Defining and Identifying Only Children: A research note on the concept and measurement illustrated with UK survey data

Jenny Chanfreau[1 2] and Alice Goisis[2]

[1] Department of Sociology and Criminology, University of Sussex, Brighton BN1 9RH

[2] UCL Centre for Longitudinal Studies, University College London, London WC1H 0AA.

Funding

This work was supported by the Economic and Social Research Council grant numbers ES/S002103/1 and ES/M001660/1.

Data availability

The data that support the findings of this study are available to registered users from the UK Data Service (https://ukdataservice.ac.uk/) with End User Licence, and from MRC Unit for Lifelong Health and Ageing at UCL (LHA; https://www.nshd.mrc.ac.uk/) with the permission of LHA. Stata code to derive a variable to identify only children in each of the datasets has been deposited by the authors with the UK Data Service ReShare archive (https://reshare.ukdataservice.ac.uk/855087/).

Abstract

Despite increasing interest in the circumstances and outcomes of only children in the demographic literature, the conceptualisation of this group has received limited scholarly attention. This Research Note argues for greater engagement by demographers and social scientists in the conceptualisation and identification of only children by addressing the following three aims. First, we outline different potential definitions of only children, presenting a framework to guide researchers' decisions, and evaluate whether only children can reliably be identified using the British birth cohort studies. Second, we show that the prevalence estimates are contingent on the timing of measurement in childhood, indicating the need for caution when deriving only child status from cross-sectional household grid data. Third, we demonstrate that both the size and the characteristics of the only child group may differ across definitions, highlighting that the accurate operationalisation of some definitions is particularly restricted by survey design tending to prioritise mothers for data collection about children and families. We argue that researchers interested in the outcomes of only children need to consider the choice of the most appropriate measure for a given research question and, given that many datasets limit how accurately any indicator of only children can capture the chosen definition, reflect on how the operationalisation of their measure may affect the results observed.

Keywords: Only children, siblings, British cohort studies, families, survey data

Introduction

With shrinking family sizes and one-child families a small but growing family form in many global north populations (Breton and Prioux 2009; Frejka 2008; Frejka, Jones and Sardon 2010), researchers have become increasingly interested in the circumstances and outcomes – such as education and health – of only children (e.g. Baranowska-Rataj, Barclay and Kolk 2017; Beaujouan and Solaz 2019; Keenan, Barclay and Goisis 2022; Laybourn 1990; Mancillas 2006; Rainer and Siedler 2012).

There has been surprisingly little discussion of the conceptualisation and measurement of only children (for similar points regarding sibling research see Präg, Choi and Monden 2020; Riswick and Engelen 2018). Prior studies have defined 'only children' in different ways, some focusing solely on the absence of full siblings, others on the absence also of half and social siblings while others specified the absence of co-resident siblings (Baranowska-Rataj et al. 2017; Beaujouan and Solaz 2019; Bobbitt-Zeher, Downey and Merry 2016; Downey and Condron 2004; Keenan et al. 2022; Lersch 2019; Lin and Falbo 2022; Präg et al. 2020). Some studies have not stated their definition (Chen and Liu 2014; Falbo 2012; Mancillas 2006; Mellor 1990), implicitly assuming readers share their 'only child' conceptualisation. The lack of reflection and common understanding of how to define only children is problematic as conceptual clarity is essential for identification of only children and for understanding the characteristics and circumstances of this sub-population group.

The aim of this paper is to encourage greater critical engagement and transparency about the conceptualisation and measurement of only children in population studies. We reflect on the necessary research decisions through new descriptive evidence from four large nationally representative British birth cohort studies. Präg et al. (2020) show that the prevalence of only children in the UK increased among cohorts born in the first five decades of the 20th century,

peaking among those born in the 1930s and 1940s at about 13%, and after a decline in the 1950s birth cohort has held relatively steady (the most recent cohort they considered was born 1981-1990). According to the same study, for individuals born in the 1970s and 1980s, the prevalence of only children in the UK is similar to Japan, somewhat lower than in France and somewhat higher than in the US and Sweden. However, although there will be particularities specific to the UK as well as to each national context and data source, the issues we raise in this research note are neither unique to the UK context nor to British birth cohort studies.

The paper is structured around three aims. First, we outline different potential definitions of only children and evaluate whether only children can reliably be identified using the British birth cohort studies. Second, we show that the prevalence estimates are contingent on the timing of measurement in childhood, indicating the need for caution when deriving only child status from household grid variables using cross-sectional data. Third, we demonstrate that both the size and the characteristics of the only child group may differ across definitions, highlighting that the accurate operationalisation of some definitions is particularly restricted by survey design tending to prioritise mothers for data collection about children and families.

Possible definitions of only children

Conceptually we may define an only child based on a stylized view, namely that of an individual without a sibling of any kind from any parent at any point in their life. In practice, this definition may be both difficult to operationalise (as it requires identifying all kinds of siblings over the whole life course) and may not be theoretically the most appropriate to adopt. Demographers and survey data collections have tended to focus on biological relations, and especially on siblings as the mother's other 'biological' children - a reflection of the focus on female fertility in demographic research (Zhang 2011). However, relevant

explanatory theories of only children's distinctiveness or effects of sibship size on outcomes tend to also include 'social' siblings – at least implicitly (resource dilution theory; Blake 1989; Downey 1995) and/or focus on co-residence with siblings irrespective of biological or social relatedness (socialisation theory; Goetting 1986). Both theories focus on explanatory processes and circumstances during childhood. Resource dilution theory require the researcher to consider whether there are any siblings who also need the index child's parents' financial and other resources, whether or not they reside with the index child. Socialisation theory, on the other hand, focuses attention on co-residence and thus a child with full or half siblings living in another household might be considered an 'only child' but a child co-residing with a social sibling might not.

In Table 1 we propose a comprehensive framework bringing these perspectives together for potential definitions based on the intersection of biological/social siblings' presence and co-residence. We assume that children tend to co-reside with their mother because 87% of parents/guardians receiving Child Benefit are female, 50-50 shared care following parental separation remains rare in the UK, and among one-parent families the proportion that are lone mother families has remained stable since the mid-90s at about 90% (Haux, McKay and Cain 2017; ONS 2022; ONS 2023). The table aims both to explain the only child definitions used in this paper and to offer researchers a guiding framework when deciding how to operationalize 'only children' in their work.

The top left and bottom right cells are relatively straight-forward. Most would agree that someone without siblings of any kind is an only child (i), and similarly that anyone with at least one full sibling (a) is not. However, cells (b)-(h) require decisions about which siblings to consider, ideally informed by theory and research questions.

Table 1 Considerations for defining 'only children'

		Biological father of the index child has:		
		Other biological children	Social children (step/ adopted/ foster)	No social or other biological children
Biological mother of the index child has:	Other biological children	(a) **Full siblings**, likely co-resident	(b) **Half siblings**, likely co-resident, + **social siblings**, likely non-resident	(c) **Half siblings**, likely co-resident
	Social children (step/ adopted/ foster)	(d) **Social siblings**, possibly co-resident + **Half siblings**, likely non-resident	(e) **Social siblings,** some possibly co-resident (maternal) + some likely non-resident (paternal)	(f) **Social siblings**, possibly co-resident
	No social or other biological children	(g) **Half siblings**, likely non-resident	(h) **Social siblings**, likely non-resident	(i) **No full, half or social siblings**

Notes: The grid shows how co-residence intersects with the two dimensions of biological and social siblings, separately for maternal and paternal sibling relations. Full siblings are the other biological children of both parents. Half siblings are the other biological children of one parent only. Social siblings include step, adoptive and foster siblings. Darker shading indicates greater likelihood of sibling co-residence. Based on ONS (2022), we assume that the index child co-resides with their biological mother following parental separation and hence also with the full and half siblings they share a biological mother with, but not the half-siblings they share a biological father with. Following this logic, we assume the index child's biological mother's step children would likely mostly live in another household (i.e. with their biological mother), as would social siblings who are the index child's biological father's step children. We assume adopted and foster children reside with the social parent who reports them in the data and the likelihood of their co-residence with the index child parallels that of half siblings. Thus, social siblings who are the mother's social children *possibly* co-reside with the index child while those who are the father's social children are *likely* non-resident with the index child.

An important layer of complexity not reflected in Table 1, but relevant for empirical operationalisation, is the age of the index child. Conceptually, only childness can be thought of as a stable characteristic; an individual growing up without siblings is an only child from birth and throughout the life course. In practice, with the exception of twins, all firstborns spend the initial period of their life as an only child. The decision whether these individuals ought to be classified as only children or as firstborn siblings will conceptually depend on the theoretical framework, the research question and possibly the age at which the outcome variable is measured (e.g. if the outcome is measured before the birth/ household entry of a sibling). The age gap between siblings may also be relevant for the identification and classification of last-born children, since co-residence will vary according to not only the age

of the index child but also the ages of any relevant sibling(s), especially as older siblings reach adulthood and move out. Additional temporal aspects of co-residence may also require careful consideration, such as the time children with separated parents spend residing with each parent (and potential half/social siblings) over the course of their childhood.

In practice, the ultimate operationalisation of only child status is often restricted by the information available in the data, with particular implications when using cross-sectional household surveys and in light of increasing family complexity, as we discuss further below.

Methods

Data

We analyse four British cohort studies which have collected detailed maternal fertility history information and/or documented the residence of children other than the cohort member (CM) in the childhood home. The longest-running study now covers the life course from birth to age 70, thus enabling identification of only children using childhood data as well as investigation of outcomes beyond childhood (e.g. Chanfreau and Goisis 2022). In contrast, surveys of adults tend to focus on contemporary household members and do not ask about respondents' siblings (e.g. Understanding Society or European Social Survey).

The National Survey of Health and Development (NSHD; Douglas, Wadsworth and Kuh 2015), has followed a subsample of the individuals born in a given week in 1946 (5,362 single births of the initially surveyed 13,687 births). The 1958 National Child Development Study (NCDS; CLS 2020) and the 1970 British Cohort Study (BCS70; Butler, Bynner and CLS 2016) follow cohorts of initially approximately 17,000 people born in a particular week in 1958 and 1970, respectively. The 2001 Millennium Cohort Study (MCS) has regularly surveyed a representative sample of nearly 19,000 individuals born between September 2000 and January 2002 (CLS 2017).

We compare estimates of the prevalence of only children obtained from these surveys with estimates from official statistics, derived using cohort fertility statistics by birth year of the woman published by the Office for National Statistics (ONS 2020), for women born approximately a generation prior to the survey cohorts.

Analyses

First, we use two methods for identifying only children which are as comparable as possible across the four surveys and compare prevalence estimates with estimates based on official statistics data from the Office for National Statistics (ONS). The *shared mother* method draws on fertility history data commonly used in demographic research, corresponding to cells d-i in Table 1, while the *co-residence* method is intended to capture the experience of growing up without other children in the parental home (Table 1: cells g-i, possibly some d-f). To compare the two sets of survey estimates with ONS data, we selected the years of birth matching the mean maternal age at first birth (MAFB) for births occurring in the years corresponding to the survey birth cohorts (ONS 2019). We note, however, that the ONS estimates reflect a shared-mother definition of only children, as neither social siblings nor paternal half-siblings are recorded in the data. Table 2 summarises the conceptual definition and variables used for operationalisation, making clear which decisions researchers are required to make.

Second, using the 2001 cohort which collected detailed household grid information at every sweep, we show how prevalence estimates differ across childhood and highlight the potential for misclassification due to siblings absent from the household when using cross-sectional household grid data. Here we combine the no shared mother and no co-residence definitions, identifying only children as CMs with neither co-resident (full, half or social) siblings nor siblings with the same mother but living elsewhere (Table 1: cells g – i).

Table 2 Data derivation of two definitions of only children

	Shared mother	Co-residence	Issues/ decisions of note
Conceptual definition	Mother has no other biological children	No other children sharing the childhood/parental home during the CM's childhood and adolescence	
NSHD 1946	Mother reports both interval to previous and subsequent birth (relative to CM) as not applicable • Derived variables made available by data team. Mothers were asked about prior births at the birth sweep and about subsequent births when the CM was aged 2, 4, 6, 8, 11 and 15 years.	No siblings reported in household grid when CM was aged 15. • Derived variable made available by data team. *'Home & family grid' records "parents and their children" living in this household.*	Household grid information specifies biological siblings and may misclassify as 'only children' CMs who are the youngest child if older siblings have moved out.
NCDS 1958	CM is mother's first birth reported at first sweep, not a twin/triplet birth and mother reports no subsequent live birth at age 11 sweep	CM is not a twin/triplet. No siblings reported in direct questions at age 16 about siblings with the same mother (the wording implies living anywhere) • *"The following questions about brothers and sisters are to be taken as referring to all children who have the same natural mother as the study child. Include in the answers any that you have already listed as being in the household. Do not count a twin of the study child as a younger or older brother or sister (the twin should have a separate form): How many older brothers? How many younger brothers? How many older sisters? How many younger sisters?"*	Given that children tend to grow up living with their mother we assume it is likely that siblings reported at age 16 shared the childhood home with the CM.

	Shared mother	Co-residence	Issues/ decisions of note
BCS70 1970	Mother reported no prior live (and surviving) births at first sweep. CM not a twin/triplet birth. Mother reports no subsequent live birth at age 5 sweep. Natural mother reports no younger siblings (living anywhere) at age 10 sweep.	No biological (including twin/triplet), step or adoptive siblings reported in household grid at age 10 sweep.	Maternal fertility histories not updated at age 10 sweep. The age 10 sweep did ask the parent respondent (96% were the mother) about any other family members in the household or living elsewhere and their relationship to the cohort member. We assume that younger siblings reported by the CM's natural mother at age 10 are her own.
MCS 2001	Mother reports no prior births at first sweep. CM is not a twin/triplet. Mother reports no subsequent live births, nor own children living elsewhere, at any sweep up to age 11.	No full, half, step or adoptive siblings reported as living in the household at any sweep up to age 11.	Foster siblings are also recorded in the household grid but we do account for them as 'siblings' due to the likely more transient nature of foster sibling relations.
ONS	Derived based on aggregate cohort fertility data for women in England and Wales born in years corresponding to average maternal age at first births (MAFB) occurring in the years of the cohort studies (i.e. MAFB reported by ONS for births occurring in 1946 is 26 years so we use ONS cohort fertility data for women born 1920). To estimate the prevalence of only children we first adjust the ONS reported proportion of all women with one child first to account for the proportion childless and then for the average family size among women with children.	n/a	ONS estimates are based on information collected at birth registration. Prior to 2012 the question on previous births was asked only to married women, ONS therefore adjusted the birth registrar information using General Household Survey (GHS). In the surveys maternal age at birth among women who had an only child was 3-4 years higher, on average, than MAFB among women who had more than one child.

Notes: CM=cohort member.

Finally, we examine the dimensions set out in Table 1 in more detail, again using the 2001 study. We disaggregate the 'co-residence' definition: no full siblings (1), no full or half siblings (2), or no full, half or social siblings (3) in the household. Combining household grid and questions about parents' non-resident biological children we derive separate shared-mother (4) and shared-father (5) definitions. Lastly, we identify those without either co-resident or shared-parent siblings (6). For each definition, we report the prevalence as well as headline socio-demographic characteristics and the number of siblings, comparing to the overall sample at age 11 (all CMs in the sweep).

Results

Prevalence: comparing definitions and data sources

Table 3 shows the prevalence estimates of only children in each study, for the co-residence and shared mother definitions. The prevalence was highest in the 1946 cohort, lower among those born in 1958 and 1970 and higher again among those born in 2001. The ONS data on the proportions of women who had one child confirm the U-shaped trend (21% of women born in 1920, 16% and 14% among women born in 1933 and 1946, 18% of women born in 1974).

Overall, the survey estimates are broadly of a comparable magnitude to the ONS fertility data. Although the 1946 cohort estimate (13.6% vs 11%) and the 1970 cohort 'shared mother' estimate (10.5% vs 6%) are higher, for the other cohorts and the 1970 co-residence definition the estimates are very similar to the ONS data estimates. Despite limitations in deriving only child prevalence estimates from aggregated ONS data, we are reassured by the similarity of the general pattern and the substantive magnitude that it is possible in the British birth cohorts to identify and externally validate a group that corresponds to a definition of only children.

Table 3 Prevalence estimates of only children at age 10/11 from four British cohort studies and ONS fertility estimates

Survey data	Only Child: % of all children		Only Child: % of first-borns			**ONS Data – England & Wales**				
Birth cohort	Co-reside	Shared mother	Co-reside	Shared mother	n	Mother's birth year	Mean age at 1st birth	% women with 1 child	*% of mothers with 1 child*	*Only as % of children*
NSHD 1946	13.6	13.6	33.1	32.7	4,154	1920	26	21	27	11
NCDS 1958	6.3	6.8	15.9	19.0	13,606	1933	25	16	18	7
BCS70 1970	6.9	10.5	16.5	25.0	13,836	1946	24	14	15	6
MCS 2001	9.2	9.6	21.3	23.1	12,997	1974	27	18	22	9

Notes: Survey data includes England, Scotland and Wales (1946, 1958, 1970) and England, Scotland, Wales and Northern Ireland (2001). 2001 cohort weighted using Sweep 5 analytical weight for whole-UK analyses to adjust for sample design and attrition; other cohorts unweighted. As a sensitivity analysis, we estimated the results using MAFB obtained from the 1958, 1970 and 2001 surveys (instead of ONS estimates) which allows us to obtain separate estimates between women who had an only child and women who had more than one child. The ONS estimated prevalence rates of only children were highly similar when using MAFB obtained using survey data: 8%, 6% and 10% for the 1958, 1970 and 2001 cohorts respectively when using MAFB for women who had an only child compared with 7%, 7% and 9% when using MAFB among those who had two or more children. This sensitivity analysis is not possible for the 1946 cohort study where maternal age is only provided in 5-year categories.

Except for the apparent over-estimation of only child prevalence for the 1970 cohort shared-mother definition (likely due to the lack of maternal birth history update at age 10), the similarity of the two sets of survey estimates shows the considerable overlap between the co-residence and shared mother definitions. We may interpret this as sibling status derived from maternal fertility history being a reasonable proxy for co-residence if detailed household grid information is unavailable (and vice versa). This may be context-specific however, as UK children usually reside with their mother and thus also with her other children. Yet lack of detail also limited the full operationalisation of the conceptual ideal definition, for example where the question-wording of items used to operationalise the *co-residence* method specified shared mother (1958) or biological siblings (1946).

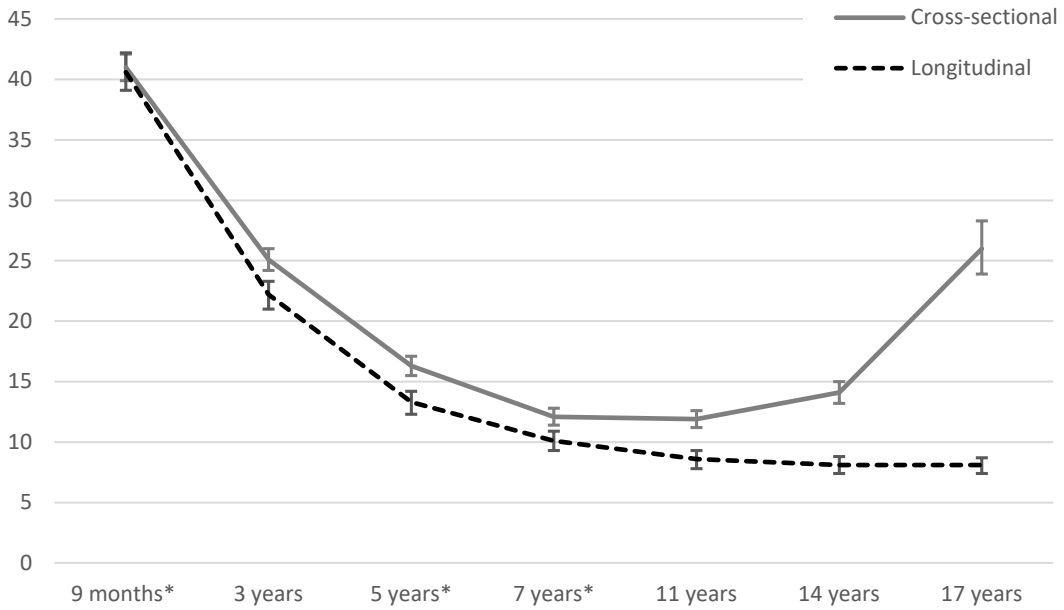
Prevalence: Timing in childhood

Timing matters for the operationalisation of only child status. Especially when using cross-sectional data household grids, the age of observing an absence of siblings should balance the chance of subsequent younger sibling births against the likelihood of older siblings still co-residing. To illustrate, using the 2001 cohort, we compare cross-sectional estimates at each age with a longitudinal derivation incorporating information from prior sweeps (Figure 1). The initial proportion of only children (at age 9 months) relates to singleton first-borns without co-resident social siblings, approximately 41% of the cohort.

After the first sweep, the proportion of children without siblings declines with age as younger siblings are born, with the largest reductions seen up to age 7. At each age, the cross-sectional estimate results in a higher prevalence of only children, but differences are small up to age 11. In adolescence, the cross-sectional estimate rises again, indicating misclassification as 'only children' CMs whose older siblings have left the parental home. The longitudinal derivation results in monotonically reducing proportions of only children, with minimal change after age 11 as such large age gaps are rare. In fact, among firstborns who had a co-resident sibling at the age 14 sweep less than 2% had an age gap greater than 11 years to their (oldest) younger sibling.

Taken together, although resulting in somewhat higher prevalence estimates than derivations based on longitudinal information, cross-sectional data collected between the ages of about 7 and 11 result in relatively more accurate identification of only children than data collected at younger or older ages. Whilst this indicates the need for caution when deriving only child status from household grid variables using cross-sectional data, the extent to which this is problematic depends on the study aims and, for studies comparing outcomes by only child status, at what age the outcome is measured.

Figure 1 Percentage of only children at different ages - 2001 cohort



Notes: Cross-sectional n 18,550 (9m); 15,589 (3y); 15,246 (5y); 13,857 (7y); 13,286 (11y); 11,725 (14y); 10,622 (17y). Longitudinal n 7,838. All analyses weighted for non-response.
* Indicates sweeps where the mother was asked about own children living elsewhere.

Group characteristics: Comparing definitions

Finally, we turn to how both the prevalence and the characteristics of the group identified may differ across derivations (Table 4). Our aim here is to demonstrate the consequences, in terms of group composition, of adopting different definitions of who is an only child and, implicitly, the biases which might be introduced by data constraints which might impose a partial identification. Definition (1), *no full* siblings, results in the highest prevalence of only children at nearly 20%. Compared with the overall sample and the other only child groups, this definition captures a disproportionately disadvantaged group. Therefore, data reporting solely full biological siblings will likely over-estimate the size and, at least in the UK, focus on a negatively selected only child group. If study outcomes are socio-economically patterned this might have implications for conclusions about only children's distinctiveness.

Table 4 Definitions and characteristics of only children using the MCS (2001 cohort)

	Only child definition						
		Co-residence		Shared parent		Combined	
	No full sibling (1)	No full/half siblings (2)	No full/ half/ social[1] siblings (3)	Mother's only child (4)	Father's only child[2] (5)	No known siblings[3] (6)	Sample overall
	%	%	%	%	%	%	%
Prevalence	19.9	9.6	9.1	9.2	11.8	7.6	100
Gender: Girl	47.5	48.1	48.5	48.9	48.4	48.9	48.5
Low income (9 months)	44.9	32.4	32.0	32.2	26.9	32.6	35.6
Mother's education: Degree	28.3	36.2	36.8	36.7	37.0	37.6	38.1
Parental social class: Managerial/ professional	28.8	37.6	38.0	38.1	36.6	38.2	38.7
Both parents in household (age 11)	31.6	39.6	40.0	41.2	43.3	40.2	60.6
	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Number of siblings[3]	0.7	0.1	0.0	0.1	0.6	0.0	1.6

Notes: [1] Here a small number of foster siblings are included under the umbrella term social sibling, along with step and adoptive siblings, although their co-residence with the cohort member is likely to be shorter. [2] Reported by biological father while co-resident with the cohort child, excludes siblings fathered by non-resident fathers. [3] Includes all co-resident (full, half or social siblings) and siblings with a shared parent living elsewhere at age 7 (the latest sweep that records non-resident siblings). [4] No co-resident siblings or siblings with a shared (co-resident) parent but living elsewhere.

Definitions (2), (3) and (4) all result in similar prevalence estimates of only children (9-10%). Except for the lower proportion of children living with both parents at age 11, the characteristics of these groups reflect the overall sample. Despite the definitional differences, siblings (identifiable in the data) are well accounted for under each version. Since 97% of all 11-year-olds in the study lived with their mother, this similarity reiterates the shared residence/mother overlap.

Definition (5) uses information about other children fathered by the CM's biological father, with nearly 12% classified as only children. Father interviews occur while residing with the CM, so any reported siblings living elsewhere are likely from the father's prior relationship. Due to lack of paternal follow-up after leaving the survey household, children fathered after

separation from the CM's mother are not identifiable as siblings. The group has 0.6 recorded siblings on average, more than the shared mother definition (4), because (5) neglects co-resident half/social siblings following maternal re-partnering. Definition (6), results in the lowest prevalence estimate of only children, at below 8%. This version is closest to cell (i) in Table 1 and by excluding all identifiable siblings might be considered the 'purest' derivation of only children. Importantly, all six derivations overlook non-resident siblings following paternal re-partnering.

Across all six derivations, being an only child is strongly related to parental separation; 32-43% of only children reside with both parents at age 11, compared with 61% overall – highlighting the need for data to reflect the complexity of family life. Complex families are of course not specific to the UK. According to the OECD family database the 'main home' living arrangements of 11-15 year olds in England – where 70% live with both parents, 18% with a lone parent and 11% in a step-family – are similar to the OECD average (74%, 16% and 9% respectively; OECD n.d.). In our data, among all 11-year-olds with separated parents and living with their mother in the overall sample: 67% had some contact with the non-resident father and among those who did: 63% had contact at least weekly and 73% reported overnight stays (including 40% staying overnight 'often'). If there are social or half siblings living with the non-resident father, children who are their mother's only child may nonetheless grow up with the social experience of having siblings.

Conclusion

In this paper we have reflected on the conceptualisation and measurement of only children by presenting different definitions and derivations based on nationally representative British cohort studies and providing a framework to guide future research aimed at understanding only children and their outcomes. The comparison of prevalence estimates derived from

surveys around age 10/11 and from official statistics show a similar trend across birth cohorts. Researchers using cross-sectional data collected around age 7-11, may also be reasonably confident about the accuracy of identifying only children using household grid information. However, identification of only children using cross-sectional data becomes more tenuous for adolescents as older siblings may have left the parental home and therefore not be recorded in the survey's household grid.

We also show the size and composition of the 'only child' group identified can vary by definition and that both conceptualization and measurement of 'only childness' are complicated by increasing family complexity. Since identifying only children is usually an initial step to investigating whether their (e.g. social or health) outcomes differ from those of children who grow up with siblings, we argue for greater transparency and critical reflection. An indicator that identifies a disproportionately disadvantaged sub-group may bias the study conclusions unless there is a strong theoretical basis for and robust operationalisation of the chosen definition. Limited information about fathers likely under-counts half and step siblings living elsewhere, somewhat overestimating only child prevalence. Depending on the research question, data gaps on non-resident siblings (including contact amount) might, with increasing shared care post-separation, become increasingly problematic and affect the results. These issues are specific neither to the UK nor to the datasets we have used for illustrative purposes.

As family life continues to diversify some changes to data collection practices, such as more complete parental fertility histories or a network approach to collecting family information, would improve both the identification of only children according to a given definition and enhance our understanding of the variety of sibling experiences. We emphasise that the most appropriate measure is not merely a question of data quality and accuracy but also of conceptual fit for a given research focus. Tailoring the definition of only children to the

research question can improve our understanding of whether, and if so why, only children's outcomes and experiences differ from those of individuals with siblings. Researchers ought to not only report how they identify only children in their data but also reflect on the ideal definition for their research question and how the operationalisation may influence results.

References

Baranowska-Rataj, A., K. Barclay, and M. Kolk. (2017). The effect of number of siblings on adult mortality: Evidence from Swedish registers for cohorts born between 1938 and 1972. *Population Studies* 71(1):43-63.

Beaujouan, E.and A. Solaz. (2019). Is the Family Size of Parents and Children Still Related? Revisiting the Cross-Generational Relationship Over the Last Century. *Demography* 56(2):595-619.

Blake, J. (1989). *Family size and achievement*. Berkeley: University of California Press.

Bobbitt-Zeher, D., D.B. Downey, and J. Merry. (2016). Number of Siblings During Childhood and the Likelihood of Divorce in Adulthood. *Journal of Family Issues* 37(15):2075-2094.

Breton, D.and F. Prioux. (2009). The one-child family: France in the European context. *Demographic Research* 20:657-691.

Butler, N., J. Bynner, and CLS. (2016*). 1970 British Cohort Study: Ten-Year Follow-Up, 1980. 6th Edition.* University of London, Institute of Education, Centre for Longitudinal Studies. UK Data Service. [data collection].

Chanfreau, J.and A. Goisis. (2022). Patterns of help and care by adult only children and children with siblings. *Ageing & Society*.

Chen, Z.-Y.and R.X. Liu. (2014). Comparing Adolescent Only Children with Those Who Have Siblings on Academic Related Outcomes and Psychosocial Adjustment. *Child Development Research* 2014:1-10.

CLS. (2017). *Millennium Cohort Study: First Survey, 2001-2003. [data collection]. 12th Edition.* University of London, Institute of Education, Centre for Longitudinal Studies. UK Data Service. SN: 4683. .

CLS. (2020). *National Child Development Study: Childhood Data from Birth to Age 16, Sweeps 0-3, 1958-1974.* [data collection]. 3rd Edition. University of London, Institute of Education, Centre for Longitudinal Studies. UK Data Service. SN: 5565, .

Douglas, J., M. Wadsworth, and D. Kuh. (2015). *MRC NSHD 1946-2005 Data.* MRC Unit for Lifelong Health and Ageing at UCL. Doi:10.5522/NSHD/Q101.

Downey, D.B. (1995). When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance. *American Sociological Review* 60(5):746-761.

Downey, D.B.and D.J. Condron. (2004). Playing Well with Others in Kindergarten: The Benefit of Siblings at Home. *Journal of Marriage and Family* 66(2):333-350.

Falbo, T. (2012). Only Children: An Updated Review. *Journal of Individual Psychology* 68(1):38-49.

Frejka, T. (2008). Parity distribution and completed family size in Europe: Incipient decline of the two-child family model? *Demographic Research* 19:47-71.

Frejka, T., G.W. Jones, and J.P. Sardon. (2010). East Asian Childbearing Patterns and Policy Developments. *Population and Development Review* 36(3):579-606.

Goetting, A. (1986). The Developmental Tasks of Siblingship over the Life Cycle. *Journal of marriage and the family* 48(4):703–714.

Haux, T., S. McKay, and R. Cain. (2017). Shared care after separation in the United Kingdom: limited data, limited practice? *Family court review* 55(4):572-585.

Keenan, K., K. Barclay, and A. Goisis. (2022). Health outcomes of only children across the life course: An investigation using Swedish register data. *Population Studies*:1-20.

Laybourn, A. (1990). Only children in Britain: popular stereotype and research evidence. *Children & Society* 4(4):386-400.

Lersch, P.M. (2019). Fewer Siblings, More Wealth? Sibship Size and Wealth Attainment. *European Journal of Population* 35(5):959-986.

Lin, S.and T. Falbo. (2022). Sibling absence and body mass index: From adolescence to adulthood. *Pediatric Obesity*:e12899-e12899.

Mancillas, A. (2006). Challenging the Stereotypes About Only Children: A Review of the Literature and Implications for Practice. *Journal of Counseling & Development* 84(3):268-275.

Mellor, S. (1990). How Do Only Children Differ from Other Children? *The Journal of Genetic Psychology* 151(2):221-230.

OECD. N.d. Table SF1.3.D: Living arrangements of young adolescents, 2013-14. In *OECD Family Database*.

ONS. (2019). Births by parents' characteristics. In *Statistical Bulletin: Population*: Office for National Statistics.

ONS. (2020). Childbearing for women born in different years, England and Wales, 2019. In *Statistical Bulletin: Population*: Office for National Statistics.

ONS. (2022). *Families and households Dataset, Table 2.*

ONS. (2023) *Child Benefit Statistics: annual release, data at August 2022 Child Benefit Statistics: annual release, data at August 2022*.

Präg, P., S. Choi, and C. Monden. (2020). The sibsize revolution in an international context: Declining social disparities in the number of siblings in 26 countries. *Demographic Research* 43(17):461-500.

Rainer, H.and T. Siedler. (2012). Family Location and Caregiving Patterns from an International Perspective. *Population and Development Review* 38(2):337-351.

Riswick, T.and T. Engelen. (2018). Siblings and life transitions: investigating the resource dilution hypothesis across historical contexts and outcomes. *The History of the Family* 23(4):521-532.

Zhang, L. (2011). Introduction and Overview. Pp. 3-10 in *Male fertility patterns and determinants*. Dordrecht ; London: Dordrecht ; London : Springer.