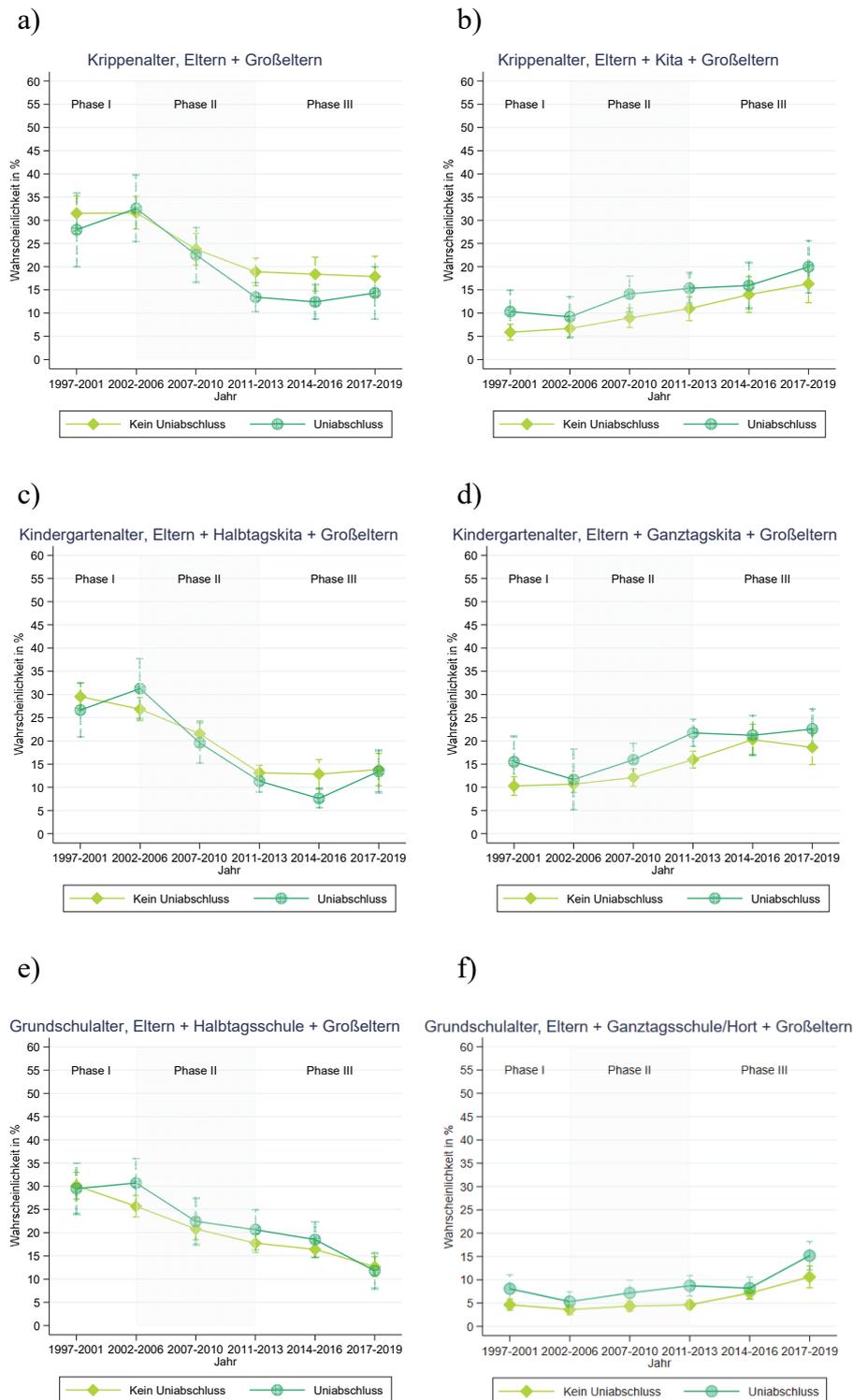
Betrachten wir die Entwicklung der Wahrscheinlichkeiten in Abbildung A4 über die Zeit, bestätigen diese das Gesamtbild: Eine Zunahme der großelterlichen Kinderbetreuung zusammen mit einer Kita-Betreuung und dem schulischen Ganztags können wir für alle Altersgruppen festmachen, und zwar unabhängig vom Bildungshintergrund der Mutter. Die Verläufe je nachdem, ob die Mutter einen Universitätsabschluss hat oder nicht bleiben im gesamten Zeitverlauf und für alle Altersgruppen parallel und sind im statistischen Sinne nicht signifikant unterschiedlich.

**Veränderungen nach Migrationshintergrund der Eltern.** Die Nutzungswahrscheinlichkeiten differenziert nach dem Migrationshintergrund der Eltern werden in Abbildung A5 dargestellt. Hier zeigen sich durchaus unterschiedliche Entwicklungen. Die oben beschriebenen Entwicklungen einer zunehmenden Kombination der Großelternbetreuung mit einer Kita-Nutzung bei Kindern im Krippenalter, einer ganztägigen Kita-Nutzung bei Kindern im Kindergartenalter und einer Nutzung des Ganztags bei Grundschulkindern lässt sich im statistischen Sinne nur für Kinder beobachten, bei denen maximal ein Elternteil einen Migrationshintergrund hat. Für Kinder, deren Eltern beide einen Migrationshintergrund haben, ist diese Tendenz nicht festzumachen.

---

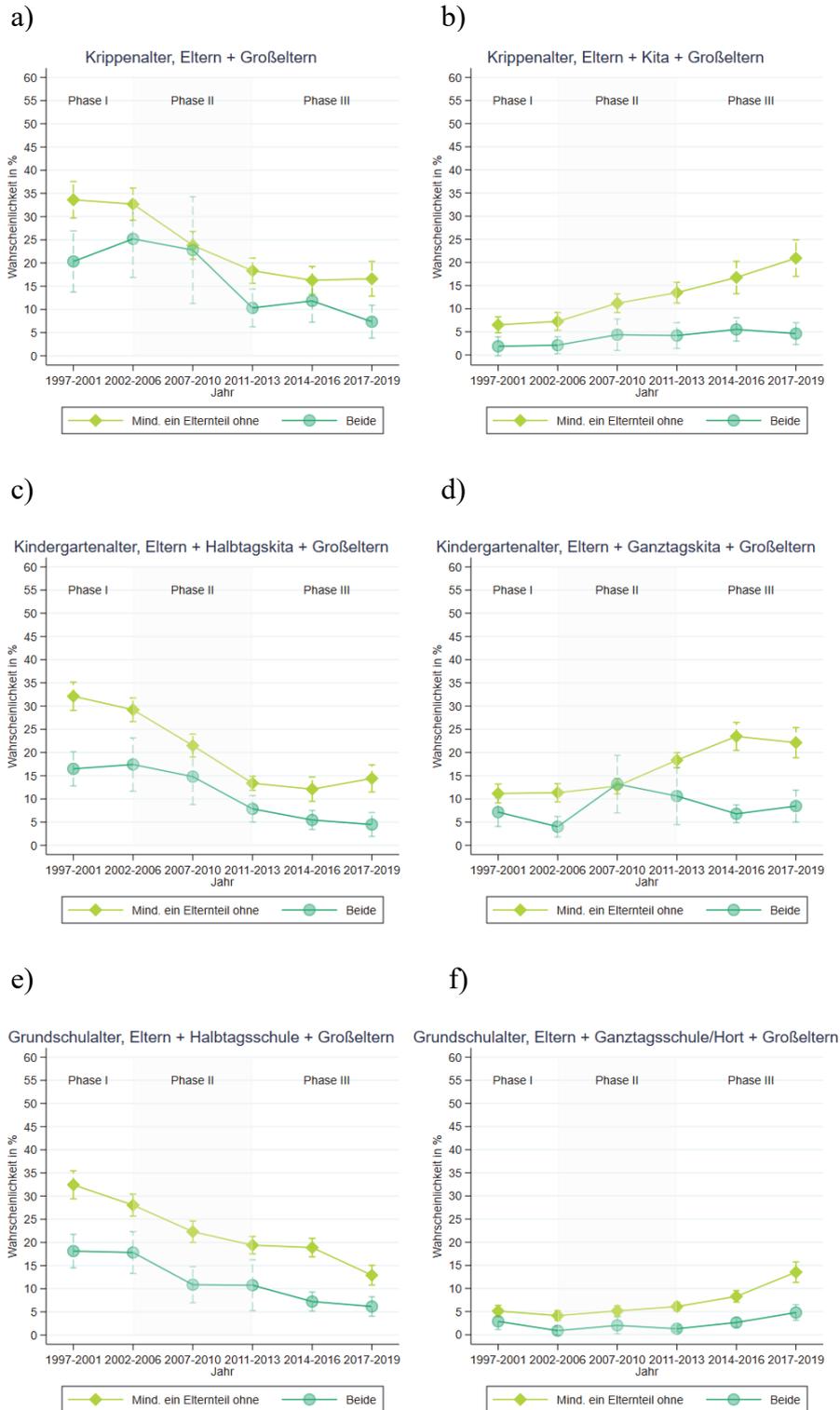
<sup>1</sup> Eine Differenzierung nach Region wird hier nicht durchgeführt, was aufgrund der Fallzahlen im Analysesample nicht möglich ist.

Abbildung A 4: Entwicklung der Nutzungswahrscheinlichkeiten von Betreuungskombinationen, nur Kinder, die von den Großeltern betreut werden, nach Bildungshintergrund der Mutter und Alter des Kindes



Anmerkung: Logistische Regression, gewichtet. Die gestrichelten Linien zeigen die 95 Prozent-Konfidenzintervalle.  
 Quelle: Eigene Berechnungen auf Basis von SOEP (1997-2019).

Abbildung A 5: Entwicklung der Nutzungswahrscheinlichkeiten von Betreuungskombinationen, nur Kinder, die von den Großeltern betreut werden, nach Migrationshintergrund der Eltern und Alter des Kindes



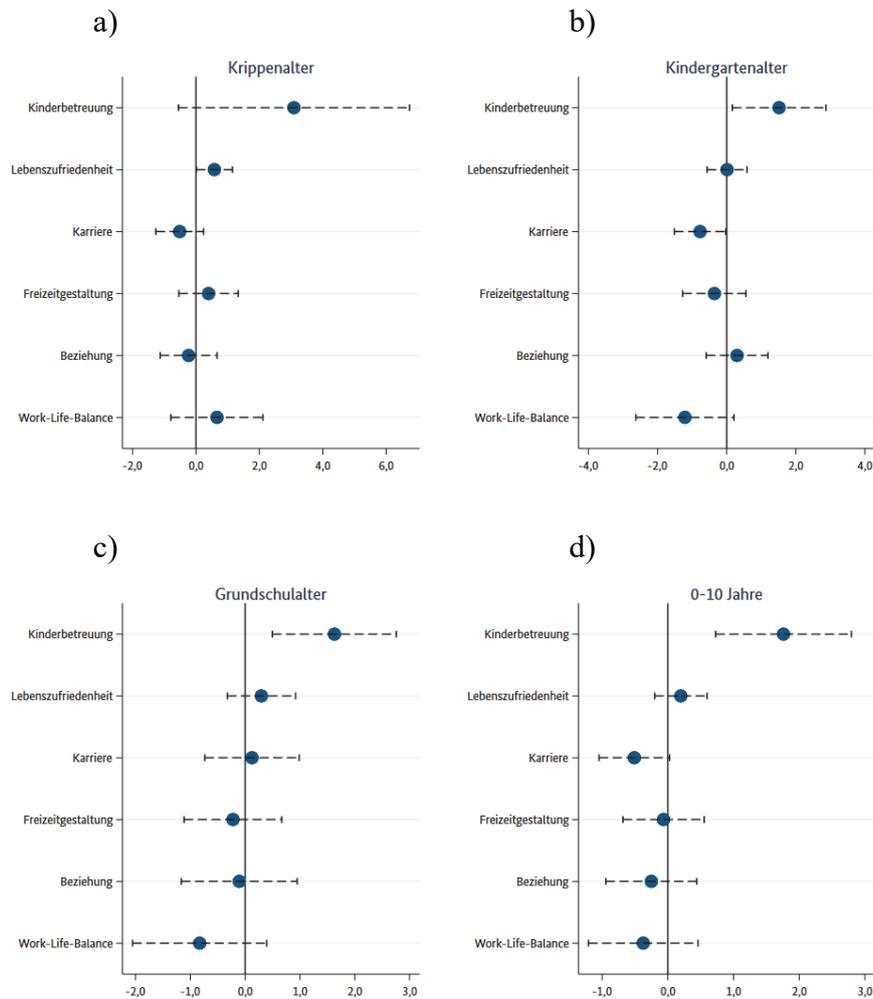
Anmerkung: Logistische Regression, gewichtet. Die gestrichelten Linien zeigen die 95 Prozent-Konfidenzintervalle.

Quelle: Eigene Berechnungen auf Basis von SOEP (1997-2019).

## A2. Differenziertere Analysen zu Kapitel 5

Im Folgenden zeigen wir zusätzliche Analysen der Effekte der Großelternbetreuung auf das väterliche Wohlbefinden. Abbildungen A6-A8 werden jeweils in Kapitel 5.2 beschrieben.

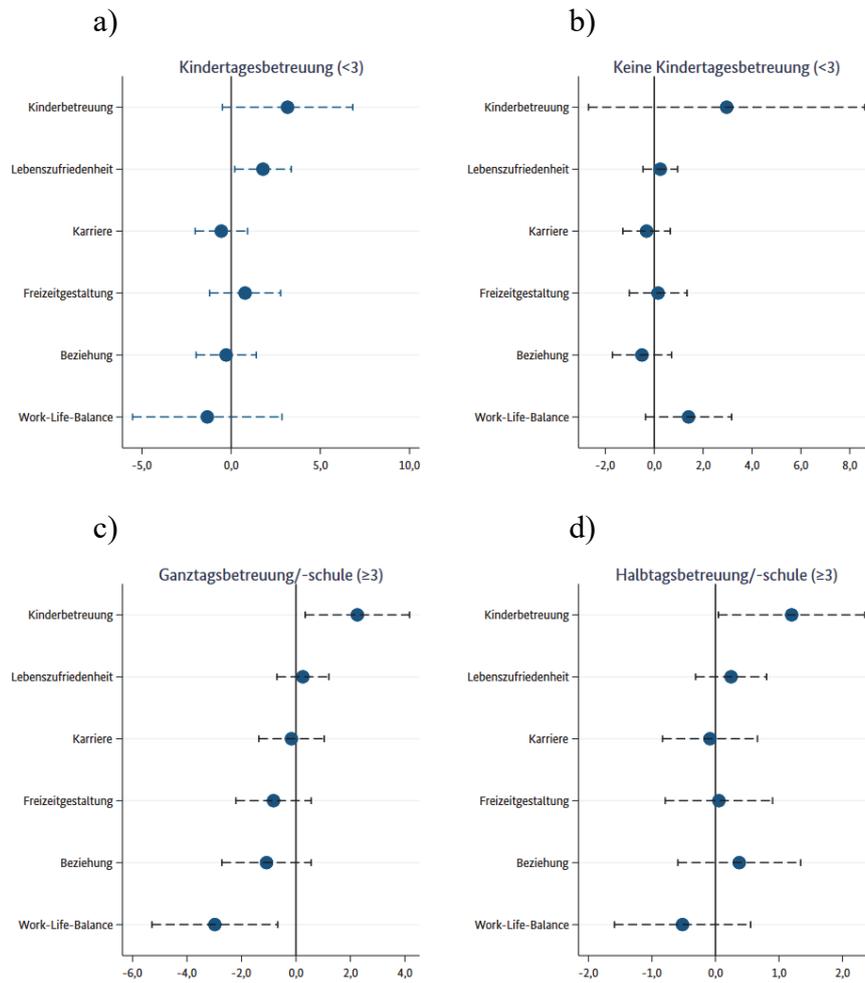
Abbildung A 6: Auswirkungen der Großelternbetreuung auf die väterliche Zufriedenheit nach Alter des Kindes



Anmerkung: Die gestrichelten Linien zeigen die 95 Prozent-Konfidenzintervalle. Die Zufriedenheit ist gemessen auf einer Skala von 1 (sehr unzufrieden) bis 11 (sehr zufrieden).

Quelle: Eigene Berechnungen auf der Basis von pairfam (2010-2020), gewichtet.

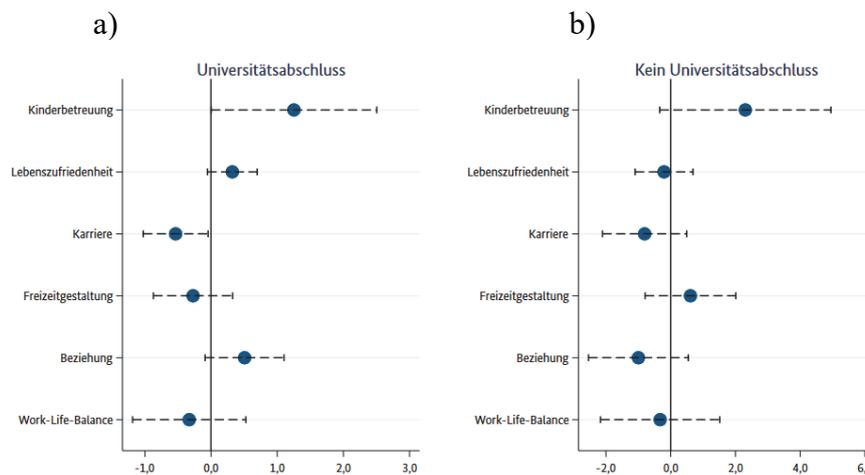
Abbildung A 7: Auswirkungen der Großelternbetreuung auf die väterliche Zufriedenheit nach der Nutzung von Kindertagesbetreuung und Kita-Ganztagsbetreuung/schulischen Ganztag



Anmerkung: Die gestrichelten Linien zeigen die 95 Prozent-Konfidenzintervalle. Die Zufriedenheit ist gemessen auf einer Skala von 1 (sehr unzufrieden) bis 11 (sehr zufrieden).

Quelle: Eigene Berechnungen auf der Basis von pairfam (2010-2020), gewichtet.

Abbildung A 8: Auswirkungen der Großelternbetreuung auf die väterliche Zufriedenheit nach Bildungshintergrund des Vaters



Anmerkung: Die gestrichelten Linien zeigen die 95 Prozent-Konfidenzintervalle. Die Zufriedenheit ist gemessen auf einer Skala von 1 (sehr unzufrieden) bis 11 (sehr zufrieden).

Quelle: Eigene Berechnungen auf der Basis von pairfam (2010-2020), gewichtet.

## **B. Differenziertere Beschreibung der Datensätze**

### **B1. Allgemeine Beschreibung aller analysierten Datensätze**

Im Folgenden beschreiben wir detaillierter die Datensätze, welche wir für unsere eigenen Analyse ausgewertet haben.

**Sozio-oekonomisches Panel (SOEP).** Das Sozio-oekonomische Panel (*SOEP*) ist eine repräsentative Haushaltsbefragung, die seit 1984 jährlich durchgeführt wird. Das Sozio-oekonomischen Panels (*SOEP*) ist eine der weltweit am längsten laufenden und größten jährlichen Wiederholungsbefragungen von Haushalten und Personen (Goebel et al. 2019), für die derzeit jährlich etwa 30.000 Menschen in knapp 15.000 Haushalten befragt werden. Seit 2003 werden für die Geburtskohorten ab 2002 sogenannte altersspezifische Fragebögen erhoben, die auch detailliertere Angaben über die Betreuung der Kinder in unterschiedlichen Altersgruppen enthalten (Spiess 2011). Inzwischen liegen altersspezifische Fragebögen für neugeborene Kinder, Kinder im Alter von 2-3 Jahren, von 5-6 Jahren, 7-8 Jahren, 9-10 Jahren, 11-12 Jahren, 13-14 Jahren und 16-17 Jahren vor. Diese Fragebögen werden von den Eltern ausgefüllt mit der Ausnahme der Fragebögen für Kinder im Alter von 11-12 Jahren und 13-14 Jahren, welche von den Kindern selbst ausgefüllt werden. Das *SOEP* umfasst Informationen über alle in einem Haushalt lebenden Personen. Neben den individuellen Fragebögen, die von allen Erwachsenen im Haushalt ausgefüllt werden, gibt es einen Haushaltsfragebogen, der dem Haushaltsvorstand Fragen zu allen im Haushalt lebenden Kindern stellt. Im Gegensatz zu *pairfam* werden die Großeltern selbst nur dann befragt, wenn sie im gleichen Haushalt wie ihre erwachsenen Kinder leben oder wenn die Eltern als Kind in einem *SOEP*-Haushalt gelebt haben und inzwischen einen eigenen Haushalt gegründet haben. Die Stichprobe, für die detaillierte Informationen über die Großeltern vorliegen, ist also eine kleine und sehr spezifische Stichprobe, die sehr selektiv ist. Sie wird von uns für weitere Analysen nicht verwendet.

Die *SOEP* Daten werden zur Beantwortung unserer ersten Forschungsfrage verwendet, da das *SOEP* der Datensatz ist, welcher den längsten Zeitraum abdeckt, der eine Betreuung durch Großeltern erfasst.

**DJI-Kinderbetreuungsstudie (DJI-KiBS).** Die DJI-Kinderbetreuungsstudie wird seit 2012 im Auftrag des Bundesministeriums für Familie, Senioren, Frauen und Jugend (BMFSFJ) jährlich erhoben und ist der Nachfolger der KiföG-Elternbefragung und KiföG-Länderstudie des DJI. Die Kinderbetreuungsstudie ist eine repräsentative und jährlich durchgeführte Befragung, die private Haushalte mit Kindern erfasst. Die Analysen basieren auf den Daten der Jahre 2014 bis 2017 und beschränken sich auf Haushalte mit Kindern unter zehn Jahren. Die *DJI-KiBS*-Daten erfassen für alle Wellen, ob das Kind normalerweise von den Großeltern betreut wird. Insbesondere erfassen die Daten in den

Wellen 4-6 (2014-2017) ob diese Betreuung „regelmäßig“ oder „Nach Bedarf“ stattgefunden hat. Darüber hinaus wird für jedes Kind abgefragt, ob sich die Eltern eine stärkere Einbindung der Großeltern in die Kinderbetreuung wünschen. Für alle Personen, die sich eine stärkere Einbindung der Großeltern wünschen, werden schließlich Gründe erfragt, weshalb eine stärkere Betreuung durch die Großeltern nicht stattfinden kann (Deutsches Jugendinstitut 2018).

**Panel Analysis of Intimate Relationships and Family Dynamics (*pairfam*).** Ein Datensatz, der für die Analyse der Auswirkungen der großelterlichen Betreuung auf das subjektive elterliche Wohlbefinden, das sozio-emotionale Verhalten und die Gesundheit der Kinder verwendet wird, ist das "Panel Analysis of Intimate Relationships and Family Dynamics" (*pairfam*). Die Teilnehmer von *pairfam* werden jährlich befragt (Huinink et al. 2011).

*Pairfam* ist eine von der Deutschen Forschungsgemeinschaft (DFG) geförderte, multidisziplinäre Längsschnittstudie zur Untersuchung partnerschaftlicher und familialer Lebensformen (Huinink et al. 2011). Die Studie startete 2008/09 mit einer Ausgangsstichprobe von 12.402 zufällig ausgewählten Ankerpersonen der Geburtsjahrgänge 1971-1973, 1981-1983 und 1991-1993, die jährlich wiederholt interviewt werden, und umfasst inzwischen 12 Wellen. Ungefähr die Hälfte der Ankerpersonen ist männlich, die andere Hälfte ist weiblich. Im Rahmen eines sogenannten „Multi-Actor-Designs“ werden parallel auch die jeweiligen Partner, die Eltern (bis Welle 8) und Kinder dieser Ankerpersonen befragt. Die Antwortquote bei den Partnern liegt bei etwa 52 Prozent. Ab der zweiten Welle werden die Eltern der Ankerpersonen mit einem eigenen Fragenkatalog befragt, der u.a. die Häufigkeit des Kontakts zum Enkelkind, gemeinsamer Unternehmungen sowie finanzielle Ausgaben für das Enkelkind umfasst. Allerdings werden diese Fragen ausschließlich für das jüngste Kind der Ankerperson zwischen 8 und 15 Jahren gestellt. Ab der dritten Welle wurde der Fragenkatalog auch noch um detailliertere Fragen zu der Großeltern-Enkel-Beziehung, bspw. Gesprächsthemen und -häufigkeit mit ausgewähltem Enkelkind (auch im Vergleich zu allen anderen Enkelkindern) sowie Fragen zur emotionalen Nähe erweitert. Ein sogenannter Kinderfragebogen umfasst von Welle 2 bis Welle 8 in jeder zweiten Welle die Kindesperspektive dieses Enkelkindes und fragt ebenfalls Gesprächs-, Geschenks-, und Kontakthäufigkeit mit den Großeltern ab. Da die Kinderbefragung nur Kinder ab 7 Jahren umfasst und die Befragung der Großelternbefragung keine hohen Antwortquoten aufweist, d.h. nur ein geringer Anteil der Großeltern an der Befragung teilnimmt, konzentrieren wir uns in unserer Analyse auf Informationen aus den sogenannten „Anker- und Partnerfragebögen“. Die Ankerperson wird in allen *pairfam*-Wellen zur emotionalen und räumlichen Nähe in Bezug auf ihre Eltern befragt. Darüber hinaus wird die Ankerperson gefragt, wie die eigenen Kinder vormittags und nachmittags betreut werden und inwiefern Großeltern in die Betreuung involviert sind. Neben der

Erfassung der regelmäßigen Betreuung durch die Großeltern wird auch abgefragt wer im Notfall (z.B. im Krankheitsfall oder bei Kita-/Schulschließung) bei der Kinderbetreuung einspringen kann. Die *pairfam*-Daten sind von besonderem Interesse, um die Ausgestaltung der generationenübergreifenden Beziehungen detaillierter darzustellen.

Für die Analyse der schulischen Leistungen von Enkeln verwenden wir einen zweiten Datensatz, da *pairfam* keine Schulnoten oder ähnliches erfasst. Dabei handelt es sich um das Sozio-oekonomische Panel (*SOEP*) (siehe oben). Die Analyse der schulischen Ergebnisse muss aufgrund der Datenverfügbarkeit im *SOEP* auf Kinder im Alter von 7 bis 10 Jahren beschränkt werden. Das *SOEP* und *pairfam* enthalten Informationen über die geografische Entfernung zwischen dem Eltern-Haushalt und allen vier Großeltern (sofern diese noch leben) in mehreren Wellen. Diese Information ist ein zentraler Bestandteil unserer empirischen Methode zur Beantwortung der zweiten Forschungsfrage („Instrumentenvariablen-Ansatz“, siehe Anhang C.2).

## **B2. Detaillierte Beschreibung der analysierten Stichprobe im zweiten Studienschwerpunkt (Kapitel 5)**

Wir führen die Analysen des zu den Wirkungen der Großelternbetreuung auf der Ebene der Kinder und der Eltern durch. Um die Auswirkungen der Großelternbetreuung auf Kinder und Eltern zu bewerten, stellt jedes Kind und alle Eltern eine Beobachtung dar. Die Analysestichprobe für die Eltern ist auf alle Personen beschränkt, die mindestens ein Kind in der entsprechenden Altersgruppe haben. Dabei werden Altersgruppen von Kindern im sogenannten Krippenalter (0-2 Jahre), im sogenannten Kindergartenalter (3-5 Jahre) und im Grundschulalter (6-10 Jahre) durchgeführt.

### **Beschreibung der analysierten Stichproben für den zweiten Studienschwerpunkt**

Wir grenzen unsere Analysestichprobe auf Basis der *pairfam*-Daten auf **Eltern ein, bei denen mindestens ein Elternteil keinen Migrationshintergrund** hat. Die Gründe dafür sind wie folgt: Tabelle A1 zeigt, dass beide Elternteile der großen Mehrheit der Kinder in unserer *pairfam*-Stichprobe (84 Prozent) keinen Migrationshintergrund haben. Nur etwa 11 Prozent der Kinder haben ein Elternteil mit Migrationshintergrund und nur knapp 5 Prozent haben zwei Elternteile mit Migrationshintergrund. Es wird auch ersichtlich, dass eine Großelternbetreuung überproportional häufig in Familien ohne Migrationshintergrund genutzt wird. Ebenso wohnen Familien ohne Migrationshintergrund im Durchschnitt näher bei den Großeltern als Familien mit Migrationshintergrund – was zu erwarten ist, da vielfach die Großeltern noch im Herkunftsland leben. Bei Familien, in denen beide Elternteile nicht in Deutschland geboren wurden ist es sehr wahrscheinlich, dass alle vier Großeltern nicht in Deutschland leben und daher nicht für eine regelmäßige Kinderbetreuung zur Verfügung stehen. Da

diese Gruppe von Eltern im Hinblick auf eine potentielle Großelternbetreuung mit anderen Rahmenbedingungen und Möglichkeiten konfrontiert ist, beschränken wir unsere Analyse auf Familien in denen mindestens ein Elternteil keinen Migrationshintergrund hat. Dies trifft auch auf die *SOEP*-Teilstichprobe zu, welche wir im Kontext der zweiten Forschungsfrage analysieren.

*Tabelle A 1: Großelternbetreuung nach Migrationsstatus – Basis: pairfam*

	Beide Eltern ohne Migrationshintergrund (Anteil Haushalte in Prozent)	Ein Elternteil ohne Migrationshintergrund (Anteil Haushalte in Prozent)	Beide Eltern mit Migrationshintergrund (Anteil Haushalte in Prozent)
Alle Haushalte	84,05	11,27	4,69
Haushalte, die Großelternbetreuung nutzen	88,14	8,67	3,18
Haushalte, die keine Großelternbetreuung nutzen	83,05	10,95	6,00
Haushalte, die maximal 30 Minuten von den Großeltern entfernt wohnen	87,93	8,64	3,44
Haushalte, die mindestens 30 Minuten von den Großeltern entfernt wohnen	70,21	20,20	9,59

Quelle: Eigene Berechnungen auf der Basis von *Pairfam* (2009-2020)

Eine **Charakterisierung der beiden Analysetichproben** enthält Tabelle A2. Sie fasst die Mittelwerte ausgewählter sozioökonomischer Merkmale sowohl für das *pairfam*- und *SOEP*-Analyze-samples zusammen. Der Vergleich der beiden Datensätze deutet auf Unterschiede in den sozioökonomischen Merkmalen hin. Neben Unterschieden in sozioökonomischen Merkmalen ist auch der Anteil von Kindern in großelterlicher Betreuung im *SOEP* fast doppelt so hoch wie in *pairfam*. Dies könnte auf die unterschiedliche Formulierung der Frage und die Art und Weise zurückzuführen sein, wie die großelterliche Betreuung gemessen wird. Außerdem fällt auf, dass die *pairfam*-Stichprobe im Durchschnitt höher gebildet, da der Anteil der Haushalte, in denen mindestens ein Partner einen Hochschulabschluss hat, etwa 12 Prozentpunkte höher ist als im *SOEP* (37 Prozent vs. 49 Prozent).<sup>2</sup> In Bezug auf Migrationshintergrund, Haushaltseinkommen, Alter der Kinder und Mütter, Geschlecht der Kinder und Anzahl der Kinder im Haushalt sind die Stichproben gut vergleichbar. Die dennoch vorhandenen Unterschiede bei den sozioökonomischen Merkmalen unterstreichen, wie wichtig es ist, in der Analyse die Einflüsse dieser Merkmale miteinzubeziehen, was wir in all unseren Analysen tun.

<sup>2</sup> Generell umfasst *pairfam* eine etwas höher gebildete Stichprobe als die deutsche Bevölkerung (Wetzel et al., 2021).

Tabelle A 2: Sozioökonomische Merkmale der Befragten in den Analysestichproben

	<i>Pairfam</i> : Mittelwert (Standardabweichung) Zeitraum: 2009-2020	<i>SOEP</i> : Mittelwert (Standardabweichung) Zeitraum: 2005-2017
Anteil die Großeltern- betreuung nutzen	23,323 Prozent	45,860 Prozent
Anteil der Großeltern, die 30 Minuten oder näher / in der selben Stadt oder näher wohnen	68,853 Prozent	51,856 Prozent
Erwerbsbeteiligung der Mutter		
- Mutter nicht erwerbstätig	- 36,332 Prozent	- 43,159 Prozent
- Mutter Teilzeit erwerbstätig	- 42,669 Prozent	- 43,472 Prozent
- Mutter Vollzeit erwerbstätig	- 18,879 Prozent	- 13,369 Prozent
Höchster Bildungsabschluss im Haushalt		
- Kein oder niedriger Schul- abschluss (z.B. Haupt- oder Realschulabschluss)	- 5,923 Prozent	- 6,429 Prozent
- Höherer Schulabschluss / Berufsausbildung	- 45,509 Prozent	- 55,933 Prozent
- Universitätsabschluss	- 48,569 Prozent	- 37,639 Prozent
Mindestens ein Elternteil hat Migrationshintergrund	11,899 Prozent	12,304 Prozent
Haushaltsnettoeinkommen (im Monat)	3416,561€ (2430,786€)	3298,097€ (1850,606)
Alter der Mutter	34,024 Jahre (7,898 Jahre)	36,286 Jahre (6,007 Jahre)
Anteil männlicher Kinder im Haushalt	50,880 Prozent	52,380 Prozent
Anzahl Kinder im Haushalt	2,043 (0,989)	1,989 (0,915)
Alter der Kinder	4,904 Jahre (3,101 Jahre)	4,885 Jahre (3,173)
Anteil der Eltern, die mit Partner/in zusammenleben	91,068 Prozent	81,816 Prozent
N	29.169	12.690

Quelle: Eigene Berechnungen auf der Basis von *pairfam* (Spalte 1) und *SOEP* (Spalte 2). Die Ergebnisse in Spalte 1 beruhen auf der *Pairfam*-Stichprobe auf Kinder-Ebene.

Die **Erfassung der Großelternbetreuung** erfolgt in den beiden Studien sehr unterschiedlich. Im Rahmen der *pairfam*-Befragung wird die Betreuungsform für den Vor- und Nachmittag separat abgefragt und zwar mit folgender Formulierung: „Wenn Sie einmal an eine normale Woche ohne Urlaubs- und Ferienzeiten denken: Von welchen Personen oder Einrichtungen wird Ihr Kind vormittags/nachmittags betreut?“ Hier gibt es insgesamt 14 Antwortmöglichkeiten, in denen u.a. Partner/in, Großeltern, Geschwister und Kita inkludiert sind. Eine Differenzierung der Betreuung durch Großmütter und/oder Großväter ist nicht möglich.

Im *SOEP* wird die Betreuungsform von Kindern zum einen im Haushaltsfragebogen erfasst. Hier wird zunächst gefragt, ob das Kind in einer Kita oder Schule betreut wird. Im Anschluss wird folgende

Frage gestellt: „Gibt es daneben noch Personen (ohne Haushaltsangehörige), die das Kind regelmäßig betreuen?“. Es gibt folgende Antwortmöglichkeiten: „Ja, Tagesmutter außerhalb des Haushalts“, „Ja, bezahlte Betreuungsperson kommt ins Haus“, „Ja, Verwandte“, „Ja, Freunde/Bekannte/Nachbarn“ und „Nein“. Die Großelternbetreuung ist somit in dieser Frage nicht explizit abgefragt, sondern fällt unter Verwandte. Zum anderen werden in den altersspezifischen Fragebögen unterschiedliche Betreuungspersonen für Kinder erfasst. In diesem Kontext wird die Großelternbetreuung, wie folgt abgefragt: „Wenn Sie einmal an eine normale Woche denken: Gibt es außer Ihnen andere Personen oder Einrichtungen, die die Betreuung des Kindes zeitweise übernehmen? Wenn ja, welche Personen oder Einrichtungen sind das und wie viele Stunden pro Woche übernehmen sie die Betreuung?“ Es gibt 10 Antwortmöglichkeiten, die u.a. Ehe(-)Partner/in, Großeltern des Kindes, ältere Geschwister des Kindes und Krippe, Kindergarten, Kindertageseinrichtung (Kita) umfassen. Eine Differenzierung der Betreuung durch Großmütter und/oder Großväter ist auch im *SOEP* nicht möglich. Ein Vergleich der Antworten im Haushalts- bzw. altersspezifischen Fragebogen bezüglich der Verwandten/Großelternbetreuung zeigt jedoch, dass die Großeltern die wichtigste Gruppe unter den genannten Optionen sind. Somit kann Großelternbetreuung auch sinnvoll im Haushaltsfragebogen abgeschätzt werden.

Vor dem Hintergrund dieser Unterschiede in der Fragestellung ist es nicht überraschend, dass die Anteile der durch die Großeltern betreuten Kinder in den beiden Stichproben voneinander abweichen. Zum einen muss konstatiert werden, dass im *SOEP* Betreuungsstunden pro Woche abgefragt werden und in *Pairfam* „nur“ die Betreuung am Vor- und Nachmittag. Beispielsweise könnten sich Familien, in denen die Großelternbetreuung nicht jede Woche gleich ist oder nicht einen ganzen Vor- oder Nachmittag andauert, auf Betreuungsmöglichkeiten mit wöchentlich festen Zeiten, die einen ganzen Vor-/Nachmittag andauert beschränken. Zum anderen wird im *SOEP* von einer „normalen Woche“ gesprochen, während in *Pairfam* explizit nach einer „normalen Woche ohne Urlaubs-/Ferienzeit“ gefragt wird. Es ist möglich, dass Großeltern in einigen Familien insbesondere in den Ferien oder in Notfällen einspringen. Da in *pairfam* die Ferienzeit explizit ausgeschlossen wird, geben Eltern, deren Kinder hauptsächlich in den Ferien von den Großeltern betreut werden, dies vermutlich nicht an. Dies könnte ebenfalls zu den niedrigeren Zahlen in den *pairfam*-Daten beitragen. Darüber hinaus ist es möglich, dass einige Großeltern die „Über-Mittag-Betreuung“ übernehmen. Beispielsweise wenn ein Elternteil nur vormittags arbeitet, könnte es sein, dass Großeltern ihr Enkelkind von der Kita abholen und über Mittag betreuen, bis das Elternteil nach Hause kommt. Durch die Unterscheidung vor-/nachmittags könnte die „Über-Mittag-Betreuung“ in der *pairfam*-Befragung nicht erfasst sein und somit ebenfalls zu den geringeren Zahlen in den *Pairfam*-Daten beitragen.

## C. Detailliertere Beschreibung des methodischen Vorgehens

### C1: Beschreibung des methodischen Vorgehens im ersten Studienschwerpunkt (Kapitel 4.1)

In Kapitel 4.1.1, Abbildung 4 verwenden wir das Logistische Regressionsmodell um zu bestimmen, ob sich die Wahrscheinlichkeit von den Großeltern über die Zeit und nach Region signifikant verändert hat. Das Logistische Regressionsmodell (kurz Logit-Modell) ist ein ökonometrisches Schätzverfahren, mit dem der Einfluss mehrerer erklärender Variablen auf eine Zielgröße (die abhängige Variable) geschätzt wird. Es wird in der Regel verwendet, wenn die Zielgröße eine binäre Variable („0/1 Variable“) ist. Ein Vorteil des Logit-Verfahrens gegenüber etwa dem linearen Regressionsmodell liegt darin, dass die Verteilung binärer Variablen modelliert werden kann (Hensher und Greene 2003). Die Ergebnisse werden als durchschnittliche marginale Effekte (AMEs) dargestellt, da diese leicht interpretierbare Informationen über die absoluten regionalen, sowie die Unterschiede über die Zeit geben (Buis 2010).

Zunächst werden also logistische Regressionsmodelle geschätzt, mit denen die Wahrscheinlichkeit berechnet wurde, dass die Kinder von den Großeltern betreut werden. Dabei berücksichtigen die Modelle das Alter des betreuten Kindes, die Anzahl der Kinder unter 16 Jahren im Haushalt, sowie das Alter der Mutter. Außerdem werden in die Modelle jeweils Interaktionsterme zwischen regionaler Zugehörigkeit und den beobachteten Zeit-Intervallen eingefügt. In dieser Schätzung werden  $\chi^2$ -Tests durchgeführt, um zu prüfen, ob sich die Betreuungsnutzung von Großeltern in den definierten Zeit-Intervallen signifikant von der Nutzung in *Phase I (1997-2003)*, also vor dem Ausbau unterscheiden. Die Schätzungen der Wahrscheinlichkeit von den Großeltern betreut zu werden, werden dann nach Region und Zeit-Intervall als durchschnittliche marginale Effekte (AMEs) in Abbildung 4 dargestellt.

In den multivariaten Analysen aus Kapitel 4.1.2, Abbildungen 12-14 haben wir ebenfalls logistische Regressionen berechnet. In diesem Fall haben wir dazu immer zwei Wahrscheinlichkeiten geschätzt, und zwar (1) dass Kinder, die von den Großeltern betreut werden, ansonsten ausschließlich von den Eltern (Krippenalter), in der Halbtagskita (Kindergartenalter) oder der Halbtagschule (Grundschulalter) betreut werden (linke Spalten der Abbildungen); und (2) dass Kinder, die von den Großeltern betreut werden, zusätzlich auch eine Kita (Krippenalter) besuchen, auch eine Ganztagskita (Kindergartenalter) oder eine Ganztagschule/Hort (Grundschulalter) besuchen (rechte Spalten der Abbildungen). Zunächst werden auch hier gepoolte logistische Regressionsmodelle geschätzt, mit denen die beiden Wahrscheinlichkeit berechnet werden. Dabei werden in diesem Fall Interaktionen zwischen den Zeit-Intervallen und (1) der Region, in der der Haushalt lebt (Abbildung A3), (2) der mütterlichen

Bildung (Abbildung A4); und (3) dem Migrationshintergrund der Eltern (Abbildung A5) berechnet. In diesem Modell werden ebenfalls  $\chi^2$ -Tests durchgeführt, um zu prüfen, ob sich die Nutzung der beschriebenen Betreuungsformen zwischen den jeweilig betrachteten Gruppen in den Zeit-Intervallen der *Phasen II und III* signifikant von der Nutzung in *Phase I (1997-2019)*, also vor dem Ausbau unterscheiden. Die Schätzungen der Wahrscheinlichkeit in einer jeweiligen Gruppe nicht kombiniert oder kombiniert von den Großeltern betreut zu werden, werden dann nach Zeit-Intervall und Region (Abbildung A3), Bildungshintergrund der Mutter (Abbildung A4) und Migrationshintergrund der Eltern (Abbildung A5) als durchschnittliche marginale Effekte (AMEs) dargestellt.

## **C2: Beschreibung des methodischen Vorgehens im zweiten Studienschwerpunkt**

Um den kausalen Effekt der großelterlichen Betreuung auf die verschiedenen Befunde bei Kindern und Eltern zu identifizieren, muss aus einer methodischen Perspektive das sogenannte „Endogenitätsproblem“ gelöst werden. Was bedeutet dies? Sofern dieses nicht bedacht wird, könnte in einem simplen „Ordinary-Least-Squares“ (OLS) Modell direkt der Effekt einer Großelternbetreuung auf die Befunde von Kindern und Eltern unter Berücksichtigung wichtiger anderer Merkmale, sogenannter Kontrollvariablen (z.B. Bildung der Eltern) geschätzt werden. Formal sieht ein entsprechendes OLS-Modell folgendermaßen aus:

$$y_{it} = \beta_1 + \beta_2 \times GPC_{it} + X'_{it} \times \beta_3 + \mu_{it}$$

wobei  $y_{it}$  die verschiedenen Befunde der Kinder und Eltern darstellen.  $i$  steht für jedes einzelne Kind oder Elternteil und  $t$  für den Messzeitpunkt. Unser Hauptinteresse gilt der Variable Großelternbetreuung ( $GPC_{it}$ ), welche – basierend auf dem *pairfam*-Datensatz - den Wert 1 annimmt für Kinder, die regelmäßig vor- oder nachmittags von den Großeltern betreut werden und null für Kinder, die nicht von den Großeltern betreut werden.  $X_{it}$  ist ein Vektor, der alle wichtigen Kontrollvariablen auf der Ebene der (Groß-)Eltern, des Kindes und des Haushalts erfasst. Im Allgemeinen umfassen wir sozio-ökonomische Merkmale der Eltern wie Bildung, Alter, Einkommen, Erwerbsstatus, Geschlecht, Bundesland des Wohnsitzes und Migrationshintergrund (sofern einer der Partner einen solchen hat). Zusätzlich nehmen wir detaillierte Informationen über die Situation des Haushalts auf (z.B. Anzahl der Kinder im Haushalt und Alter des jüngsten Kindes). Eine Übersicht über die Kontrollvariablen für jede Ergebnisvariable findet sich in Tabelle A3. In weiteren Modellen prüfen wir ob in Abhängigkeit der einbezogenen Kontrollvariablen (z.B. durch Ausschluss potenziell endogener Variablen wie dem Erwerbsstatus der Mutter und dem Einkommen) die Ergebnisse der ersten Modellschätzung robust bleiben.

Tabelle A 3: Übersicht Kontrollvariablen

Variable	Definition	Schätzung der Effekte der Großelternbetreuung auf die				
		Gesundheit	Kinder		Eltern	
	Sozio-emotionales Verhalten		schulische Leistungen	Zufriedenheit mit der Kinderbetreuung	Zufriedenheit mit anderen Dimensionen	
<b>Eltern-Ebene</b>						
post-sekundäre Bildung	Höchster Abschluss im HH	✓	✓	✓		
	Eigene Bildung				✓	✓
Erwerbstätigkeit Mutter	Eltern-Ebene	✓	✓	✓	✓	✓
Vater		✓	✓		✓	✓
Alter	Alter der Mutter	✓	✓	✓		
	Eigenes Alter				✓	✓
Religion	Ein Elternteil ist religiös	✓	✓			
	Eigene Religiosität				✓	✓
Migrationshintergrund	Ein Elternteil hat direkten Migrationshintergrund	✓	✓	✓		
	Eigener direkter Migrationshintergrund				✓	✓
Partnerinformation	Partner hat den Partnerfragebogen beantwortet	✓	✓			
Erziehungsziele	Wichtigkeit von Bewegung und Ernährung	✓				
Gesundheit	Mindestens ein Elternteil krank	✓	✓	✓		
	Eigene Gesundheit				✓	✓
Übergewicht	Mindestens ein Elternteil übergewichtig	✓	✓			
	Eigenes Übergewicht				✓	✓
Schwangerschaft	Mutter ist schwanger	✓	✓		✓	✓
Wohnsituation	Eltern leben zusammen	✓	✓	✓	✓	✓
Verwitwet	Ein Elternteil verwitwet	✓	✓	✓	✓	
	Eigener Status					✓
Einzelkind	Mindestens ein Elternteil ist Einzelkind	✓	✓			
	Eigener Einzelkindstatus				✓	✓
Zufriedenheit mit Kinderbetreuung	Kinder-Ebene		✓			



Fortsetzung Tabelle A 3: Übersicht Kontrollvariablen

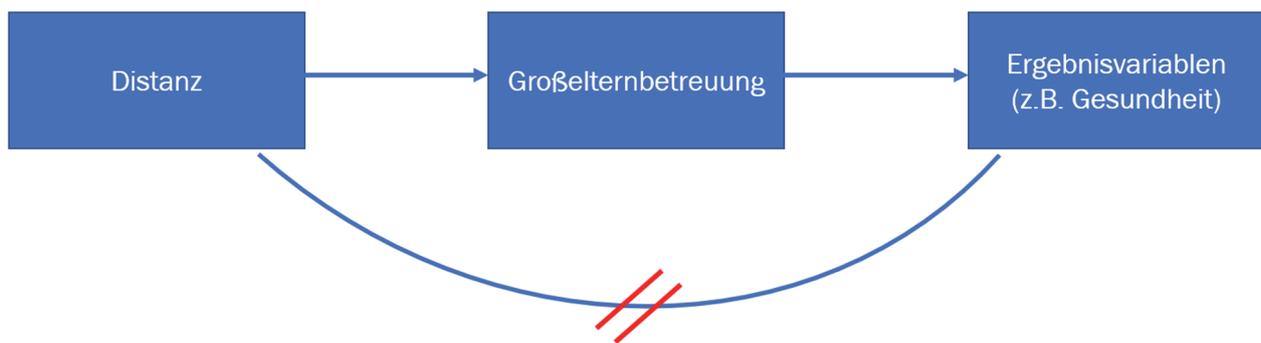
Variable	Definition	Schätzung der Effekte der Großelternbetreuung auf die				
		Gesundheit	Kinder		Eltern	
			Sozio-emotionales Verhalten	schulische Leistungen	Zufriedenheit mit der Kinderbetreuung	Zufriedenheit mit anderen Dimensionen
<b>Kinder-Ebene</b>						
Geschlecht	Geschlecht des Kindes	✓	✓	✓	✓	
	Geschlechterkomposition der Kinder im Haushalt					✓
Alter Kind	Alter des Kindes	✓	✓	✓	✓	
	Alter des jüngsten Kindes					✓
Anzahl Kinder im Haushalt	Insgesamt	✓	✓	✓	✓	
	Anzahl zwischen 0 und 2 Jahren					✓
	Anzahl zwischen 3 und 5 Jahren					✓
	Anzahl zwischen 6 und 10 Jahren					✓
	Anzahl älterer Kinder					✓
Geburtsreihenfolge	Alter im Vergleich zu Geschwistern	✓	✓	✓	✓	
Kita-Nutzung	Kind (0-5 Jahre) in der Kita	✓	✓		✓	
	Anzahl Kinder (0-5 Jahre) in der Kita					✓
Temperament	Temperament Kinder 0-6 Jahre		✓			
<b>Großeltern-Ebene</b>						
Schulbildung	Mutter der Befragungsperson	✓	✓		✓	✓
	Vater der Befragungsperson	✓	✓		✓	✓
	Großmutter mütterlicherseits			✓		
	Großmutter väterlicherseits			✓		
	Großvater mütterlicherseits			✓		
	Großvater väterlicherseits			✓		
Alter	Durchschnitt aller lebenden Großeltern	✓	✓	✓	✓	✓
<b>Haushalts-Ebene</b>						
Einkommen	logarithmiert	✓	✓	✓	✓	✓
Jahr	Befragungsjahr	✓	✓	✓	✓	✓
Bundesland	Bundesland, indem HH lebt	✓	✓	✓	✓	✓
Gemeindegröße	Größe Gemeinde, in der HH lebt	✓	✓	✓	✓	✓

Die Verwendung des OLS-Modells führt jedoch nicht unbedingt zu Schätzungen, die kausal interpretiert werden können, d.h. dass sie tatsächlich nur den Einfluss der Großelternbetreuung auf die Befunde bei Kindern und Eltern messen. Die Identifizierung eines kausalen Effekts der großelterlichen Betreuung auf die erfassten Befunde von Kindern und Eltern ist potenziell mit einem „Endogenitätsproblem“ behaftet: Die Entscheidung für eine großelterliche Betreuung könnte endogen sein, da sie von Eltern und Großeltern getroffen wird und möglicherweise durch unbeobachtete Merkmale beeinflusst werden, die auch die erfassten Befunde beeinflussen. Da solche nicht zu beobachteten Merkmale nicht berücksichtigt werden können, könnte eine solche Schätzung zu einer Verzerrung durch ausgelassene Variablen führen. Ein Beispiel für eine solche unbeobachtete Variable sind die Präferenzen der Großeltern für die Betreuung ihres Enkelkindes. Diese beeinflussen wahrscheinlich den Umfang der Unterstützung, die die Großeltern anbieten, und könnten sich auch direkt auf die Befunde der Kinder und Eltern auswirken. Ein weiteres Problem, welches zu dem „Endogenitätsproblem“ führen könnte, ist die sogenannte „umgekehrte Kausalität“. Dieses Problem liegt vor, wenn die Richtung des Effektes nicht eindeutig ist, also nicht nur Großelternbetreuung die erfassten Befunde beeinflusst, sondern auch die erfassten Befunde die Entscheidung für oder gegen eine Großelternbetreuung beeinflussen. Beispielsweise könnte das Wohlbefinden der Eltern Einfluss darauf haben, wie viel Unterstützung sie von den Großeltern benötigen und bei ihnen „nachfragen“. Auch der Gesundheitszustand oder das sozio-emotionalen Verhalten der Kinder dürfte die Entscheidung, die Großeltern um Unterstützung zu bitten, beeinflussen. So könnten Eltern mit Kindern, die unter einem schlechten Gesundheitszustand leiden, befürchten, dass die Betreuung dieser Kinder eine zu große Belastung für die Großeltern darstellen würde. Andererseits könnten auch genau solche Eltern auf die Großeltern angewiesen sein, da keine andere nicht-elterliche Betreuungsform für kranke Kinder in Frage kommt.

Daher könnte die Anwendung des simplen OLS-Modells zu verzerrten und inkonsistenten Schätzungen des Effektes der großelterlichen Betreuung führen und würde keinen kausalen Effekt widerspiegeln. Es gibt somit viele Gründe sowohl nach oben als auch nach unten verzerrte OLS-Schätzungen zu erwarten. Wenn z.B. nur gesunde und sozio-emotional stabile Kinder in großelterlicher Betreuung sind, erwarten wir, dass der OLS-Schätzer nach oben verzerrt ist, das heißt einen positiveren Effekt als den „wahren“ Effekt widerspiegelt. Wenn wir hingegen davon ausgehen, dass Eltern mit geringem subjektivem Wohlbefinden eher Großeltern um Unterstützung bei der Kinderbetreuung bitten, wäre der OLS-Schätzer nach unten verzerrt, das heißt würde einen negativeren Effekt als den „wahren“ Effekt darstellen. Wir können das „Endogenitätsproblem“ nicht lösen, indem wir alle Störfaktoren als Kontrollvariablen einbeziehen, da einige von ihnen in den vorliegenden Daten nicht beobachtet werden oder generell nicht messbar und somit nicht berücksichtigt werden können.

Um das „Endogenitätsproblem“ zu überwinden und den kausalen Effekt der großelterlichen Betreuung auf die verschiedenen untersuchten Faktoren zu identifizieren, wenden wir eine Instrumentenvariablenstrategie (IV) an. Mit Hilfe eines Instruments, das mit dem endogenen Regressor (Großelternbetreuung) korreliert ist, aber die abhängigen Variablen („Ergebnisvariablen“, welche die Befunde bei Kindern und Eltern erfassen) nur durch seine Wirkung auf die unabhängige Variable (Großelternbetreuung) beeinflusst, können wir die vorhandene Variation in der Großelternbetreuung bestimmen. Zu diesem Zweck verwenden wir die Entfernung zu den Großeltern als ein Instrument für weitere kausale Modellschätzungen. Der Zusammenhang zwischen Instrument (Distanz), endogenem Regressor (Großelternbetreuung) und der abhängigen Variable (Ergebnisvariablen) ist in Abbildung A9 dargestellt. Dieses Instrument wurde bereits von anderen Autoren in der Literatur genutzt (z.B. Del Boca et al. 2018 und Compton und Pollak 2014).

Abbildung A 9: Zusammenhang zwischen Instrument, endogener Variable und abhängiger Variable



Quelle: Eigene Darstellung.

Die Distanz zu den Großeltern wird in *SOEP* und *pairfam* unterschiedlich gemessen. Beide Datensätze enthalten Informationen über die geografische Entfernung zwischen den Erwachsenen im Haushalt und allen vier Großeltern (falls diese noch leben) in mehreren Wellen. In *pairfam* messen wir dies, indem wir Familien, die 30 Minuten oder weniger von den Großeltern entfernt wohnen, mit solchen vergleichen, die weiter entfernt wohnen, während im *SOEP* das „Instrument“ das Wohnen in derselben Stadt mit dem Wohnen in einer anderen Stadt vergleicht. Genauer gesagt wird die Entfernung zu den Großeltern in *pairfam* mit der folgenden Frage erhoben: "Wie viel Zeit benötigen Sie, um zur Wohnung Ihrer Mutter zu gelangen? (an einem normalen Tag, mit normalen Verkehrsmitteln)". In den Fällen, in denen die Eltern nicht in einem Haushalt leben, wird ihnen die gleiche Frage nach der Entfernung zur Wohnung des Vaters gestellt. Die Entfernung wird als kategoriale Variable mit sechs Kategorien gemessen. Die sechs Kategorien sind: "wir leben in einem Haus", "weniger als 10 Minuten", "10 Minuten bis weniger als 30 Minuten", "30 Minuten bis weniger als 1 Stunde", "1 Stunde bis weniger als 3 Stunden" und "3 Stunden oder mehr".

Auf dieser Grundlage konstruieren wir eine binäre Variable, die gleich Eins ist, wenn mindestens ein Großelternanteil näher als 30 Minuten entfernt wohnt, und ansonsten gleich Null ist. Wir verwenden diese binäre Variable, weil die Beziehung zwischen der Entfernung und dem Umfang der großelterlichen Betreuung wahrscheinlich nicht linear ist. So dürfte der Unterschied zwischen einer Entfernung von 10 oder 30 Minuten einen größeren Einfluss haben als der Unterschied zwischen 3 Stunden und 3 Stunden und 20 Minuten. Wir verwenden 30 Minuten als Grenzwert, da dies eine realistische Entfernung ist, die eine tägliche Betreuung eines Enkelkinds erlaubt.<sup>3</sup> Die Verteilung der Distanzvariable, die zur Konstruktion unseres Instruments verwendet wurde, und der großelterlichen Betreuungsvariable in *pairfam* ist in Kapitel 5, Kasten 4, Abbildung A20 abgebildet. Diese Abbildung zeigt den Anteil der Kinder, die von den Großeltern betreut werden, nach der Mindestentfernung des Kindes zu den Großeltern. Es zeigt sich, dass die meisten Kinder in der Nähe von mindestens einem Großelternanteil leben. In der gesamten Stichprobe leben etwa 69 Prozent der Haushalte weniger als 30 Minuten von mindestens einem Großelternanteil entfernt. Darüber hinaus ist zu erkennen, dass der Anteil der Haushalte, die großelterliche Betreuung in Anspruch nehmen, mit abnehmender Entfernung nichtlinear ansteigt.<sup>4</sup>

Im *SOEP* wird die Entfernung zu den Großeltern in den einzelnen Fragebögen der Eltern mit der folgenden Frage erhoben: "Welche und wie viele der folgenden Verwandten haben Sie? Bitte geben Sie auch an, wo sie wohnen." Die Entfernung wird als kategoriale Variable mit sieben Kategorien gemessen, die alle 5 Jahre erhoben wird. Die sieben Kategorien sind: "im gleichen Haushalt", "im gleichen Haus, aber in einem anderen Haushalt", "in der gleichen Nachbarschaft", "in der gleichen Stadt, aber mehr als 15 Minuten Fußweg entfernt", "in einer anderen Stadt, aber innerhalb einer Autostunde", "weiter weg, aber in Deutschland" und "im Ausland". In unserer Analyse verwenden wir die in den Jahren 2011 und 2016 ermittelten Entfernungen. Um eine größere Stichprobe für unsere Analyse verwenden zu können, imputieren wir die Entfernung im Jahr vor und nach der Befragung. Das bedeutet, dass unsere *SOEP*-Analyse sowohl auf den Jahren 2010-2012 als auch auf 2015-2017 basiert. Wie bei *pairfam* definieren wir eine binäre Variable für die Entfernung, die dem Wert eins entspricht, wenn mindestens ein Großelternanteil in derselben Stadt wie der Haushalt lebt (aber möglicherweise mehr als 15 Minuten zu Fuß entfernt ist), und ansonsten dem Wert Null. 52 Prozent der Haushalte in der *SOEP*-Stichprobe leben in der gleichen Stadt wie mindestens eines der Großelternanteile.

---

<sup>3</sup> In Robustheitsprüfungen testen wir, ob unsere Ergebnisse sich verändern, wenn wir andere Definitionen des Instruments nutzen (Verwendung eines ordinalen Instruments und Verwendung einer Stunde als Grenzwert).

<sup>4</sup> Es scheint, dass von den Haushalten in unserer Stichprobe, die weiter als drei Stunden von allen Großeltern entfernt leben, immer noch etwas mehr als 5 Prozent angeben, regelmäßig großelterliche Betreuung in Anspruch zu nehmen. Da dies unwahrscheinlich erscheint, schließen wir diese Haushalte in einer Robustheitsprüfung aus, was unsere Ergebnisse nicht verändert.

Der Instrumentenvariablen-Ansatz besteht aus zwei Stufen. In der ersten Stufe wird der Effekt der Distanz auf die Großelternbetreuung, die wir als endogen annehmen, unter Beachtung der Kontrollvariablen geschätzt. Formal sieht die Gleichung der ersten Stufe folgendermaßen aus:

$$GPC_{it} = \gamma_1 + \gamma_2 \times D_{it} + X'_{it} \times \gamma_3 + \varepsilon_{it}$$

wobei  $D_{it}$  gleich Eins ist, wenn der Haushalt weniger als 30 Minuten von mindestens einem Großeltern teil entfernt wohnt, und andernfalls gleich Null. Für die auf dem *SOEP* basierenden Analysen ist dies als Eins für Haushalte definiert, die in derselben Stadt wie die Großeltern leben, und andernfalls als Null.  $X_{it}$  beschreibt denselben Vektor von Kontrollvariablen wie in der OLS-Gleichung zuvor. Die abhängige Variable  $GPC_{it}$  ist die binäre Großelternbetreuungs-Variable.

In der zweiten Stufe des IV-Ansatzes werden die geschätzten Werte der Großelternbetreuung ( $\widehat{GPC}_{it}$ ) aus der ersten Stufe genutzt und damit der Effekt von Großelternbetreuung auf die Ergebnisvariablen unter Berücksichtigung der Kontrollvariablen geschätzt. Formal sieht die zweite Stufe folgendermaßen aus:

$$y_{it} = \beta_1 + \beta_2 \times \widehat{GPC}_{it} + X'_{it} \times \beta_3 + \mu_{it}$$

In dieser Regression sind  $y_{it}$  die verschiedenen kindlichen und elterlichen Ergebnisvariablen.  $X_{it}$  ist derselbe Vektor von Kontrollvariablen wie in der ersten Stufe.  $\beta_2$  schätzt den durchschnittlichen Effekt von Großelternbetreuung auf unsere Kontrollvariablen und ist somit für uns von besonderem Interesse.

Damit der Abstand zu den Großeltern als valides Instrument gelten kann, muss er eine Reihe von Bedingungen erfüllen. Besonders wichtig sind die Annahmen zur sogenannten „Relevanz“ und „Exogenität des Instruments“. Relevanz bedeutet, dass das Instrument ausreichend mit dem endogenen Regressor Großelternbetreuung korreliert sein muss. Die Entfernung zu den Großeltern erfüllt die Relevanzbedingung, da eine geringere Entfernung die großelterliche Betreuung erleichtert. Die Korrelation zwischen unserem Instrument und der großelterlichen Betreuung ist in Abbildung 20 zu sehen. Diese Abbildung zeigt den Anteil der Kinder, die von den Großeltern betreut werden, nach der Mindestentfernung des Kindes zu den Großeltern. Außerdem ist zu erkennen, dass der Anteil der Haushalte, die die großelterliche Betreuung in Anspruch nehmen, mit abnehmender Entfernung nicht-linear zunimmt. Die Korrelation zwischen dem Instrument und dem endogenen Regressor wird auch

im Rahmen des IV-Ansatzes in der ersten Stufe getestet. Hier wird der Effekt der Distanz auf die Großelternbetreuung unter Beachtung der Kontrollvariablen geschätzt. Dieser Zusammenhang ist sehr stark: Eine Entfernung von maximal 30 Minuten zu den Großeltern erhöht die Wahrscheinlichkeit einer Betreuung durch die Großeltern um ca. 23 Prozentpunkte.

Die kritischere Annahme ist die Exogenitätsannahme des Instruments, die voraussetzt, dass das Instrument die Ergebnisvariable nur indirekt über die Großelternbetreuung beeinflusst. Es scheint plausibel, dass die Entfernung die Ergebnisse der Kinder nur durch die großelterliche Betreuung beeinflusst. Es kann jedoch argumentiert werden, dass das Wohnen in der Nähe der Großeltern das elterliche Wohlbefinden nicht nur durch die großelterliche Betreuung beeinflusst, sondern auch durch die Beziehung zu den Großeltern und die Zeit, die Eltern und Großeltern miteinander verbringen können. Um sicherzustellen, dass sich die Entfernung nur durch die großelterliche Betreuung auf die elterlichen Ergebnisse auswirkt, kontrollieren wir in einer Robustheitsprüfung für die emotionale Nähe zwischen Eltern und Großeltern. Darüber hinaus kann argumentiert werden, dass die Notwendigkeit einer Kinderbetreuung die Wahrscheinlichkeit erhöht, dass die Familien näher bei den Großeltern leben (z.B. Chen und Zhang 2018). Um die Exogenität der Entfernung zu den Großeltern weiter zu testen, untersuchen wir, ob die Entfernung zwischen Eltern und Großeltern um die Geburt herum abnimmt, was darauf hinweisen würde, dass entweder die Eltern näher zu den Großeltern oder die Großeltern näher zu den Eltern gezogen sind. Der Grund für ein systematisches Umzugsverhalten könnte die Erleichterung durch eine großelterliche Kinderbetreuung sein, was die Entfernung zu den Großeltern zu einer endogenen Variable machen würde. Untersuchungen des Umzugsverhaltens im Jahr vor und nach der Geburt des Erstgeborenen beziehungsweise aller Kinder zeigen allerdings keine systematische Mobilität hin zum Wohnort der Großeltern. In einer Robustheitsprüfung schränken wir die Stichprobe weiter auf die Haushalte ein, die während des Beobachtungszeitraums nicht umgezogen sind, und schließen somit alle Haushalte aus, die möglicherweise näher zu den Großeltern gezogen sind, um Unterstützung bei der Kinderbetreuung zu haben. Die Ergebnisse ändern sich jedoch nicht.