



Mehrwert durch Geo-Milieu-Daten?

Die Nutzung kleinräumiger kommerzieller
Daten für kommunale Fragestellungen

Katharina Knüttel, Thomas Groos, Volker Kersting,
Friederike-Sophie Niemann, Rebekka Vollmer-Baratashvili

Im Jahr 2011 haben die Landesregierung Nordrhein-Westfalen und die Bertelsmann Stiftung das Modellvorhaben „Kein Kind zurücklassen! Kommunen in NRW beugen vor“ ins Leben gerufen. Ziel dieser Initiative war und ist es, gemeinsam mit den beteiligten Modellkommunen allen Kindern und Jugendlichen bestmögliche Chancen für ein gelingendes Aufwachsen und gesellschaftliche Teilhabe zu ermöglichen – und das unabhängig von ihrer Herkunft. Die Initiative wurde von Beginn an wissenschaftlich begleitet. Das Ziel: Ansatzpunkte und Mechanismen gelingender Präventionsarbeit zu identifizieren. Die Bertelsmann Stiftung verantwortet die Begleitforschung gemeinsam mit ihren wissenschaftlichen Partnern. Größter Partner ist das Zentrum für interdisziplinäre Regionalforschung ZEFIR an der Ruhr-Universität Bochum. In der vorliegenden gemeinsamen Schriftenreihe des ZEFIR und der Bertelsmann Stiftung werden in unregelmäßigen Abständen Einblicke und Erkenntnisse aus der gemeinsamen Begleitforschung veröffentlicht.

In 2011, the state government of North Rhine-Westphalia and the Bertelsmann Stiftung launched the initiative, “Kein Kind zurücklassen! Kommunen in NRW beugen vor” (“Leave No Child Behind! Municipalities in North Rhine-Westphalia providing equal opportunities for all children”). The goal of this initiative remains unchanged: To partner with the participating model municipalities to enable every child and young person to have the best possible chance to participate in society and have a successful upbringing – regardless of their background. The initiative has been guided by academic research since its inception. The goal is to identify the approaches and mechanisms that result in successful prevention. Together with its partners from academia, the Bertelsmann Stiftung is overseeing the research that accompanies the initiative. One of the principal academic partners is the Centre for Interdisciplinary Regional Studies (ZEFIR) at the University of Bochum. From time to time, insights and findings from the collaborative accompanying research will be published in this series of joint papers by the ZEFIR and the Bertelsmann Stiftung.

Katharina Knüttel, Thomas Groos, Volker Kersting,
Friederike-Sophie Niemann, Rebekka Vollmer-Baratashvili

Mehrwert durch Geo-Milieu-Daten?

Die Nutzung kleinräumiger kommerzieller Daten für kommunale Fragestellungen

Arbeitspapiere wissenschaftliche Begleitforschung „Kein Kind zurücklassen!“
Werkstattbericht

Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie;
detaillierte bibliographische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Katharina Knüttel, Thomas Groos, Volker Kersting,
Friederike-Sophie Niemann, Rebekka Vollmer-Baratashvili

Mehrwert durch Geo-Milieu-Daten?

Die Nutzung kleinräumiger kommerzieller Daten
für kommunale Fragestellungen

Schriftenreihe Arbeitspapiere wissenschaftliche Begleitforschung „Kein Kind zurücklassen!“
Erscheinungsort Gütersloh
Band 16 (Juni 2019)

© Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
Telefon 05241 81-81 285
www.bertelsmann-stiftung.de
Dr. Kirsten Witte, Director Programm LebensWerte Kommune, Bertelsmann Stiftung

© Zentrum für interdisziplinäre Regionalforschung (ZEFIR)
Fakultät für Sozialwissenschaft, Ruhr-Universität Bochum
LOTA 38, 44780 Bochum
Telefon 0234 32-24 675
www.ruhr-uni-bochum.de/zefir
Prof. Dr. Jörg Bogumil

Redaktion
Dr. Regina von Görtz, Projektleitung „Kein Kind zurücklassen!“, Bertelsmann Stiftung
Dr. David H. Gehne, Forschungskordinator „Kein Kind zurücklassen!“, ZEFIR Bochum

Autoren
Katharina Knüttel, Thomas Groos, Volker Kersting, Friederike-Sophie Niemann,
Rebekka Vollmer-Baratashvili

Koordination Vera Hanke, Bertelsmann Stiftung
Titelbild © Christian Schwier – stock.adobe.com
Gestaltung Dietlind Ehlers, Bielefeld
Lektorat Rudolf Jan Gajdacz, team 4media&event, München
Druck Druckhaus Rihn, Blomberg

ISSN-Print 2199-6393
ISSN-Internet 2199-6407

Mit finanzieller Unterstützung des Landes Nordrhein-Westfalen und des Europäischen Sozialfonds.

Inhalt

1	Einleitung	10
2	Die „horizontale“ Differenzierung: Milieuforschung in der (Stadt-)Soziologie	12
3	Datengrundlage: Datenabgleich in Mülheim an der Ruhr	17
4	Sinus-Geo-Milieus® in Mülheim an der Ruhr	38
5	(Milieuspezifische) Kita-Segregation	51
6	Fazit	59
7	Literatur	61
	Anhang	65

Abbildungen

ABBILDUNG 1:	Abweichungen zwischen empirischen und erwarteten Werten bei „Haushalten mit Migrationshintergrund“	32
ABBILDUNG 2:	Kartographischer Vergleich „Haushalt mit Migrationshintergrund“	34
ABBILDUNG 3:	Vergleich der SGB-II-Quoten und der Kaufkraft in Rasterzellen	36
ABBILDUNG 4:	Sinus-Milieus® in Deutschland 2017 – Soziale Lage und Grundorientierung („Kartoffelgraphik“)	39
ABBILDUNG 5:	Pole der Leitmilieus – Konservativ-Etablierte und Expeditiv	44
ABBILDUNG 6:	Wahrscheinlichste und überrepräsentierte Milieus	48
ABBILDUNG 7:	Ortslagen in Mülheim an der Ruhr, Statistische Bezirke	65

Tabellen

TABELLE 1:	Vergleich der Bevölkerungszahlen amtliche und kommunale Statistik, 2014 bis 2016	20
TABELLE 2:	Vergleich der Haushaltszahlen amtliche und kommunale Statistik der Jahre 2014 bis 2016	22
TABELLE 3:	Anteil der Menschen mit Migrationshintergrund in Mülheim an der Ruhr, 2012 bis 2016	24
TABELLE 4:	Vergleich der Bevölkerung und der Haushalte auf Rasterebene	29
TABELLE 5:	Korrelationskoeffizienten zwischen verschiedenen Konstruktionen von „Haushalten mit Migrationshintergrund“	31
TABELLE 6:	Vergleich der Milieuwahrscheinlichkeiten in Prozent, Deutschland und Mülheim an der Ruhr	41
TABELLE 7:	Vergleich der Milieuwahrscheinlichkeiten in den 100x 100-Meter-Rasterzellen	42
TABELLE 8:	Korrelationsmatrix der Milieuwahrscheinlichkeiten in Rasterzellen	46
TABELLE 9:	Typisierungsvergleich – kommunale versus kommerzielle Daten	54
TABELLE 10:	Durchschnittliche Sozialgeldquote, Migrationshintergrundanteil und Kaufkraft nach Kita-Typ	55
TABELLE 11:	Korrelationen kommerzieller und kommunaler Typisierungsvariablen in den Kitas	56
TABELLE 12:	Spannweiten der Milieuwahrscheinlichkeiten auf verschiedenen Ebenen, in Prozentpunkten	58

1 Einleitung

Arme und Reiche, Menschen mit und Menschen ohne Migrationshintergrund leben in Städten oft nicht am gleichen Ort, sondern in unterschiedlichen Vierteln – die Soziologie spricht hier von „Segregation“, der räumlichen Ungleichverteilung von Bevölkerungsgruppen nach bestimmten Eigenschaften.

In den vorangegangenen Studien im Bereich der Mikrodatenanalyse der wissenschaftlichen Begleitforschung von „Kein Kind zurücklassen! Kommunen in NRW beugen vor“ (KeKiz) konnte gezeigt werden, dass nicht nur Stadtviertel hochgradig segregiert sind, sondern auch Institutionen wie Kitas und Schulen (Groos 2015 und 2016; Groos und Jehles 2015; Groos, Trappmann und Jehles 2018): In einigen Kitas sammeln sich tendenziell die Kinder der Privilegierten, in anderen eher die Kinder der Benachteiligten. Insbesondere im Bereich der frühkindlichen Bildung treten dabei negative Effekte durch Armutssegregation auf. So erzielen Kinder aus „sozial benachteiligten“ Kitas beispielsweise in der Schuleingangsuntersuchung schlechtere Ergebnisse, wobei dieser Effekt *unabhängig* von den individuellen Eigenschaften der Kinder besteht (Groos und Jehles 2015).

Bislang wurden zur Durchführung von Segregationsanalysen vorrangig demographische und sozialstrukturelle Variablen genutzt. In der Sozialstrukturanalyse haben sich hier verschiedene Faktoren etabliert, so z.B. das Alter von Kindern und deren Eltern, ihre Konfessionszugehörigkeit sowie der Bezug von Transferleistungen.

Zunehmend werden in der Forschung jedoch auch weitere Faktoren in die Analysen einbezogen. Seit den 1980er Jahren erfahren Lebensstile und Milieus in der soziologischen Debatte eine stärkere Beachtung. Das in diesem Kontext entstandene Milieumodell des Sinus-Instituts hat sich dabei zu einem wichtigen fachlichen Bezugspunkt entwickelt

(Otte und Rössel 2011: 9) und wurde in der kommerziellen Markt- und Wahlforschung bereits sehr erfolgreich eingesetzt (Geißler 2014: 114). Soziale Milieus umfassen nach diesem Modell Menschen mit ähnlicher Lebensführung, Grundhaltung und Wertvorstellung. Sie erweitern das traditionelle Konzept sozialer Schichtung, welches vor allem auf Indikatoren wie Einkommen, Berufsstatus und Bildungsabschluss basiert, um kulturelle, ästhetische und psychographische Aspekte wie Wertorientierungen, Lebensweisen und Lebensstile (Barz und Tippelt 2007: 12). Inwiefern Lebensstile und Milieus allerdings tatsächlich eigenständige oder ergänzende Erklärungskraft für Segregationsanalysen besitzen, ist in der Forschung umstritten (Otte und Rössel 2011: 11), die Forschungslage uneinheitlich (Rössel und Hoelscher 2012) und wird hier für den spezifischen Gegenstand empirisch auszuloten sein.

Dies erfolgt anhand der Sinus-Geo-Milieus®, die durch die Firma microm Micromarketing-Systeme und Consult GmbH in Zusammenarbeit mit dem Sinus-Institut Ende der 1990er Jahre entwickelt wurden. Mithilfe einer Verknüpfung repräsentativer Befragungsdaten der Sinus-Milieuforschung mit microm-Adressdaten für jedes Haus in Deutschland wird dabei die statistische Wahrscheinlichkeit von Sinus-Milieus berechnet. Die Daten stehen dabei auf verschiedenen kleinräumigen Ebenen (so beispielsweise für Straßenabschnitte, Baublöcke, 100x100-Meter-Raster usw.) zur Verfügung (vgl. Küppers 2018).

Im vorliegenden Werkstattbericht werden die bisherigen Segregationsanalysen innovativ mit dem Milieuansatz verknüpft: Durch eine Übertragung der Sinus-Geo-Milieus® und weiterer soziodemographischer kleinräumiger microm-Daten auf die Kinder in Mülheim an der Ruhr sowie die anschließende Aggregation auf Kitas soll eruiert werden, inwieweit (1) die Daten geeignet sind, Kita-Segregation abzubilden, und ob sie (2) in der Lage sind, Kitas schärfer zu profilieren, als dies anhand sozialstruktureller Daten möglich ist.

Dazu wird nach einer Einführung in die Theorie sozialer Milieus (Kapitel 2) zunächst die kleinräumige kommerzielle Datengrundlage auf ihre grundsätzliche Plausibilität überprüft (Kapitel 3). Anschließend wird erforscht, inwiefern sich mithilfe der „harten“ kommerziellen Variablen der Kaufkraft und des Migrationshintergrunds benachteiligte und privilegierte Kitas identifizieren lassen und ob sich neben der sozialstrukturellen Kita-Segregation eine milieuspezifische Dimension abbilden lässt (Kapitel 4). Die Ergebnisse werden in Kapitel 5 zusammengefasst.

2 Die „horizontale“ Differenzierung: Milieuforschung in der (Stadt-)Soziologie

In der Forschung besteht Unklarheit darüber, wie die Begriffe „Lebensstil“ und „Milieu“ exakt zu bestimmen sind, wobei große Schnittmengen zwischen beiden Konzepten existieren (Otte und Rössel 2011: 12 ff.). Konsens besteht weitestgehend darüber, dass

- „Lebensstile“ sich auf wiederkehrende Muster und Strukturen von Verhaltensweisen und Entscheidungen beziehen und damit stärker Individuen fokussieren;
- der Begriff „Milieu“ hingegen den Fokus weg von den Individuen hin zu dem sie umgebenden Umfeld verschiebt.

Die wissenschaftliche Beschäftigung mit sozialen Milieus reicht bis in die Anfänge der Soziologie zurück: Die Tradition kann hier bis in die Mitte des 19. Jahrhunderts zurückverfolgt werden (Hradil 2006: 3). Hinsichtlich der Beschäftigung mit Lebensstilen werden darüber hinaus als frühe Kronzeugen Georg Simmel (2014) und Max Weber (Otte und Rössel 2011: 9) sowie Thorstein Veblens „Theorie der feinen Leute“ aus dem Jahr 1899 genannt (vgl. Barz 2018: 156).

In der deutschen Soziologie nach 1945 lag jedoch zunächst der Schwerpunkt auf der *vertikalen* Differenzierung der Gesellschaft: Die Sozialstruktur wurde vorrangig in Klassen- oder Schichtmodellen analysiert (vgl. Burzan 2011: 12). Erst seit den 1980er Jahren ist eine zunehmende Beschäftigung mit Fragen des Lebensstils festzustellen. Diskursiv befeuert wurde dies im deutschsprachigen Raum durch das Erscheinen des von Reinhard Kreckel herausgegebenen Sonderbands Soziale Welt mit dem Titel „Soziale Ungleichheiten“ (1983). Kreckel moniert in der Einleitung, dass „die ausschließliche Deutung gesellschaftlicher Ungleichheitsverhältnisse unter dem Gesichtspunkt des

Höher und Tiefer, also: die vertikale Klassifikation selbst, zum Erkenntnismoment werden [drohe]“ (a.a.O.: 5). Vor allem zwei Beiträge des Bandes entfalteten wissenschaftlich eine große Wirkung und wurden zahlreich und intensiv diskutiert: einerseits Pierre Bourdieus „Ökonomisches Kapital, kulturelles Kapital, soziales Kapital“ und andererseits Ulrich Becks „Jenseits von Klasse und Stand“.

Bourdieu gilt als moderner Klassiker der Soziologie. In seiner umfangreichen und komplexen Gesellschaftstheorie spielen soziale Milieus und Lebensstile eine wichtige Rolle. Nichtsdestotrotz sind Lebensstile bei Bourdieu keinesfalls unabhängig von dem zu verstehen, was er als „soziale Klasse“ bezeichnet. Vielmehr sind beide Konzepte – Lebensstil und soziale Klasse – bei Bourdieu eng miteinander verschränkt. Für die französische Gesellschaft der 1960er Jahre identifiziert er aufgrund ihrer Nähe im sozialen Raum¹ drei „theoretische“ Klassen und mit ihnen korrespondierende Geschmacksformen/Lebensstile:

- Die herrschende Klasse, die den *legitimen Geschmack* repräsentiert und das Ziel der sozialen Distinktion verfolgt,
- die Mittelklasse mit dem *mittleren oder prätentiosen Geschmack* und dem Ziel einer Imitation des legitimen Geschmacks und schließlich
- die beherrschte Klasse mit dem *populären Geschmack* und einer Orientierung an der Notwendigkeit (vgl. Bourdieu 1987: 37 f.).

Während Lebensstile in Bourdieus Gesellschaftsdiagnose auch eine Klassenfrage sind, ergibt sich ihre zunehmende Bedeutung in der Theorie Ulrich Becks geradezu durch den von ihm unterstellten Bedeutungsverlust der Klassen. Im Rahmen des von ihm

¹ Bourdieus Konzeption des sozialen Raums beinhaltet zwei miteinander verschränkte Dimensionen: den Raum der Positionen („Objektivität erster Ordnung“) und den Raum der Lebensstile/Perspektiven („Objektivität zweiter Ordnung“). Die Positionen im ersten sind geprägt durch „Kapitalvolumen, Kapitalstruktur und zeitliche Entwicklung dieser beiden Größen (ausgedrückt in der vergangenen wie potenziellen sozialen Laufbahn)“ (Bourdieu 1987: 195 f.). Im Raum der Lebensstile hingegen geht es um symbolisches Kapital, um Legitimität: „Die soziale Welt ist ein Ort ständiger Kämpfe um den Sinn dieser Welt“ (Bourdieu und Wacquant 2006: 101). Die symbolische Gewalt, die in der Objektivität zweiter Ordnung ausgeübt wird, benötigt immer anerkennende Akteure: „[...] man [kann] die symbolische Gewalt nicht verstehen [...], wenn man nicht den Gegensatz von Zwangsausübung und Zustimmung, äußerem Zwang und innerem Trieb aufhebt“ (a.a.O.: 209). Das soll nicht bedeuten, dass die Positionsunterschiede der Objektivität erster Ordnung negiert werden: Mittels der symbolischen Gewalt wird aber über ihre *Legitimität* entschieden, sie sorgt dafür, dass diejenigen, die nicht von der gegebenen Ordnung profitieren, diese trotzdem anerkennen und nicht hinterfragen.

identifizierten „Fahrstuhleffekte“ werde „die „Klassengesellschaft“ [...] insgesamt eine Etage höher gefahren“ (Beck 2016: 122). Das „kollektive Mehr“ an Wohlstand führe zu einem Bedeutungsverlust sozialer Klassen und Schichten und setze einen Prozess der Individualisierung und Diversifizierung von Lebenslagen und Lebensstilen in Gang (ebd.).

Mit dem Erscheinen dieser beiden Beiträge und den darin aufgezeigten theoretischen Überlegungen wurde der Impuls für eine Forschungsdynamik gesetzt, die neben Einkommen und Beruf auch Lebensstile und Milieus in den Blick nimmt. Diese Dynamik prägte auch die sozialwissenschaftliche Stadt- und Raumforschung. Seit den 1990er Jahren werden hier besonders Fragen nach dem Zusammenhang zwischen Lebensstilen und Wohnstandortwahl, Wohnungsausstattung und räumlicher Mobilität bearbeitet (Rössel und Hoelscher 2012: 306). Der Erklärungswert von Lebensstilen ist dabei durchaus umstritten: So untersucht Klee (2003) in Nürnberg zwar die Ungleichverteilung und das Nebeneinander verschiedener Lebensstile im Raum, kontrolliert aber nicht systematisch Effekte der Sozialstruktur. Scheuner (2006) modelliert bei seinen Analysen des Verkehrsverhaltens in der Region Köln den Effekt soziodemographischer Merkmale (bei ihm als „Lebenslage“ bezeichnet) in seinen Strukturgleichungsmodellen mit und stellt *teilweise* eigenständige Effekte des Lebensstils auf das Verkehrsverhalten fest, wobei die Effekte der Lebenslagen stärker sind (a.a.O.: 60). Auch in Bezug auf eine raumstrukturelle Gegenüberstellung städtischen versus ländlichen Lebens „zeigen sich Effekte eines großstädtischen bzw. dörflichen Kontextes auf die Lebensstile, doch diese Effekte [sind] deutlich geringer als die der sozialstrukturellen Größen“ (Spellerberg 2014: 228).

Die skizzierten Ergebnisse der sozialwissenschaftlichen Stadt- und Raumforschung lassen vermuten, dass Lebensstilansätze zwar ergänzend zu soziodemographischen Ansätzen interessante Einblicke liefern, die stärkeren Effekte jedoch durch „klassische“ soziodemographische Faktoren wie Alter, Einkommen oder Familiensituation zustande kommen. Selbst dort, wo eine stärkere Erklärungskraft der Lebensstile festgestellt wird, wie z. B. bei Dangschat (2018), erscheint die *zusätzliche* Aufklärungsleistung im Vergleich mit der sozialen Lage tatsächlich eher gering (vgl. a. a. O.: Abbildung 1 auf Seite 144). Der Blick auf den aktuellen Forschungsstand zeigt aber auch, dass der Erklärungsgewinn durch Lebensstile und Milieus je nach Gegenstand stark variiert (vgl. ebd.). Beispielsweise legen Blasius und Friedrichs (2011) in Bezug auf Wohnstandortentscheidungen dar, dass die unteren Klassen sich schlicht keine Lebensstilentscheidung leisten kön-

nen: Sie müssen an bezahlbaren Wohnraum kommen.² Dangschat (2018: 148) resümiert mit diesen Überlegungen übereinstimmend, dass zusätzliche Erklärungsgewinne deshalb vor allem dort erwartet werden können, wo die Freiheitsgrade groß sind.

Die Betrachtung von Lebensstilen und Milieus kann also zusammenfassend die „klassische“ vertikale Sozialstrukturanalyse sinnvoll ergänzen. Es besteht darüber hinaus empirische Evidenz, dass die zusätzliche Erklärungskraft sich (1) je nach Gegenstandsbereich (im Rahmen der Stadtforschung z.B. das Thema „Suburbanisierung“ oder „Mobilitätsverhalten“) unterscheidet und (2) von der Gestaltungsfreiheit der Akteure im jeweiligen Bereich abhängt, die wiederum von soziodemographischen Faktoren strukturiert wird.

Im vorliegenden Werkstattbericht wird lediglich der Milieubegriff verwendet. Auf der einen Seite erscheint es inhaltlich bei Kindern unter zehn Jahren nicht sinnvoll, von einem individuellen Lebensstil zu sprechen, da ihre „Freiheitsgrade“ bei der Lebensgestaltung stark familiär und institutionell beschränkt sind. Auch wenn Kinder schon Charaktereigenschaften und Vorlieben für Sport, Musik oder anderes mitbringen – ob sie diesen nachgehen können und inwiefern sich Talente entfalten, ist in diesem Alter stark durch ihr Umfeld bestimmt. Sofern hier von einem Lebensstil gesprochen werden kann, ist dieser also als familiär und weniger als individuell zu bezeichnen. Auf der anderen Seite liegt es aus forschungspragmatischer Sicht nahe, mit dem Milieubegriff zu operieren. Denn die empirische Annäherung erfolgt über die Verwendung der Sinus-Geo-Milieus®. Da hier also Kinder und ihr Kita-Besuch anhand der in ihrer Wohnumgebung geschätzten Sinus-Geo-Milieus® analysiert werden sollen, wird die Rolle des räumlichen und sozialen Umfelds ihrer Wohnumgebung untersucht. Damit ist der Milieubegriff der Begriff der Wahl.

Im Folgenden soll der Frage nachgegangen werden, inwieweit durch die Nutzung und Übertragung der Sinus-Geo-Milieus® auf Kinder eine milieuspezifische Kita-Nachfrage sichtbar wird. Das heißt, es wird der Frage nachgegangen, inwieweit sich Kinder bestimmter Milieus in bestimmten Kitas konzentrieren und inwieweit die Milieuzugehörigkeit möglicherweise die Auswahl einer spezifischen Kita beeinflusst. Mit der Übertragung der kleinräumigen kommerziellen Geodaten auf kommunale Individual-

2 Die aktuell steigenden Immobilienpreise (Ache 2018) tragen hinsichtlich der Wohnstandortentscheidung sicher nicht zu einer Entschärfung dieses Zustands bei.

daten³ wird dabei Neuland beschritten. Daher ist besondere Aufmerksamkeit hinsichtlich der Datengrundlage geboten: Bevor in Kapitel 4 ausgelotet wird, welche Ergebnisse aus der Übertragung auf die spezifische Subpopulation der Kita-Kinder resultieren, soll in Kapitel 3 zunächst die Datengrundlage genauer betrachtet werden.

3 Die „kleinräumigen kommerziellen Geodaten“ sind hier die von microm auf der 100x100-Meter-Rasterebene zur Verfügung gestellten Variablen zu den Sinus-Geo-Milieus®, der Kaufkraft und des Migrationshintergrunds. Diese werden auf bestimmte kommunale Individualdaten übertragen, in diesem Fall auf die Daten derjenigen Kinder, die in Mülheim an der Ruhr eine Kita besuchen.

3 Datengrundlage: Datenabgleich in Mülheim an der Ruhr

Mit dem technischen Fortschritt der letzten Jahrzehnte gewinnen Daten eine immer größere Bedeutung. In der öffentlichen Verwaltung hat sich dies in den letzten Jahren durch die Etablierung der Idee eines „Neuen Steuerungsmodells“ (NSM) bemerkbar gemacht: Hierzu gehört unter anderem die Idee der Input-Output-Steuerung, die mit einem größeren Bedarf an empirischen Daten zur Vorbereitung von Entscheidungen einhergeht (vgl. bspw. Grohs 2012). Gleichzeitig unterliegen amtliche und kommunale Daten strengen gesetzlichen Datenschutzvorschriften und werden sparsam erhoben. Dies erschwert kommunalen Fachplanungen und kommunaler Politik die Erstellung datenbasierter Entscheidungsgrundlagen. Der aktuelle gesellschaftliche Entwicklungsstand hinsichtlich des Umgangs mit Daten sorgt dafür, dass private Unternehmen für die Vermarktung ihrer Waren immer mehr Informationen über die Bevölkerung erhalten und so z. T. über bessere Informationen verfügen als diejenigen Institutionen, die vor Ort die öffentliche Sorge für die Bevölkerung tragen (wie z. B. das Jugend- oder Gesundheitsamt sowie weitere kommunale Institutionen). Wenn man davon ausgeht, dass für diese Schieflage keine zeitnahen, einfachen Lösungen abzusehen sind, ist es für Kommunen grundsätzlich möglich, die Daten der kommerziellen Anbieter einzukaufen und zu nutzen.

Tatsächlich zeichnet sich in den letzten Jahren ab, dass neben privatwirtschaftlichen Akteuren auch zunehmend Wissenschaft und Forschung kommerzielle kleinräumige Geodaten nutzen. An prominenter Stelle ist hier das Deutsche Institut für Wirtschaftsforschung zu nennen, das mit der microm Micromarketing-Systeme und Consult GmbH (microm) zusammenarbeitet. Für die dort durchgeführte Analyse zu „Ausmaß und Trends sozialer Segregation in Deutschland“ im Rahmen der Armuts- und Reichtums-

berichterstattung der Bundesregierung (Goebel und Hoppe 2015) wurden kleinräumige Daten von microm genutzt. Die Autoren verweisen allerdings darauf, dass,

„[w]ährend in vielen europäischen Ländern die amtliche Statistik bereits auf die besonderen Potenziale sozialräumlicher Analysen reagiert hat, indem sie raumbezogene Informationen der wissenschaftlichen Öffentlichkeit zugänglich macht, bislang für den deutschen Kontext kein vergleichbares Angebot an Forschungsinfrastruktur [existiert]“ (a.a.O.: 30).

Hier ist also der Rückgriff auf kommerzielle Daten zu Zwecken der Sozialberichterstattung keinesfalls die Methode der Wahl, sondern auch der Not geschuldet, dass die amtliche Statistik kein umfangreiches Datenangebot bereithält. Auch in der Sozialberichterstattung des Landes Nordrhein-Westfalen 2016 (Jeworutzki, Schräpler und Schweers 2016) wurden kleinräumige microm-Daten verwendet.

Im vorliegenden Bericht konnten zwei Datenquellen zusammengeführt und genutzt werden:

- Kommerzielle Daten: Kleinräumige Daten der microm Micromarketing-Systeme und Consult GmbH zu Kaufkraft, Personen- und Haushaltszahl, Migrationshintergrund und Wahrscheinlichkeit der Zugehörigkeit zu den Sinus-Milieus®. Diese liegen auf einer 100 x 100-Meter-Rasterebene mit Zeitbezug 2016 vor.
- Kommunale Daten: Daten der kommunalen Statistik der Stadt Mülheim an der Ruhr aus dem Einwohnermeldewesen, der Sozialagentur sowie dem Kinder- und Jugendamt.

Kommunale (prozessproduzierte) und kommerzielle Daten bieten jeweils unterschiedliche Vor- und Nachteile. Bei kommunalen Daten ist der Prozess der Entstehung in der Regel gut nachvollziehbar: Die Einwohnermeldedaten werden beispielsweise von den zuständigen Behörden auf gesetzlicher Grundlage (seit Januar 2015 dem Bundesmeldegesetz) erhoben. Das bedeutet nicht, dass sie „fehlerfrei“ sind, wie unter anderem im Zensustest zur Umstellung auf registergestützte Verfahren untersucht wurde: Für den 5. Dezember 2001 wurde für Deutschland insgesamt ein Fehlbestand von 1,7 Prozent und ein Karteileichenbestand von 4,1 Prozent geschätzt, wobei die Fehlerquote in größeren Kommunen höher ist (Statistische Ämter des Bundes und der Länder 2004: 816). Auch wenn das beim Zensus 2011 angewendete Korrekturverfahren zur Melderegister-

bereinigung seither Bestandteil von Widersprüchen und Klagen ist, so ist doch das Verfahren der *Datenerhebung* ein formal geregelter Verwaltungsakt.

Bei kommerziellen Datenanbietern hingegen ist sowohl der Entstehungs- als auch der Bereinigungsprozess der Daten intransparent. Welche Daten genau aus den verschiedenen Quellen von Amtlicher Statistik über die Deutsche Post bis hin zu Immobilienscout auf welche Weise miteinander zusammengeführt werden, bleibt unklar. Die Daten selbst sind jedoch professionell aufbereitet und für jeden, der sie käuflich erwirbt, nutzbar.

Bei Ämtern und Behörden hingegen ist dies, insbesondere wenn keine Statistikstelle eingerichtet wurde,⁴ oftmals schwierig. Einerseits spielen Unsicherheiten über Fragen des Datenschutzes eine Rolle, andererseits Fragen der Datenaufbereitung oder der Kooperation zwischen verschiedenen Ämtern. Möchte beispielsweise das Amt für Kinder und Jugend wissen, aus welchen Stadtteilen die Kinder der verschiedenen Kitas kommen, dann kommt es schon für eine derart einfache Fragestellung ohne eine Zuordnung der (ihm vorliegenden) Adressen zu Stadtteilen nicht weiter. Es muss zunächst mit einem Katasteramt (oder einem anderen Amt, das diese kleinräumige Gliederung/Zuordnung von Adressen zu Stadtteilen verwaltet) kooperieren. Wenn dort ebenfalls Unsicherheiten in Bezug auf die Weitergabe von Daten bestehen, erscheinen langwierige bürokratische Prozesse vorprogrammiert.

Im vorliegenden Fall lag die glückliche Ausgangssituation vor, dass im Rahmen der KeKiz-Begleitforschung sowohl kommunale Daten in einer abgeschotteten Statistikstelle als auch die kommerziellen Daten genutzt werden konnten. Der erste Schritt ist daher zunächst ein Abgleich der beiden Datenquellen bezüglich ähnlicher Merkmale: Es gilt hier die Frage zu beantworten, inwiefern kommunale Statistik und kommerzielle Daten der Firma microm übereinstimmen. Dies erfolgt anhand der Personen- und Haushaltszahl, anhand von Nationalität und Migrationshintergrund sowie schließlich anhand des SGB-II-Bezugs und der Kaufkraft.

4 Die Einrichtung einer abgeschotteten Statistikstelle gewährleistet zunächst, dass personenbezogene Daten dort von fachkundigem Personal aufbereitet werden *können* und zu statistischen und/oder wissenschaftlichen Zwecken genutzt werden *können*.

Datenabgleich auf kommunaler Ebene

Personen- und Haushaltszahl

Laut microm-Daten befinden sich in den Rasterzellen für Mülheim an der Ruhr insgesamt 168.455 Personen, die von microm als sogenannte „Prognoseeinwohner“ geliefert werden. Der Wert liegt damit sowohl nah an den amtlichen Daten (Fortschreibung Zensus 2011 durch das Statistische Landesamt IT.NRW) als auch nah an den kommunalen Daten im Zeitraum 2014 bis 2016.⁵ Tabelle 1 stellt die Bevölkerungszahlen mit dem Stichtag 31.12. für die Jahre 2014 bis 2016 dar.⁶ In Klammern ist jeweils dargestellt, wie stark der Wert der Prognoseeinwohner davon abweicht.

Insgesamt liegen die sogenannten „Prognoseeinwohner“ leicht unter den amtlichen und kommunalen Angaben – ob dies daran liegt, dass bei ersteren unterschätzt oder bei letzteren überschätzt wird, kann an dieser Stelle nicht geklärt werden, beides ist grundsätzlich möglich.

TABELLE 1: Vergleich der Bevölkerungszahlen amtliche und kommunale Statistik, 2014 bis 2016

Zeitpunkt	Statistisches Landesamt IT.NRW	Kommunale Statistik Einwohnermeldewesen		
	Fortschreibung Zensus 2011	Erstwohnsitz	Zweitwohnsitz	Insgesamt
31.12.2014	167.108 (+0,81%)	168.690 (-0,14%)	1.933	170.623
31.12.2015	169.278 (-0,49%)	170.901 (-1,43%)	760	171.661
31.12.2016	170.936 (-1,45%)	172.593 (-2,40%)	753	173.346

In Klammern: Prozentualer Vergleich zu den microm-Prognoseeinwohnern 2016.

Daten: IT.NRW; Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr, eigene Berechnung. Stichtag 31.12.

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

5 Zensus 2011: 166 804 (Tabelle 12410-01i IT.NRW), Kommunalstatistik: 168.566 Personen der Bevölkerung am Ort der Hauptwohnung.

6 Während amtliche und kommunale Daten in der Regel mit dem Stichtag 31.12. dargestellt werden, gibt es für die kommerziellen Daten lediglich ein Bezugsjahr, in diesem Fall 2016.

Neben der Angabe zu sogenannten „Prognoseeinwohnern“ wurden von microm auch Informationen zu Haushalten bezogen. Dies ist insofern interessant, als dass in der amtlichen oder kommunalen Statistik keine prozessproduzierten Haushaltsdaten vorliegen. Die kommunale Statistik kann jedoch ein sogenanntes „Haushaltegenerierungsverfahren“ durchführen und aus den Daten des Einwohnermeldewesens Schätzungen zu Haushalten auf kleinräumiger Ebene erhalten.⁷

Die amtliche Statistik führt im Rahmen des Zensus (und damit nach aktueller Planung alle zehn Jahre) ebenfalls eine Haushaltegenerierung durch. Auf Kreisebene werden dabei anhand der Kombination aus Zensusergebnissen, Bevölkerungsfortschreibung und der 1-Prozent-Bevölkerungsstichprobe des Mikrozensus entsprechende Hochrechnungen vorgenommen. Eine kleinräumige Haushaltegenerierung gibt es hier jedoch nicht. Aufgrund des Stichprobencharakters des Mikrozensus sind die Hochrechnungen zudem mit Unsicherheit behaftet und über die Jahre unter Umständen nicht vergleichbar.⁸ Hinzu kommt, dass die Veröffentlichung der Schätzung in verhältnismäßig groben Tausenderschritten erfolgt.

Microm schätzt 83.984 Haushalte für 2016 – Tabelle 2 stellt die von IT.NRW veröffentlichten Hochrechnungen des Mikrozensus und die generierten Haushaltszahlen der kommunalen Statistik der Stadt Mülheim an der Ruhr gegenüber. In Klammern findet sich die prozentuale Abweichung des microm-Wertes zur jeweiligen Angabe.

Auffällig ist hier keinesfalls die Abweichung von kommerziellen und amtlichen Daten, sondern vielmehr die offensichtlich „gute“ Anpassung der kommerziellen an die kommunalen Daten sowie die davon verhältnismäßig stark abweichenden hochgerechneten Mikrozensuswerte.

7 Beispielsweise hat der Verbund Kommunales Statistisches Informationssystem (KOSIS-Verbund) das Haushaltegenerierungsverfahren HHGEN entwickelt, das von vielen der beteiligten Kommunen genutzt wird. Es bestehen darüber hinaus alternative Verfahren. Die Grundzüge sind online auf der gemeinsamen Seite von KOSIS und dem VdSt (Verband Deutscher Städtestatistiker) dargestellt: <http://www.staedtestatistik.de/284.html>. Mülheim an der Ruhr führt seit mehreren Jahren ein dort entwickeltes alternatives Verfahren zur Haushaltegenerierung durch.

8 Die aktuelle Tabelle der Landesdatenbank beinhaltet beispielsweise folgenden Hinweis: „Ergebnisse des Mikrozensus; ab 2011 basiert die Hochrechnung auf den fortgeschriebenen Ergebnissen des Zensus 2011. Durch Umstellung auf eine neue Stichprobe sowie durch Sondereffekte im Kontext der Bevölkerungsentwicklung ist die Vergleichbarkeit der Mikrozensusergebnisse für das Berichtsjahr 2016 mit den Vorjahren eingeschränkt.“

TABELLE 2: Vergleich der Haushaltszahlen amtliche und kommunale Statistik der Jahre 2014 bis 2016

Zeitpunkt	Statistisches Landesamt IT.NRW Hochrechnung Mikrozensus*	Kommunale Statistik Einwohnermeldewesen Insgesamt
31.12.2014	91.000 (– 7,71 %)	83.590 (+ 0,47 %)
31.12.2015	89.000 (– 5,63 %)	85.577 (– 1,86 %)
31.12.2016	83.000 (+ 1,19 %)	86.516 (– 2,93 %)

* auf 1.000 gerundet | In Klammern: Prozentualer Vergleich zu den microm-Prognoseeinwohnern 2016.

Daten: IT.NRW; Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr, eigene Berechnung. Stichtag 31.12.

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

Migrationshintergrund

Die Notwendigkeit einer statistischen Erfassung des „Migrationshintergrunds“ und die entsprechende Datenverfügbarkeit ist in den letzten Jahrzehnten zunehmend diskutiert worden. Lange Zeit ist in der amtlichen und kommunalen Statistik lediglich die Staatsbürgerschaft erhoben worden (vgl. VDSt 2013). Da Integrationsbedarfe, Diskriminierung und Sprachbarrieren sich allerdings nur bedingt über die Staatsangehörigkeit messen lassen, wurde ein neues statistisches Konzept benötigt. Mittlerweile existiert in unterschiedlichen Bereichen eine Vielzahl an Erhebungsformen (über Geburtsland und / oder Nationalität der Eltern, Verkehrssprache etc., für einen Überblick vgl. a. a. O.). Nach amtlicher Definition des Statistischen Bundesamtes besitzt eine Person einen Migrationshintergrund,

„[...] wenn sie selbst oder mindestens ein Elternteil nicht mit deutscher Staatsangehörigkeit geboren wurde. Im Einzelnen umfasst diese Definition zugewanderte und nicht zugewanderte Ausländerinnen und Ausländer, zugewanderte und nicht zugewanderte Eingebürgerte, (Spät-) Aussiedlerinnen und (Spät-) Aussiedler sowie die als Deutsche geborenen Nachkommen dieser Gruppen“ (Statistisches Bundesamt 2018).

An dieser Definition orientiert sich auch microm und stellt Daten zum Migrationshintergrund des Haushaltsvorstandes auf kleinräumiger Ebene zur Verfügung.

Die amtliche und kommunale Verfügbarkeit dieses Merkmals ist (beispielsweise im Vergleich zu den Merkmalen Alter und Nationalität) schwierig: Auf kommunaler Ebene gibt

es Ergebnisse des Zensus 2011 als Hochrechnung aus einer Haushaltsstichprobe.⁹ Für Kreise und kreisfreie Städte können diese Ergebnisse anhand des Mikrozensus aktualisiert werden, aber auch hier führt die 1-Prozent-Stichprobenziehung des Mikrozensus möglicherweise zu Verzerrungen.

Auf kommunaler und kleinräumiger Seite kann aus dem Einwohnermeldewesen der Migrationshintergrund z. B. über das Verfahren MigraPro (wie HHGEN ein KOSIS-Produkt, vgl. Fußnote 7) generiert werden. Allerdings setzt dieses Verfahren voraus, dass die Geburtsländer beider Elternteile dokumentiert sind – eine Ausgangslage, die in der Praxis mit Problemen behaftet ist.¹⁰ In Mülheim an der Ruhr wurde zum Zeitpunkt der Datenauswertung kein MigraPro durchgeführt, sodass lediglich Informationen zur Staatsbürgerschaft zugänglich waren.

Für den gesamtstädtischen Wert gab microm 2006 an, dass 27,5 Prozent der Haushalte einen Vorstand mit Migrationshintergrund hatten.

Auf *Personenebene* stellt Tabelle 3 als Vergleichsgrößen (1) die Werte des Integrationsmonitorings des Ministeriums für Kinder, Familie, Flüchtlinge und Integration NRW (MKFFI) und (2) die kommunalen Angaben zu Staatsbürgerschaften dar. Über die Staatsbürgerschaften (ausländische oder deutsche und eine weitere) erfasst man lediglich einen Teilbereich der Personen mit Migrationshintergrund – der Abstand von sechs Prozentpunkten zum microm-Wert erscheint daher nicht überraschend und grundsätzlich plausibel.

Beim letzten verfügbaren Wert aus dem Integrationsmonitoring (Jahr 2015) hingegen besteht ein Abstand von drei Prozentpunkten zum microm-Wert.¹¹ Ein Teil dieses Abstands kann durch die Anzahl der Menschen mit Migrationshintergrund erklärt

9 Dann allerdings mit einer zusätzlichen Einschränkung durch den Zeitraum: „Als Personen mit Migrationshintergrund werden alle zugewanderten und nicht zugewanderten Ausländer/-innen sowie alle nach 1955 auf das heutige Gebiet der Bundesrepublik Deutschland zugewanderten Deutschen und alle Deutschen mit zumindest einem nach 1955 auf das heutige Gebiet der Bundesrepublik Deutschland zugewanderten Elternteil definiert“ (<https://ergebnisse.zensus2011.de/#Glossary:>).

10 Beispielsweise sind diese vor allem bei älteren Datensätzen/älteren Personen oder bei unbekannter Vaterschaft nicht oder nicht vollständig dokumentiert.

11 Im entsprechenden Dokument wird als Jahresgrenze der Zuwanderung allerdings das Jahr 1950 angegeben, was der Konvention des Statistischen Bundesamtes entspricht, allerdings nicht der Definition aus dem Zensus 2011.

TABELLE 3: Anteil der Menschen mit Migrationshintergrund in Mülheim an der Ruhr, 2012 bis 2016

Zeitpunkt	Ministerium für Kinder, Familie, Flüchtlinge und Integration NRW (MKFFI) Integrationsmonitoring NRW	Kommunale Statistik: Personen mit ausländischer oder doppelter Staatsangehörigkeit am 31.12.
2012	24,4 %	18,3 %
2013	23,7 %	19,0 %
2014	21,5 %	19,9 %
2015	24,5 %	21,5 %
2016	26,2 %	22,6 %

Aufgrund einer Umstellung der Stichprobe 2016 sind die Werte 2016 nicht mit den vorherigen Werten vergleichbar.
Daten: MKFFI; Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

werden, die im Zuge der jüngsten Zuwanderungsbewegungen aus Krisengebieten gestiegen ist. Allerdings kann diese Steigerung nicht die Größenordnung von drei Prozentpunkten erklären – so sind im Jahr 2016 lediglich 905 Geflüchtete neu nach Mülheim an der Ruhr gezogen. Dies entspricht einem Bevölkerungsanteil von circa einem halben Prozent. Microm liegt damit circa zwei bis zweieinhalb Prozent höher als die amtlichen Angaben. Diese leiden ihrerseits unter dem Stichprobencharakter des Mikrozensus, sodass der Abstand vernachlässigbar erscheint.

Eine weitere mögliche Erklärung für den Abstand zwischen den Werten ist die unterschiedliche Messebene: Haushaltsvorstände versus Personen. Dies würde eine Überschätzung aber nur dann erklären, wenn überproportional viele Menschen mit Migrationshintergrund als Haushaltsvorstand fungieren würden – eine wenig plausible Annahme.

Um den Unterschied zwischen Haushalte- und Personenverteilung greifbarer zu machen, wurden aus der Mülheimer Haushaltegenerierung des Stichtags 31.12.2014 experimentell für die Unterscheidung nach Staatsbürgerschaft drei Szenarien geschätzt:

- mindestens eine Person im Haushalt hat eine ausländische oder doppelte Staatsbürgerschaft: 19,8 Prozent der Haushalte
- alle Personen im Haushalt haben eine ausländische oder doppelte Staatsbürgerschaft: 13,1 Prozent der Haushalte

- mindestens die Hälfte der im Haushalt lebenden Personen hat eine ausländische oder doppelte Staatsbürgerschaft: 18,4 Prozent der Haushalte

Die „schwächste“ Konstruktion eines „Migrationshintergrunds“ auf Haushaltsebene, bei der nur eine der im Haushalt lebenden Personen eine ausländische Staatsbürgerschaft hat, liegt nahe dem Personenanteilswert, alle anderen Varianten deutlich darunter. Die Unterschiedlichkeit der Messebenen (Haushalte versus Personen) taugt damit nicht zur Erklärung des Abstands.

Im Ergebnis zeigt sich eine insgesamt gute Passung der Datensätze mit Blick auf den Migrationshintergrund; möglicherweise wird der Migrationshintergrund in den kommerziellen Daten leicht überschätzt – die Abweichung ist aber gering.

Kaufkraft: Armut und Reichtum

Die (kleinräumige) Kaufkraftinformation liefert großes Analysepotenzial für die Stadtforschung. In der amtlichen Statistik gibt es bislang vor allem zwei Quellen für kleinräumige Berichterstattung zu Armut und Reichtum:

- Daten aus dem Sozialleistungsbezug (z. B. SGB II/SGB XII/AsylbLG): Diese sind in der Sozialberichterstattung etabliert und lassen sich für den Bereich des SGB II für Kommunen und Kreise von der Bundesagentur für Arbeit (BA) beziehen.¹² Im Gegensatz zur Kaufkraft bilden diese Daten jedoch lediglich eine Seite der Medaille räumlich strukturierter Einkommensverteilung ab, nämlich die der Armut sowie hierunter nur den Transferleistungsbezug. Damit wird folglich weder die sogenannte „Dunkelziffer der Armut“¹³ abgebildet noch der Teil der

¹² Es sei denn, es handelt sich um „Optionskommunen“ (auch: „zugelassene kommunale Träger“). So werden Kommunen bezeichnet, die die alleinige Trägerschaft der SGB-II-Leistungen besitzen. Mit der Bundesagentur für Arbeit tauschen sie Informationen und Daten aus, die Vermittlung und Eingliederung erwerbsfähiger Hilfebedürftiger liegt jedoch alleine in ihrer Verantwortung (eine Aufgabe, die in den „Gemeinsamen Einrichtungen“ kommunaler Träger und Bundesagentur für Arbeit die Letztgenannte übernimmt). Neben der größeren kommunalen Gestaltungsfreiheit ist damit auch der kommunale Träger „Herr der Daten“, wobei für Verknüpfung und Analyse hohe datenschutzrechtliche Standards erfüllt werden müssen (vgl. Groos et al. 2018). Da die Daten vor Ort vorliegen, müssen sie dementsprechend nicht erst von der Bundesagentur bezogen werden, sondern können in kommunaler Eigenregie kleinräumig aufbereitet werden.

¹³ Dieser Begriff beschreibt das Dunkelfeld der Nichtinanspruchnahme von Sozialleistungen. Hiermit werden somit alle Personen umfasst, die Sozialleistungen beziehen könnten, diese aber nicht beantragen (vgl. z. B. Becker 2007).

Bevölkerung, der *relativ* arm ist, aber noch oberhalb des gesetzlich garantierten Existenzminimums lebt. Tendenziell wird das Ausmaß an Armut mit dieser Messmethode daher insgesamt unterschätzt.

- Daten aus der Lohn- und Einkommenssteuerstatistik: Diese sind aus der amtlichen Statistik „die einzige verfügbare Datenquelle, mit der Einkommensstrukturen und Verteilungen auch in tiefer regionaler Verteilung abgebildet werden können“ (Schräpler und Seifert 2008: 8). In Nordrhein-Westfalen konnten die Kreise und Kommunen diese für einige Jahre auch kleinräumig (mit gewissen Einschränkungen im oberen Einkommensbereich) von IT.NRW beziehen. Allerdings haben sich diese Daten für die kleinräumige kommunale Sozialberichterstattung bislang nicht etabliert. Gründe hierfür könnten in der Komplexität der Datenlage, der aufgrund der Bearbeitungszeiten geringen Aktualität und den Problemen bei der zeitlichen Vergleichbarkeit der Daten liegen (vgl. a. a. O.: 8 ff.). Für aktuelle Jahre werden die Daten derzeit nicht mehr zur Verfügung gestellt.

Auf der größeren Ebene der Kommunen und Kreise liegt als angemessene Vergleichsgröße in Nordrhein-Westfalen für die Kaufkraft pro Person das verfügbare Einkommen pro Person vor, die Definitionen laut Sozialberichterstattung und microm sind hierbei äquivalent.¹⁴ Das verfügbare Einkommen pro Person wurde im Rahmen der Landes-sozialberichterstattung in Nordrhein-Westfalen vom Arbeitskreis „Volkswirtschaftliche Gesamtrechnungen der Länder“ zur Verfügung gestellt und durch das Ministerium für Arbeit, Gesundheit und Soziales für die Jahre 2009 bis 2015 veröffentlicht.¹⁵ Das statistische Landesamt in Nordrhein-Westfalen (IT.NRW) hat im September 2018 aktualisierte Daten, auch auf Gemeindeebene, veröffentlicht.

In Mülheim an der Ruhr stieg das verfügbare Einkommen pro Person in den letzten Jahren kontinuierlich: Gemäß IT.NRW betrug es 22.460 Euro im Jahr 2014, 22.718 Euro

14 Die Definition aus der Sozialberichterstattung lautet exakt: „Das verfügbare Einkommen der privaten Haushalte ergibt sich, wenn dem Primäreinkommen der privaten Haushalte einerseits diejenigen monetären Sozialleistungen (z. B. Alters- und Hinterbliebenenversorgung, Arbeitslosengeld) und sonstigen laufenden Transfers hinzugefügt werden, welche die privaten Haushalte überwiegend seitens des Staates empfangen, und andererseits die laufenden Transfers, die von den privaten Haushalten selbst zu leisten sind (z. B. Einkommen- und Vermögensteuern, Sozialbeiträge) abgezogen werden. Das verfügbare Einkommen der privaten Haushalte entspricht somit den Einkommen, die den privaten Haushalten letztendlich zur Verfügung stehen, um es für Lebensunterhalt, soziale Teilhabe und Sparzwecke zu verwenden“ (http://www.sozialberichte.nrw.de/sozialindikatoren_nrw/indikatoren/4_einkommensentwicklung/indikator4_5/index.php.)

15 vgl. Sozialberichterstattung online (http://www.sozialberichte.nrw.de/sozialindikatoren_nrw/indikatoren/4_einkommensentwicklung/indikator4_5/index.php).

im Jahr 2015 und im Jahr 2016 schließlich 23.091 Euro. Die von microm auf Rasterebene zur Verfügung gestellte Kaufkraft kann als mit den Bevölkerungszahlen in den Zellen gewichteter Mittelwert berechnet werden und beträgt dann 23.732,68 Euro. Microm schätzt die Kaufkraft also etwas höher ein als die amtliche Statistik. Der bereits im Frühjahr 2017 von microm zur Verfügung gestellte Wert liegt dabei nah am kürzlich publizierten amtlichen Wert.

Ergebnis des Datenabgleichs auf kommunaler Ebene

Für die vier Variablen der Bevölkerung, der Haushalte, des Migrationshintergrunds und der Kaufkraft liegen die kommerziellen Daten sehr nah an den amtlichen Vergleichsgrößen. Der Migrationshintergrund wird möglicherweise leicht überschätzt – bedenkt man allerdings die Abweichungen bei der Haushaltszahl zwischen Mikrozensus-schätzungen von IT.NRW und der kommunalen Haushalte, erscheinen Abweichungen von circa zwei Prozentpunkten auf der Ebene der Gesamtstadt als relativ genau. Abgesehen von der Bevölkerungszahl aus dem Einwohnermeldewesen liegt keine der betrachteten Variablen zwangsläufig in der Kommune vor: Migrationshintergrund und Haushalte müssen zunächst aus den Einwohnermeldedaten generiert werden, die Kaufkraftinformation gibt es generell nicht. Kommunen haben für Analysen der Armuts- und Reichumsverteilung zwar die Möglichkeit, kleinräumige Daten zum Transferleistungsbezug von der Bundesagentur für Arbeit (BA) – und in Nordrhein-Westfalen auch in vergangenen Jahren die Möglichkeit gehabt –, kleinräumige Daten aus der Lohn- und Einkommensteuerstatistik zu beziehen, müssen diese aber ebenfalls zunächst extern beschaffen.¹⁶ Dabei besteht bei Ersteren das Problem, dass sie mit dem Transferleistungsbezug Armut generell unterschätzen, indem sie nur einen Teilbereich abbilden. Bei Letztgenannten ist zudem auf das Problem fehlender Aktualität und ungewissen Datenzugangs hinzuweisen.

Bislang ging es darum, zu eruieren, inwiefern die Daten auf der Ebene der Stadt Mülheim an der Ruhr passende Beschreibungen liefern. Hier scheint ein stimmiges Bild gezeichnet zu werden. Das Besondere ist jedoch die *Kleinräumigkeit* der Daten. Im Folgenden soll daher skizziert werden, inwieweit die Ergebnisse auch auf der Ebene der Rasterzellen plausibel erscheinen.

¹⁶ Im Fall der BA-Daten entfällt dies, insofern die Kommune zu den sogenannten „Optionskommunen“ gehört, also zu denjenigen Kommunen, die selbst die Trägerschaft für die Grundsicherung für Arbeitssuchende innehaben.

Datenabgleich auf Rasterebene

Wie im vorangegangenen Teil sollen die Variablen zu Bevölkerung, Haushalten, Migrationshintergrund und Kaufkraft auf ihre Stimmigkeit hin untersucht werden. Bei den Bevölkerungs- und Haushaltszahlen ist dies vergleichsweise einfach, da passende Vergleichsvariablen vorhanden sind. Für den Vergleich des Migrationshintergrunds wurde die Bevölkerung mit nicht deutscher Staatsangehörigkeit gewählt, für die Kaufkraft der SGB-II-Bezug. Aus Gründen der Datenverfügbarkeit wurden für sämtliche folgende Auswertungen die Werte vom 31.12.2014 gewählt.

Die Punktkoordinaten der Bevölkerung, der Haushalte, der Nationalität und des SGB-II-Bezugs aus der abgeschotteten Statistikstelle des Referats V.1 – Stadtforschung und Statistik – wurden mithilfe der freien Software QGIS in die Rastergeometrie von microm aggregiert, um miteinander verglichen werden zu können. Dabei ist anzumerken, dass die Geometrien nicht vollständig übereinstimmen, d.h., dass nicht alle Einwohneradresskoordinaten auch in einer der Rasterzellen liegen. Der Ausfall ist allerdings minimal: Bezüglich der Einwohner handelt es sich um 273 Personen, die nicht zugeordnet werden konnten. Dies entspricht einem Anteil von 0,16 Prozent. Auf der Haushaltsebene konnten von den 83.590 Haushalten 83.460 zugeordnet werden. Dies entspricht einem Anteil von 0,2 Prozent der Haushalte, die nicht in den Rasterzellen verortet werden konnten.

Tabelle 4 stellt für die Rasterzellen den Vergleich typischer statistischer Kennwerte zusammen. Es ist eine sehr gute Passung erkennbar – die Mittelwerte sind nahezu identisch, auch Summe und Streuung liegen eng beieinander. Die vergleichsweise große Abweichung bei den Maximalwerten kommt durch die Rasterzelle zustande, in der der Campingplatz „Entenfang“ liegt – dort haben sich entgegen gesetzlicher Bestimmungen einige Menschen häuslich niedergelassen. Tatsächlich war gerade das Jahr 2016, also das Bezugsjahr der microm-Daten, auch das Jahr, in dem die Stadt Mülheim an der Ruhr die dortigen Siedler zum Aufgeben des Erstwohnsitzes aufgefordert hat. Sowohl der zeitliche Verzug als auch die ungewöhnliche Konstellation von Dauersiedlern auf einem Campingplatz erklären also die zunächst groß erscheinenden Abweichungen beim Maximalwert und der Spannweite.

Der statistische Zusammenhang zwischen den kommunalen und den kommerziellen Daten ist dementsprechend auch sehr stark: Der Korrelationskoeffizient beträgt

TABELLE 4: Vergleich der Bevölkerung und der Haushalte auf Rasterebene

	Bevölkerung Einwohnermeldewesen Mülheim an der Ruhr	Bevölkerung microm (Prognoseeinwohner)	Haushalte aus Haushaltegenerierung Mülheim an der Ruhr	Haushalte microm
Minimum	0	0	0	0
Maximum	536	471	168	256
Spannweite	536	471	168	256
Summe	168.417	168.455	83.460	83.984
Median	32,0	32,0	15,0	16,0
Mittelwert	42,5	42,6	21,1	21,2
Standard- abweichung	43,7	41,7	22,3	20,9
gültige Werte	3.960	3.954	3.960	3.960
fehlende Werte (null)	337	31	352	37
fehlende Werte (keine Angaben)	0	6	0	0

Daten: microm; Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr; eigene Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

$r = 0,93$ (Bevölkerung) bzw. $r = 0,92$ (Haushalte) bei einem rechnerisch möglichen Maximalbetrag von 1 und einem Minimalbetrag von 0.

Die Passung der kleinräumigen Datengrundlage hinsichtlich Bevölkerungs- und Haushaltszahlen kann damit als sehr gut bezeichnet werden.

Schwieriger abzugleichen sind hingegen Migrationshintergrund und Kaufkraft, weil die von der Firma microm gelieferten Daten hier inhaltlich über das hinausgehen, was Kommunen zur Verfügung steht. Zudem ist der Migrationshintergrund als Haushaltsmerkmal zunächst kontraintuitiv und die Rolle eines „Haushaltsvorstandes“, über den man diesen festmachen sollte, nicht im Einwohnermeldewesen festgehalten. Beim Migrationshintergrund auf Haushaltsebene gibt es auf Grundlage der durch das Referat V.1 in Mülheim an der Ruhr durchgeführten Haushaltegenerierung die bereits genannten drei Möglichkeiten der Konstruktion eines „Haushaltes mit Migrationshintergrund“:

- mindestens eine Person im Haushalt hat eine ausländische oder doppelte Staatsbürgerschaft: 19,8 Prozent der Haushalte
- alle Personen im Haushalt haben eine ausländische oder doppelte Staatsbürgerschaft: Anteil 13,1 Prozent der Haushalte
- mindestens die Hälfte der im Haushalt lebenden Personen haben eine ausländische oder doppelte Staatsbürgerschaft: 18,4 Prozent der Haushalte

Möchte man die Rasterwerte miteinander vergleichen, ist dies anhand absoluter oder relativer Werte möglich. Bei den absoluten Werten ergeben sich trotz der konzeptionellen Unterschiede noch verhältnismäßig hohe Korrelationskoeffizienten von ungefähr 0,8. Allerdings spiegeln absolute Zahlen auch immer die Bevölkerungs- bzw. Haushaltsdichte in den Zellen wider. Insofern erscheint ein Vergleich der relativen Zahlen angemessener. Hier besteht jedoch der Nachteil, dass sich bei einer feinen Gliederung (durchschnittlich circa 20 Haushalte pro Zelle) schon kleine absolute Änderungen sehr stark auswirken. Wenn in einer Rasterzelle beispielsweise nur fünf Haushalte leben, macht ein einziger „Haushalt mit Migrationshintergrund“ mehr oder weniger schon eine Änderung des Anteils von 20 Prozentpunkten aus. Bei so kleinen Fallzahlen wirkt sich unter Umständen auch die Problematik des unterschiedlichen Messkonzepts („Migrationshintergrund“ nach der Definition des Statistischen Bundesamtes versus Staatsbürgerschaft als Indikator für Migrationshintergrund) stärker aus. Deshalb wurden mit den relativen Werten verschiedene Zusammenhangsberechnungen unter Ausschluss der Zellen mit geringen Bevölkerungszahlen durchgeführt. Tabelle 5 fasst die Ergebnisse zusammen.

Während die absoluten Werte starke statistische Zusammenhänge zeigen, offenbart der Vergleich mit den Anteilswerten, dass dies zu einem Großteil auf die gleichzeitige Messung der Bevölkerungsdichte zurückzuführen ist. Bei den relativen Werten wiederum ist der Zusammenhang eher schwach – Werte von unter 0,5 bilden einen nur mittelstarken statistischen Zusammenhang ab. Legt man allerdings nur die Zellen zugrunde, in denen mindestens 20 Personen gemeldet sind, wächst der Korrelationskoeffizient auf über 0,6 an.

Das heißt: Der von microm zur Verfügung gestellte „Migrationshintergrund auf Haushaltsebene“ hängt in den kleinräumigen Rasterzellen zwar durchaus statistisch mit kommunalen Konstruktionsvarianten eines „Migrationshintergrunds auf Haushaltsebene“, die lediglich über die Staatsbürgerschaften abgebildet werden können, zusammen. Es gibt kleinräumig allerdings auch deutliche Abweichungen: Der Zusammenhang

TABELLE 5: Korrelationskoeffizienten zwischen verschiedenen Konstruktionen von „Haushalten mit Migrationshintergrund“

Konstruktion aus der kommunalen Haushaltegenerierung	Anteil Kommune 31.12.2014	Korrelation absolute Werte in Zellen	Korrelation absolute Werte in Zellen	Korrelation absolute Werte in Zellen mit n>10	Korrelation absolute Werte in Zellen mit n>20
... mindestens eine Person im Haushalt hat eine ausländische oder doppelte Staatsbürgerschaft	19,8 %	0,81	0,45	0,57	0,61
... alle Personen im Haushalt haben eine ausländische oder doppelte Staatsbürgerschaft	13,1 %	0,79	0,46	0,57	0,62
... mindestens die Hälfte der im Haushalt lebenden Personen hat eine ausländische oder doppelte Staatsbürgerschaft	18,4 %	0,81	0,47	0,58	0,63

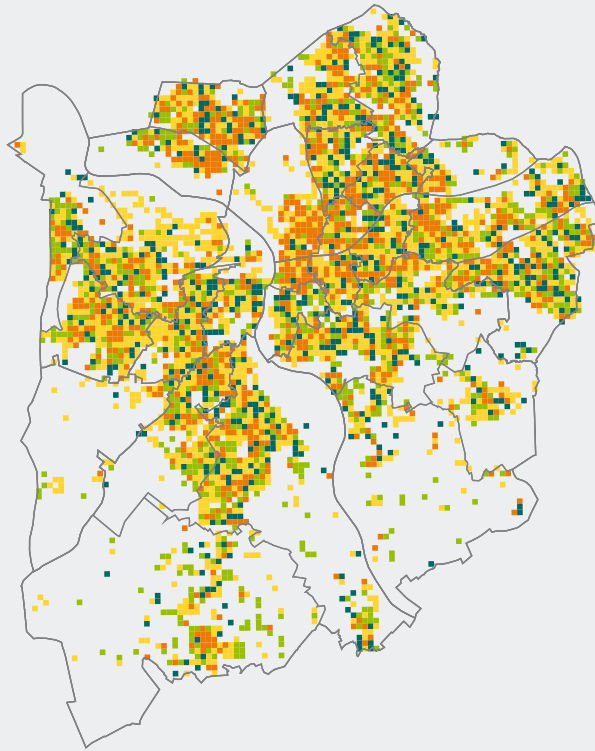
Daten: microm; Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr; eigene Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

ist nicht sehr stark. Dies kann daran liegen, dass verschiedene Konzepte (Staatsbürgerschaft/Migrationshintergrund) zugrunde gelegt wurden.

Ergänzend soll daher eingeschätzt werden, inwiefern die Abweichungen und Zusammenhänge plausibel sind. Einen Hinweis hierauf kann man erhalten, indem die Abweichungen in den Prognosewerten der beiden Variablen hinsichtlich ihrer innerstädtischen Verteilung betrachtet werden. Dabei wird der Anteil an Haushalten mit Migrationshintergrund nach kommunaler Konstruktion durch die microm-Variable statistisch vorhergesagt und anschließend untersucht, inwiefern die statistische Vorhersage vom tatsächlichen Wert abweicht. Das heißt, die beiden Variablen werden auf diese Art und Weise miteinander verglichen, um die Güte der räumlichen Verortung der „Haushalte mit Migrationshintergrund“ einschätzen zu können. Das Ergebnis ist Abbildung 1 zu entnehmen. Abweichungen im Rahmen von fünf Prozentpunkten auf dieser Ebene wurden als geringe Abweichung klassifiziert. Abweichungen zwischen fünf und 15 Prozentpunkten wurden als Über- beziehungsweise Unterschätzungen sowie alles über 15 Prozentpunkten als „stark“ über- beziehungsweise unterschätzt klassifiziert.

ABBILDUNG 1: **Abweichungen zwischen empirischen und erwarteten Werten bei „Haushalten mit Migrationshintergrund“**



Abweichungen ■ stark unterschätzt ■ unterschätzt ■ wenig Abweichung ■ überschätzt ■ stark überschätzt

Daten: microm; Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr; eigene Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

Hierbei zeigt sich eine deutliche Systematik: Die dunkelbraunen Bereiche finden sich vermehrt im Stadtteil Styrum und der Altstadt.¹⁷ Dies sind in Mülheim an der Ruhr Stadtteile, in denen tatsächlich überdurchschnittlich viele Ausländer und Doppelstaatler leben. Aufgrund des weicher definierten „Migrationshintergrunds“ resultieren also deutlich niedrigere Erwartungswerte für die Stadtteile, die bekanntermaßen die höchsten Ausländeranteile aufweisen. Umgekehrt gibt es in den südöstlichen Stadtteilen gleichwohl Menschen mit Migrationshintergrund, die schon länger in Deutschland leben, sodass es bei den Staatsbürgerschaften sehr viel weniger sind, als sich anhand des Migrationshintergrunds erwarten ließe. Das Muster der Abweichungen zwischen Migrationshintergrund der microm-Daten einerseits und dem über Staatsbürgerschaft konstruierten Migrationshintergrund andererseits erscheint damit durchaus plausibel.

Für einen Eindruck der kleinräumigen Darstellung eignen sich beide Variablen: Abbildung 2 zeigt links eine thematische Karte mit den microm-Daten, rechts die aus den Haushalten konstruierte Variante. In thematischen Karten wird ein bestimmtes Merkmal – hier die unterschiedlich konstruierten Migrationshintergrundanteile der Haushalte – abgebildet. Bei der Verbildlichung gilt es grundsätzlich, zwei Entscheidungen zu treffen:

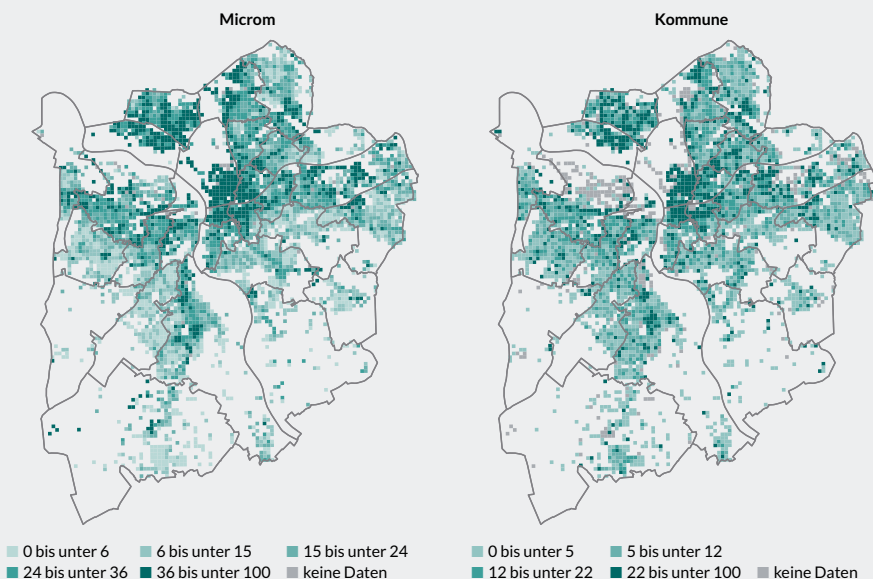
- Wie sollen die Einheiten, in diesem Falle Rasterzellen, unterschieden werden? Da Anteilswerte viele verschiedene Ausprägungen (alle Prozentwerte zwischen 0 und 100) annehmen können, gilt es hier, eine Entscheidung für eine Klassifizierung zu treffen. Es können beispielsweise alle Rasterzellen, in denen der Anteilswert kleiner als 10 Prozent ist, zu einer Klasse zusammengefasst werden, dann alle diejenigen, die Anteilswerte zwischen 10 bis unter 20 Prozent aufweisen zu einer weiteren Klasse usw. Anstelle einer solchen „willkürlichen Grenzziehung“ können jedoch auch „verteilungsorientierte“ Verfahren angewendet werden. Verteilungsorientiert bedeutet, dass sich die Klassenbildung daran orientiert, wie das Merkmal in den Rasterzellen empirisch ausgeprägt ist. Ein Verfahren, das sich aufgrund seiner vergleichsweise einfachen Konstruktion und Interpretierbarkeit in den letzten Jahren durchgesetzt hat, ist die Darstellung in Quintilen. Häufig findet man hierfür auch die Bezeichnung „gleiche Klassenbesetzung“. Hier wird so gruppiert, dass diejenigen 20 Prozent der Rasterzellen zusammengefasst werden, die die geringsten Anteilswerte aufweisen. Anschließend werden diejenigen 20 Prozent der Zellen zusammengefasst, die die zweit-

17 Zur Lage der Stadtteile vgl. Abbildung 7, Anhang.

geringsten Anteilswerte besitzen, bis schließlich die 20 Prozent der Zellen mit den höchsten Anteilswerten zur letzten Gruppe zusammengefasst werden. Die Anzahl an Gruppen ist dabei variabel. Allerdings gilt zu beachten, inwiefern die Zellen empirisch tatsächlich unterschiedlich sind, sodass eine feine Unterteilung sinnvoll erscheint.

- Welche Farbskala soll ausgewählt werden? In diesem Fall liegen Anteilswerte vor, also ein Merkmal, das „sequenziell“ ist: Es kann gering oder stark, mehr oder weniger ausgeprägt sein – im Gegensatz z.B. zur am stärksten gewählten Partei, dies kann beispielsweise die SPD, CDU/CSU oder irgendeine andere Partei sein, jedoch nicht „mehr“ oder „weniger“ Partei. Für die Darstellung wurde daher auch eine sequenzielle Skala aus Grüntönen gewählt. Eine geringe Farbdichte zeigt an, dass der Anteil von Haushalten mit Migrationshintergrundanteil gering ist, eine dunklere Farbgebung bedeutet, dass er hier vergleichsweise hoch ist. Da es

ABBILDUNG 2: Kartographischer Vergleich „Haushalt mit Migrationshintergrund“



Daten: microm; Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr; eigene Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

sich um zwei unterschiedliche Merkmale mit unterschiedlichem Niveau handelt, wurde separat in Quintilen klassifiziert.

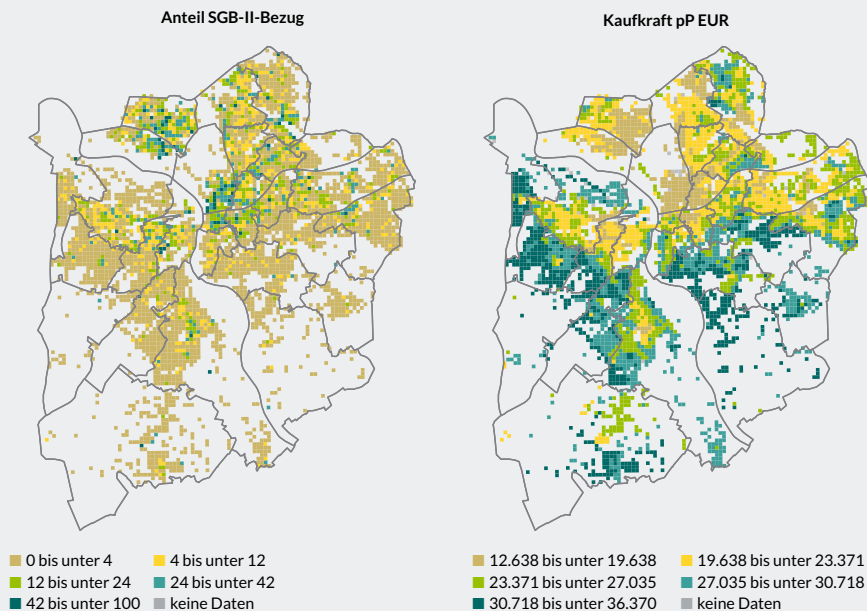
Bei der so vorgenommenen Klassifizierung der kommunalen Daten ergab sich die Besonderheit, dass in deutlich mehr als einem Fünftel der Rasterzellen überhaupt keine „Haushalte mit Migrationshintergrund“ ansässig waren. Hier wird eine klare räumliche Schieflage deutlich („ethnische Segregation“): Während einzelne Lagen Anteilswerte von bis zu 100 Prozent erreichen, gibt es an sehr vielen Orten überhaupt keine Haushalte mit Migrationshintergrund (gemessen nach Staatsangehörigkeit). Da die beiden untersten Kategorien beide den Wert Null umfassen, wurden sie zusammengefasst. Die unterste Kategorie bei der Karte mit kommunalen Daten (in der Darstellung rechts) enthält also circa 40 Prozent der Rasterzellen.

Bezüglich der „Hot Spots“ kristallisieren sich in beiden Karten dieselben Gebiete heraus: Der nordwestlich gelegene Stadtteil Styrum sowie die zentral liegende Altstadt¹⁸ fallen auf beiden Karten als Gebiete mit potenziell großen Migrantenanteilen auf. Auch wenn es im Einzelfall also durchaus Abweichungen gibt und die beiden Merkmale nicht besonders stark miteinander korrelieren, sind beide Datenquellen jedoch für die Identifizierung von „Hot Spots“ und einen Eindruck kleinräumiger Verteilungen in der Kommune, der notwendigerweise mit Vereinfachungen (Klassifizierung) einhergeht, nutzbar. Die microm-Variable liegt dabei tendenziell auf einem höheren Niveau als die Konstruktionen aus den Einwohnermeldedaten.

Die letzte kleinräumige Variable, die einer vergleichenden Betrachtung unterzogen werden soll, ist die Kaufkraft. Diese steht sowohl auf Personen- als auch auf Haushaltsebene zur Verfügung. Die Angaben auf beiden Ebenen scheinen aus sehr ähnlichen Schätzungen zu resultieren, da der Korrelationskoeffizient beider Variablen 0,93 beträgt. Mit der Kaufkraft steht eine Variable zur Verfügung, die einen kleinräumigen Blick auf „die andere Seite der Medaille“ von Armut erlaubt: Hier wird der Blick nicht nur auf einen bestimmten Teil von Armut, nämlich den in der kleinräumigen Sozialberichterstattung etablierten Indikator des SGB-II-Bezugs, gelegt, sondern das gesamte Spektrum der Einkommensverteilung betrachtet. Insofern ist auch hier aufgrund der unterschiedlichen Konzepte ein statistischer Zusammenhang zu erwarten, der allerdings nicht allzu stark ausfallen dürfte. Tatsächlich ergibt sich ein Korrelationskoeffizi-

18 Zur Lage der Stadtteile vgl. Abbildung 7, Anhang.

ABBILDUNG 3: Vergleich der SGB-II-Quoten und der Kaufkraft in Rasterzellen



Daten: microm; Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr; eigene Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

ent in der erwarteten Richtung von $-0,52$, was als mittelstarker negativer Zusammenhang zu interpretieren ist. Dies bedeutet: Je höher die SGB-II-Quote in einer Rasterzelle, desto niedriger die geschätzte Kaufkraft pro Person (Abbildung 3).

Die beiden Karten verbildlichen die Unterschiedlichkeit der Konzepte: In denjenigen Bereichen, in denen die SGB-II-Quoten stärker differenzieren, differenziert die Kaufkraftvariable nicht – dort liegt der „helle“ Bereich, die Kaufkraft, in den unteren Kategorien. Die Kaufkraft differenziert vielmehr innerhalb der Stadtteile, die insgesamt niedrige SGB-II-Quoten aufweisen.

Resümee: Abgleich der „harten“ Daten

Im vorangegangenen Teil wurden am Beispiel der Stadt Mülheim an der Ruhr die kommerziellen Daten der Firma microm Micromarketing-Systeme und Consult GmbH mit den verfügbaren Daten aus amtlichen und kommunalen Quellen hinsichtlich der Eigenschaften Bevölkerung und Haushalte, Migrationshintergrund und Kaufkraft und auf den beiden Ebenen der Kommune sowie der 100x100-Meter-Rasterzellen gegenübergestellt. Insbesondere Haushalts- und Bevölkerungswerte sind dabei auf beiden Ebenen nahezu identisch. Auf der Ebene der Kommune zeigte sich, dass die microm-Werte bei Migrationshintergrund und Kaufkraft im Vergleich zu denjenigen Werten, die aus den verfügbaren Angaben der amtlichen Statistik erwartbar gewesen wären, leicht erhöht sind. Beachtet werden sollte außerdem, dass der von microm zur Verfügung gestellte Migrationshintergrund auf Haushaltsebene ausgewiesen wird, was die Vergleichbarkeit erschwert und auch konzeptionell nicht unbedingt eingängig ist. Bei den kleinräumigen Vergleichen von Migrationshintergrund und Kaufkraft lagen keine direkt vergleichbaren kleinräumigen Indikatoren vor. Die Korrelationen zu Konstruktionen über die Staatsbürgerschaft (Migrationshintergrund) und den SGB-II-Anteil (Kaufkraft) lagen in der erwarteten Richtung und stellten sich als mittelstark heraus. Die Abweichungen erscheinen angesichts der unterschiedlichen Konzepte plausibel.

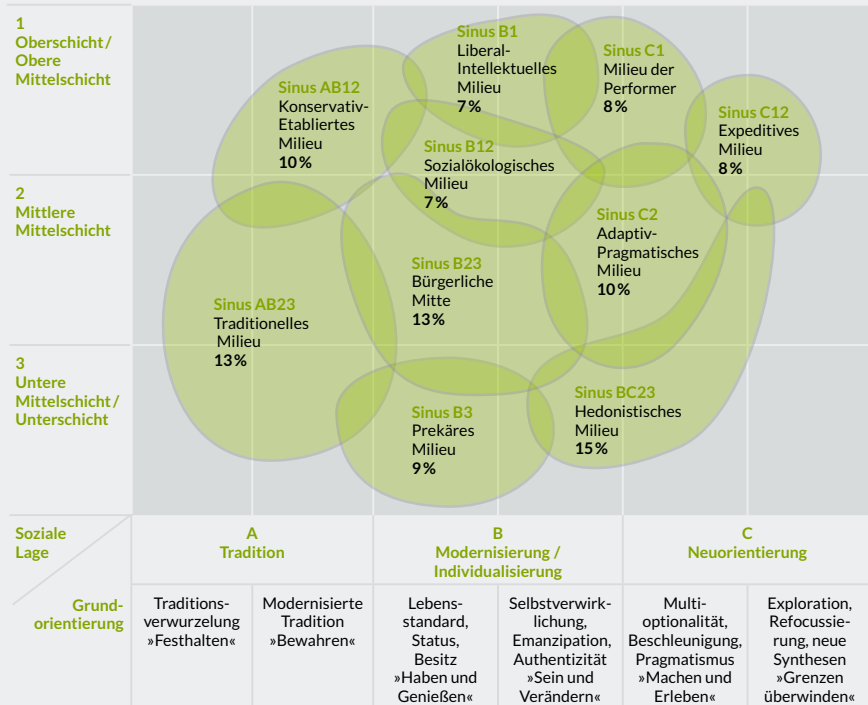
4 Sinus-Geo-Milieus® in Mülheim an der Ruhr

Während die bislang betrachteten Variablen sich mit Bevölkerungs- und Haushaltszahl, Migrationshintergrund/Staatsangehörigkeit sowie Kaufkraft/SGB-II-Bezug auf vergleichsweise „handfeste“ Eigenschaften beziehen, sind Milieus mit ihrem Bezug zu Einstellungen, Verhalten und Lebenslagen (vgl. Kapitel 2) „weiche“ Variablen. Das verwendete Sinus-Modell wurde im Rahmen der in Kapitel 2 geschilderten theoretischen Diskussionen in den 1970er und 1980er Jahren vom Heidelberger Forschungsinstitut Sinus Sociovision entwickelt. Seither gibt es auch die in Abbildung 4 dargestellte sogenannte „Kartoffelgraphik“, die regelmäßig anhand aktueller Daten und gesellschaftlicher Entwicklungen aktualisiert und unter anderem auch auf der Homepage des Instituts zur Verfügung gestellt wird (vgl. für eine frühe Version z. B. Flaig und Barth 2018: 7). Die Grundannahme der Entwicklung der Sinus-Milieus ist, dass die rein vertikale Differenzierung der Gesellschaft an Erklärungskraft verloren habe und durch eine kulturelle Dimension ergänzt werden müsse. In den Anfängen wurden die Milieus durch rein qualitative Erhebungsverfahren wie beispielsweise narrative Interviews erfasst. Später wurden die Ergebnisse quantitativ nachmodelliert (vgl. a. a. O.: 4 ff.).

Die beiden Grunddimensionen der Sinus-Milieu®-Konstruktion sind (1) die „soziale Lage“, die ein traditionelles vertikales Schichtsystem abbildet. Hinzu kommt (2) die horizontale Dimension, die anhand der Pole „Tradition“ und „Moderne“ die Wertorientierungen abbildet. In dem durch diese beiden Dimensionen aufgespannten sozialen Raum lassen sich dann unterschiedliche Milieus identifizieren.

Aktuell differenziert das Sinus-Institut zehn unterschiedliche Milieus in Deutschland. Vier Milieus gehören zu den sozial gehobenen (Konservativ-Etablierte, Liberal-Intellektuelle, Performer und Expeditiv) sowie drei zu denen der Mitte (Bürgerliche Mitte,

ABBILDUNG 4: Sinus-Milieus® in Deutschland 2017 – Soziale Lage und Grundorientierung („Kartoffelgraphik“)



Quelle: Sinus Sociovision, <https://www.sinus-institut.de/sinus-loesungen/sinus-milieus-deutschland/>

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

Adaptiv-Pragmatische und Sozialökologische). Drei weitere zählen zu den Milieus der unteren Mitte bzw. Unterschicht (Traditionelle, Prekäre und Hedonisten).

Für die Übertragung in den Raum wird die umfangreiche microm-Datenbank genutzt und aus „über eine[r] Milliarde Einzelinformationen“ (Küppers 2018: 98) für jedes Haus in Deutschland die Wahrscheinlichkeit geschätzt, mit der ein bestimmtes Milieu vorkommt, sowie ein „dominantes“ Milieu ausgewiesen (vgl. ebd.). Die Ergebnisse

Gewichtung der Rasterdaten

- Die Rasterzellen sind unterschiedlich „besetzt“: In einigen leben viele, in anderen wenige Menschen.
- Für Berechnungen und Analysen über die Rasterzellen hinweg (z. B. Gesamtmittelwerte oder Zusammenhangsmaße) empfiehlt es sich daher, Gewichtungsverfahren einzusetzen.
- Da die Milieus Aussagen über Haushalte treffen, ist es sinnvoll, die Milieuwahrscheinlichkeiten mit den Haushaltszahlen zu gewichten.

können auf unterschiedliche räumliche Ebenen bezogen werden. Wie genau die Werte zustande kommen, bleibt intransparent – ein Umstand, der auch für die Sinus-Milieus® im Allgemeinen (ohne Geoinformation) zutrifft und daher häufig Kritik unterzogen wird.

Die sozialen Milieus sind bundesweit regional nicht gleichermaßen vertreten. Tabelle 6 stellt die Milieuzugehörigkeitswahrscheinlichkeiten für Mülheim an der Ruhr¹⁹ im Vergleich zum Bundesdurchschnitt dar.

Im Vergleich zum Bundesdurchschnitt gibt es in Mülheim an der Ruhr Abweichungen bezüglich der Bürgerlichen Mitte, des sozialökologischen Milieus sowie der Hedonisten. Während die Bürgerliche Mitte leicht unterrepräsentiert wird, sind das hedonistische sowie das sozialökologische Milieu leicht überrepräsentiert.

Anhand der Tabelle 6 soll kurz auf die „Natur“ der Daten eingegangen werden, die sich deutlich von den üblichen in der Stadtforschung oder kommunalen Statistik verwendeten Daten unterscheidet: Es handelt sich hier um aus Regressionsmodellen geschätzte statistische Wahrscheinlichkeiten, die sich auf 100 Prozent summieren. Beispielsweise beträgt in Mülheim an der Ruhr für eine zufällig ausgewählte Person die Wahrscheinlichkeit, dass diese dem konservativ-etablierten Milieu angehört, etwa 10 Prozent. Gleichzeitig hat diese Person dann eine Wahrscheinlichkeit von rund 7 Prozent, dem liberal-intellektuellen Milieu anzugehören. Schließlich besitzt die Person eine Wahrscheinlichkeit von 100 Prozent, *irgendeinem* der zehn Milieus anzugehören.

¹⁹ Berechnung: Gewichteter Durchschnitt der Rasterzellen. Gewichtung über die von microm zugrunde gelegte Haushaltszahl.

TABELLE 6: Vergleich der Milieuwahrscheinlichkeiten in Prozent, Deutschland und Mülheim an der Ruhr

Milieu	Deutschland	Mülheim an der Ruhr
Konservativ-Etablierte	10,0	10,2
Liberal-Intellektuelle	7,0	7,2
Performer	8,0	8,2
Expeditiv	8,0	7,8
Bürgerliche Mitte	13,0	11,1
Adaptiv-Pragmatische	10,0	9,7
Sozialökologische	7,0	8,4
Traditionelle	13,0	12,7
Prekäre	9,0	8,4
Hedonisten	15,0	16,4
	100,0	100,1

Daten: Sinus Sociovision und microm

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

Die Wahrscheinlichkeiten wurden aus einer Vielfalt georeferenzierter Einzeldaten generiert. Dabei bestehen bezüglich der Datengenerierung mindestens zwei Unsicherheitsaspekte:

1. Die Generierung der Sinus-Milieus®: Sinus Sociovision verortet anhand von Milieuindikatoren die Befragten in einer großen quantitativen Studie – dieses Vorgehen „führt natürlich auch zu Fehlzuordnungen“, funktioniert im Aggregat aber gut (Flaig und Barth 2018: 6). Nichtsdestotrotz bleibt hier Unsicherheit darüber bestehen, wie die quantitativen Milieus eigentlich gebildet werden.
2. Die Generierung der Sinus-Geo-Milieus®: Weiß man schon bei den großen Stichprobenmilieuzuordnungen nicht genau, wie sie zustande kommen und mit welchem Grad von Unsicherheit zu rechnen ist, so kommt bei der Generierung der Wahrscheinlichkeiten auf Geoebene noch eine weitere Unsicherheit hinzu: Es bleibt unklar, welche Variablen in welchem Ausmaß in die Schätzung einfließen.

Im Ergebnis erhält man für jedes der Milieus eine geschätzte Wahrscheinlichkeit, zu welchem der Milieus die Personen, die in einer Rasterzelle leben, gehören. Diese Wahr-

scheinlichkeit ist in keiner Rasterzelle null – selbst in einem Villenviertel gibt es also für die dort lebenden Personen noch eine bestimmte statistische Wahrscheinlichkeit, zu einem der unteren Milieus zu gehören. Dies ist inhaltlich zwar nicht ausgeschlossen – die soziale Welt ist bunt und bringt manchmal überraschende Konstellationen mit sich. Aber alleine aufgrund der hohen Mieten dürfte diese Wahrscheinlichkeit sehr viel geringer sein, als es die Sinus-Geo-Milieu[®] schätzen, wie im Folgenden exemplarisch gezeigt wird.

Tabelle 7 stellt zentrale statistische Kennwerte für die Verteilung der einzelnen Milieuwahrscheinlichkeiten in den Rasterzellen zusammen.

TABELLE 7: Vergleich der Milieuwahrscheinlichkeiten in den 100 x 100-Meter-Rasterzellen

Milieu	Minimum	Maximum	Durchschnitt	Median	1. Quartil	3. Quartil	Standardabweichung	Spannweite	Quartilsabstand
Konservativ-Etablierte	5,1	20,0	11,4	10,9	9,4	13,2	2,7	14,9	3,8
Liberal-Intellektuelle	3,0	14,8	8,1	8,1	6,3	9,8	2,3	11,8	3,5
Performer	3,2	19,2	9,3	9,1	7,0	11,1	2,8	16,0	4,1
Expeditive	2,2	17,4	7,9	7,6	6,7	8,9	1,7	15,2	2,2
Bürgerliche Mitte	4,5	19,5	11,3	11,3	10,0	12,6	2,0	15,0	2,6
Adaptiv-Pragmatische	3,7	26,2	9,6	9,3	8,2	10,6	2,2	22,5	2,4
Sozial-ökologische	3,9	15,5	9,0	8,9	7,7	10,2	1,8	11,6	2,5
Traditionelle	2,0	25,0	11,8	11,4	9,2	14,1	3,4	23,0	4,9
Prekäre	1,6	18,6	7,7	7,7	6,1	9,4	2,2	17,0	3,3
Hedonisten	5,6	36,8	14,0	13,2	10,0	17,2	5,1	31,2	7,2

Daten: microm; eigene Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

Zu Beginn sind die statistischen Minima der Milieuwahrscheinlichkeiten aufgeführt: Dies ist jeweils der geringste Wahrscheinlichkeitswert, der in einer der 3.960 Rasterzellen für das jeweilige Milieu auftritt. Die kleinste auftretende Wahrscheinlichkeit in einer Zelle über alle Milieus hinweg betrifft das prekäre Milieu: Die Wahrscheinlichkeit beträgt hier in einer Rasterzelle nur 1,6 Prozent.

Betrachtet man diese Rasterzelle in Mülheim an der Ruhr, so findet sich dort ein von hochpreisigen Immobilien geprägtes Umfeld. Laut microm-Daten sind hier zwei Haushalte mit einer überdurchschnittlichen Kaufkraft ansässig. Aber auch, wenn microm die Wahrscheinlichkeit für das prekäre Milieu hier mit 1,6 Prozent schätzt, so beträgt die geschätzte Wahrscheinlichkeit für die unteren Milieus zusammengenommen noch fast 10 Prozent, die für ein mittleres Milieu noch insgesamt ein Drittel. Das heißt: Auch wenn nach Ortskenntnis die *subjektive* Wahrscheinlichkeit für eines der unteren Milieus dort nahe null liegt, beträgt die *statistische* Wahrscheinlichkeit für eines der unteren der Sinus-Geo-Milieus® immer noch knapp 10 Prozent. Zu beachten ist hier ebenfalls, dass der Wert der Rasterzellen keinesfalls als empirische Häufigkeit interpretiert werden kann. Empirisch gehören die beiden Haushalte einem oder zwei der zehn Milieus an, dort leben hingegen nicht „10 Prozent der Bevölkerung aus den Milieus der unteren Mitte“.

In der nächsten Spalte sind die statistischen Maxima der Rasterzellen für die jeweiligen Milieus angegeben. Das liberal-intellektuelle Milieu weist hier insgesamt den kleinsten Wert auf, das hedonistische den größten. Dies bedeutet: In keiner der 3.960 Zellen ist die Wahrscheinlichkeit für das liberal-intellektuelle Milieu größer als 14,8 Prozent. Es lässt sich also vermuten, dass sich in Mülheim an der Ruhr insgesamt circa 7 Prozent Vertreterinnen und Vertreter dieses Milieus befinden, in keiner der Raumeinheiten auf Rasterebene wird jedoch eine Wahrscheinlichkeit von über 14,8 Prozent dafür erreicht, dass die dortigen Bewohner tatsächlich diesem Milieu zuzurechnen sind.

Die höchste Wahrscheinlichkeit, mit der man ein bestimmtes Milieu an einem Ort mithilfe dieser Daten vermuten kann, liegt bei knapp 37 Prozent des hedonistischen Milieus in einer innenstadtnahen Lage. Die von microm geschätzte Wahrscheinlichkeit dafür, dass hier allgemein untere Milieus vertreten sind, liegt bei circa 61 Prozent. Angesichts der dortigen SGB-II-Quote von 93 Prozent ist dies eher wenig. Dennoch beträgt die statistische Wahrscheinlichkeit für eines der Milieus der Oberschicht dort immer noch mehr als 20 Prozent.

Es liegt allerdings im statistischen Verfahren der Regressionsanalyse begründet, dass Durchschnitts- / Erwartungswerte und keine Extremwerte geschätzt werden. Gleichwohl sollte man sich beim Umgang mit derartig geschätzten Werten vor Augen halten, dass die Sinus-Geo-Milieus® tendenziell ein richtiges Bild zeichnen. Einen besonderen oder extremen Einzelfall, wie äußerst exklusive Wohnlagen oder aber sogenannte „Armutsquartiere“ mit nahezu vollständigem SGB-II-Bezug, bildet man mit diesen Wahrscheinlichkeiten dennoch nicht ab. Für die Segregationsforschung bedeutet dies, dass man das Ausmaß der Segregation anhand solcher Daten unterschätzt.

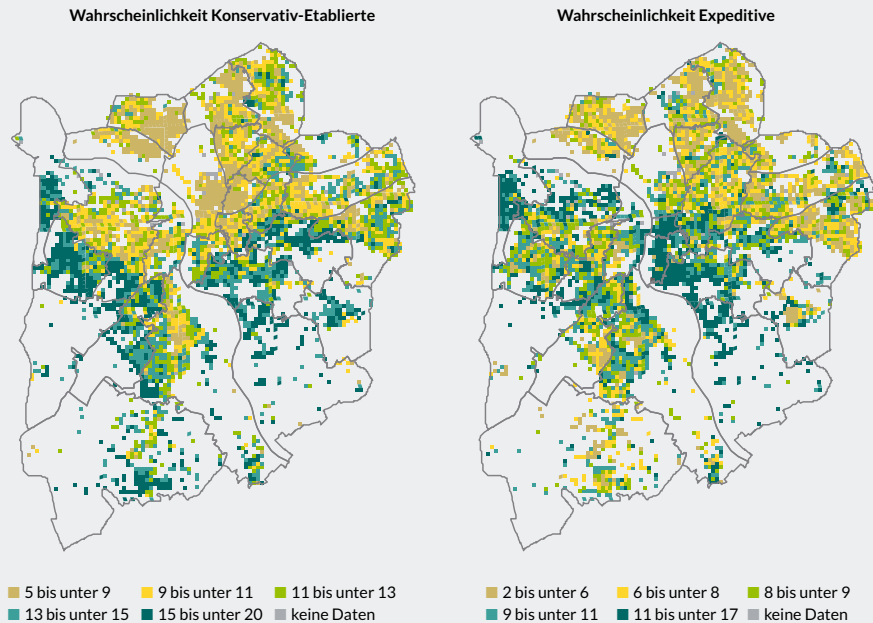
Hinzu kommt eine weitere Problematik, die in den letzten drei Spalten der Tabelle ersichtlich wird: Die Streuung der Wahrscheinlichkeiten (angegeben in Prozentpunkten) zwischen den Zellen ist relativ gering. Die Spannweite gibt dabei die Differenz zwischen der größten und der kleinsten Wahrscheinlichkeit an. Beim hedonistischen Milieu findet sich hier zwar eine Spannweite von über 30 Prozentpunkten, bei den insgesamt weniger vertretenen Milieus wie dem sozialökologischen oder dem liberal-intellektuellen hingegen betragen die Spannweiten nur noch knapp zwölf Prozentpunkte. Dies ist der Vergleich der Extrema – schaut man sich den Bereich an, in dem die mittleren 50 Prozent der Werte streuen (den sogenannten „Quartilsabstand“), so ist die Streuung lediglich beim hedonistischen Milieu größer als sieben Prozentpunkte.

Da Schätzungen aus Regressionen im Allgemeinen mit Unsicherheit behaftet sind, ist es für die Interpretation als problematisch zu bewerten, dass eine tendenziell geringe Streuung bei einer nicht quantitativ zu beurteilenden Unsicherheit der Schätzung vorliegt. Ungeachtet dessen ergeben sich insgesamt plausible Bilder, wenn man die Milieuwahrscheinlichkeitsverteilung in der Stadt kartographisch betrachtet.

Um die Differenzierung auch innerhalb einer sozialen Schicht sichtbar zu machen, kartiert Abbildung 5 exemplarisch die Wahrscheinlichkeitsverteilung zweier „Leitmilieus“: das konservativ-etablierte und das expeditiv Milieu.

Bei beiden Leitmilieus ist zu erkennen, dass sie tendenziell im Süden der Stadt zu finden sind. Darüber hinaus gibt es lediglich zwei Viertel, in denen beide Milieus mit sehr hoher Wahrscheinlichkeit auftreten:

ABILDUNG 5: Pole der Leitmilieus – Konservativ-Etablierte und Expeditive



Daten: microm; eigene Darstellung und Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

1. Prinzenhöhe, das im Nordwesten der Stadt in Speldorf gelegene Villenviertel
2. Der direkt an Ruhrauen und Ruhrstrand gelegene Bereich zwischen Mendener Brücke und Steinknappen bei den Bezirken Menden/Ickten und Holthausen²⁰

Innerhalb der Altstadt oder des nördlichen Bereichs sind beide Milieus tendenziell nicht zu finden. Bei der restlichen Verteilung unterscheiden sich die beiden Milieus: Die Konservativ-Etablierten sind sehr viel eher im südwestlichen Bereich Mülheims zu finden, die Expeditiven innenstadtnah zwischen Altstadt und Menden.

²⁰ Zur Lage der Stadtteile vgl. Abbildung 7, Anhang.

Expeditive und Konservativ-Etablierte haben bei den Leitmilieus die geringste Wahrscheinlichkeit, am gleichen Ort zu leben: Die Korrelation der Milieuwahrscheinlichkeiten in den Rasterzellen beträgt lediglich 0,18. Hier zeichnet sich also auch bei der Oberschicht eine räumliche Trennung zwischen dem jüngeren, innovationsorientierten Milieu und dem konservativen Establishment ab. Dies stützt die in Kapitel 2 begründete Annahme, dass die horizontale Differenzierung vor allem dort von Bedeutung ist, wo entsprechende Freiheitsgrade bestehen. Die beiden genannten Milieus besitzen die finanziellen Mittel, Wohnort und -lage selbst zu wählen. Beim traditionellen, hedonistischen und prekären Milieu hingegen sind die statistischen Zusammenhänge deutlich stärker. Dies kann darauf zurückzuführen sein, dass Personen, die diesen Milieus ange-

TABELLE 8: Korrelationsmatrix der Milieuwahrscheinlichkeiten in Rasterzellen

Milieu	Konservativ-Etablierte	Liberal-Intellektuelle	Performer	Expeditive	Bürgerliche Mitte	Adaptiv-Pragmatische	Sozialökologische	Traditionelle	Prekäre	Hedonisten
Konservativ-Etablierte	1,00									
Liberal-Intellektuelle	0,84	1,00								
Performer	0,67	0,83	1,00							
Expeditive	0,18	0,39	0,61	1,00						
Bürgerliche Mitte	0,25	0,19	-0,16	-0,62	1,00					
Adaptiv-Pragmatische	-0,17	-0,03	0,25	0,44	-0,24	1,00				
Sozial-ökologische	0,76	0,69	0,69	0,26	-0,04	-0,05	1,00			
Traditionelle	-0,51	-0,69	-0,85	-0,73	0,32	-0,48	-0,61	1,00		
Prekäre	-0,63	-0,79	-0,85	-0,69	0,18	-0,28	-0,63	0,78	1,00	
Hedonisten	-0,87	-0,86	-0,69	-0,13	-0,47	-0,03	-0,69	0,46	0,58	1,00

Mit Haushaltszahl gewichtete Korrelationen. Daten: microm; eigene Darstellung und Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

hören, aufgrund ihrer finanziellen Lage auf dem Wohnungsmarkt in ihren Auswahlmöglichkeiten eingeschränkter sind.

Tabelle 8 stellt sämtliche Korrelationen (statistische Zusammenhänge) zwischen den Milieuwahrscheinlichkeiten auf Rasterebene dar. Hohe Korrelationen (Betrag $> 0,75$) wurden rot und grün markiert. Grün bedeutet, dass eine hohe *positive* Korrelation vorliegt: In den Rasterzellen, in denen die Milieuwahrscheinlichkeit für das in den Zeilen benannte Milieu hoch ist, ist auch die Wahrscheinlichkeit für das in den Spalten benannte Milieu hoch. Dies lässt sich am Beispiel der zweiten Zeile veranschaulichen: Dort, wo die Wahrscheinlichkeit für das liberal-intellektuelle Milieu (Zeilenbeschriftung) relativ hoch (oder niedrig) ist, ist auch die Wahrscheinlichkeit für das konservativ-etablierte Milieu (Spaltenbeschriftung) relativ hoch (oder niedrig). Rot hingegen steht für eine *negative* Korrelation, wie es beispielsweise beim ersten Wert der letzten Zeile der Fall ist: Dort, wo die Wahrscheinlichkeit für das hedonistische Milieu relativ *hoch* ist, ist die Wahrscheinlichkeit für das konservativ-etablierte Milieu relativ *niedrig*.

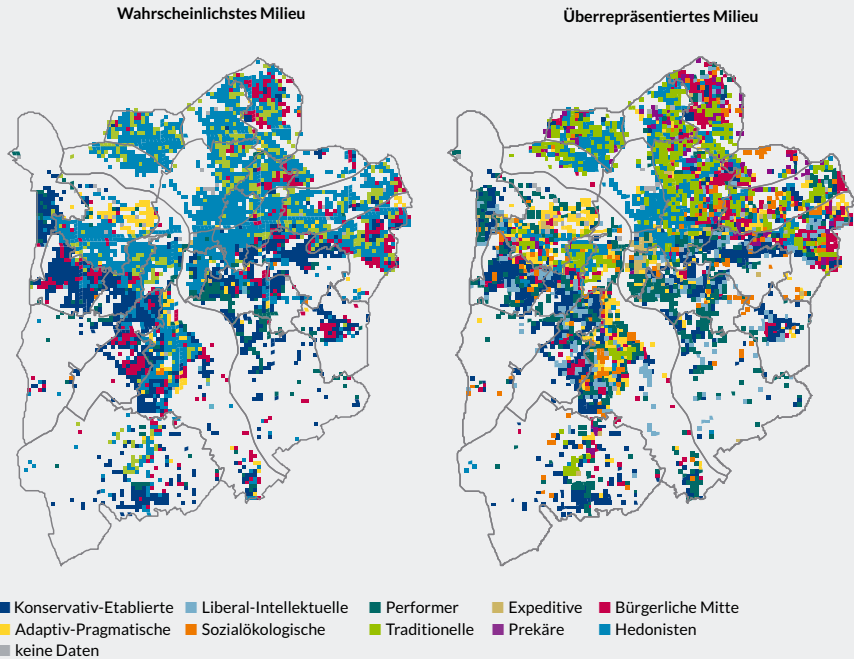
Die stärksten Zusammenhänge sind dabei die negativen zwischen Leitmilieus und Unterschicht: Hedonisten und Konservativ-Etablierte wohnen tendenziell nicht am gleichen Ort, ähnlich wie Traditionelle und Performer. Die *sowohl* hinsichtlich der sozialen Lage *als auch* hinsichtlich der Wertorientierung entgegengesetzten Milieus teilen die Wohnlage also nicht miteinander.

Bei der zusammenfassenden Kartierung der Milieus muss zunächst entschieden werden, wie die Milieuzuordnung der Rasterzellen erfolgen soll. Weiter oben wurden bereits exemplarisch das konservativ-etablierte und das expeditiv Milieu dargestellt. Die einzelne Kartierung aller zehn Milieus ist allerdings wenig übersichtlich – für einen Überblick über *alle* Milieus müssen die Informationen nochmals verdichtet werden. Möglich ist dies mithilfe der Verwendung des von microm geschätzten „dominanten Milieus“ oder aber anhand einer eigenen Konstruktion, bei der für jede Rasterzelle *ein* Milieu als Richtwert festgelegt wird.

Beispielhaft für Mülheim an der Ruhr wurden in Abbildung 5 zwei Zusammenfassungen kartiert:

- Variante 1: Kartierung des wahrscheinlichsten Milieus. Das wahrscheinlichste Milieu in einer Rasterzelle ist dasjenige, das im Vergleich der zehn Milieu-

ABBILDUNG 6: **Wahrscheinlichste und überrepräsentierte Milieus**



Der gesamtstädtische Wert wurde als gewichteter Mittelwert der Rasterzellen berechnet.
Zur Gewichtung wurde die von microm zugrunde gelegte Haushaltszahl verwendet.
Daten: microm; eigene Darstellung und Berechnung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

wahrscheinlichkeiten miteinander den höchsten Wahrscheinlichkeitswert aufweist. Das Problem hierbei ist, dass Milieus mit einer insgesamt geringen Ausgangswahrscheinlichkeit „unsichtbar“ werden, da sie nur sehr selten als sehr wahrscheinlich geschätzt werden.

- Variante 2: Kartierung eines „überrepräsentierten“ Milieus. Das überrepräsentierte Milieu ist dasjenige, das im Vergleich zu seinem gesamtstädtischen Wert die höchste (positive) Abweichung aufweist. Ähnlich ist auch das „dominante Milieu“ konstruiert. Der gesamtstädtische Wert wurde als gewichteter Mittelwert der Rasterzellen berechnet. Zur Gewichtung wurde die von microm zugrunde gelegte Haushaltszahl verwendet.

Es fällt auf, dass bei der Zuordnung anhand des wahrscheinlichsten Milieus das Stadtbild sehr viel einheitlicher erscheint als bei der Kartierung der überrepräsentierten Milieus. Ursächlich ist die bereits erwähnte „Unsichtbarmachung“ der Milieus mit insgesamt geringen Ausgangswahrscheinlichkeiten: Das liberal-intellektuelle, das sozialökologische und das expeditiv Milieu haben jeweils in weniger als zehn Rasterzellen die höchste Wahrscheinlichkeit, das prekäre Milieu sogar in keiner einzigen. Die Stadt wird von den Farben des konservativ-etablierten und des hedonistischen Milieus dominiert. Das konservativ-etablierte ist dasjenige der Leitmilieus mit der höchsten Wahrscheinlichkeit – fast überall dort, wo sich die unteren Milieus z.B. die Mieten nur schwer leisten können (im Südwesten und Südosten der Stadt), besitzt dieses Milieu die höchste Wahrscheinlichkeit. Im nördlichen Teil der Stadt hingegen hat fast überall das hedonistische Milieu die größte Wahrscheinlichkeit. Allerdings gibt es auch einige weitere interessante Farbinseln: Im Stadtteil Speldorf-Nordost gibt es ein kleines Gebiet nahe der Ruhr, wo die Gruppe der Adaptiv-Pragmatischen ihr Refugium zu haben scheint – die „moderne Mitte der Gesellschaft“. An Ruhrauen und Ruhrstrand gelegene Bereiche des Stadtteils Menden haben hohe Wahrscheinlichkeiten, dass sich dort Performer niederlassen – die Leistungselite auf dem Wertespektrum unter dem Motto des „Machens und Erlebens“. Und schließlich fallen auch Inseln der bürgerlichen Mitte auf: Im südlich gelegenen Saarn, im östlich gelegenen Heißen sowie die darunter größte „Insel“ im nördlich gelegenen Dümpten.²¹

Die Karte mit den überrepräsentierten Milieus differenziert das Farbspektrum weiter: So mehren sich z.B. die gelben Flecken des liberal-intellektuellen Milieus in den süd-östlichen Teilen der Stadt.

Welche Darstellung man bevorzugt, hängt im Wesentlichen vom Erkenntnisinteresse ab. Möchte man sich an ein bestimmtes Milieu wenden, ist es sinnvoll, dessen Verteilung separat zu kartieren. Bei einer zusammenfassenden Darstellung aller Milieus kann es sinnvoll sein, zu untersuchen, wen man in den einzelnen Regionen mit der höchsten Wahrscheinlichkeit trifft. Angesichts der unterschiedlichen Ausgangswahrscheinlichkeiten sind dies häufig Menschen aus dem konservativ-etablierten oder dem hedonistischen Milieu. Interessiert man sich für diejenigen Milieus, die insgesamt wenig vertreten sind, ist es sinnvoll, zu betrachten, wo welches Milieu *überrepräsentiert* ist.

21 Zur Lage der Stadtteile vgl. Abbildung 7, Anhang.

Bei der Darstellung der Milieus im Raum besteht abgesehen von den bereits angeführten Problemen (Intransparenz der Milieugenerierung, Intransparenz der Übertragung in den Raum, fehlende Angaben über die statistische (Un-)Sicherheit der Schätzung, Schätzung von Erwartungswerten ohne Abbildungsmöglichkeit von Extremen) wenig Anlass, die Qualität der Daten infrage zu stellen. Extreme Werte werden zwar unterschätzt, tendenziell aber richtig zugeordnet. Die Darstellung im Raum ist plausibel.

Welchen Mehrwert erhält man nun durch das Wissen über soziale Milieus? Für zielgruppenorientiertes Marketing sind die Vorteile von Geoinformationen offensichtlich: Ob es sich nun um Plakate, Postwurf oder Ähnliches handelt – Produkte sind oft für bestimmte Zielgruppen interessant. Die Sinus-Geo-Milieus® helfen dabei, diese zu finden. Aber wie sieht es mit der Stadtforschung, Raumplanung und kommunalen Beteiligungsprozessen aus? Der vhw – Bundesverband für Wohnen und Stadtentwicklung e.V. beschäftigt sich seit einigen Jahren mit sozialen Milieus in der Stadtgesellschaft (vgl. z. B. vhw 2013 oder Kuder 2016). Einer der wesentlichen Ausgangspunkte ist dabei, Bürgerbeteiligung „gerechter“ zu verwirklichen. Da Beteiligungsverfahren sozial selektiv sind – z. B. nehmen mit steigender Bildung und steigenden Einkommen deutlich mehr Personen an Beteiligungsverfahren teil (vgl. Kuder 2016: 3) –, besteht durch das Wissen über soziale Milieus und ihre Wohnorte die Möglichkeit, gezielt bestimmte Gruppen anzusprechen und zu beteiligen. So kann gewährleistet werden, dass Partizipation in der Stadt umfassender ist, eine Vielfalt an Stimmen und Meinungen abbildet und sich nicht auf Eliten beschränkt. Dabei sollte immer auch reflektiert werden, wen man über welche Kanäle erreichen möchte – das Amtsblatt ist beispielsweise als Informationsmedium über städtische Prozesse einem Großteil der Bevölkerung unbekannt und/oder die Aufmachung ist so „amtlich“ und wenig ansprechend, dass sie nur schwer verstanden werden kann (vgl. für eine anschauliche Darstellung vhw 2013: 270).

Die Sinus-Geo-Milieus® können also dabei helfen, bestimmte Milieus durch Kenntnis ihrer Wohnlage zu erreichen und Kommunikationsformen und -kanäle zu reflektieren sowie inklusiver zu gestalten.

Im Folgenden soll ein weiterer Anwendungstyp ausgelotet werden: Lassen sich die Daten auch dazu verwenden, auf eine bestimmte Bevölkerungsgruppe – in diesem Fall noch nicht schulpflichtige Kinder – übertragen zu werden? Können darüber hinaus die Inanspruchnahme von Angeboten (hier: Kita-Besuch) anhand der Datenverteilung erklärt und die Settings (hier: Kitas) typisiert werden?

5 (Milieuspezifische) Kita-Segregation

Bildungs- und damit auch Lebens- und Teilhabechancen von Kindern sind in großem Maße durch deren soziale Herkunft strukturiert. Dies wird auch im Rahmen der sogenannten „PISA-Studien“ regelmäßig in den öffentlichen Diskurs getragen. Das schlechte Abschneiden Deutschlands bei der OECD-Studie, die alle drei Jahre mit unterschiedlichen Schwerpunkten die Leistungen von 15-jährigen Schülerinnen und Schülern erhebt und vergleicht, ist mittlerweile unter dem Schlagwort „PISA-Schock“ bekannt geworden. Der Effekt der sozialen Herkunft, der seit der ersten Studie diskutiert wird, ist in Deutschland dabei weiterhin überdurchschnittlich (vgl. OECD 2016: 237).

Allerdings wirkt sich ein früher Kita-Besuch positiv auf die Kompetenzen der Kinder aus – so zeigten Groos und Jehles (2015) die signifikanten Effekte auf die Ergebnisse der Schuleingangsuntersuchung, Bach, Koebe und Peter (2018) auf wesentliche – auch für den Arbeitsmarkt relevante – Persönlichkeitseigenschaften wie Kommunikationsfähigkeit und Durchsetzungsvermögen. Frühe Bildung, und hierzu gehört schließlich auch der Kita-Besuch, hat dabei unabhängig von der sozialen Herkunft positive Effekte und bietet damit eine Chance für mehr Gerechtigkeit im Bildungssystem.

Sieht man davon ab, dass der Betreuungsbedarf von Kindern bislang nicht durch das Angebot gedeckt wird (BMFSFJ 2018: 12), sind sowohl der Kita-Besuch im Allgemeinen als auch die Wahl einer spezifischen Kita hochgradig selektiv (vgl. Groos und Jehles 2015 und Groos, Trappmann und Jehles 2018). Aber auch die soziale Zusammensetzung der Kita hat einen *eigenständigen* Effekt auf das spätere Kompetenzniveau der Kinder (vgl. Groos und Jehles 2015). Auf dem Weg zu einer gerechteren Gesellschaft und fairen Lebens- und Teilhabechancen bestehen hier also zentrale gesellschaftliche „Stellschrauben“.

Um an diesen Schrauben zu drehen, muss bekannt sein, wo konkreter Handlungsbedarf besteht. Groos, Trappmann und Jehles (2018) haben einen Sozialindex für Kitas vorgeschlagen, der anhand der SGB-II-Dichten an den Wohnadressen der Kinder berechnet wird. Exemplarisch haben sie dies für die Städte Mülheim an der Ruhr, Hamm und Gelsenkirchen durchgeführt und analysiert. Allerdings sind hierzu kleinräumige Daten der Bundesagentur für Arbeit bzw. der Jobcenter erforderlich, die nur in abgeschotteten Statistikstellen mit den Kita-Daten verknüpft werden dürfen. Mülheim an der Ruhr und Hamm verfügten als Optionskommunen²² über Individualdaten. Gelsenkirchen bezog die Daten von der Bundesagentur für Arbeit auf Baublockebene. Letzteres ist zum Zeitpunkt des Erscheinens dieses Berichtes nicht mehr möglich: Die Bundesagentur für Arbeit stellt kleinräumige SGB-II-Daten an Kommunen aktuell nur noch für Raumeinheiten mit einer Bevölkerungszahl von mindestens 1.000 Personen zur Verfügung, sodass Zuordnungen über die Wohnadresse nur noch ein unscharfes Bild ergeben. Sollten sich die kommerziellen Daten auf die Kinder übertragen lassen und damit ein gutes Abbild von Kita-Segregation und der Identifizierung von sozial benachteiligten Kitas leisten, würden sie damit eine wertvolle Option für Fachplanungen außerhalb von Optionskommunen und ohne abgeschottete Statistikstelle darstellen. Weiterhin wird auszuloten sein, inwieweit die Differenzierung anhand der Milieudaten neue Perspektiven für die Kommunikation mit Eltern und die Entwicklung pädagogischer Konzepte bietet.

Kita-Typisierung

Um Kitas mit besonderen Unterstützungsbedarfen zu identifizieren, sind verschiedene statistische Verfahren denkbar. Die folgenden Typisierungen wurden anhand eines clusteranalytischen Verfahrens vorgenommen. Clusteranalysen gruppieren Fälle – hier Kitas – derart, dass sich die Fälle innerhalb der Gruppen möglichst ähnlich sind, wohingegen sich die Gruppen möglichst stark voneinander unterscheiden sollen.

Hier sollen zwei Typisierungsergebnisse miteinander verglichen werden:

- die „kommunale“ Variante: Typisierung der Kitas anhand der Sozialgeldquote und der Migrationsquote, und
- die „kommerzielle“ Variante: Typisierung anhand der microm-Variablen zur

²² Vgl. zum Optionsmodell Fußnote 12.

Kaufkraft der Haushalte und des Migrationshintergrunds, die über die Wohnorte auf die Kinder übertragen wurden.

Dabei sind beide Varianten hinsichtlich der Datenvorbereitung voraussetzungsfull. Die kommunale Variante ist in der durchgeführten Form eigentlich nur die Variante für Optionskommunen, da die Basis für die Sozialgeldquoten der Kitas die SGB-II-Einzel-daten waren (zu den strengen datenschutzrechtlichen Besonderheiten der zugrunde liegenden Auswertung vgl. Groos et al. 2018). Die Verknüpfung der SGB-II-Daten mit den Daten der Kinder ist dabei nicht unproblematisch: Zentrale Fehlerquellen sind z. B. Umzüge zwischen den Meldezeitpunkten oder Tippfehler bei den Namen. Erst danach können die Werte für die Kitas so zusammengefasst werden, dass für jede Kita der Anteil an Kindern, die Sozialgeld beziehen, angegeben werden kann. Beim Migrationshintergrund ist es ähnlich kompliziert: Das Referat V.1 „Stadtforschung und Statistik“ der Stadt Mülheim an der Ruhr führt zwar kein allgemeines standardisiertes Verfahren zur Ermittlung des Migrationshintergrunds durch. Allerdings wird dort durchaus aus den Einwohnermeldedaten für die Personen unter 21 Jahren ein Migrationshintergrund generiert. Dabei wird neben Staatsbürgerschaft und Geburtsland auch ein Verfahren verwendet, welches Geburtsland und Staatsbürgerschaft der Eltern berücksichtigt. Hierdurch können für die Kitas wiederum Anteile an Kindern mit Migrationshintergrund generiert werden. Für die folgenden Auswertungen sind diese Variablen für die 86 Mülheimer Kitas des Jahres 2014/2015 mit insgesamt 4.586 Kindern erfolgt.

Für die „kommunale Variante“ erfordert die Datenaufbereitung der Sozialgeldquote und des Migrationshintergrunds folglich einen erheblichen Aufwand, der nur unter besonderen datenschutzrechtlichen Voraussetzungen geleistet werden kann.

Für die Variante mit den kommerziellen Daten werden als Ausgangsvoraussetzung ebenfalls die Adressdaten derjenigen Kinder, die eine Kita besuchen, benötigt. Um diese mit den kleinräumigen Rasterdaten verknüpfen zu können, müssen sie entweder mit den Koordinaten einer „kleinräumigen Gliederung“²³ der Kommune zusammengeführt oder auf anderem Wege georeferenziert werden, um den Kindern daraufhin die Raster-

23 Als „kleinräumige Gliederung“ wird zumeist die Zuordnung von allen in einer Kommune existierenden Adressen zu Koordinaten, Sozialräumen, statistischen Bezirken und anderen vorliegenden räumlichen Unterteilungen bezeichnet.

daten mit einer GIS-Software²⁴ zuzuweisen. Anschließend können die Werte der Kinder für jede Kita zusammengefasst werden, sodass für jede Kita ein durchschnittliches Haushaltseinkommen der Kinderwohn Umgebung und ein Migrationshintergrundanteil geschätzt werden können.

Um der unterschiedlichen Verteilung der Variablen gerecht zu werden, wurden die Werte vor der Durchführung der Clusteranalyse z-standardisiert²⁵. Es gibt mittlerweile eine Vielzahl clusteranalytischer Verfahren. Eine praktikable Lösung für die vorliegende Situation ist das k-means-Verfahren (vgl. zur Einführung beispielsweise Wiedenbeck und Züll 2010). Bei einer Lösung für jeweils vier Gruppen ergeben sich die in Tabelle 9 dargestellten Zuordnungen. Bei 60 der insgesamt 86 untersuchten Kitas kommt man zu exakt der gleichen Zuordnung, dies entspricht einem Anteil von knapp 70 Prozent. In den restlichen 30 Prozent der Fälle besteht ein Unterschied von einer Kategorien-

TABELLE 9: Typisierungsvergleich – kommunale versus kommerzielle Daten

		Typisierung mit microm-Variablen			
		benachteiligt	eher benachteiligt	eher privilegiert	privilegiert
Typisierung mit kommunalen Daten	benachteiligt	13	4	0	0
	eher benachteiligt	1	19	6	0
	eher privilegiert	0	7	14	4
	privilegiert	0	0	4	14

Daten: microm; Referat V.1 Stadtforschung und Statistik der Stadt Mülheim an der Ruhr; eigene Berechnung und Darstellung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

- 24 GIS steht für „Geoinformationssystem“ – als GIS-Software werden Programme/Apps bezeichnet, die räumliche Daten verarbeiten und analysieren können.
- 25 Bei einer z-Standardisierung wird von jedem Wert eines Merkmals der Mittelwert des Merkmals abgezogen und durch die Standardabweichung dividiert. Die resultierenden Werte haben den Mittelwert 0 und die Standardabweichung 1. So wird beispielsweise verhindert, dass insgesamt größere Unterschiede beim Migrationshintergrund zwischen den Kitas die Typisierung stärker beeinflussen als die Unterschiede bei der Sozialgeldquote, die vielleicht durchschnittlich geringer ausfallen. Beide Merkmale werden durch die z-Standardisierung vergleichbar skaliert.

stufe, konträre Zuordnungen werden über die Verfahren nie vorgenommen. Legt man den Fokus auf die benachteiligten Kitas, wird sogar ein noch höherer Anteil (13 von 17, also gut drei Viertel) dieser Kitas auch durch das Vorgehen mit den kommerziellen Daten als solche identifiziert.

Die Typisierung anhand kommunaler Daten stellt keine unumstößliche empirische „Gewissheit“ dar – erstens, weil auch hier Ausfälle bei der Datenaufbereitung zu Problemen und anderen Zuordnungen führen können. Zweitens, weil Sozialgeldquote und Migrationshintergrundquote zwar als Indikatoren für potenziell schwierige Situationen in Einrichtungen taugen, aber nicht die tatsächliche Problemlage objektiv darstellen können. Drittens sind statistische Typisierungen zwar generell objektiv nachvollziehbar, es existiert aber eine Vielzahl möglicher Algorithmen, die auch leicht abweichende Ergebnisse produzieren können.

Führt man sich diese Unschärfen bei der Typisierung und die Tatsache vor Augen, dass Kaufkraft und Sozialgeldquote auch konzeptionell unterschiedliche Merkmale sind, so ist eine übereinstimmende Typisierung in 70 Prozent der Fälle als sehr hoch zu bewerten.

Bei der Betrachtung der Gruppenmittelwerte fällt allerdings auf, dass das Vorgehen der Übertragung anhand der Wohnorte der Kinder dazu führt, dass das Niveau der Variablen nicht mehr sinnvoll interpretiert werden kann. Tabelle 10 stellt die Durchschnittswerte der Ausgangsvariablen für die Kita-Typen (Ergebnisse der kommerziellen Variante)

TABELLE 10: Durchschnittliche Sozialgeldquote, Migrationshintergrundanteil und Kaufkraft nach Kita-Typ

Typisierung (microm)	Sozialgeldquote	Migrationshintergrund Kommune	Kaufkraft pro Haushalt	Migrationshintergrund (microm)
benachteiligt	54,5	76,3	39.162,48	21,5
eher benachteiligt	27,2	47,5	43.861,62	16,6
eher privilegiert	16,1	36,0	48.284,09	12,9
privilegiert	5,8	23,7	54.224,23	9,1

Daten: microm; Referat V.1 Stadtforschung und Statistik der Stadt Mülheim an der Ruhr; eigene Berechnung und Darstellung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

dar. Während bei den benachteiligten Kitas durchschnittlich ein Anteil von Kindern mit Migrationshintergrund von 76,3 Prozent aufzufinden ist, gibt der anhand der Wohnorte der Kinder über die Rasterzellen übertragene Schätzer lediglich 21,5 Prozent aus. Auch die durchschnittliche Kaufkraft pro Haushalt ist mit knapp 40.000 Euro bei einer Sozialgeldquote von circa 55 Prozent vermutlich deutlich zu hoch geschätzt.

Die Hauptursache wird darin zu suchen sein, dass Kinder sich nicht wie der Rest der Bevölkerung im Stadtgebiet verteilen: Typischerweise sind Familien mit Kindern in einigen Quartieren vermehrt und in anderen weniger oft anzutreffen. Die Richtung der Daten stimmt dabei durchaus, was sich auch an den statistischen Zusammenhängen in Tabelle 11 zeigt.

TABELLE 11: Korrelationen kommerzieller und kommunaler Typisierungsvariablen in den Kitas

	Anteil Sozialgeld	Anteil Migrationshintergrund
Kaufkraft pro Haushalt	- 0,82	
Migrationshintergrund (microm)		0,82

Daten: microm; Referat V.1 Stadtforschung und Statistik der Stadt Mülheim an der Ruhr; eigene Berechnung und Darstellung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

Da kommerzielle Anbieter wie microm auch Segmentierungsmöglichkeiten²⁶ anbieten, könnte es perspektivisch für Anwendungsziele wie die Typisierung von Kitas sinnvoll sein, diese auszuloten und zu erschließen, inwiefern dies zu besseren Schätzergebnissen, z.B. für den Anteil von Kindern mit Migrationshintergrund in Kitas, führt. Die vorangegangenen Auswertungen zeigen allerdings, dass sich auch ohne eine solche Segmentierung über die Wohnorte der Kinder benachteiligte und privilegierte Einrichtungen anhand der Geodaten von microm gut identifizieren lassen.

²⁶ Segmentierung bedeutet in diesem Zusammenhang, eine bestimmte Teilgruppe („Zielgruppensegment“) der gesamten Bevölkerung herauszugreifen. In diesem Fall wären z.B. Daten für Kinder unter sechs Jahren von Interesse.

Kita-Milieus

Abschließend soll darauf eingegangen werden, inwieweit durch das beschriebene Verfahren der Übertragung der Sinus-Geo-Milieus® auf diejenigen Kinder, die eine Kita besuchen, auch eine *milieuspezifische* Kita-Segregation abgebildet werden kann. Differenzieren sich die Kitas weiter aus, sodass sich Ansatzpunkte für zielgruppenorientierte Kommunikation bieten? Mithilfe des beschriebenen wohnortbasierten Verfahrens lassen sich dafür leider keine soliden Ansatzpunkte finden: Betrachtet man die Kitas anhand der höchsten Milieuwahrscheinlichkeit, so finden sich sieben Kitas, bei denen diese das konservativ-etablierte Milieu ist, und 79 Kitas, in denen das hedonistische Milieu die höchste Wahrscheinlichkeit hat. Die acht anderen Milieus weisen in keiner einzigen Kita die höchste Wahrscheinlichkeit auf.

Über das Verfahren der „überrepräsentierten“ Milieus lässt sich zwar Vielfalt erzeugen,²⁷ allerdings ist die Spannweite bei vielen der so berechneten Milieuwahrscheinlichkeiten in Kitas so klein, dass die Abweichungen als zu gering betrachtet werden müssen, als dass man ihnen angesichts der dem Verfahren zugrunde liegenden Unsicherheiten entsprechende Bedeutung zuweisen sollte. Tabelle 12 fasst die Spannweiten zwischen den Milieus auf der Ebene der Rasterzellen, der 28 statistischen Bezirke Mülheims, zwischen den 86 Kitas sowie schließlich die durchschnittliche Spannweite innerhalb der Kitas zusammen. Bei der durchschnittlichen Spannweite innerhalb der Kitas wurde für jede Milieuwahrscheinlichkeit in jeder Kita die Differenz zwischen dem Kind mit der jeweils höchsten und der jeweils niedrigsten Wahrscheinlichkeit gebildet und diese Differenz über alle Kitas gemittelt. Sie kann somit als Indikator für die Heterogenität der Kitas hinsichtlich des Milieus herangezogen werden. Dabei wird ersichtlich, dass innerhalb der einzelnen Kitas größere Unterschiede bestehen als zwischen den Kitas. Von milieuspezifischer Kita-Segregation in dem Sinne, dass sich z. B. die Kinder mit hohen Wahrscheinlichkeiten für das sozialökologische Milieu in einer Kita zusammenfinden würden, kann anhand des Vorgehens mit den Sinus-Geo-Milieus® also keine Rede sein. Dies bedeutet nicht, dass milieuspezifische Kita-Segregation generell nicht auftritt – es bedeutet aber, dass man diese nicht mithilfe des Verfahrens der Wohnortübertragung der Rasterwerte auf die Kinder sichtbar machen kann.

27 Analog zum Vorgehen in Abschnitt 3.4 wird für jedes Milieu in jeder Kita berechnet, inwieweit die Wahrscheinlichkeit über dem Kita-Durchschnitt liegt. Das Milieu, das die stärkste positive Abweichung zum Mittelwert aufweist, wird dann als das „überrepräsentierte“ gewählt.

TABELLE 12: Spannweiten der Milieuwahrscheinlichkeiten auf verschiedenen Ebenen, in Prozentpunkten

Milieu	100x100-Meter-Rasterzellen	Statistische Bezirke	Zwischen Kitas	Durchschnittlich innerhalb der Kitas
Konservativ-Etablierte	14,9	6,1	7,2	8,5
Liberal-Intellektuelle	11,8	6,7	5,1	7,1
Performer	16,0	7,3	5,4	8,2
Expeditive	15,2	5,1	3,7	6,3
Bürgerliche Mitte	15,0	4,9	3,8	7,8
Adaptiv-Pragmatische	22,5	3,8	3,3	7,2
Sozialökologische	11,6	3,3	3,5	5,8
Traditionelle	23,0	8,0	7,6	10,9
Prekäre	16,9	6,1	5,4	7,0
Hedonisten	31,1	15,0	13,7	17,7

Daten: microm; Referat V.1 Stadtforschung und Statistik der Stadt Mülheim an der Ruhr; eigene Berechnung und Darstellung

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.

Auch die Zusammenfassung der Milieuwahrscheinlichkeiten (nicht tabellarisch dargestellt) innerhalb der Kitas führt hier nicht weiter: Hinsichtlich der vertikalen Dimension lassen sich zwar deutliche Differenzen zwischen Leitmilieus und unteren Milieus zwischen den Kitas ausmachen, in Bezug auf die horizontale Dimension in die Pole „traditionell/modern“ wird die Streuung aber sogar geringer als bei der Betrachtung einzelner Milieus.

Dies kann auch darauf hindeuten, dass Kita-Segregation sich hauptsächlich anhand der vertikalen sozialstrukturellen Dimension vollzieht und die horizontale Dimension nur eine geringe Rolle spielt. Da Milieueffekte vor allem bei großen Freiheitsgraden in der Entscheidung zu erwarten sind und Kita-Plätze ein knappes Gut darstellen, ist dies auch eine theoriekonforme Interpretation. Aufgrund der benannten zahlreichen Unsicherheiten im Verfahren der Übertragung räumlicher Daten auf eine spezifische Subpopulation und der Intransparenz der Datengenerierung bleibt dies aber eine Hypothese.

6 Fazit

Der vorliegende Werkstattbericht sollte den (Mehr-)Wert kommerzieller kleinräumiger Daten für kommunale Fachplanungszwecke exemplarisch ausloten. In einem ersten Analyseschritt wurde dafür die kommunale Passung kommerzieller Daten an die verfügbaren amtlichen und kommunalen Statistiken überprüft. Es zeigte sich eine hohe Übereinstimmung – Migrationshintergrund und Kaufkraft werden dabei seitens der kommerziellen Daten leicht höher eingeschätzt als amtliche und kommunale Statistik erwarten lassen würden. Auf einer Rasterebene mit 100x100-Meter-Zellen ist die kleinräumige Passung hinsichtlich Bevölkerungs- und Haushaltszahl ausgesprochen gut. Für Kaufkraft und Migrationshintergrund stehen keine direkten Vergleichswerte zur Verfügung, ein Abgleich nach SGB-II-Quoten und Ausländeranteilen ergab ein insgesamt plausibles Bild. Damit steht kommunalen Akteuren ohne Zugang zu kleinräumigen Verwaltungsdaten eine vielversprechende kommerzielle Option offen.

Auch die „weichen“ Milieudaten zeichnen schlüssige Bilder der räumlichen Verteilung der Stadtgesellschaft, die für gezielte Ansprachen z. B. bei Bürgerbeteiligungsverfahren genutzt werden können (vgl. Kapitel 4). In ihrer Eigenschaft als Schätzwerte multinomialer logistischer Regressionsmodelle bilden sie allerdings extreme Lagen nicht beziehungsweise nur unzureichend ab: Auch in einem Armutsquartier mit einer SGB-II-Quote von über 90 Prozent werden noch Wahrscheinlichkeiten von gut 20 Prozent für die Leitmilieus geschätzt. Anwenderinnen und Anwender in der Stadtforschung sollten sich dies bewusst machen, da diese Eigenschaft zu einer Unterschätzung milieuspezifischer Segregation führt, auch wenn die Daten die Tendenz richtig zeigen.

Da Institutionen und Settings sowohl eine hohe soziale Wirksamkeit erzielen als auch politische und pädagogische Ansatzpunkte bilden, wurde zuletzt der Frage nachgegangen,

inwieweit sich die räumlichen Daten auf die Bevölkerung übertragen und dann wieder auf Einrichtungen aggregieren lassen. Exemplarisch wurde dies anhand der Kinder, die Kitas in Mülheim an der Ruhr besuchen, durchgeführt. Es zeigte sich, dass mithilfe dieses Vorgehens sehr gut benachteiligte und privilegierte Kitas identifizieren konnte, die Differenzierung hinsichtlich der horizontalen Dimension und der Milieus hingegen sehr schwach ausfiel. Ob dies dem methodischen Vorgehen geschuldet ist oder die Wahl einer Kita empirisch sehr viel stärker durch vertikale sozialstrukturelle Eigenschaften als durch Lebensstile und Milieus strukturiert wird, muss durch zukünftige Forschung geklärt werden. Perspektivisch kann das Wissen um milieuspezifische Kita-Segregation zu höherer Sensibilität bezüglich der Vielfalt frühkindlicher Förderungsbedarfe und der Entwicklung passgenauer Angebote führen, um jedem Kind die Chance auf eine gute Entwicklung und gesellschaftliche Teilhabe zu bieten. Aber auch so ist es mithilfe der genutzten Daten immerhin schon möglich, Kitas mit höherem Problemdruck zu identifizieren und so politische Maßnahmen zu ihrer Entlastung zu ergreifen.

7 Literatur

-
- Ache, Peter (2018). „Immobilienmarktbericht Deutschland 2017 bestätigt verbreitet steigende Immobilienpreise“. *Immobilienpreisentwicklungen in Deutschland. Erkenntnisse aus dem Arbeitskreis Immobilienpreise am 16. Oktober 2017 in Bonn. BBSR-Berichte KOMPAKT 1/2018*. Bonn. Hrsg. BBSR. 12–13.
- Bach, Maximilian, Josefine Koebe und Frauke Peter (2018). „Früher Kita-Besuch beeinflusst Persönlichkeitseigenschaften bis ins Jugendalter“. *DIW Wochenbericht* 15. 289–297.
- Barz, Heiner (2018). „Bildungsforschung mit den Sinus-Milieus®“. *Praxis der Sinus-Milieus®. Gegenwart und Zukunft eines modernen Zielgruppenmodells*. Hrsg. Bertram Barth, Berthold Bodo Flaig, Norbert Schäuble und Manfred Tautscher. Wiesbaden. 155–169.
- Barz, Heiner, und Rudolf Tippelt (2007). *Weiterbildung und soziale Milieus in Deutschland – Praxishandbuch Milieumarketing*. 2. Aufl. Bertelsmann Verlag. Bielefeld.
- Beck, Ulrich (2016). *Risikogesellschaft. Auf dem Weg in eine andere Moderne*. 23. Auflage. Frankfurt am Main.
- Becker, Irene (2007). „Verdeckte Armut in Deutschland. Ausmaß und Ursachen“. *Fachforum. Analysen & Kommentare* Heft 2. Hrsg. Friedrich-Ebert-Stiftung. Berlin.
- Blasius, Jörg, und Jürgen Friedrichs (2011). „Die Bedeutung von Lebensstilen für die Erklärung von sozialräumlichen Prozessen“. *Lebensstilforschung*. Hrsg. Jörg Rössel und Gunnar Otte. Wiesbaden.
- BMFSFJ – Bundesministerium für Familie, Senioren, Frauen und Jugend (2018). *Kindertagesbetreuung Kompakt. Ausbaustand und Bedarf 2017*. Ausgabe 3. Berlin. (Auch online unter http://www.fruehe-chancen.de/fileadmin/PDF/Fruehe_Chancen/Betreuungszahlen/Kita_Kompakt_Dritte_Ausgabe-BF.PDF, Download 2.10.2018.)

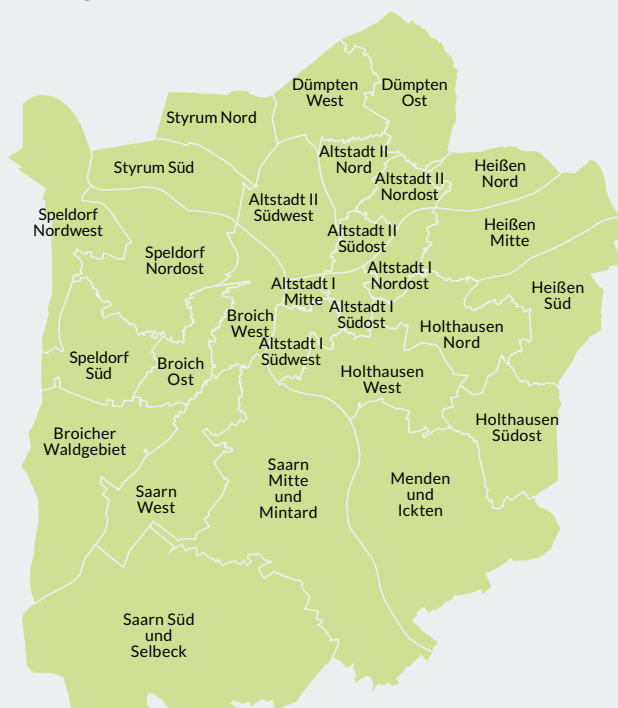
- Bourdieu, Pierre (1987). *Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft*. Frankfurt am Main.
- Bourdieu, Pierre, und Loïc J. D. Wacquant (2006). *Reflexive Anthropologie*. Frankfurt am Main.
- Burzan, Nicole (2011). *Soziale Ungleichheit. Eine Einführung in die zentralen Theorien*. 4. Auflage. Wiesbaden.
- Dangschat, Jens S. (2018). „Soziale Milieus in der Mobilitätsforschung“. *Praxis der Sinus-Milieus®. Gegenwart und Zukunft eines modernen Zielgruppenmodells*. Hrsg. Bertram Barth, Berthold Bodo Flaig, Norbert Schäuble und Manfred Tautscher. Wiesbaden. 139–153.
- Flaig, Berthold Bodo, und Bertram Barth (2017). „Hoher Nutzwert und vielfältige Anwendung: Entstehung und Entfaltung des Informationssystems Sinus-Milieus®“. *Praxis der Sinus-Milieus®. Gegenwart und Zukunft eines modernen Zielgruppenmodells*. Hrsg. Bertram Barth, Berthold Bodo Flaig, Norbert Schäuble und Manfred Tautscher. Wiesbaden. 3–21.
- Geißler, Rainer (2014). *Die Sozialstruktur Deutschlands*. 7., grundlegend überarbeitete Auflage. Wiesbaden.
- Goebel, Jan, und Lukas Hoppe (2015). *Lebenslagen in Deutschland. Ausmaß und Trends sozialräumlicher Segregation in Deutschland. Armuts- und Reichtumsberichterstattung der Bundesregierung im Auftrag des Bundesministeriums für Arbeit und Soziales*. Berlin. (Auch online unter http://www.armuts-und-reichtumsbericht.de/SharedDocs/Downloads/Service/Studien/abschlussbericht-sozialraeumliche-segregation.pdf?__blob=publicationFile&v=3, Download 2.10.2018.)
- Grohs, Stephan (2012). „Die Umsetzung des Neuen Steuerungsmodells: eine empirische Bestandsaufnahme“. *Modernisierung der kommunalen Sozialverwaltung: soziale Arbeit unter Reformdruck?*. Hrsg. Julia Hagn, Peter Hammerschmidt und Juliane Sagebiel. Neu-Ulm. 103–125.
- Groos, Thomas (2015). *Gleich und gleich gesellt sich gern. Zu den sozialen Folgen freier Grundschulwahl. Arbeitspapiere wissenschaftliche Begleitforschung „Kein Kind zurücklassen!“*, Band 5. Gütersloh. (Auch online unter http://www.zefir.rub.de/mam/content/kekiz_wb_5_gr%C3%BCn_final.pdf, Download 2.10.2018.)
- Groos, Thomas, und Nora Jehles (2015). *Der Einfluss von Armut auf die Entwicklung von Kindern. Ergebnisse der Schuleingangsuntersuchung. Arbeitspapiere wissenschaftliche Begleitforschung „Kein Kind zurücklassen!“*, Band 3. Gütersloh. (Auch online unter http://www.zefir.ruhr-uni-bochum.de/mam/content/03_werkstattbericht_einfluss_von_armut_final_opt.pdf, Download 2.10.2018.)

- Groos, Thomas (2016). *Schulsegregation messen*. Arbeitspapiere wissenschaftliche Begleitforschung „Kein Kind zurücklassen!“, Band 6. Gütersloh. (Auch online unter http://www.zefir.rub.de/mam/content/kekiz_wb_6_gr%C3%BCn_final.pdf, Download 2.10.2018.)
- Groos, Thomas, Carolin Trappmann und Nora Jehles (2018). „Keine Kita für alle“. *Zum Ausmaß und zu den Ursachen von Kita-Segregation*. Arbeitspapiere wissenschaftliche Begleitforschung „Kein Kind zurücklassen!“, Band 12. Gütersloh. (Auch online unter http://www.zefir.ruhr-uni-bochum.de/mam/content/kekiz_wb_12_gruen_final.pdf, Download 2.10.2018.)
- Groos, Thomas, Carolin Trappmann, Nora Jehles und Volker Kersting (2018). *Die Illusion der Chancengleichheit – Ausgewählte Ergebnisse der kommunalen Wirkungsforschung mit Mikrodaten*. Arbeitspapiere wissenschaftliche Begleitforschung „Kein Kind zurücklassen!“, Band 13. Gütersloh.
- Hradil, Stefan (2006). „Soziale Milieus – eine praxisorientierte Forschungsperspektive“. *Aus Politik und Zeitgeschichte* 44–45. 3–9.
- Jeworutzki, Sebastian, Jörg-Peter Schräpler und Stefan Schweers (2016). „Soziale Segregation – Die räumliche Ungleichverteilung von SGB-II-Bezug in NRW“. *Sozialbericht NRW 2016. Armuts- und Reichtumsbericht*. Hrsg. MAIS NRW. Düsseldorf.
- Klee, Andreas (2003). „Lebensstile, Kultur und Raum: Anmerkungen zum Raumbezug soziologischer Gesellschaftsformationen“. *Geographische Zeitschrift* (91) 2. 63–74.
- Kreckel, Reinhard (1983). „Theorien sozialer Ungleichheit im Übergang“. *Soziale Ungleichheiten. Soziale Welt, Sonderband 2*. Hrsg. Ders. Göttingen. 3–12.
- Kuder, Thomas (2016). „Starke Lokale Demokratie: Leitlinien für eine hochwertige, inklusive Bürgerbeteiligung“. *vhw werkSTADT* 8, Oktober 2016. Hrsg. vhw – Bundesverband für Wohnen und Stadtentwicklung. Berlin.
- Küppers, Rolf (2018). „Übertragung in den Raum: Die Sinus-Geo-Milieus®“. *Praxis der Sinus-Milieus®. Gegenwart und Zukunft eines modernen Zielgruppenmodells*. Hrsg. Bertram Barth, Berthold Bodo Flaig, Norbert Schäuble und Manfred Tautscher. Wiesbaden. 95–102.
- OECD – Organisation for Economic Co-operation and Development (2016). *PISA 2015 Ergebnisse. Exzellenz und Chancengerechtigkeit in der Bildung*. Band I. (Auch online unter https://read.oecd-ilibrary.org/education/pisa-2015-ergebnisse-band-i_9789264267879-de, Download 2.10.2018.)
- Otte, Gunnar, und Jörg Rössel (2011). „Lebensstile in der Soziologie“. *Kölner Zeitschrift für Soziologie und Sozialpsychologie Sonderheft* 51/2011. Wiesbaden. 7–34.

- Rössel, Jörg, und Michael Hoelscher (2012). „Lebensstile und Wohnstandortwahl“. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 2. Wiesbaden. 303–327.
- Scheuner, Joachim (2006). „Wohnen und Aktionsraum: Welche Rolle spielen Lebensstil, Lebenslage und Raumstruktur?“. *Geographische Zeitschrift* Bd. 94, 1. 43–62.
- Schräpler, Jörg-Peter, und Wolfgang Seifert (2008). *Kleinräumige Einkommensstrukturen und ihr sozialer Kontext in Nordrhein-Westfalen*. Essen und Düsseldorf.
- Simmel, Georg (2014)[1903]. „Die Großstädte und das Geistesleben“. *Die Großstadt. Vorträge und Aufsätze zur Städteausstellung. Jahrbuch der Gehe-Stiftung Dresden*. Band 9, 3. Auflage. Hrsg. Theodore Petermann. Dresden und Berlin, 185–206.
- Spellerberg, Annette (2014). „Was unterscheidet städtische und ländliche Lebensstile?“. *Urbane Ungleichheiten. Neue Entwicklungen zwischen Zentrum und Peripherie*. Hrsg. Peter A. Berger, Carsten Keller, Andreas Klärner und Rainer Neef. Wiesbaden. 199–232.
- Statistische Ämter des Bundes und der Länder (2004). „Ergebnisse des Zensus 2011“. *Wirtschaft und Statistik* 8. 813–833. https://www.zensus2011.de/SharedDocs/Downloads/DE/Publikationen/Aufsaeetze_Archiv/2004_%20Destatis_Ergebnisse_des_Zenustests.pdf?__blob=publicationFile&v=25 (Download 2.10.2018).
- Statistisches Bundesamt (2018). „Glossar. Migrationshintergrund“. <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/MigrationIntegration/Glossar/Migrationshintergrund.html?view=getColorboxEntry> (Download 2.10.2018).
- VDSt – Verband Deutscher Städtestatistiker (Hrsg.) (2013). *Migrationshintergrund in der Statistik – Definitionen, Erfassung und Vergleichbarkeit. Materialien zur Bevölkerungsstatistik*. Heft 2. Erfurt.
- vhw – Bundesverband für Wohnen und Stadtentwicklung e.V. (Hrsg.) (2013). „vhw-Kommunikationshandbuch. Praxisbezogene Kommunikation mit den Milieus der Stadtgesellschaft“. *vhw Schriftenreihe* 4. Berlin.
- Wiedenbeck, Michael, und Cornelia Züll (2010). „Clusteranalyse“. *Handbuch der sozialwissenschaftlichen Datenanalyse*. Hrsg. Christof Wolf und Henning Best. Wiesbaden. 525–552.

Anhang

ABBILDUNG 7: Ortslagen in Mülheim an der Ruhr, Statistische Bezirke



Daten: Referat V.1, Stadtforschung und Statistik, Stadt Mülheim an der Ruhr

© Bertelsmann Stiftung und ZEFIR, mit finanzieller Unterstützung des Landes NRW und des Europäischen Sozialfonds.



Sozial- und Bildungsmonitoring für kommunale Prävention

Bildungs- und Gesundheitschancen von Kindern und Jugendlichen sind räumlich ungleich verteilt. Das Sozial- und Bildungsmonitoring ermöglicht einen Blick auf die kommunalen Unterschiede, deren Kenntnis von zentraler Bedeutung ist, um konkrete Handlungsbedarfe abzuleiten. In diesem Werkstattbericht werden Indikatoren zu den soziodemographischen und fiskalischen Rahmenbedingungen sowie zu den Übergängen in die Kita, in die Grundschule, in die weiterführende Schule, in die Sekundarstufe II und in den Arbeitsmarkt analysiert. Gegenstand der Untersuchungen sind die Kreise und kreisfreien Städte Nordrhein-Westfalens.



Keine Kita für alle

Kindertageseinrichtungen sind ein Grundpfeiler der deutschen Bildungslandschaft. Die Betreuungsplätze werden von öffentlichen und freien Trägern der Jugendhilfe angeboten. In diesem Werkstattbericht wird der Frage nachgegangen, inwieweit die Trägervielfalt Einfluss auf die Zusammensetzung der Kinder in den Kitas hat. Für drei Beispielkommunen werden lokale Kita-Strukturen analysiert und Mechanismen der Kita-Wahl identifiziert. Dabei zeigt sich, dass der Zugang in die Kitas sozial und ethnisch selektiv ist und die Konfession des Kindes eine große Rolle bei der Kita-Wahl spielt.

Die Nutzung kommerziell angebotener kleinräumiger Daten gewinnt in Privatwirtschaft und Wissenschaft zunehmend an Bedeutung. Auch für Kommunen und deren Fachplanungen ist die Nutzung solcher Daten eine attraktive Option. Bislang liegen jedoch kaum (dokumentierte) Validierungen hierzu vor. Der vorliegende Werkstattbericht setzt an dieser Stelle an und möchte den (Mehr-)Wert kommerzieller kleinräumiger Daten für kommunale Analysen und Fachplanungszwecke ausloten. Hierfür wurde exemplarisch untersucht, inwiefern sich die Daten der Firma microm dafür eignen, Segregation von Kitas in Mülheim an der Ruhr abzubilden. Beim Abgleich der amtlichen und kommerziellen Daten zeigt sich zunächst sowohl auf der kommunalen als auch auf der kleinräumigen Ebene eine gute Passung. Überträgt man die Daten wohnortbasiert auf die Kinder der Stadt Mülheim an der Ruhr und betrachtet die Unterschiede beim Kita-Besuch, lassen sich soziostrukturell benachteiligte und privilegierte Kitas gut identifizieren. Milieuspezifische Kita-Wahlen lassen sich hingegen (zumindest) mit diesem Verfahren nicht valide abbilden, da die Unterschiede in den Wahrscheinlichkeiten zu gering sind.

The use of commercially available small-scale data is becoming increasingly important in private business and in science. The use of such data is also an attractive option for municipalities and their sectoral planning. Up to now, however, there have been scarcely any (documented) validations available for this purpose. The current workshop report draws on this point and aims to explore the (added) value of commercial, small-scale data for municipal analysis and sectoral planning purposes. For this, the suitability of data of the company microm was examined for depicting the segregation of nurseries in Mülheim an der Ruhr in North Rhine-Westphalia. When comparing official and commercial data, a good fit was initially found at both the municipal and small-scale levels. When the data is applied to the children of Mülheim an der Ruhr according to their place of residence, differences in nursery attendance are observed and socio-structurally disadvantaged as well as privileged nurseries are easy to identify. However, a milieu-specific selection of a nursery by the parents cannot be validly mapped with this method, since the differences in probabilities are too small.

www.bertelsmann-stiftung.de/kekiz
www.zefir.ruhr-uni-bochum.de

ISSN-Print 2199-6393
ISSN-Internet 2199-6407

Neher-Durbin

Geonimien-Durbin

Neher-Durbin

Geonimien-Durbin

Neher-Durbin

Geonimien-Durbin

Neher-Durbin

Geonimien-Durbin

Neher-Durbin

Geonimien-Durbin

Neher-Durbin

Geonimien-Durbin

Neher-Durbin

Geonimien-Durbin

Neher-Durbin

Geonimien-Durbin

Neher-Durbin

Geonimien-Durbin