Hogrebe, Nina; Pomykaj, Anna; Schulder, Stefan

Segregation in early childhood education and care in Germany. Insights on regional distribution patterns using national educational studies

Diskurs Kindheits- und Jugendforschung 16 (2021) 1, S. 36-56

Segregation in Early Childhood Education and Care in Germany: Insights on Regional Distribution Patterns Using National Educational Studies

Nina Hogrebe, Anna Pomykaj, Stefan Schulder

Abstract
Early Childhood Education and Care (ECEC) is believed to contribute to educational equality and to serve social inclusion and democracy. Segregation in day-care centres counteracts these aims but has hardly been researched in Germany so far. We describe ethnic/linguistic and social segregation at different regional levels (federal states as well as East and West Germany more generally) using data from the Early Childhood Education and Care Quality Study in the Socio-Economic Panel (K^2ID-SOEP) and the National Educational Panel Study (NEPS). We find pronounced differences in distribution patterns of ECEC settings' composition especially between West and East Germany and discuss the research implications of our findings.

Keywords: Early Childhood Education and Care, Educational Inequality, Segregation, Distribution Patterns, National Educational Studies

Segregation in der Kindertagesbetreuung in Deutschland: Ergebnisse zu regionalen Verteilungsmustern auf Basis nationaler Bildungsstudien

Zusammenfassung
An frühkindliche Bildung und Betreuung wird die Erwartung herangetragen, zum Abbau von Bildungsungleichheit beizutragen und Inklusion zu fördern. Segregation in Kindertageseinrichtungen (Kitas) wirkt diesen Zielen grundsätzlich entgegen. Bislang gibt es aber für Deutschland nur wenig empirische Erkenntnisse über das Ausmaß und mögliche Ursachen von entsprechenden Entmischungsprozessen. Die vorliegende Studie beschreibt anhand von Daten der Erhebung „Kinder und Kitas in Deutschland" im Rahmen des Sozio-ökonomischen Panels (K^2ID-SOEP) und des Nationalen Bildungspanels (NEPS) das Ausmaß und die Varianz der Kita-Zusammensetzung auf regionaler Ebene. Wir finden unterschiedliche Verteilungsmuster insbesondere zwischen Ost- und Westdeutschland und diskutieren die Forschungsimplikationen unserer Ergebnisse.

Schlagwörter: Kindertagesbetreuung, Bildungsungleichheit, Segregation, Verteilungsmuster, Nationale Bildungsstudien

1 Introduction

Early Childhood Education and Care (ECEC)[1] is believed to contribute to educational equality by being especially effective for disadvantaged children, but it is also considered to serve social inclusion and democracy more generally (*European Commission* 2011). *Vandenbroeck* (2015) differentiates between individual achievements through ECEC and ECEC as a context for learning how people live together. As to the individuals' perspective, the author states that the demographic make-up of ECEC settings is linked to children's development (child development perspective); the most positive effects are realized in mixed groups whereas a concentration of children from under-privileged or disadvantaged families disparages learning outcomes (for such composition effects in ECEC see, for example, *Hogrebe/Pomykaj* 2019; *Niklas/Tayler* 2018). Additionally, he considers enrolment in ECEC as a first step into society and argues that it is important that children and families are socialized as early as possible in contexts of diversity (democratic theory perspective). Similarly, *Gans* (2007) describes educational institutions as symbols and institutions of democratic pluralism. Institutionalized ECEC is important to provide "opportunities to experience meaningful intergroup contact (…) because children's early life experiences can have long-term consequences for their developing intergroup attitudes and beliefs" (*Tropp/Saxena* 2018, p. 1).

Segregation – defined as an unequal distribution of population groups resulting from spatial differentiation, sorting, and separation processes – counteracts both aims. The spatial concentration of population groups has originally been observed in urbanism, but the phenomenon also applies to educational institutions resulting in different demographic make-ups (compositions) of learning environments. Segregated ECEC relates to individual learning outcomes but also prevents experiences of diversity and, thereby, contradicts the idea of social inclusion and democracy. From that perspective, any kind of homogeneity limits exposure to and experiences of diversity for all children. From a child development perspective, a concentration of disadvantaged or minority children undercuts equal educational opportunities for those children as it is negatively related to child care quality and child development (*Fram/Kim* 2012).

Despite its relevance, little is known about the peer-related contexts in early education institutions so far, and only a few studies explicitly address segregation in ECEC. To conduct reliable research in this still under-researched area, we need to find information on relevant factors that should be considered in respective analyses. While it seems immediately understandable that local population and supply structures influence the demographic makeup of ECEC settings, it is less clear to what extent higher levels might play a role. Against this background, we explore two regional layers that are generally important in the German ECEC landscape: the difference between East and West Germany as well as federal states.

2 Findings on Segregation in Early Childhood Education and Care

Studies on segregation usually include dimensions of social and/or ethnic/racial/linguistic segregation but might use different measures or definitions. The most easily understandable measures are composition, i.e. the relative proportions of population groups in ob-

served units, and concentration, i.e. the extent to which students are enrolled in settings with high proportions of certain population groups.[2]

In her study on preschool segregation in the U.S., *Frankenberg* (2016) defines racially isolated (non-)white settings as those with enrolments of 90 percent or more of students of colour or white students respectively. She shows that nationwide about 40 percent of preschoolers in public school-based programs attend racially isolated preschools; most of them (28.2 percent) visit racially isolated non-white schools. With more than 50 percent, this especially applies to black and Hispanic children as well as children lacking English proficiency (LEP). By contrast, about 20 percent of white children and eleven percent of non-LEP students visit racially isolated white preschools. Albeit there have been slight declines in racial isolation by a few percentage points, a recent update of the study presents quite similar results (*Piazza/Frankenberg* 2019). In this report, the authors additionally show that non-white sub-groups are also generally quite segregated from each other. Besides these national-level segregation patterns, both reports show pronounced differences between the states. In relation to these inter-state differences, *Piazza* and *Frankenberg* (2019) find it hardly surprising that "highly segregated non-white preschools are less likely to exist in places with relatively low non-white populations" (p. 27). The authors conclude that "preschool diversity is challenged by an overwhelmingly white statewide composition" (p. 28).

Another recent study adds to these research results by not only looking at public school-based programs but including data on other centre-based as well as home-based programs (*Urban Institute* 2019). Generally, the finding of a U-shaped distribution of Black or Hispanic enrolment shares supports the impression of highly segregated ECEC in the USA. Only a few programs (less than 20 percent) have moderate black or Hispanic enrolment shares (i.e. 30 to 70 percent). Rather, almost half of the programs have either less than 10 (about one third of the programs) or more than 90 percent (almost 20 percent) Black or Hispanic children. Differential data analyses show state-level as well as regional variation as follows: ECEC is segregated the most in the Northeast and the least in the Midwest, and programs in suburban communities are less segregated than those in urban or rural communities. Additionally, ECEC is more segregated than later grades of schooling. For example, services in early childhood are twice as likely to be nearly 100 percent black or Hispanic as in kindergarten and first grade (nearly 20 percent versus ten percent). The authors attribute this finding almost exclusively to home-based programs which are more segregated than centre-based ECEC.

Most of the studies on segregation in the U.S. focus on racial segregation, and there is only little research on social segregation. In relation to poverty, *Reid* et al. (2015) describe that children are often clustered in preschool classrooms that are both high-minority and high-poverty. The authors define high-minority classrooms as classrooms with 70 to 100 percent minority children. They find that almost half of the children in their sample (47.1 percent) attend such high-minority classrooms in which about 75 percent of the children are poor as well. Only very few children (17 percent) visit classrooms that are racially diverse (i.e. 30 to 70 percent minority) and medium-high income (i.e. with an average family income of about 30,000 USD). To the authors, such diversity in preschool programs is "more the exception than the rule" (p. 7).

The above depicted research problematizes pronounced racial (and social) segregation in ECEC in the USA. However, these results cannot per se be transferred to the German context due to differences in history, demographic structures, and programs. First, in the U.S. segregation in education has been a hotly debated topic for more than 50 years now;

efforts of de-segregation in the school sector have been followed by processes of re-segregation (*Frankenberg/Lee* 2002). Second, population structures in the U.S. with a majority of non-white (i.e. mostly Hispanic, Black, and Asian) children (*Frankenberg* 2016, p. 10) are quite different to the respective demographic structures in Germany. Third, the childcare landscape in America is quite fragmented with low-income children, for example, being enrolled in subsidized programs, and upper-income children visiting fee-for-service programs (*Kagan* 1990). Targeted programs like Head Start, for example, are not very heterogeneous as they primarily serve economically disadvantaged students unlike universal programs in Germany.[3]

So far, however, there has not been much research on segregation in ECEC for the German context. Some educational effectiveness studies consider peer context information. They report means of about 20 to 30 percent of children with German as a second language or migration background in settings or groups. A variation from zero up to 90 or even 100 percent and standard deviations of about 20 to 30 percent indicate pronounced differences in compositions and, thus, ethnic segregation (e.g. *Anders* et al. 2012; *Kuger/Kluczniok* 2008; *Tietze* et al. 2013).

Additionally, monitoring reports say that children with a migration background concentrate in certain settings; about one third of the children who do not predominantly speak German at home visit ECEC settings in which the majority of the children do not speak German at home either. The reports also point to regional differences. A concentration of these children in certain day-care centres can be found especially in urban agglomerations (*Authoring Group Educational Reporting* 2016) where in some districts the institutions are challenged in a special way (*Authoring Group Educational Reporting* 2020). By contrast, the respective proportions are remarkable low in rural areas in East Germany (*Olszenka/Meiner-Teubner* 2020). Nevertheless, such regional patterns do not necessarily follow a linear logic: There are some districts that have comparatively low proportions of children with a migration background but still a high concentration of this population group in some day-care centres; others have rather high numbers of migrant children who are more evenly distributed among the facilities (*Hüsken* 2011).

The National Educational Report in 2016 states that segregation in the school system is similar to that in early education but more closely linked to performance and socio-economic risk (*Authoring Group Educational Reporting* 2016). However, individual local studies support the impression that ethnic segregation in ECEC might exceed segregation in primary schools and increase over time when the proportions of the respective population groups grow and vice versa (*Hogrebe* 2014; 2016a; 2016b). Some of those studies also shed light on social segregation. With regard to children living in poverty, for example, proportions of children receiving social benefits range from zero to a good 70 percent. High proportions of poor children in day-care centres correlate with the respective proportions of migrant children (*Strohmeier* et al. 2014) and mostly affect children living in poverty themselves; according to one study, half of them are cared for in the fifth of the settings with above-average poverty rates (*Hock/Holz/Kopplow 2014*).

One has to keep in mind, though, that this local research is mostly conducted in the federal state of North-Rhine-Westphalia (and Hesse), and the results cannot be automatically be applied to other cities, regions or states. Also, the research presented above implicates that small-scale population structures are important parameters but not sufficient to fully explain segregation patterns (*Hogrebe* 2014; *Drange/Telle* 2020); other factors discussed are, for example, supply structures and administrative regulations (*Hock/*

Holz/Kopplow 2014; Hogrebe 2016a; *Strohmeier* et al. 2014). As aspects relating to political regulation and systemic structures are set at the level of federal states in Germany, this analysis level might be a relevant context information forming segregation patterns. Due to different population structures as well as different historic traditions of the respective ECEC systems, a more general differentiation between East and West Germany should be considered, too (*Becker/Schober* 2017).

3 Exploring Regional Differences in Segregation in German ECEC: Data and Methods

Against the background of the research on segregation depicted above, we explore ethnic and social segregation in German ECEC using national-level data, and, in doing so, especially look for differences between regions and federal states that have become visible in other countries.[4] Our main research question is whether such regional differences in segregation patterns can be observed in Germany, too. In line with international research, we expect that there are no day-care centres with high concentrations of migrant children in East Germany as overall population structures show low percentages of these children (9 percent). For West Germany, the respective percentages are higher (32 percent) so that we also assume more pronounced ethnic segregation here (*Lochner/Jähnert* 2020). We do not believe to find any pronounced differences in social segregation as the proportions of persons receiving basic income support, for example, are quite similar in both regions (about 9 percent and 11 percent respectively) (*Destatis* 2019). Additionally, we analyse if segregation patterns further vary between federal states as this might be mirroring different implementation laws and subsequent political regulations.

In order to allow for a comprehensive picture that includes several dimensions and indicators of ECEC setting composition, we draw our analyses on two data sources: The Early Childhood Education and Care Quality Study in the Socio-Economic Panel (K^2ID-SOEP) (*Schober* et al. 2017) and the National Educational Panel Study (NEPS) (*Blossfeld/Roßbach/von Maurice* 2011). Both studies originally gather individual level data but additionally conduct surveys in day-care centres. We use these datasets that contain setting level information for our analyses.

3.1 Data and Samples[5]

Based on the household sample of the SOEP subsamples A to K and the panel study Families in Germany (FiD) from 2014, the K^2ID-SOEP study[6] gathered additional data from the directors and educators of the respective children's ECEC settings. This information on group (educators) and overall setting (directors) level is provided in two separate datasets. In 2015, a second wave was conducted starting from the SOEP subsample M (migration sample) in order to compensate for an underrepresentation of day-care centres visited by children from migrant households.

Similarly, NEPS focuses on individuals as target persons, but additionally offers context information of ECEC settings provided by the child-care centres' directors and educators. The information at both levels is provided by one person respectively. To use this information that also contains the day-care centres' or groups' compositions, we reshaped

the datasets: Each case in our analysis represents one day-care centre from Starting Cohort 1 (SC 1; Newborns) and Starting Cohort 2 (SC 2; Kindergarten).[7] Within each starting cohort we focused on one sample wave which includes the most relevant information in terms of surveyed variables and case numbers: wave 1 in SC 2 (2011) and wave 4 in SC 1 (2015). Information on SC 2 rely on a representative nationwide sampling process. Day-care centres were sampled indirectly via primary schools. For SC 1 the data was collected via interviewed mothers handing out questionnaires to day-care centre personnel and is affected by sample selectivity: 63 percent of the mothers who forwarded the survey had a tertiary education in contrast to only about one third of the women aged 30 to 34 in Germany (*Destatis* 2015a). It is a common phenomenon that people with lower formal education are less likely to participate in such surveys (*Jacob/Heinz/Décieux* 2013).

After data cleansing, we used information from 165 to 657 directors and 172 to 685 educators who answered questions with regard to the day-care centres' composition at setting (directors) respectively group level (educators) (see Table 1).

Table 1: Samples and Sample Sizes

Total number of cases of day-care centres	K^2ID-SOEP		NEPS	
	Wave 1 (2014)	Wave 2 (2015)	SC 1 (2015)	SC 2 (2011)
Directors (= setting level)	657	167	492	165
Educators (= group level)	685	172	505	207

Note: The table provides information on case numbers of the respective K^2ID-SOEP and NEPS samples.

3.2 Variables

Both data sources contain information which can be used to describe ethnic/linguistic and social segregation, but no single sample covers all relevant indicators of the respective segregation dimensions. Thus, to provide a comprehensive picture, we include variables from all data sources. For some variables, respondents were directly asked to estimate the relative proportions of different population groups in their day-care centres or groups. In other cases, they had to state the exact quantity of certain children which we converted into proportions based on the total number of children in the respective groups or settings. The wording of the questions slightly varies between the studies, samples and respondents. However, as the K^2ID-SOEP instruments are based on NEPS questionnaires, we believe these differences to be negligible and assume that the respective operationalizations allow a high degree of comparability. Appendix 1 provides detailed information about the questions and variables which are shortly described below.

Ethnic/Linguistic Segregation

Three variables are suitable for investigating setting compositions that relate to ethnic or linguistic segregation. First, both studies provide information on children having learned or are mainly speaking a *language other than German* at home at group and setting level. Second, in the K^2ID-SOEP study data on the number of children *in need of language support* was offered by the educators. Third, in NEPS educators and directors stated the numbers/proportions of children having a *migration background*, i.e. children or at least one of their parents having been born abroad.

Social Segregation

We used three variables to describe social segregation: In the K^2ID-SOEP study directors were asked to state the number of children *exempt from fees*. This variable includes families who pay the lowest possible fee for childcare services or no fee at all on the basis of a low household income and can be used as a proxy for poverty. NEPS encompasses information about the proportion of children with at least one parent who has completed a higher educational program (*parents with higher education*) and the respondents' assessment of the share of children from families from a rather *lower social class.*

3.3 Analysis Strategy

To explore regional differences in segregation in German ECEC, we use composition (and concentration) as measures. In a first step, we focus on national-level segregation in order to establish the reference frame for the subsequent comparisons. In a second step, we investigate differences in segregation at regional level (West Germany versus East Germany). Finally, we compare distribution patterns in the federal states. For the latter attempt, we only use K^2ID-SOEP data due to small sample sizes in NEPS.

In the following, we present differences in the variation of day-care centres' compositions by region and states using descriptive statistics. Graphs are used to further illustrate distribution patterns at regional and state level. The main focus lays on differences in the variances between setting composition in the regions and states which are statistically analysed using Analysis of Variances (ANOVA) and additional post-hoc tests. To analyse the K^2ID-SOEP data, SPSS Version 26 and Stata/IC 16.1 was used. The analyses of the NEPS data were performed with the statistics software SPSS 25, Stata/SE 15.1, and R version 3.6.3 (*R Core Team* 2020) in the RemoteNEPS environment from the Leibniz Institute for Educational Trajectories (LIfBi).

4 Results

For reasons of simplicity, we mainly present results from our analyses of the K^2ID-SOEP data in the following because they are based on the bigger sample sizes. Additional information gained from the NEPS data are provided if the respective segregation dimension or indicator is not covered by the K^2ID-SOEP data. Also, due to statistically significant correlations between group and centre level compositions (ranging from $r = .74$ to $r = .91$; see Appendix 2), we only draw on information at group level if information at setting level is not available. When a variable was covered by more than one data set and/or at both levels, a comparison of the respective results reveals no substantial differences so that we present this information only once (all information not presented here can be obtained from the authors).

4.1 Segregation at National Level

Table 2 displays results for setting composition in relation to ethnic and linguistic segregation in Germany overall. The displayed mean values – i.e. the average proportion calculated over all settings – (19 percent for the proportion of children with a *language other than German,* 16 percent of *children in need of language support,* and 26 to 27 percent for children having a *migration background*) are consistent with other studies (e.g. *Anders* et al. 2012; *Biedinger/Becker* 2010; *Hogrebe/Pomykaj* 2019; *Kuger/Kluczniok* 2008; *Tietze* et al. 2013) and in line with average proportions of these children participating in German ECEC (i.e. 17 percent when German is not primarily spoken in the family and 26 percent when at least one parent is not of German origin) (*Destatis* 2015b, own calculations). Standard deviations of similar sizes as well as an overall distribution ranging nearly always from zero to 100 percent point at pronounced differences between the day-care centres and, thus, marked ethnic and linguistic segregation.[8]

Table 2: Ethnic and linguistic segregation at national level

	n	M	SD	Mdn	Min	Max
Language other than German						
K^2ID-SOEP	769	19	21	11	0	100
Wave 1 (A-K & FiD)	609	16	19	9	0	100
Wave 2 (migration)	160	30	23	27	0	95
In need of language support						
K^2ID-SOEP	583	16	19	11	0	100
Wave 1 (A-K & FiD)	467	16	19	11	0	100
Wave 2 (migration)	116	19	22	11	0	100
Migration background						
NEPS SC 1	504	26	25	19	0	98
NEPS SC 2	134	27	23	23	0	94

Note: The table presents descriptive statistics for linguistic segregation at national level based on K^2ID-SOEP and NEPS data. *n* = number of cases, *M* = Mean, *SD* = standard deviation, *Mdn* = Median, *Min* = Minimum, *Max* = Maximum. Except for the number of cases, all values refer to proportions in percent.

The extent of social segregation is displayed in Table 3. In line with previous research, social segregation is also quite high albeit a little bit smaller than ethnic/linguistic segregation: The proportions of children *exempt from fees,* with *parents with higher education* and *from lower social classes* also vary from zero to at least about 76 percent or even 100 percent.

Table 3: Social segregation at national level

	n	M	SD	Mdn	Min	Max
Exempt from fees						
K^2ID-SOEP	412	17	19	11	0	100
Wave 1 (A-K & FiD)	330	17	20	10	0	100
Wave 2 (migration)	82	18	19	13	0	94
Parents with higher education						
NEPS SC 2	132	25	23	18	0	93
Lower social class						
NEPS SC 2	140	21	21	11	0	76

Note: The table presents descriptive statistics for social segregation at national level based and K^2ID-SOEP and NEPS data. *n* = number of cases, *M* = Mean, *SD* = standard deviation, *Mdn* = Median, *Min* = Minimum, *Max* = Maximum. Except for the number of cases, all values refer to proportions in percent.

4.2 Regional Differences in Segregation

Further analyses of ethnic/linguistic and social segregation in West and East Germany (except Berlin) reveal that there are striking differences between those two regions especially for ethnic/linguistic segregation (see Table 4). In this regard, we find average proportions of children with a *language other than German*, *in need of language support* and *migration background* of approximately 20 to 30 percent in West Germany. With values of ten percent at the most, these proportions are much lower in East Germany. These differences reflect the average participation rates of the respective population groups in German ECEC settings according to official statistics (*Destatis* 2015b).

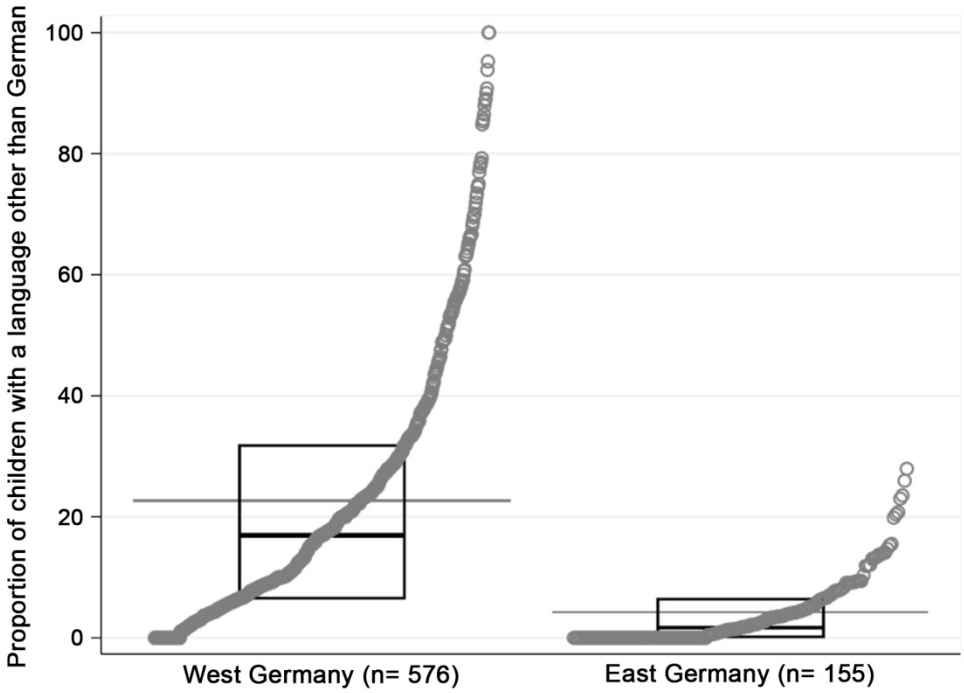
Table 4: Linguistic and ethnic segregation at regional level

	n	M	SD	Mdn	Min	Max
Language other than German						
K^2ID-SOEP						
West Germany*	576	23	22	17	0	100
East Germany*	155	4	6	2	0	28
				Welch's F(1, 697) = 335.617, p = .000		
In need of language support						
K^2ID-SOEP						
West Germany*	428	18	21	13	0	100
East Germany*	122	10	13	7	0	100
				Welch's F(1, 225) = 20.863, p = .000		
Migration background						
NEPS SC1						
West Germany*	397	28	25	20	0	98
East Germany*	58	8	11	4	0	50
				Welch's F(1, 161) = 100.849, p = .000		

Note: The table presents descriptive and nonparametric test statistics for ethnic and linguistic segregation between West and East Germany based on K^2ID-SOEP and NEPS data. Cases from Berlin are excluded from the data. *n* = number of cases, *M* = Mean, *SD* = standard deviation, *Mdn* = Median, *Min* = Minimum, *Max* = Maximum, *Welch's* $F(v_1, v_2)$ = Welch's ANOVA F-ratio with v_1 and v_2 (rounded) degrees of freedom, *p* = probability value, * = indicates statistically significant differences. Except for the number of cases and the information on statistical tests, the values refer to proportions in percent.

The lower mean proportions in East Germany translate into smaller differences between day-care centres as well. This is reflected by both comparatively small standard deviations and lower maximum values. While the latter consistently reach 98 to 100 percent in West Germany, they are between 28 and about 50 percent at the most in East Germany (except for the proportion of children *in need of language support*; a reason for this might be that this variable is independent from a child's ethnic background). Analyses of variance show that there are significant differences between East and West Germany in all variables. The strip plots in Figure 1 illustrate the different distributions using the example of children with a *language other than German*. Every dot in each strip plot stands for the respective proportion of one centre. Here we can see that there are hardly any day-care centres with high concentrations of this population group in East Germany whereas West Germany covers the whole range.

Figure 1: Proportions of children with a language other than German at regional level



Note: The y-axis shows the proportion of children with a language other than German in West and East Germany based on K^2ID-SOEP data. Cases from Berlin are excluded from the data. The interpretation of the box is similar to boxplots: The median is the line in the centre of the box, the lower frame defines the first quartile and the upper the third quartile. Additionally, the thin long line represents the mean. *n* = number of cases.

As expected, differences between West and East Germany are smaller and not statistically significant for social segregation (see Table 5).

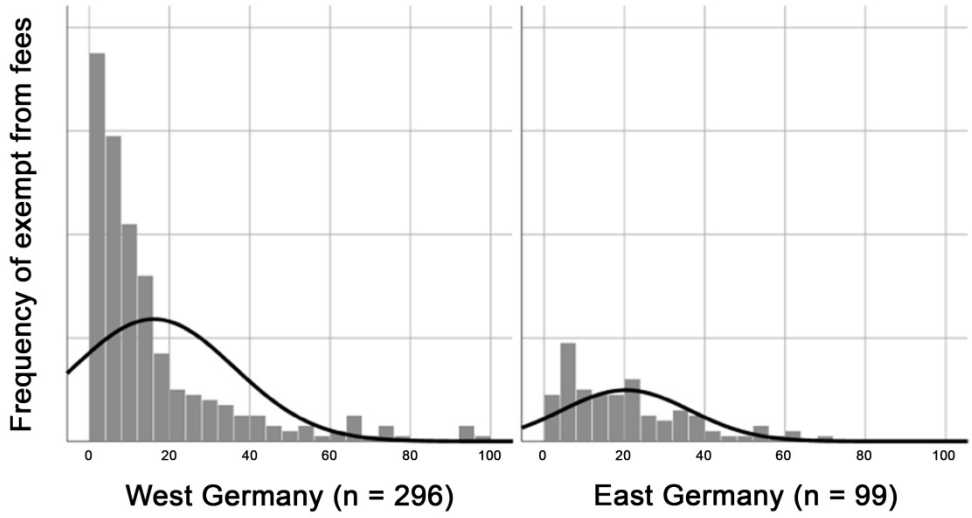
Table 5: Social segregation at regional level

	n	M	SD	Mdn	Min	Max
Exempt from fees						
K^2ID-SOEP						
West Germany	296	16	20	9	0	100
East Germany	99	20	16	17	0	71
						Welch's F(1, 158) = 2.954, p = .088

Note: The table presents descriptive and nonparametric test statistics for social segregation between West and East Germany based on K^2ID-SOEP data. Cases from Berlin are excluded from the data. *n* = number of cases, *M* = Mean, *SD* = standard deviation, *Mdn* = Median, *Min* = Minimum, *Max* = Maximum, *Welch's* $F(v_1, v_2)$ = Welch's ANOVA F-ratio with v_1 and v_2 (rounded) degrees of freedom, *p* = probability value. Except for the number of cases and the information on statistical tests, the values refer to proportions in percent.

Figure 2 illustrates that the overall distribution patterns between the two regions do not differ as clearly from each other as for ethnic segregation. In both East and West Germany, children mostly attend centres with rather low proportions of children *exempt from fees*. In East Germany, however, the curve flattens out at around 50 percent and in West Germany at over 60 percent. Here, we also find most settings having rather low proportions and only some day-care centres with high concentrations, i.e. proportions of up to 100 percent. In East Germany, the respective proportions seem to be more equally distributed.

Figure 2: Distributions of children exempt from fees at regional level

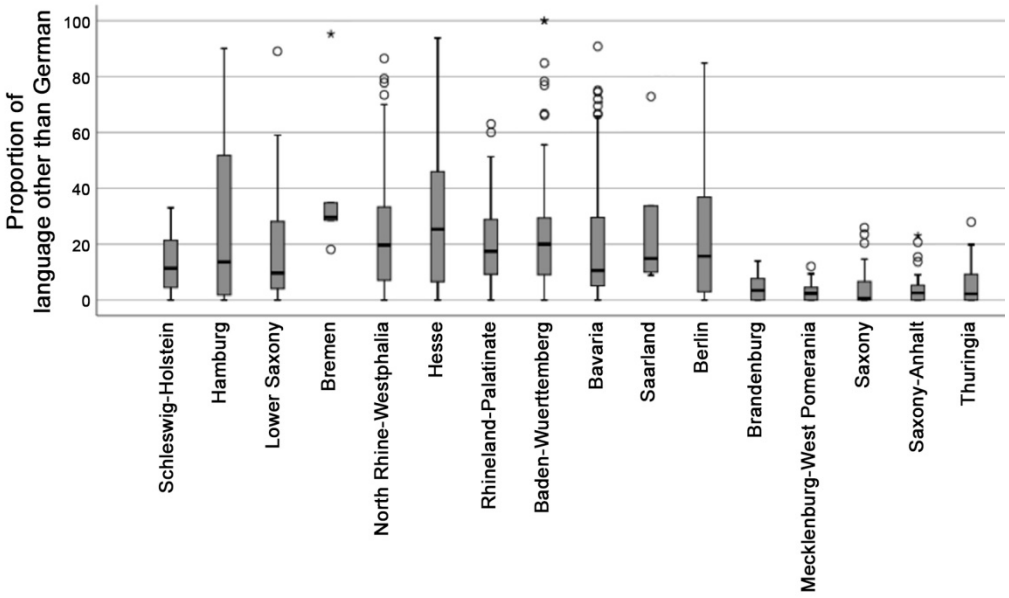


Note: These histograms present the distribution of the frequencies for the proportions of children exempt from fees in East and West German day-care centres based on K^2ID-SOEP data. Cases from Berlin are excluded from the data. *n* = number of cases.

4.3 Segregation Patterns at Federal States Level

A closer look at differences between the federal states supports the above presented results on regional differences. The proportions of children with a *language other than German* significantly differ between the federal states (*Welch's F(15, 101)* = 24.378, p = .000), but post-hoc tests indicate that these differences in linguistic segregation are attributable to general differences between West and East German federal states. The boxplots in Figure 3 illustrates this graphically. Here we can see low proportions throughout all East German states. In addition, it becomes visible that most West German states have a rather widespread distribution and, thus, large differences between settings at the upper and lower end of the spectrum. Only Schleswig-Holstein seems to have a relatively even distribution of the children with a language other than German between the day-care settings.

Figure 3: Proportions of children with a language other than German at federal states level

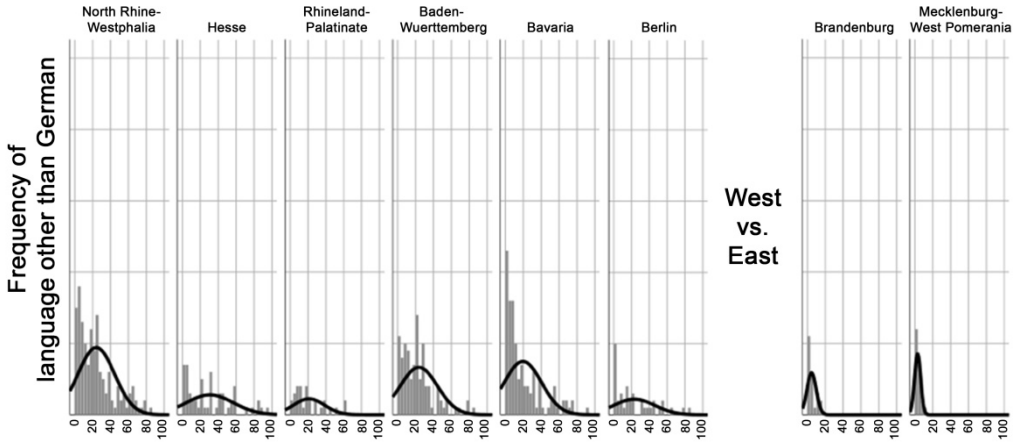


Note: The y-axis shows the proportion of children with a language other than German at setting level based on K^2ID-SOEP data. Boxplots: The median is the line in the centre of the box, the lower frame defines the first quartile and the upper the third quartile. Circles and stars indicate outliers and extreme outliers, respectively. *n* = number of cases: Schleswig-Holstein, n = 28; Hamburg, n = 24; Lower Saxony, n = 64; Bremen, n = 5; North Rhine-Westphalia, n = 145; Hesse, n = 55; Rhineland-Palatinate, n = 29; Baden-Wuerttemberg, n = 104; Bavaria, n = 117; Saarland, n = 5; Berlin, n = 38; Brandenburg, n = 23; Mecklenburg-West Pomerania, n = 22; Saxony, n = 62; Saxony-Anhalt, n = 26; Thuringia, n = 22. Due to low case numbers, the results for Bremen and Saarland are not interpreted.

In terms of distribution patterns, the East German states show a right-skewness which flattens out at higher proportions (see Figure 4). As already said before, here we do not find high concentrations of certain population groups in individual day-care centres.

Berlin rather seems to have a distribution that resembles other West German states. In this regard, however, some further – albeit small – differences can be described. Some federal states, like Bavaria, Baden-Wuerttemberg and North Rhine-Westphalia show a broad right-skewness with many settings having lower proportions of children with a *language other than German* and only some settings where those children concentrate. Others, like Hesse, Rhineland-Palatinate or Berlin, have a flat curve indicating a more equal distribution. Similar findings can be found for the distribution patterns of the proportions of children *in need of language support*, which are lower in East German states. Rhineland-Palatinate or Baden-Wuerttemberg, for example, show a broader distribution.

Figure 4: Distributions of children with a language other than German in selected federal states



Note: These histograms present the distribution of the frequencies for language other than German in selected federal states based on K^2ID-SOEP data. *n* = number of cases: North Rhine-Westphalia, n = 145; Hesse, n = 55; Rhineland-Palatinate, n = 29; Baden-Wuerttemberg, n = 104; Bavaria, n = 117; Berlin, n = 38; Brandenburg, n = 23; Mecklenburg-West Pomerania, n = 22.

With regard to social segregation, the proportion of *children exempt from fees* in ECEC centres also significantly differs between the federal states (*Welch's F(15, 41)* = 2.810, *p* = .004), but these differences exclusively result from regional differences between East and West Germany.

5 Discussion

ECEC is expected to reduce educational inequality by providing beneficial learning contexts especially for disadvantaged children and to serve social inclusion and democracy more generally by providing diverse learning contexts (*Vandenbroeck* 2015). However, our research findings indicate that the actual distribution of children in ECEC settings in Germany might provide a stark contrast to this perspective. Segregation in ECEC seems highly developed, and, consequently, the peer-related makeup of day-care centres varies considerably. As context studies focusing on composition effects show, this has implica-

tions for children's competence development, which is especially the case for disadvantaged children (e.g. *Hogrebe/Pomykaj* 2019; *Niklas/Taylor* 2018). Moreover, high concentrations of certain population groups in particular settings as well as their absence in others contradict the idea of social inclusion and integration inherent to ECEC. Consequently, experiences of diversity are restricted for all of children in these day-care centres.

As most of the research explicitly addressing segregation in ECEC is conducted in the USA and cannot per se be applied to other contexts, we aimed at exploring segregation in institutionalized early education in Germany at regional and federal state level. Our reason for doing so was to get a better understanding about the analysis levels at which differences in segregation patterns become visible. We expected pronounced differences between West and East Germany especially in ethnic segregation. We also assumed that segregation might take different forms in the federal states as regulations and structures are set at this level.

Our results support the impression that segregation is a relevant issue in German ECEC as well. At national level, differences of up to 100 percentage points of children with certain ethnic or linguistic respectively social background characteristics show extreme differences between day-care centres. However, unlike in the USA where the majority of children in isolated settings experience high concentrations of non-white children, in Germany most children visit rather isolated non-migrant day-care centres. We do not find a u-shaped distribution as research did for the U.S. (*Urban Institute* 2019). Rather, in Germany there are more settings with low proportions and few settings with very high proportions. In this context, the goal of promoting democracy through social mixing is particularly at risk, which appears problematic especially against the backdrop of increasing processes of social division that can be observed. However, we must not forget that we also find settings with high concentrations of disadvantaged children – albeit to a lesser extent – which puts these children at a double risk.

Despite the general pattern of large differences between the ECEC settings' compositions throughout Germany, we mainly find regional variation. Segregation is more pronounced in West Germany, which is not surprising given the different population structures in the respective states. Consequently, we do not find high concentrations of, for example, migrant children in East German settings. Less segregation in this regard, however, does not automatically mean more diversity. On the contrary, in line with the findings of *Piazza* and *Frankenberg* (2019) the rather homogeneous compositions of day-care centres of East German states show that diversity is challenged by its overwhelmingly non-migrant state-wide composition.

However, higher proportions of the respective population groups in the West German sates do not necessarily result in their equal allocation to centres. Consequently, the general presence of diversity does not automatically lead to more mixed educational settings. In this regard, the different distribution patterns of the West-German states are interesting. With regard to the state level, there are only subtle differences visible: In some federal states, the majority of centres has rather low proportions of disadvantaged children; these mostly concentrate in only very few settings. Other states have flatter curves indicating a more uniform distribution. As to the reasons for the different distribution patterns, more research is needed on state and also local levels to understand the mechanisms of segregation processes in the early years. Further studies should also include home-based childcare which we did not consider in this paper but which seems to intensify segregation in America.

Generally, it is rather unlikely that diversity in ECEC "will happen without explicit efforts to create and sustain them" (*Frankenberg* 2016, p. 21). Due to the importance of diversity in day-care centres for both individual children and society, policy makers should try to find ways to avoid segregation and develop concepts for a socially and ethnically inclusive system of ECEC. The result of our study imply that German policy makers should adopt a dual strategy: First, programmes that support day-care centres with a concentration of disadvantaged children should be continued and expanded in order to answer to the challenges these settings and children face. Second, such measures should be complemented by incentives for more diversity that address the many settings with low proportions of disadvantaged children. While the first approach is already implemented in different forms at national and federal state level, the second idea should probably be given more attention in future.

However, in order to be able to implement meaningful governance strategies that might lead to more diversity, more research is needed that sheds light on the reasons for segregation. Our study implicates that regional differences in segregation patterns should be considered in such future analyses. It seems reasonable that different approaches are needed for different mechanisms causing segregation (i.e. a lack of certain populations groups in East Germany and de-mixing processes in West Germany where those population groups are present). However, at this stage we cannot yet say whether regional populations structures just translate to populations structures in local catchment areas or adds further explanation potential (i.e. different traditions in provider structures and political regulation). Similarly, further research on segregation patterns in and between federal states might provide valuable insights. Although differences between states are quite subtle, those states that show a more uniform distribution in West Germany might provide an interesting starting point to identify conditions for more diversity in all day-care centres. Another analytic perspective that should be considered is types of agglomeration areas. It seems reasonable that segregation is more pronounced in urban agglomerations with more demand than supply than in suburban or rural areas. Overall, more research that includes the different levels and layers of segregation is needed to get a better understanding of its causes.

The differences other research reports on school and ECEC segregation implies that segregation in early childhood education is not yet performance-linked and – at least from its institutional idea – less exclusive. The high degree of segregation in this rather egalitarian education sector therefore raises the question of how social inequalities are (re)produced here. So far, little is known, for example, about the consequences and effects of the plural provider structures with regard to possible segregation effects (*Authoring Group Educational Reporting* 2020).

Notes

1 For out-of-home care of children aged one to six in child and youth welfare institutions we use terms like early education, ECEC settings, or day-care centres. The term preschool is used in the context of research from the USA and usually refers to children aged two to five.

2 Another measure is the exposure index that illustrates the extent to which children of a particular group are exposed to other children in their setting. If all settings were perfectly integrated, all children would be exposed to the same composition of children. The index of dissimilarity focuses on the sorting of population groups across units to understand their distribution in a certain geographic area and, therefore, is only applicable if the available data include all units (*Frankenberg* 2016). For an overview and discussion of different segregation measures see *White* (1983, 1986).

3 However, universal programs are not necessarily more equal. Despite its universal approach, in New York, for example, half of all preschool classrooms have more than 70 percent of children from a single racial or ethnic group; in 15 percent of the classrooms this is the case for more than 90 percent of the children (*Potter* 2016). Similarly, in a recent paper analysing segregation in Oslo (Norway), *Drange* and *Telle* (2020, p. 1) conclude that a universal child care system is not sufficient per se "to reduce social inequality and improve opportunities of disadvantaged children."

4 The analyses are part of a project in cooperation with colleagues from the Martin-Luther-University in Halle-Wittenberg funded by the Federal Ministry of Education and Research under the funding code 01NV1809B. The responsibility for the content of this publication lies with the authors.

5 We would like to thank Hannah M. Heister from the TU Dortmund for her support in the data preparation.

6 This paper uses data from the K^2ID-SOEP extension study (doi: J0.5684/k2id-soep-2013-15/vl). The K^2ID-SOEP extension study was funded by the Jacobs Foundation.

7 This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Newborns, doi:10.5157/NEPS:SC1:7.0.0 and Starting Cohort Kindergarten, doi:10.5157/NEPS:SC2:8.0.1. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

8 While we do not observe any differences in minimum and maximum values for the two waves of the K^2ID data, a comparison of means reveals that the day-care centers that are visited by children with a migration background are characterized by average proportions of children with a *language other than German* which are nearly two times higher. Comparisons of means show that these differences are statistically significant (Mann Whitney U test statistic: $U = 30660.000$, $z = -7.234$, $p = .000$). This indicates that these children on average visit day-care centres with much higher concentrations of children who do not speak German at home.

References

Anders, Y./Roßbach, H.-G./Weinert, S./Ebert, S./Kuger, S./Lehrl, S./von Maurice, J. (2012): Home and preschool learning environments and their relations to the development of early numeracy skills. Early Childhood Research Quarterly, 27, 2, pp. 231-244. https://doi.org/10.1016/j.ecresq.2011.08.003

Authoring Group Educational Reporting (2016): Education in Germany 2016. An indicator-based report including an analysis of education and migration. – Bielefeld.

Authoring Group Educational Reporting (2020): Education in Germany 2020. An indicator-based report including an analysis of education in a digitalized world. – Bielefeld.

*Becker, B./Schober, P. S. (*2017): Not just any child care center? Social and ethnic disparities in the use of early education institutions with a beneficial learning environment. Early Education and Development. https://doi.org/10.1080/10409289.20 *17.1320900.*

Biedinger, N./Becker, B. (2010): Frühe ethnische Bildungsungleichheit: Der Einfluss des Kindergartenbesuchs auf die deutsche Sprachfähigkeit und die allgemeine Entwicklung. In Becker, B./Reimer, D. (eds.): Vom Kindergarten bis zur Hochschule. Die Generierung von ethnischen und sozialen Disparitäten in der Bildungsbiographie. – Wiesbaden, pp. 49-79. https://doi.org/10.1007/978-3-531-92105-1_3

Blossfeld, H.-P./Roßbach, H.-G./von Maurice J. (Hrsg.) (2011): Education as a Lifelong Process – The German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft, Special Issue 14. – Heidelberg.

Destatis (2015a): Internationale Bildungsindikatoren im Ländervergleich. Online available: https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00017749/Internationale_Bildungsindikatoren_2015.pdf, 16.07.2020.

Destatis (2015b): Statistiken der Kinder- und Jugendhilfe. Kinder und tätige Personen in Tageseinrichtungen und in öffentlicher geförderter Kindertagespflege am 01.03.2015. Online available: https://www.bvktp.de/media/tageseinrichtungenkindertagespflege5225402157004.pdf, 14.07.2020.

Destatis (2019): Statistisches Jahrbuch – Deutschland und Internationales. Online available: https://www.destatis.de/DE/Themen/Querschnitt/Jahrbuch/statistisches-jahrbuch-2019-dl.pdf?__blob=publicationFile, 14.12.2020.

Drange, N./Telle, K. (2020). Segregation in a universal child care system: Descriptive findings from Norway. *European Sociological Review*. https://doi.org/10.1093/esr/jcaa026

European Commission (2011): Early childhood education and care: Providing all our children with the best start for the world of tomorrow. Online available: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0066:FIN:EN:PDF, 14.07.2020.

Fram, M. S./Kim, J. (2012): Segregated from the start: Peer context in center-based child care. Children & Schools, 34, 2, pp. 71-82. https://doi.org/10.1093/cs/cds011

Frankenberg, E. (2016): Segregation at an early age. Pennsylvania State University. Center for Education and Civil Rights; National Coalition on School Diversity. Online available: https://cecr.ed.psu.edu/sites/default/files/Segregation_At_An_Early_Age_Frankenberg_2016.pdf, 14.07.2020.

Frankenberg, E./Lee, Ch. (2002): Race in American public schools: Rapidly resegregating school districts. The Civil Rights Project. – Cambridge, MA.

Gans, H. J. (2007): The Balanced Community: Homogeneity or Heterogeneity in Residential Areas? Journal of the American Institute of Planners, 27, 3, pp. 176-184. https://doi.org/10.1080/01944366108978452

Hock, B./Holz, G./Kopplow, M. (2014): Kinder in Armutslagen. Grundlagen für ar-mutssensibles Handeln in der Kindertagesbetreuung (WiFF Expertise Nr. 38). – München.

Hogrebe, N. (2014): Indicators for a needs-based resource allocation in early child-hood education: Regional data as valid proxies for setting level needs? Journal for Educational Research Online, 6, 2, pp. 44-65.

Hogrebe, N. (2016a): Segregation im Elementarbereich – Mobilität und Trägerschaft. Zeitschrift für Grundschulforschung, 9, 1, pp. 20-33.

Hogrebe, N. (2016b): Segregation am Übergang von der Kita zur Grundschule. In *McElvany, N./ Strietholt, R./Bos, W./Holtappels, H.-G.* (eds.): Jahrbuch der Schulentwicklung. Daten, Beispiele, Perspektiven (Band 19). – Weinheim, pp. 80-105.

Hogrebe, N./Pomykaj A. (2019): Die Schuleingangsuntersuchung als Datenquelle für Kontextstudien im Elementarbereich. Zum Zusammenhang von Kita-Komposition und kindlichen Sprachkompetenzen. Die Deutsche Schule, Beiheft 14, pp. 71–86. https://doi.org/10.31244/dds.bh.2019.14.05

Hüsken, K. (2011): Kita vor Ort: Betreuungsatlas auf Ebene der Jugendamtsbezirke 2010. Online available: https://www.dji.de/veroeffentlichungen/literatursuche/detailansicht/literatur/24454-kita-vor-ort.html, 14.07.2020.

Jacob, R./Heinz, A./Décieux, J. P. (2013): Umfrage – Einführung in die Methoden der Umfrageforschung. – München, pp. 98-120. https://doi.org/10.1524/9783486736175

Kagan, S. L. (1990): Excellence in early childhood education: Defining characteristics and next-decade strategies. Online available: https://files.eric.ed.gov/fulltext/ED318580.pdf, 14.07.2020.

Kuger, S./Kluczniok, K. (2008): Prozessqualität im Kindergarten: Konzept, Umsetzung und Befunde. Zeitschrift für Erziehungswissenschaft, Sonderheft 11, pp. 159-178. https://doi.org/10.1007/978-3-531-91452-7_11

Lochner, S./Jähnert, A.(eds.) (2020): DJI-Kinder- und Jugendmigrationsreport 2020. Datenanalyse zur Situation junger Menschen in Deutschland. https://doi.org/10.3278/6004754w

Niklas, F./Tayler, C. (2018): Room Quality and Composition Matters: Children's Verbal and Numeracy Abilities in Australian Early Childhood Settings. Learning and Instruction, 54, 114–124. https://doi.org/10.1016/j.learninstruc.2017.08.006

Olczyk, M./Will, G./Kristen, C. (2014): Personen mit Zuwanderungshintergrund im NEPS: Zur Bestimmung von Generationenstatus und Herkunftsgruppe (NEPS Working Paper No. 41b). – Bamberg.

Olszenka, N./Meiner-Teubner, C. (2020): Kinder in Kindertageseinrichtungen und Tagespflege. In: Lochner, S./Jähnert, A. (eds.): DJI Kinder- und Jugendmigrationsreport 2020. Datenanalyse zur Situation junger Menschen in Deutschland – Bielefeld, pp. 98–110.

Piazza, P./Frankenberg, E. (2019): Segregation at an early age: 2019 update. Online available: https://cecr.ed.psu.edu/sites/default/files/Segregation_At_An_Early_Age_Piazza_Frankenberg_2019.pdf, 14.07.2020.

Potter, H. (2016): Diversity in New York City's universal pre-k classrooms. Online available: https://tcf.org/content/report/diversity-new-york-citys-universal-pre-k-classrooms/, 14.07.2020.

R Core Team (2020): R: A language and environment for statistical computing. R Foundation for Statistical Computing. – Vienna. Online available: https://www.R-project.org/.

Reid, J. L./Kagan, S. L./Hilton, M./Potter, H. (2015): A better start. Why classroom diversity matters in early education Washington, D.C. and New York, NY: The Century Foundation and The Poverty & Race Research Action Council. Online available: https://www.prrac.org/pdf/A_Better_Start.pdf, 14.07.2020.

Schober, P./Spieß, K./Stahl, J./Zoch, G./Camehl, G. (2017): The Early Childhood Education and Care Quality in the Socio-Economic Panel (SOEP-ECEC Quality) Study – K2ID-SOEP Data, Data Documentation 91. – Berlin.

Strohmeier, K. P./Gehne, D. H./Groos, T./Jehles, N. (2014): Arbeitspapiere wissenschaftliche Begleitforschung „Kein Kind zurücklassen!" Die fachliche Begleitforschung. Konzept und erste Ergebnisse. Online available: http://www.kein-kind-zuruecklassen.de/modellvorhaben/evaluation/publikationen-und-infomaterial.html, 14.07.2020.

Tietze, W./Becker-Stoll, F./Bensel, J./Eckhardt, A. G./Haug-Schnabel, G./Kalicki, B., Keller, H./Leyendecker, B. (eds.) (2013): *NUBBEK.* Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit. Fragestellungen und Ergebnisse im Überblick. – Weimar.

Tropp, L./Saxena, S. (2018): Re-weaving the social fabric through integrated schools: How intergroup contact prepares youth to thrive in a multiracial society. NCSD Research Brief. Online available: https://school-diversity.org/wp-content/uploads/2018/05/NCSD_Brief13.pdf, 14.07.2020.

Urban Institute (2019): Segregated from the start: Comparing segregation in early childhood and K–12 education. Online available: https://www.urban.org/features/segregated-start, 14.07.2020.

Vandenbroeck, M. (2015): Ethnic diversity and social inclusion in ECCE in Europe. In *Marope, P. T. M./Kaga, Y.* (eds.): Investing against evidence. The global state of early childhood care and education. – Paris, pp. 105-118.

White, M. J. (1983): The measurement of spatial segregation. American Journal of Sociology, 88, 5, pp. 1008-1018. https://doi.org/10.1086/227768

White, M. J. (1986). Segregation and diversity measures in population distribution. Population Index, 52, pp. 198-221. https://doi.org/10.2307/3644339

Appendix 1: Survey Questions and Operationalization

Variable	Data Set	Respondent	Level	Question in English	Original Question in German
Language other than German	K^2ID-SOEP	Educators	Group	How many children are currently registered in your core group? Please also indicate how many girls, boys, children of non-German origin* and children with integration status are currently registered. [*Children of non-German origin means: The child speaks a language other than German in his/her family.]	Wie viele Kinder sind derzeit insgesamt in Ihrer Stammgruppe angemeldet? Bitte geben Sie auch an wie viele Mädchen, Jungen, Kinder nichtdeutscher Herkunftssprache* und Kinder, die einen Integrationsstatus** haben, derzeit angemeldet sind. [*Kinder nichtdeutscher Herkunftssprache bedeutet: Das Kind spricht in seiner Familie überwiegend eine andere Sprache als Deutsch.]
	K^2ID-SOEP	Directors	Setting	How many children are currently registered in your institution? Please also indicate how many girls, boys, children of non-German origin* and children who have an integration status are currently registered. [*Children of non-German origin means: The child speaks a language other than German in his/her family.]	Wie viele Kinder sind derzeit insgesamt in Ihrer Einrichtung angemeldet? Bitte geben Sie auch an wie viele Mädchen, Jungen, Kinder nichtdeutscher Herkunftssprache* und Kinder, die einen Integrationsstatus haben, derzeit angemeldet sind. [*Kinder nichtdeutscher Herkunftssprache bedeutet: Das Kind spricht in seiner Familie überwiegend eine andere Sprache als Deutsch.]
	NEPS SC 1, wave 4	Educators	Group	Are there any children in your group that speak a language other than German either exclusively or at least partially while in your care? If yes, how many?	Gibt es in Ihrer Gruppe Kinder, die während der Betreuung ausschließlich oder zumindest teilweise eine andere Sprache als Deutsch sprechen? Und wenn ja, wie viele?
	NEPS SC 2, wave 1	Directors	Setting	What is the approximate portion of children with a language of origin other than German* at your facility? [*children with a language of origin other than German means: The child has learned a language other than German in its family ("mother tongue").]	Wie groß ist der Anteil der Kinder nichtdeutscher Herkunftssprache* in Ihrer Einrichtung etwa? [*Kinder nichtdeutscher Herkunftssprache bedeutet: Das Kind hat eine andere Sprache als Deutsch in seiner Familie gelernt („Muttersprache“).]
In need of language support	K^2ID-SOEP	Educators	Group	How many children in your core group are currently participating in language training?	Wie viele Kinder in Ihrer Stammgruppe nehmen aktuell an einer Sprachförderung teil?
Migration background	NEPS SC 1, wave 4	Educators	Group	How many children in your core group have a migrant background*? [*Migration background means: The child or at least one of the parents were born abroad.]	Wie viele der Kinder Ihrer Stammgruppe haben einen Migrationshintergrund*? [*Migrationshintergrund bedeutet: Das Kind selbst oder mindestens ein Elternteil ist im Ausland geboren.]

	NEPS SC 1, wave 4	Directors	Setting	How large approximately is the proportion of children in your institution with a migration background*? [*Migration background means: The child or at least one of the parents were born abroad.]	Wie groß ist der Anteil von Kindern mit Migrationshintergrund* in Ihrer Einrichtung etwa? [*Migrationshintergrund bedeutet: Das Kind selbst oder mindestens ein Elternteil ist im Ausland geboren.]
	NEPS SC 2, wave 1	Educators	Group	Please indicate for each age group (in other words, in each line) (a) the number of children currently visiting your core group, (b) the number of child-care hours per day, (c) how many of them have a migration background* and (d) a disorder. [*Migration background means: the child or at least one parent was born abroad.]	Bitte geben Sie für jede Altersgruppe (also in jeder Zeile) (a) die Anzahl der Kinder an, die derzeit Ihre Stammgruppe besuchen, (b) wie lange diese pro Tag betreut werden, (c) wie viele einen Migrationshintergrund* und (d) eine Behinderung haben. [*Migrationshintergrund bedeutet: Das Kind selbst oder mindestens ein Elternteil ist im Ausland geboren.]
	NEPS SC 2, wave 1	Directors	Setting	Please indicate for each age group (in other words, in one line) (a) the number of children currently attending your facility, (b) how long they are taken care of per day, (c) how many have an immigration background* and (d) suffer from a disability. [*immigration background means: The child or at least one parent was born abroad.]	Bitte geben Sie für jede Altersgruppe (also in jeder Zeile) (a) die Anzahl der Kinder an, die derzeit Ihre Einrichtung besuchen, (b) wie lange diese pro Tag betreut werden, (c) wie viele einen Migrationshintergrund* und (d) eine Behinderung haben. [*Migrationshintergrund bedeutet: Das Kind selbst oder mindestens ein Elternteil ist im Ausland geboren.]
Exempt from fees	K^2ID-SOEP	Directors	Setting	For how many children who visit your institution do the parents pay on the basis of their household income the lowest possible fees or are exempt from paying fees? [Please do not count children who pay less or nothing because of their age or sibling status].	Für wie viele Kinder, die Ihre Einrichtung besuchen, zahlen die Eltern aufgrund ihres Haushaltseinkommens den geringstmöglichen Beitrag oder sind von den Beitragszahlungen befreit? [Bitte zählen Sie Kinder, die aufgrund ihres Alters oder Geschwisterstellung weniger oder nichts bezahlen, nicht mit.]
Parents with higher education	NEPS SC 1, wave 4	Educators	Group	Of how many children in your core group has at least one parent successfully completed a higher education program?	Bei wie vielen Kindern in Ihrer Stammgruppe hat mindestens ein Elternteil ein Studium abgeschlossen?
	NEPS SC 2, wave 1	Educators	Group	Of how many children in your core group has at least one parent successfully completed a course of study*? [*Children with at least one parent having successfully completed a course of study.]	Bei wie vielen Kindern in Ihrer Stammgruppe hat mindestens ein Elternteil ein Studium abgeschlossen*? [*Kinder mit mindestens einem Elternteil mit abgeschlossenem Studium.]
	NEPS SC 2, wave 1	Directors	Setting	What is the approximate portion of children in your facility where at least one parent has completed a course of study*? [*Children with at least one parent having successfully completed a course of study.]	Wie groß ist in etwa der Anteil der Kinder in Ihrer Einrichtung, bei denen mindestens ein Elternteil ein Studium abgeschlossen hat*? [*Kinder mit mindestens einem Elternteil mit abgeschlossenem Studium.]

Lower social class	NEPS SC 1, wave 4	Educators	Group	How many children in your core group come from families from rather lower social classes?	Wie viele Kinder Ihrer Stammgruppe kommen aus Familien aus eher niedrigen sozialen Schichten?
	NEPS SC 2, wave 1	Educators	Group	How many children in your core group come from families from rather lower social classes?	Wie viele Kinder Ihrer Stammgruppe kommen aus Familien aus eher niedrigen sozialen Schichten?
	NEPS SC 2, wave 1	Directors	Setting	What is the percentage of children in your facility coming from families from lower social classes?	Wie viel Prozent der Kinder in Ihrer Einrichtung kommen aus Familien aus eher niedrigen sozialen Schichten?

Note: For both studies the surveys of the day-care staff were carried out using questionnaires in German. However, for K^2ID-SOEP, the English questions are based on own translations. In both waves, the same questions were asked (see data documentation online: http://dx.doi.org/10.5684/k2id-soep-2013-15/v1). For research purposes English versions are available for NEPS (see data documentation online: https://www.neps-data.de/Data-Center/Data-and-Documentation/).

Appendix 2: Correlations between composition at group and setting level

	K^2ID-SOEP	NEPS SC 1	NEPS SC 2
a) Language other than German	.88* (n = 637)	–	–
b) Migration background	–	.78* (n = 333)	.91* (n = 109)
c) Parents with higher education	–	–	.77* (n = 106)
d) Lower social class	–	–	.74* (n = 111)

Note: Intercorrelation (Pearsons'r) between settings and groups in the respective settings for those characteristics where information is available at both levels for the same centre. n = number of cases, *$p < .01$.