DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 54, ARTICLE 42, PAGES 1375–1412
PUBLISHED 19 JUNE 2026
https://www.demographic-research.org/Volumes/Vol54/42/
DOI: 10.4054/DemRes.2026.54.42

Research Article

Click, collect, compare: Evaluating a nonprobability web survey for family demography

Leonie Diffené

Thomas Leopold

Marcel Raab

Zafer Buyukkececi

Contents

Demographic Research: Volume 54, Article 42
Research Article

Click, collect, compare: Evaluating a nonprobability web survey for family demography

Leonie Diffené[1]

Thomas Leopold[2]

Marcel Raab[3]

Zafer Buyukkececi[4]

Abstract

BACKGROUND
High-quality population data are expensive to collect and limited in substantive scope, constraining timely and innovative research.

OBJECTIVE
We evaluate nonprobability data from the KINMATRIX survey by comparing it with external data on kinship structure and relationships to kin.

METHODS
KINMATRIX is a web-based quota survey of over 12,000 younger adults (aged 25–35) in 9 European countries and the United States. Respondents reported egocentric kinship networks, yielding more than 250,000 ego–kin dyads. We compare indicators in three domains – family network size (number of living kin), family relationships (emotional closeness, geographical distance, coresidence, and contact frequency), and family complexity (parental separation and number of half-siblings) – with estimates from probability-based surveys, demographic projections based on official vital statistics, and national registers.

RESULTS
KINMATRIX estimates of family network size tend to be lower than demographic projections based on official data, with smaller gaps for aunts and uncles and larger gaps for siblings and cousins. Cross-national differences in the number of parents, siblings, aunts, uncles, and cousins are broadly consistent across data sources, whereas patterns differ for grandparents. Regarding family relationships, the correspondence between

---

[1] Corresponding author. University of Cologne, Cologne, Germany. Email: diffene@wiso.uni-koeln.de.
[2] University of Cologne, Cologne, Germany.
[3] State Institute for Family Research at the University of Bamberg (ifb), Bamberg, Germany.
[4] Koç University, Istanbul, Turkey, and Max Planck Institute for Demographic Research, Rostock, Germany.

KINMATRIX and probability surveys varies across indicators, but established north–south contrasts are generally preserved. Regarding family complexity, KINMATRIX yields lower rates of parental separation than probability-based data, but it broadly captures cross-national differences.

CONTRIBUTION
Our comparisons point to nonprobability data having both promise and pitfalls in family demography, highlighting the need for study-specific benchmarking and targeted validation.

1. Introduction

Family demography is characterized by high standards of data quality. Because the field depends on precise information about demographic behaviors such as marriage, fertility, mortality, leaving home, cohabitation, separation, divorce, and remarriage, scholars have traditionally relied on data sources widely regarded as accurate, including official statistics, national registers, and large-scale probability surveys (Pesando et al. 2023). Quantitative research on family relationships has similarly drawn on probability-based surveys as well as data derived from official statistics and registers (Castrén et al. 2021). Although debates about data quality are common, they typically revolve around issues within an established baseline of rigorous standards.

However, maintaining these standards entails significant costs. The most obvious is financial, as collecting representative data for cross-sectional and longitudinal surveys that meet such standards requires substantial funding. Another cost is substantive, as existing high-quality sources often do not support innovative research questions and designs. Multi-purpose surveys have limited interview time, core modules in panel study questionnaires remain largely unchanged over time, and population data from official sources are often constrained in scope, coverage, and timeliness. As a result, family demographers face considerable challenges in accessing novel types of data while maintaining the quality standards expected in their field.

A potential solution to these challenges is the use of nonprobability data, which can be collected at a much lower cost through methods such as river sampling, recruiting participants through online advertising while they are engaged in other online activities (Murray-Watters et al. 2023), quota sampling from online access panels (e.g., YouGov, Dynata, Qualtrics), or crowdsourcing platforms like Amazon MTurk (Lehdonvirta et al. 2021).

Nonprobability data have become increasingly common in subfields that prioritize internal over external validity. For example, in experimental research the emphasis is often on random assignment to support causal inference (internal validity), while generalizability to other populations or contexts (external validity) is treated as a separate, and usually open, question. In fields such as psychology and behavioral economics, nonprobability samples from MTurk are now routinely used for experimental studies (Buhrmester, Kwang, and Gosling 2011; Paolacci and Chandler 2014). Public opinion research and political science also rely on nonprobability data to gauge public sentiment on sociopolitical issues or to reach hard-to-sample populations (Batrinca and Treleaven 2015; Salvatore 2023; Vicente 2023). In some fields, nonprobability data have become the dominant source of data, enabling researchers to rapidly test new ideas by collecting tailor-made data on demand.

By contrast, family demographers may have to abandon timely and innovative research questions because data meeting their rigorous quality standards are not readily available (Pesando et al. 2023). This raises an important question: Should these researchers follow the lead of other disciplines and begin collecting nonprobability data on demographic behaviors, family structures, and family relationships?

Critics point to concerns about the quality of nonprobability data (Arndt et al. 2022; Chmielewski and Kucker 2020; Freese and Jin 2025; Peer et al. 2021). These concerns are particularly relevant in family demography, where biased data can entirely invalidate conclusions drawn at the population level. Research in survey methodology indicates that certain risks can be mitigated. For instance, broad recruitment strategies can achieve a certain degree of ‘natural randomization’ and thus improve the quality of nonprobability samples (Vehovar, Toepoel, and Steinmetz 2016). Quality controls during and after data collection, such as attention checks and response validation, can add further improvements (Callegaro et al. 2014; Freese and Jin 2025; Kennedy et al. 2016). In addition, techniques like post-stratification and calibration weighting can help reduce selection bias (Bishop, Leite, and Snyder 2018; Lee and Valliant 2009).

Several studies have indicated that findings from nonprobability samples can be robust, particularly in bivariate and multivariate analyses (Ansolabehere and Schaffner 2014; Rohr, Silber, and Felderer 2025; Yeager et al. 2011). The potential and the pitfalls of nonprobability data in family demography are already evident in research that uses data sources such as genealogies and online family trees (Colasurdo and Omenti 2024; Kaplanis et al. 2018). Other innovative data sources, such as social media platforms and online forums, offer new ways of sampling family members (Herbell and Zauszniewski 2018) or studying family-related relationships or behaviors (Mondal et al. 2023; Sivak and Smirnov 2018; Smith et al. 2021), including a recent study using Twitter data to explore regional differences in family ties (Gil-Clavel and Mulder 2024).

Nonetheless, research in this field has only started to tap this potential. Empirical comparisons of nonprobability data with established data sources remain scarce. Although survey methodologists have conducted numerous comparisons of probability and nonprobability samples for basic demographic indicators (e.g., age, gender, education, marital status), these findings do not necessarily speak to the accuracy of nonprobability data for core concepts in family demography, including family structure, family (in)stability, demographic behaviors, and family relationships.

Addressing this gap is important, as it involves balancing the advantages of innovative, flexible data collection against potential risks to data quality in the population sciences. In this study, we take a first step toward filling that gap. We compare the recently collected KINMATRIX survey (Leopold et al. 2024, 2025a; see www.kinmatrix.eu) to data from probability surveys, official statistics, and national registers. Fielded between 2022 and 2024 in 9 European countries and the United States, KINMATRIX is a web-based nonprobability survey that recruited large-scale quota samples from online access panels. The data were released to the scientific community in February 2025 and offer innovative potential for family demography, capturing kinship networks far more comprehensively than previous surveys. KINMATRIX data go beyond nuclear family ties to include a broad range of maternal and paternal relatives – parents, grandparents, siblings, aunts, uncles, and first-degree cousins – as well as ‘complex kin’ such as step- and half-relatives arising from separation, re-partnering, and multi-partner fertility. Using an egocentric design, the survey recruited younger adults aged 25 to 35 as primary respondents (‘egos’), each of whom represents a kinship network of, on average, 20 family members (‘alters’). In total, the initial release of KINMATRIX comprises data on over 12,000 egos and more than 250,000 ego–alter dyads.

The KINMATRIX survey provides an example of innovative nonprobability data that capture concepts central to family demography (e.g., kinship structure and complexity) and sociology (e.g., support, contact, closeness, proximity, and conflict). For our comparison we draw on data from the 2024 Revision of the World Population Prospects (United Nations 2024), national registers (Swedish and Dutch), and probability-based data (European Social Survey 2018, 2020, 2024; International Social Survey Programme 2017; Generations and Gender Survey 2020–2023; German Panel Analysis of Intimate Relationships and Family Dynamics 2020–2022). Our analysis centers on three domains: family network size (average numbers of living kin in various kinship categories, such as parents, siblings, grandparents, aunts, uncles, and cousins), family relationships (coresidence, geographical distance, emotional closeness, and frequency of contact with parents and siblings), and family complexity (parental separation, number of half-siblings).

In contrast to controlled comparison designs in which the sampling method is the only feature that varies, the reference data differ from KINMATRIX not only in sampling but also in mode of data collection, fieldwork timing, question wording, response scales, and other features. Moreover, we assembled the reference data from diverse sources (surveys, official statistics, and national registers) and the methods used to derive estimates range from simple univariate and bivariate measures to matrix models that generate demographic projections of kin numbers from aggregate data (Caswell 2019; Williams et al. 2021). Although all reference data are regarded as high quality, each has limitations, and none can be treated as a benchmark that precisely captures a commonly defined population value. As a result, our design cannot pinpoint the reasons for deviations from KINMATRIX estimates, which may stem, for example, from differences in measurement or harmonization rather than sampling bias in nonprobability data. Conversely, even apparent alignment may be partly coincidental (e.g., resulting from offsetting errors). Moreover, we evaluate only a single nonprobability dataset, so our results do not speak to nonprobability designs more generally. In light of these constraints, our analyses necessarily provide only a cautious assessment of KINMATRIX and represent only an initial step toward addressing the broader question of what nonprobability data can contribute to family demography.

2. Background

2.1 Strengths and weaknesses of probability and nonprobability sampling

In a probability sample, every individual in the target population has a known, non-zero probability of being selected, enabling researchers to accurately quantify sampling error and generalize findings with known levels of confidence (Cochran 1977; Levy and Lemeshow 2008). Despite these benefits, probability sampling is limited by practical challenges, notably high costs and declining response rates. Pew Research reported a drop in telephone survey response rates from 36% in 1997 to 9% in 2012 (Kohut et al. 2012). Although initial samples may be randomly selected, low response rates along with underrepresentation of specific subpopulations may undermine the benefits of probability sampling, given that factors such as employment status, socioeconomic status, age, relationship status, parenthood status, mental and physical health, and sex are linked to non-response (Lindner, Murphy, and Briers 2001; Smith 1995). Standard strategies to address non-response, such as repeated follow-ups with non-respondents, are time-consuming, costly, and of limited effectiveness (Tourangeau et al. 2017).

Meanwhile, nonprobability samples have gained traction in the age of big data, particularly where cost efficiency and convenience are prioritized (Elliott and Valliant 2017). One common nonprobability approach is the use of volunteer web panels where participants receive monetary or in-kind incentives (Callegaro et al. 2014). Because nonprobability sampling lacks a formal framework for evaluating accuracy and statistical uncertainty, researchers often question the validity of inferences drawn from these data (Baker et al. 2013).

The extent to which nonprobability samples can estimate population parameters depends on three assumptions: (1) ignorability, where the selection mechanism is independent of the variables of interest; (2) positivity, which requires that all subgroups defined by potential confounders are represented; and (3) composition, requiring that the sample composition aligns well with the population (Elliott and Valliant 2017; Mercer et al. 2017; Rosenbaum and Rubin 1983). These assumptions are often difficult to validate, and failure to meet them can bias estimates in ways that may be difficult to predict or correct.

A common strategy for evaluating the quality of nonprobability data is to compare them to probability surveys or population-level data. In univariate analyses, nonprobability samples yield less accurate population estimates than probability samples (Erens et al. 2014; Gittelman et al. 2015; Hays, Liu, and Kapteyn 2015; Jerit and Barabas 2023; Kocar and Baffour 2023; Lavrakas et al. 2022; Lehdonvirta et al. 2021; Rohr, Silber, and Felderer 2025; Yeager et al. 2011; Zack, Kennedy, and Long 2019). A comprehensive comparison shows consistently higher bias in nonprobability surveys, particularly for demographic variables (Yeager et al. 2011), a finding corroborated by a comparison of 17 nonprobability surveys against benchmark data (Gittelman et al. 2015).

Bivariate associations and multivariate model estimates may be less sensitive to nonprobability sampling than univariate (marginal) estimates (Ansolabehere and Rivers 2013; Pasek 2016; Zack, Kennedy, and Long 2019). In benchmarking studies, errors tend to be smaller for relationships than for levels, an advantage that is most evident when selection operates through observed covariates that can be adjusted for. Accordingly, results vary by context and by variable (Pasek 2016; Rohr, Silber, and Felderer 2025; see also Zack, Kennedy, and Long 2019 for mixed findings).

Online panels, a common form of nonprobability sampling and the primary focus of this study, frequently overrepresent younger, more tech-savvy respondents while underrepresenting singles, ethnic minorities, older adults, and low-income populations with limited internet access or willingness to participate (Lehdonvirta et al. 2021). Such biases can lead to inaccurate findings, particularly in research on sensitive topics like substance use or mental health (Bornstein, Jager, and Putnick 2013).

Statistical methods such as weighting, calibration, and combining data from probability and nonprobability sampling aim to reduce bias (Chen, Li, and Wu 2020; Nandram, Choi, and Liu 2021; Wiśniowski et al. 2020). Fair and uniform compensation in online panels may decrease economic self-selection (Lehdonvirta et al. 2021), and excluding respondents with high levels of missing data or irregular response patterns can further improve data quality (Hays, Liu, and Kapteyn 2015). Nevertheless, probability surveys maintain higher accuracy, regardless of data collection mode – face-to-face, online, or telephone – even when response rates are as low as 15%–20% (Lavrakas et al. 2022; MacInnis et al. 2018).

2.2 Nonprobability data in family demography

Although nonprobability data remain relatively rare in quantitative family research, studies increasingly use such data to address specific challenges and questions. A key area is hard-to-reach populations, particularly in contexts where probability sampling is impractical. For example, quantitative studies on caregiving and family assistance routinely rely on convenience and snowball sampling, recruiting participants through word of mouth, direct contact with project staff, press releases, and postings on nonprofit organization websites (Pruchno et al. 2008).

Beyond targeting specific populations, nonprobability sampling enables researchers to collect data promptly. For instance, Mocho et al. (2024) investigate parental stress and child well-being during the COVID-19 pandemic using a convenience and snowball sampling approach via Google Forms, Facebook, Instagram, and WhatsApp. Similarly, Smith et al. (2021) analyze fertility-related social networks on Twitter before and during the pandemic, collecting real-time data with NodeXL.

Nonprobability sampling also facilitates the use of structural data (e.g., information embedded in specific platforms or settings) and behavioral data (e.g., posting patterns on social media), offering new insights for family demography. An example is provided by Kaplanis et al. (2018), who created large-scale family trees using data on 86 million genealogy profiles from Geni.com. Although extensive data cleaning was required, their study suggested that the online genealogical data aligned well with official records for certain periods and cohorts. Yet, subsequent research has highlighted limitations and biases, including an overrepresentation of socially privileged groups, underreporting of early-age mortality and childlessness, as well as gender and regional biases (Chong et al. 2022; Colasurdo and Omenti 2024; Stelter and Alburez-Gutierrez 2022). Other studies have leveraged behavioral data from social media to explore family dynamics in new ways. For instance, Gil-Clavel and Mulder (2024) examine north-south differences in European family ties through Twitter user activity, and Sivak and Smirnov (2018) investigate gender inequalities in Russia by analyzing social media posts that reference ‘daughter’ or ‘son.’

While these studies illustrate potential applications of nonprobability data in family demography, they also highlight similar limitations as observed in other social science disciplines, most notably selection bias. Despite these challenges, nonprobability data are on the rise, offering new potential to study marginalized groups, implement innovative designs, and study timely phenomena. If such data also allow valid conclusions at the population level, they can expand the scope and depth of research in ways that traditional data sources do not allow.

3. Data and methods

3.1 KINMATRIX: A web-based nonprobability survey

The KINMATRIX survey, fielded between 2022 and 2024, collected egocentric family network data of young adults aged 25–35 years in 9 European countries (Denmark, Finland, Germany, Italy, the Netherlands, Norway, Poland, Sweden, the United Kingdom) and the United States (Leopold et al. 2024, 2025a). A full documentation of questionnaires, programming, web design, respondent look-and-feel, sampling, and translation procedures is available at www.kinmatrix.eu.

In the KINMATRIX web survey, respondents ('egos') created a family tree and reported on relationships with their kin ('alters'). The data encompass various family members across maternal and paternal lines, including nuclear kin (parents and siblings), extended kin (grandparents, aunts, uncles, and first-degree cousins), and complex kin (half- and step-relatives). Respondents were recruited from Dynata's online access panels. Hard (i.e., fixed) sampling quotas were employed for age (25–30 and 31–35 years). Soft (i.e., flexible) sampling quotas were used for gender, education levels (low, medium, and high), and geographical regions within each country. All country-specific quotas were obtained from national statistical offices (OECD). Since not all targeted quotas were met during data collection, weights were generated to match the quotas for gender, age, education, and region in each country. For each country sample, iterative proportional fitting was used to calibrate the sample to known population margins at the respondent level (Bergmann 2011).

During and after data collection, the project conducted quality checks and data cleaning procedures to remove implausible or invalid data, focusing on the numbers, names, and living statuses of kin included as 'alters' in respondents' kinship networks. A detailed description of the procedures is available in the method report (www.kinmatrix.eu). The cleaned data released to the scientific community in February 2025 include a total sample of 12,438 egos and 252,278 alters. Table 1 shows the sample size broken down by the 10 survey countries. For country-specific comparisons, it is important to take into consideration the very different country sample sizes in KINMATRIX, ranging from 127 egos and 2,579 alters in Norway to 5,005 egos and 113,697 alters in the United States.

Table 1: Numbers of egos and alters in KINMATRIX

Country	Total number		Mean number of alters per ego
	Egos (respondents)	Alters (kin)	
DE	1,208	19,726	16.33
DK	154	3,006	19.52
FI	241	4,695	19.48
IT	1,909	35,369	18.53
NL	394	7,793	19.78
NO	127	2,579	20.31
PL	1,732	32,926	19.01
SE	355	6,849	19.29
UK	1,313	25,638	19.53
US	5,005	113,697	22.72
Total	12,438	252,278	20.28

Note: DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom, US=United States.

3.2 KINMATRIX indicators selected for comparison

Family network size. While other surveys often record only a subset of kin, the KINMATRIX web survey instrument asked respondents to report all parents and siblings, half-siblings, grandparents, aunts, uncles, and first cousins on both paternal and maternal sides (where applicable). The instrument explicitly aimed to elicit the full set of biological kin in these categories. When respondents did not know a particular number of kin – for example, the number of children of a paternal uncle – a follow-up prompt asked them to report those kin they were able to identify.

Given that extended kin frequently outnumber nuclear kin, these data allow researchers to analyze not only the quality of relationships within kinship networks but also the quantity of kin available for social interaction, support, transmission, and influence. Results from the survey demonstrate that simultaneously examining the quality and quantity of kin relationships yields richer insights into the demography and sociology of families and kinship (Hünteler et al. 2026; Leopold et al. 2025b; Leopold and Kalmijn 2026).

Despite its analytic importance, measuring family network size is challenging. In the KINMATRIX web survey instrument, respondent burden was substantial: on average, younger adults reported around 20 kin per ego, and often many more. The large number of extended kin, particularly aunts, uncles, and cousins, required considerable motivation and knowledge on the part of respondents, creating a risk of response bias (e.g., underreporting distant kin to shorten the survey, uncertainty about the existence or exact number of certain relatives). We selected 9 count measures for comparison with external data: mean numbers of living fathers, mothers, brothers, sisters, grandfathers, grandmothers, uncles, aunts, and cousins.

Family relationships. Following the name generators eliciting an extensive set of kin as alters in younger adults' kinship networks, the KINMATRIX web survey employed multiple name interpreters to gather information on respondents' relationships with these kin, as well as on kin attributes. These questions were presented using various response formats, such as name carousels, to collect relational data and attribute data efficiently while minimizing response burden (for survey screenshots, see the method report at www.kinmatrix.eu).

Core questions on family relationships drew on the intergenerational solidarity model (Bengtson and Roberts 1991). Other measures were adapted from instruments used in probability surveys such as the European Social Survey, the Generations and Gender Survey, the Longitudinal Study of Generations, the Survey of Health, Ageing, and Retirement in Europe, and the German Family Panel. KINMATRIX data on these indicators offer novel potential for analysis, enabling researchers to evaluate multiple dimensions of solidarity from a broader perspective on family networks.

Response burden was a concern for each of these relational indicators, as respondents had to answer the same relational questions for a large number of kin. We selected three indicators for comparison with external data: (1) emotional closeness, (2) geographical distance, and (3) contact frequency.

Family complexity. The KINMATRIX survey's detailed assessment of family complexity offers further analytic potential. KINMATRIX includes complex kin such as parents' partners and step-siblings, and aims for a complete record of all half-siblings. In addition, the questionnaire gathered detailed information on both parents' union histories, including parental separation, re-partnering, coresidence, and custodial arrangements. These data have the potential to expand research on the consequences of separation and the structure of complex families beyond the usual emphasis on child outcomes and parent–child relationships. By capturing ripple effects through larger kinship networks, KINMATRIX highlights that divorce, separation, re-partnering, and multi-partner fertility can lead not only to large losses in the quality of relationships with kin but also to gaining new kin (Leopold et al. 2025b; Sytkina and Leopold 2026).

Similar to previous indicators, this detailed assessment involves a risk of response bias, as respondents may, for example, underestimate the number of half-siblings due to incomplete family knowledge. We selected two constructs for comparison with external data: (1) the prevalence of parental separation and (2) the number of younger adults' half-siblings.

3.3 Reference data

To compare these three sets of KINMATRIX indicators with reference data, we selected several established sources, including probability-based surveys, register data, and kinship projections derived from formal demographic models using United Nations (UN) vital statistics. For our assessment of family network size we drew on reference data from structural-demographic sources; for relationship quality and family complexity we relied on probability surveys. To identify the most appropriate data for comparison, we used five criteria.

- Pertinence: Domains aligned with the KINMATRIX indicators selected for comparison (i.e., family network size, family relationships, and family complexity).
- Quality: High data quality and sufficiently large country samples.
- Harmony: Indicators that could be satisfactorily harmonized to align with the KINMATRIX indicators.
- Scope: Comprehensive comparative coverage, prioritizing cross-national data where available.
- Recency: Data collected as close as possible to the field period of KINMATRIX, while also considering less recent data if no alternatives were available.

The reference data are summarized in Table 2. We used the DemoKin R package (Williams et al. 2021) to calculate estimates of family network size based on a formal demographic matrix ('DemoKin estimates') (Caswell 2019, 2020, 2022; Caswell and Song 2021). These models use a two-sex time-varying approach that takes age-specific mortality and fertility rates as input to project the expected number of living kin for individuals of specified age and sex. As the data do not include male fertility rates, we approximated male fertility by shifting female rates upward by 4 years to reflect men's later fertility. All projections of kin numbers are based on demographic rates from the 2024 Revision of the World Population Prospects (United Nations 2024) and estimate the average number of kin alive per kinship category (e.g., father, mother) for individuals aged 25 to 35 (aligned with the KINMATRIX age range) in every survey country. In addition, we included register-based estimates published in De Bel et al. (2024) and Kolk et al. (2023) for the mean number of kin alive (except for brothers and sisters, due to the lack of sex-specific sibling data) in the Netherlands in 2018 and in Sweden in 2017.

For family relationships, we used cross-national and national probability-based survey data. We drew on Round 10 (collected in 2020) of the European Social Survey (ESS) for reference data on emotional closeness, geographical distance, coresidence, and daily or weekly contact between individuals aged 25 to 35 and one of their randomly selected living parents in up to 8 European countries. In addition, we drew on the

International Social Survey Programme (ISSP) collected in 2017 as reference data for the average frequency of contact with the most frequently contacted parent or sibling. ISSP data were available for the United States and 5 European countries included in KINMATRIX. Since the ESS and ISSP cover Great Britain but not the United Kingdom, we reduced the comparison with these reference data to Great Britain by excluding Northern Ireland.

As some of the indicators were only partially harmonizable with KINMATRIX data, we additionally included national comparison data from Wave 14 (collected in 2021–2022) of the German Family Panel (pairfam), which showed a higher degree of harmony. Moreover, the pairfam data allowed us to conduct bivariate gender-specific analyses of contact frequency, emotional closeness, and geographical distance between younger adults and both of their parents (see Table 2 for details of the use of weights).

For comparison data on family complexity, we used a measure of parental separation from Wave 1 of Round 2 (collected in 2020–2023) of the Generations and Gender Survey (GGS), which indicated the share of individuals whose parents had separated, if both parents were still alive. As the GGS is a cross-national panel survey, data were available for most European countries included in KINMATRIX. As a second and supplementary indicator of family complexity, we used published estimates from national register data for the mean number of living half-siblings in the Netherlands in 2018 (De Bel et al. 2024) and in Sweden in 2017 (Kolk et al. 2023).

Table 2: Overview of reference data

Indicator	External data	Age range	Variable name / Indicator	Weight	Countries
Family network size					
Number of living kin per kinship category	UN World Population Prospects 2024 (DemoKin)	25–35	Kin-specific estimates calculated using fertility and mortality rates from the UN World Population Prospects 2024 (United Nations 2024) and two-sex matrix kinship models (Caswell 2022) implemented in the DemoKin R package (Williams and Alburez-Gutierrez 2025)	–	DE, DK, FI, IT, NL, NO, PL, SE, UK, US
	Register data (2018) from De Bel et al. (2024)	25–35	Estimates published in the online supplementary material.	–	NL
	Register data (2017) from Kolk et al. (2023)	25–35	Estimates published in online Appendix 2; excluding deceased kin.	–	SE

Table 2: (Continued)

Indicator	External data	Age range	Variable name / Indicator	Weight	Countries
Family relationships					
Emotional closeness with random parent	ESS (2020)	25–35	‘closepnt’. Taking everything into consideration, how close do you feel to {him/her}?	anweight. (composite weights with design, post-stratification, and population adjustments)	DE, FI, GB, IT, NL, NO, PL, SE
Geographical distance to random parent	ESS (2020)	25–35	‘ttminpnt’. About how long would it take you to get to where {he/she} lives, on average? Think of the way you would travel and of the time it would take door to door.	anweight. (see above)	DE, FI, GB, IT, NL, NO, PL, SE
Coresidence with random parent	ESS (2020)	25–35	‘ttminpnt’. (see above) ‘hhlipnt’. Does {he/she} live in the same household as you?	anweight. (see above)	DE, FI, GB, IT, NL, NO, PL, SE
Daily or weekly contact with random parent	ESS (2020)	25–35	‘speakpnt’. How often do you speak with {him/her} in person? Please only include occasions where you are physically in the same location. ‘scrnpnt’. How often do you speak with {him/her} such that you can see each other on a screen? ‘phonepnt’. How often do you speak with {him/her} using a phone or other device? Please include calls you make or receive, but exclude calls where you see each other on a screen. ‘compnt’. How often do you communicate with each other via text, email, or messaging apps?	anweight. (see above)	DE, FI, GB, IT, NL, NO, PL, SE
Contact with the most frequently contacted parent [sibling]	ISSP (2017)	25–35	‘v48’ [‘v49’]. Please think about the parent [brother/ sister] you have contact with most frequently: How often do you have contact with that parent [brother/sister], either face to face, by phone, internet, or any other communication device?	weight. (country-specific weights, including design and post-stratification weights)	DE, DK, FI, GB, SE, US
Emotional closeness with mother [father]	Pairfam (2021–2022)	27–31	‘igr40p1’ [‘igr40p3’]. How close do you feel emotionally to your mother [father] today?	cd2weight. (see above)	DE
Geographical distance to mother [father]	Pairfam (2021–2022)	27–31	‘igr41p1’ [‘igr41p3’]. How much time do you need to get to your mother’s [father’s] residence? (on a normal day, using normal means of transportation)	cd2weight. (see above)	DE
Contact frequency with mother [father]	Pairfam (2021–2022)	27–31	‘igr39p1’ [‘igr39p3’]. How often do you have contact with your mother [father], adding up all visits, letters, phone calls, emails etc.?	cd2weight. (calibrated design weight)	DE
Family complexity					
Parental separation	GGS (2020–2023)	25–35	‘gen38a’. Did your biological parents ever break up?	weight. (country-specific weights, including design and post-stratification weights)	DE, DK, FI, NL, NO, SE, UK
Number of living half-siblings	Register data (2018) from De Bel et al. (2024)	25–35	Estimates published in the online supplementary material.	–	NL
	Register data (2017) from Kolk et al. (2023)	25–35	Estimates published in online Appendix 2. Type reduced to half-mother and half-father.	–	SE

Note: DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States.

3.4 Variable and sample harmonization

To improve comparability across data sources, we harmonized variables and analytical samples. Variable harmonization required recoding indicators that were not identical across data sources. Detailed recoding schemes are available in the online supplement (S1. Recoding scheme). For most indicators, the recoded variables could be fully or almost fully aligned. In some cases we could achieve only partial harmonization. For example, the GGS captured parental separation in an affirmative manner (Did your biological parents ever break up?), whereas KINMATRIX asked this question only in the minority of cases in which at least one parent had died, and otherwise asked if the respondent's parents were still together.

To harmonize the samples of KINMATRIX and the probability-based comparison surveys, we selected individuals within the KINMATRIX age bounds of 25 to 35. The only exception to this was data from the pairfam study, which by design selects only a narrow range of birth cohorts. For comparisons with pairfam, we used the cohort born in 1991–1993 and aged 27 to 31 in pairfam Wave 14 (2021–2022). Register estimates from Kolk et al. (2023) and de Bel et al. (2024) pertain to individuals aged 25 to 35 in 2017 in Sweden and aged 25 to 35 in 2018 in the Netherlands.

We further reduced the KINMATRIX and reference samples to non-missing values on the respective indicator. The exception was parental coresidence, where we treated 'don't know' as no coresidence because we assumed that the kin was non-coresident if the respondent did not know the travel time to the kin. For family relationships, we excluded cases in which kin were not alive because respondents did not receive follow-up questions about their relationship with those kin. In addition, we reduced the KINMATRIX samples to respondents with non-missing values on the weight variable (missing values could occur if a respondent did not provide a valid answer on their gender, education, or region). We flagged comparisons based on fewer than 200 respondents per indicator and country (see results in Figure A-1 and Figure A-3 in the Appendix). Analytical sample sizes can be found in the online supplement (S2. Sample sizes).

We used weighted KINMATRIX data based on calibrated post-stratification weights derived via iterative proportional fitting. For the external survey data we employed the weights as indicated in Table 2. In the analysis we compared weighted and unweighted KINMATRIX means against weighted means of external survey data and unweighted means if the reference data were based on official statistics.

For family network size, we additionally calculated the symmetric mean absolute percentage error (sMAPE) and bootstrapped confidence intervals of the mean number of living kin to account for the varying scales across kinship categories. The results of this measure, shown in the Appendix (Figure A-2), indicate the degree to which (weighted or unweighted) KINMATRIX estimates deviated from reference data in terms of absolute percentages. Unlike the mean absolute percentage error (MAPE), sMAPE prevents extreme errors when actual values are close to zero and ensures that over- and underestimations are treated equally (Tofallis 2015).

4. Results

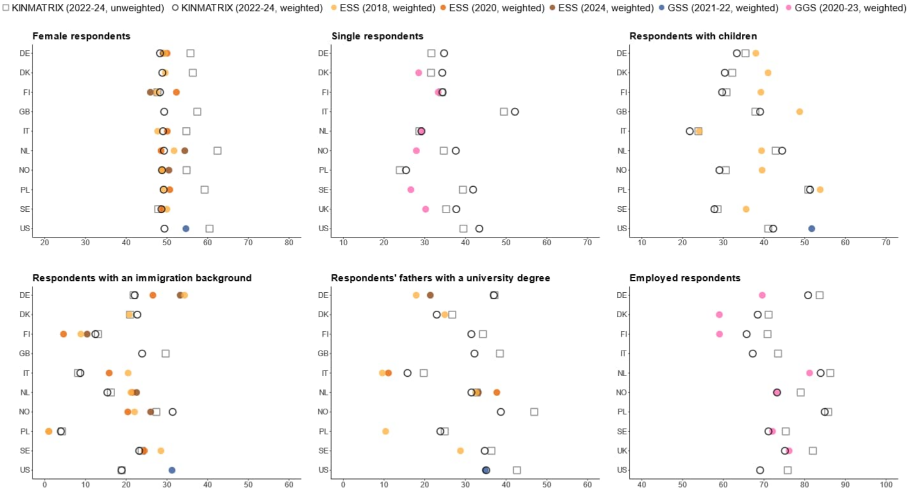
4.1 Sociodemographic characteristics

Before turning to the main analysis, we explored the sample composition of KINMATRIX in comparison to probability-based surveys across all countries. Figure 1 shows the country-specific means of sociodemographic characteristics by data source at the respondent level, focusing on gender, singlehood, parenthood, immigration background, paternal university degree, and employment status.

The comparison of weighted KINMATRIX data (marked as hollow circles) and unweighted KINMATRIX data (hollow squares) with probability-based survey data (coloured filled circles) showed that women were substantially overrepresented in the unweighted KINMATRIX data and matched probability-based data only after applying post-stratification weights. The remaining demographic variables shown in Figure 1 were not used in the calibration of weights and often display sizable deviations from probability-based data, although variance across probability-based data sources was also substantial in some of the countries. Overall, KINMATRIX data tended to underrepresent parents and immigrants, while singles, respondents from academic families, and employed respondents (including full-time, part-time, and self-employed) were overrepresented in most countries.

These compositional differences may also be relevant for the substantive comparisons that follow. In particular, the underrepresentation of immigrants and respondents with children and the overrepresentation of respondents from higher-educated family backgrounds could plausibly affect estimates of family network size and family complexity. For example, if immigrant groups tend to have more biological kin, their underrepresentation in KINMATRIX may contribute to lower kin counts. Similarly, if parental separation is less common among respondents from higher-educated family backgrounds, their overrepresentation may contribute to lower estimated levels of parental separation. Such effects cannot be separated from other potential sources of discrepancy in the present design, but they should be kept in mind when interpreting the results below.

Figure 1: Sociodemographic characteristics (mean share in %)



Note: DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States. Details on variable coding can be found in 'S1. Recoding scheme' in the online supplement. Country-specific reference data based on fewer than 200 respondents are excluded (for complete reference estimates, see Figure A-1 in the Appendix).
Source: Own calculation and illustration.

4.2 Family network size

The first step of our comparison focused on kinship network size across all KINMATRIX countries. Figure 2 presents the mean number of kin alive by data source, country, and kinship category. The comparison of weighted and unweighted KINMATRIX data with DemoKin estimates (red-filled circles) and register-based estimates (blue-filled circles) showed alignment for both fathers and mothers, including higher paternal than maternal mortality and cross-national differences such as higher paternal mortality in Poland and higher maternal mortality in the United States.

In most cases, the KINMATRIX estimates for the remaining kinship categories were lower than the reference estimates. For example, in most countries the mean number of living brothers or sisters was approximately 0.1 to 0.3 lower in KINMATRIX than that calculated by demographic projections. Weighted KINMATRIX estimates showed slightly closer alignment. Well-established cross-national fertility patterns tended to be reflected in both the reference data and the KINMATRIX data – visible in terms of lower numbers of siblings in Germany and Italy and higher numbers in Scandinavian countries and the United States (Campisi et al. 2020; Dudel and Klüsener 2021).

For grandparents, the estimates in Figure 2 represent mean values combined from both lines (i.e., the mean number of paternal and maternal grandmothers/grandfathers alive for an adult aged 25 to 35). These quantities – and cross-national differences therein – are shaped not only by grandparental mortality but also by the timing of fertility across successive generations and the resulting average length of generations. For the mean number of living grandmothers, KINMATRIX estimates were again lower than comparison estimates in most countries, whereas for grandfathers this pattern was reversed and deviations were larger. In contrast to the sibling counts, weighting had little effect on reducing discrepancies.

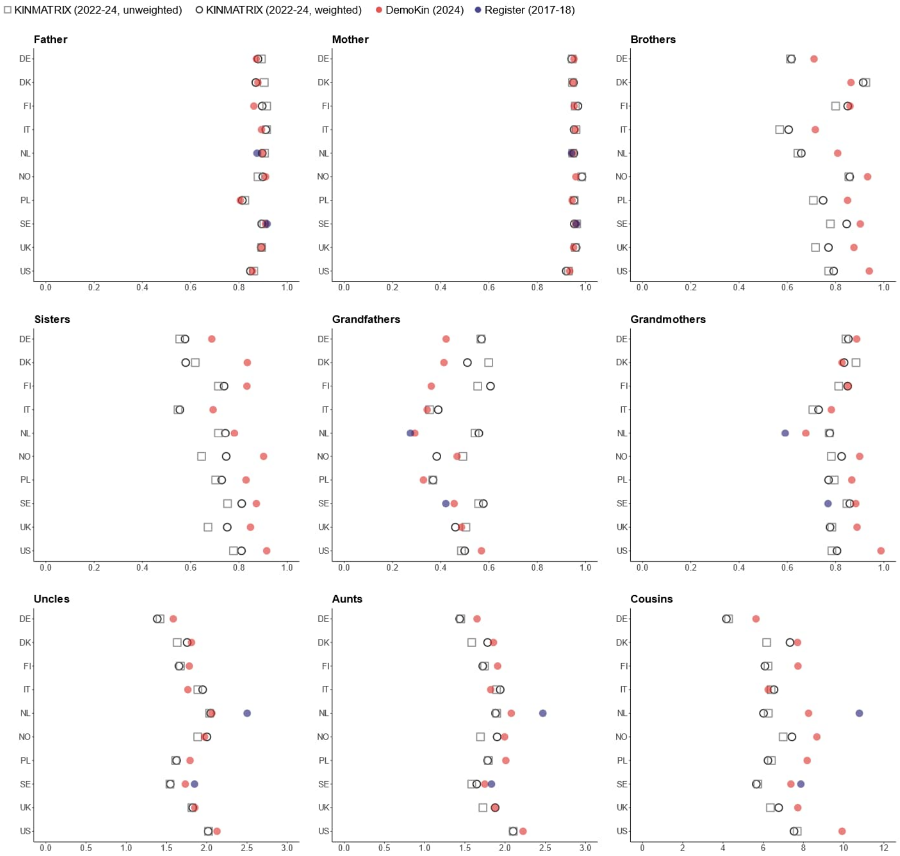
For the number of uncles and aunts, KINMATRIX estimates aligned relatively closely with the reference estimates. Again, KINMATRIX estimates were lower than those from demographic projections, with gaps of 0.2 to 0.3 in several countries. For cousins, gaps between KINMATRIX and reference estimates tended to be larger, with reference estimates exceeding those of KINMATRIX in all countries except Italy. For the number of uncles, aunts, and cousins, weighted KINMATRIX estimates tended to align slightly better with reference estimates in the smaller country samples of Northern Europe (Denmark, Finland, Norway, Sweden). Given that the kinship categories presented in Figure 2 differ substantially in terms of their quantity, the standardized sMAPE estimates presented in the Appendix (Figure A-2) allow for comparing the extent of alignment and deviation across kinship types. These results indicate that the closest overall alignment was for parents, followed by uncles and aunts. Deviations were larger for full siblings and cousins, and largest for grandfathers.

In addition to demographic projections, Figure 2 shows register-based estimates for the Netherlands and Sweden. These estimates yielded smaller numbers of living grandparents[5] and larger numbers of aunts, uncles, and cousins. Overall, KINMATRIX estimates for the small country samples of the Netherlands and Sweden tended to deviate more from the register-based estimates than from the demographic projections.

---

[5] A smaller number of grandparents might be partially due to the requirement of complete parental linkages in the register data, as well as the exclusion of kin not residing in the country or not traceable in the administrative data (De Bel et al. 2024; Kolk et al. 2023).

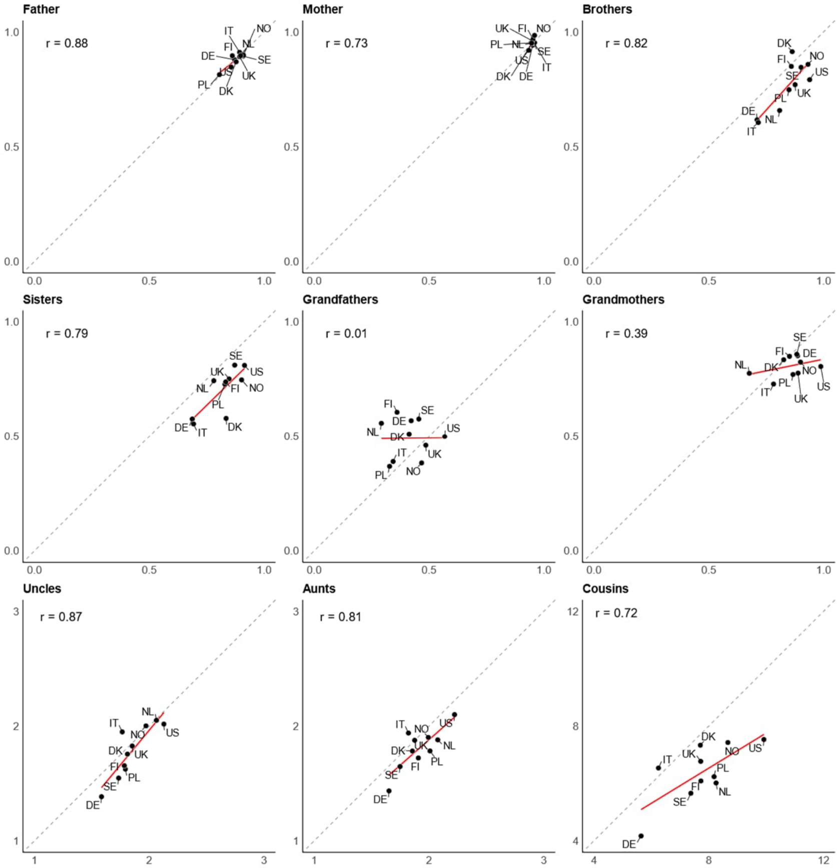
Figure 2: Family network size (mean number of kin alive)



Note: DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States. Details on variable coding can be found in ‘S1. Recoding scheme’ in the online supplement.
Source: Own calculation and illustration.

In Figure 3 we illustrate the extent to which KINMATRIX data aligned with demographic projections in terms of cross-national differences in family network size. The figure shows DemoKin-based demographic projections on the x-axis, and KINMATRIX estimates on the y-axis, along with a dashed reference line indicating complete alignment. The majority of markers are located below this reference line, indicating that DemoKin estimates often exceeded those of KINMATRIX. However, cross-national patterns appeared broadly similar across data sources for all kinship categories except grandparents. Higher-fertility countries with higher average numbers of aunts, uncles, and cousins in demographic projections – notably the United States, the Netherlands, and Norway – also tended to show higher numbers of these kin types in the KINMATRIX data. This suggests that while KINMATRIX data differed from demographic projections in terms of absolute numbers, the estimates broadly aligned in terms of relative differences between countries in family network size.

Figure 3: Cross-national differences in family network size (mean number of kin alive)



Note: KINMATRIX estimates on y-axis, DemoKin-based demographic projections on x-axis. DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States. Details of variable coding can be found in 'S1. Recoding scheme' in the online supplement.
Source: Own calculation and illustration.

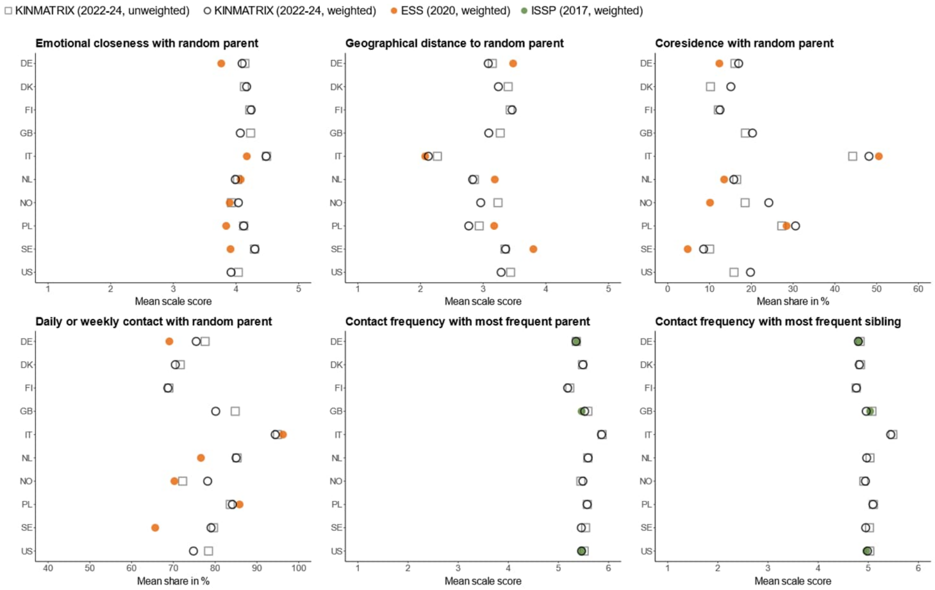
4.3 Family relationships

In a second step, we examined three dimensions of family relationships: emotional closeness, geographical distance (including coresidence), and contact frequency. In the top row and bottom left plot of Figure 4, we show the results obtained from comparing KINMATRIX to ESS data (orange-filled circles) on emotional closeness, geographical distance, coresidence, and weekly contact with one randomly selected parent. In survey countries where sufficient reference data were available, both weighted and unweighted KINMATRIX estimates aligned well with ESS data on emotional closeness and geographical distance, and on coresidence between respondents and their parents (for smaller country samples, see Figure A-3 in the Appendix). Larger discrepancies between KINMATRIX and ESS emerged for contact frequency with parents, although estimates for Italy and Poland continued to show close alignment. Overall, KINMATRIX data reflected the well-established contrast between weaker and stronger family regimes, characterized by a north-south gradient in coresidence, geographic distance, and contact frequency (Albertini and Kohli 2013; Hank 2007; Mönkediek and Bras 2014; Reher 1998, 2004).

The green reference markers in the bottom row of Figure 4 represent ISSP data on the average frequency of contact between respondents and their most frequently contacted parent (middle plot) or sibling (right-hand plot), indicators that allowed us to add Great Britain and the United States and offered a higher degree of data harmonization with KINMATRIX than those from the ESS. Both plots, although constrained in their cross-national variance by the definition of the indicators, show a close match between KINMATRIX and ISSP data across countries.

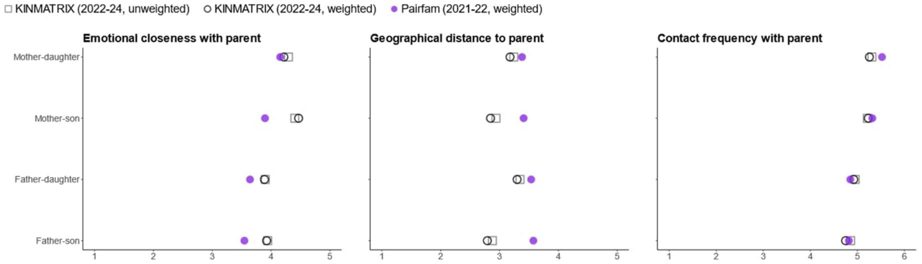
In Figure 5 we present bivariate relationships, comparing emotional closeness, geographical distance, and contact frequency for each of the four gender constellations of parent–child relationships. Due to data availability, this analysis was limited to a comparison between the German KINMATRIX sample and the German pairfam data, which provided a high degree of data harmonization. Similar to ISSP data, KINMATRIX data closely matched pairfam data in the average contact frequency for each gender constellation (right-hand plot). Moreover, the KINMATRIX data reproduced the gradient in contact frequency found in the comparison data, whereby contact frequency was highest in mother–daughter ties, followed by mother–son ties, father–daughter ties, and father–son ties (Fingerman, Huo, and Birditt 2020).

Figure 4: Family relationships



Note: DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States. Details on variable coding can be found in 'S1. Recoding scheme' in the online supplement. Country-specific reference data based on fewer than 200 respondents are excluded (for complete reference estimates, see Figure A-3 in the Appendix).
Source: Own calculation and illustration.

Figure 5: Gender-specific adult child–parent relationships in Germany



Note: Means on Likert scales (5-point or 6-point) are shown. Details on variable coding can be found in 'S1. Recoding scheme' in the online supplement.
Source: Own calculation and illustration.

The same gradient was still present for emotional closeness (left-hand plot) in the pairfam data, but not in KINMATRIX. Instead, KINMATRIX appeared to be selective towards male respondents who had closer ties to their parents. Regarding geographical distance (middle plot), KINMATRIX data remained below pairfam values for all gender constellations of parent–child dyads, while the fit for parent–daughter dyads was closer than for parent–son dyads. Weighting did not increase the fit between KINMATRIX and the pairfam data.

4.4 Family complexity

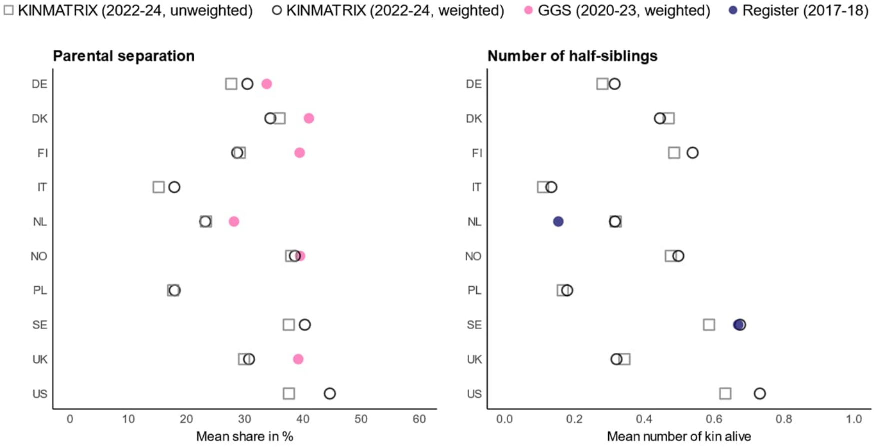
In the third and final step, we focused on family complexity. In the left-hand plot of Figure 6 we show the mean share of individuals whose living parents had separated between the KINMATRIX data and probability-based survey data from the GGS. Looking at absolute differences, both weighted and unweighted KINMATRIX data yielded lower estimates for parental separation than the comparison data. However, the country rankings aligned with known cross-national differences in union (in)stability. For Europe, KINMATRIX estimates reflected the well-established north–south gradient of higher shares of separated parents in the Nordic countries, medium shares in Central Europe, and the lowest shares in Southern Europe (Mönkediek and Bras 2014; Reher 1998, 2004).

Similarly, the high prevalence of divorce and separation in the United States (Adler 2023) was reflected in the KINMATRIX data. Beyond this overall pattern in the United States, the KINMATRIX data captured known ethnic/racial differences, whereby shares of younger adults who experienced parental separation were substantially larger in the sample of Black than of White Americans (estimates not shown).

In the right-hand plot of Figure 6, we illustrate family complexity in terms of the mean number of living half-siblings included in younger adults' kinship networks. Given the limited probability-based survey data available on half-siblings, this comparison focused on the two countries for which register-based estimates of mean numbers of half-siblings have recently been published, the Netherlands and Sweden (De Bel et al. 2024; Kolk et al. 2023).

Although the KINMATRIX samples for the Netherlands and Sweden were relatively small, results for the mean number of half-siblings aligned relatively well with the register-based estimates, especially in Sweden, where the weighted KINMATRIX estimates matched the register-based estimates. Moreover, the cross-national difference in family complexity between the Netherlands (lower) and Sweden (higher), measured by the number of half-siblings, was reflected in the KINMATRIX data, although the magnitude of the contrast was smaller, especially in the unweighted data. Beyond the comparison to register data, the KINMATRIX data reproduced known patterns such as the high degree of family complexity in the United States and lower degrees in Poland and Italy (Adler 2023; Benson 2020).

Figure 6: Family complexity



Note: DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States. Details of variable coding can be found in ‘S1. Recoding scheme’ in the online supplement.
Source: Own calculation and illustration.

5. Discussion

In this study, we aimed to evaluate newly collected web-based nonprobability data from the KINMATRIX survey by comparing it with different sources of reference data, including demographic projections from aggregate official data, estimates from national registers, and estimates from various probability-based surveys. Because the reference data differed from KINMATRIX in many ways beyond sampling and were obtained from diverse sources with diverse estimation methods, and because we evaluated only one nonprobability dataset, our findings on deviation and alignment are merely suggestive, and the evidence presented is only an initial step toward understanding what nonprobability data can add to family demography and quantitative family research more generally.

The results illustrate both the promises and pitfalls of using nonprobability data. In each of the three domains we investigated – family network size, family relationships, and family complexity – we found instances of near or full alignment between KINMATRIX and the comparison data. Yet these instances were accompanied by several deviations, some of which were large.

A current research frontier examines the size of family networks (Alburez-Gutierrez, Williams, and Caswell 2023; De Bel et al. 2024; Feng, Song, and Caswell 2025; Kolk et al. 2023; Leopold et al. 2025b; Leopold and Kalmijn 2026). In this domain, KINMATRIX estimates tend to be lower than demographic projections based on official data, with larger deviations for collateral kin (siblings and cousins), smaller deviations for aunts and uncles, and no deviations for the number of living mothers and fathers. Cross-national differences in the number of parents, siblings, aunts, uncles, and cousins are broadly consistent across sources, whereas patterns differ for the number of living grandparents.

Because KINMATRIX was designed to enumerate all biological kin in the listed categories, lower counts should be treated as underestimates when the quantity of interest is the demographic presence of kin. At the same time, omissions in a survey setting are likely to reflect not only genealogical distance but also social distance to kin. If respondents are not aware of, or cannot reliably enumerate, all of their aunts, uncles, and cousins, those who are reported will tend to be the more socially salient ones, even though social salience was not a criterion in the name generators used. The discrepancies observed in this study are therefore most consequential for demographic accounting of how many kin exist. However, for questions about how kin matter in people's lives, survey-based counts may be more informative because they tend to omit kin who exist 'on paper' but play no role in people's lives. For example, when the aim is to understand innermost family circles, support networks, or potential safety nets, kin reported in a survey may be a more pertinent measure than aggregate-data projections of kin who are demographically present but may be socially absent.

Overall, choosing and validating any approach to estimating kin numbers – projections based on aggregate data, register-based assessments based on biological and household linkages, or survey-based assessments based on name generators – remains challenging, as each captures different substantive concepts and is subject to different sources of error (coverage, linkage, recall, etc.). Because true benchmark values are often not available or hard to define, comparisons across sources should be interpreted as triangulation rather than adjudication, informing where and why estimates diverge, and clarifying which measure is best suited to a substantive construct of interest.

In the domain of family relationships, we found that the correspondence between KINMATRIX and probability-based estimates varies across indicators, but established north–south contrasts corresponding to 'weak' and 'strong' family regimes (Reher 1998) were broadly retained. We note that deviations tended to be larger in comparisons with the ESS, for which harmonization was limited, and smaller in comparisons with the

German family panel, for which harmonization was near complete. Our comparison with pairfam on parent–child gender constellations did not indicate that bivariate measures aligned better with comparison data than univariate measures.

Finally, regarding family complexity, measures of parental separation and the number of half-siblings captured known cross-national patterns but, in some instances, still deviated substantially from probability-based estimates. Notably, parental separation was consistently less prevalent in KINMATRIX than in probability-based data from the GGS. At the same time, probability-based surveys themselves exhibit substantial variation in estimates of complex kin constellations. Kuhnt and Steinbach (2014), for example, document that estimates of stepfamily prevalence in Germany range from 7% to 17% across different probability-based data sources. This suggests that measuring complex family structures poses methodological challenges regardless of sampling approach.

What do these results mean for the tension between high data-quality requirements and the desire for timely and cost-effective designs? As our study evaluated only a single nonprobability dataset and cannot be generalized to other nonprobability designs, additional assessments across datasets and designs are needed to situate our results within a larger evidence base. Especially desirable are controlled comparison studies that allow attributing deviations to the sampling method while holding other potential sources for discrepancies constant. In our study, several factors beyond nonprobability sampling may have contributed to the differences observed, including limited harmonization of constructs, response bias in a context of high task difficulty and respondent burden, and small sample sizes in the Nordic countries. A criterion for future research is defining clear benchmark values against which estimates from different data sources can be compared.

Although any interpretations of our findings are merely suggestive in light of the limitations of our design, they do indicate that nonprobability data cannot be generally dismissed as incompatible with the high standards of data quality in the population sciences. For several indicators of family structure and family relationships, we demonstrated that such data can align closely with widely accepted comparison data and reproduce well-established patterns across countries. This potential is particularly relevant in innovative areas within the field of family demography, where existing probability-based surveys may not fully accommodate the needs of analysts.

At the same time, caution is warranted when drawing population-level conclusions from nonprobability surveys such as KINMATRIX. Given the discrepancies observed across indicators and countries, close scrutiny and a case-by-case evaluation are necessary when collecting or analyzing nonprobability data in family demography. We recommend that researchers using KINMATRIX conduct study-specific benchmarking against high-quality external data whenever feasible. Innovative data sources such as KINMATRIX typically have both sweet spots and weak spots, and their identification requires targeted validation – a type of additional analysis not routinely required when working with established data sources but crucial for interpreting results from nonprobability data. Examples of this can be found in two recent studies using the United States sample of KINMATRIX (Leopold and Kalmijn 2026, Sytkina and Leopold 2026). In these studies, benchmarking against the American Community Survey (ACS) and comparison with the General Social Survey (GSS) and the National Survey of Families and Households (NSFH) yielded encouraging results: KINMATRIX aligned well with several ACS benchmarks, Black–White differences in sibship size corresponded closely with the GSS, and differences by parental separation in mother–child and father–child ties were similar in direction and magnitude to those in the NSFH. Yet, as illustrated in the present study, any such sweet spots in nonprobability data may be accompanied by weak spots. A careful and balanced approach is therefore needed to make full and responsible use of such data.

In conclusion, our study suggests that family demography can benefit from incorporating nonprobability approaches. A fuller evaluation of this potential requires subsequent studies to advance our understanding of the trade-offs involved, refine best practices, and systematically validate nonprobability data in the population sciences.

6. Acknowledgements

For KINMATRIX data access and documentation, see www.kinmatrix.eu. This project received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 848861) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC 2126/2 – 390838866.

References

Adler, M.A. (2023). The complexities of family diversity in the contemporary US. In: Adler, M.A. and Lenz, K. (eds.). *The changing faces of families*. Routledge: 187–212. doi:10.4324/9781003193500-10.

Albertini, M. and Kohli, M. (2013). The generational contract in the family: An analysis of transfer regimes in Europe. *European Sociological Review* 29(4): 828–840. doi:10.1093/esr/jcs061.

Alburez-Gutierrez, D., Williams, I., and Caswell, H. (2023). Projections of human kinship for all countries. *Proceedings of the National Academy of Sciences* 120(52): e2315722120. doi:10.1073/pnas.2315722120.

Ansolabehere, S. and Rivers, D. (2013). Cooperative survey research. *Annual Review of Political Science* 16(1): 307–329. doi:10.1146/annurev-polisci-022811-160625.

Ansolabehere, S. and Schaffner, B.F. (2014). Does survey mode still matter? Findings from a 2010 multi-mode comparison. *Political Analysis* 22(3): 285–303. doi:10.1093/pan/mpt025.

Arndt, A.D., Ford, J.B., Babin, B.J., and Luong, V. (2022). Collecting samples from online services: How to use screeners to improve data quality. *International Journal of Research in Marketing* 39(1): 117–133. doi:10.1016/j.ijresmar.2021.05.001.

Baker, R., Brick, J.M., Bates, N.A., Battaglia, M., Couper, M.P., Dever, J.A., Gile, K.J., and Tourangeau, R. (2013). Summary report of the AAPOR task force on non-probability sampling. *Journal of Survey Statistics and Methodology* 1(2): 90–143. doi:10.1093/jssam/smt008.

Batrinca, B. and Treleaven, P.C. (2015). Social media analytics: A survey of techniques, tools and platforms. *AI and Society* 30(1): 89–116. doi:10.1007/s00146-014-0549-4.

Bengtson, V.L. and Roberts, R.E.L. (1991). Intergenerational solidarity in aging families: An example of formal theory construction. *Journal of Marriage and the Family* 53(4): 856–870. doi:10.2307/352993.

Benson, H. (2020). Divorce rates are falling across Europe and the UK is leading the way. Romford: Marriage Foundation. https://marriagefoundation.org.uk/wp-content/uploads/2020/01/MF-analysis-Divorces-rates-falling-across-Europe-1.pdf.

Bergmann, M. (2011). IPFWEIGHT: Stata module to create adjustment weights for surveys [Computer software]. (Statistical Software Components S457353). Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s457353.html.

Bishop, C.D., Leite, W.L., and Snyder, P.A. (2018). Using propensity score weighting to reduce selection bias in large-scale data sets. *Journal of Early Intervention* 40(4): 347–362. doi:10.1177/1053815118793430.

Bornstein, M.H., Jager, J., and Putnick, D.L. (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. *Developmental Review* 33(4): 357–370. doi:10.1016/j.dr.2013.08.003.

Buhrmester, M., Kwang, T., and Gosling, S.D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science* 6(1): 3–5. doi:10.1177/1745691610393980.

Callegaro, M., Baker, R., Bethlehem, J., Göritz, A.S., Krosnick, J.A., and Lavrakas, P.J. (2014). *Online panel research: Data quality perspective* (1st ed.). Wiley. doi:10.1002/9781118763520.

Campisi, N., Kulu, H., Mikolai, J., Klüsener, S., and Myrskylä, M. (2020). Spatial variation in fertility across Europe: Patterns and determinants. *Population, Space and Place* 26(4): e2308. doi:10.1002/psp.2308.

Castrén, A.-M., Česnuitytė, V., Crespi, I., Gauthier, J.-A., Gouveia, R., Martin, C., Moreno Mínguez, A., and Suwada, K. (2021). *The Palgrave handbook of family sociology in Europe* (corrected publication). Palgrave Macmillan. doi:10.1007/978-3-030-73306-3.

Caswell, H. (2019). The formal demography of kinship: A matrix formulation. *Demographic Research* 41(24): 679–712. doi:10.4054/DemRes.2019.41.24.

Caswell, H. (2020). The formal demography of kinship II: Multistate models, parity, and sibship. *Demographic Research* 42(38): 1097–1146. doi:10.4054/DemRes.2020.42.38.

Caswell, H. (2022). The formal demography of kinship IV: Two-sex models and their approximations. *Demographic Research* 47(13): 359–396. doi:10.4054/DemRes.2022.47.13.

Caswell, H., and Song, X. (2021). The formal demography of kinship III: Kinship dynamics with time-varying demographic rates. *Demographic Research* 45(16): 517–546. doi:10.4054/DemRes.2021.45.16.

Chen, Y., Li, P., and Wu, C. (2020). Doubly robust inference with nonprobability survey samples. *Journal of the American Statistical Association* 115(532): 2011–2021. doi:10.1080/01621459.2019.1677241.

Chmielewski, M. and Kucker, S.C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science* 11(4): 464–473. doi:10.1177/1948550619875149.

Chong, M., Alburez-Gutierrez, D., Alexander, M.J., and Zagheni, E. (2022). Correcting historical mortality rate bias in big crowd-sourced online genealogies. (Max Planck Institute for Demographic Research Working Paper WP-2022-005). Rostock: MPIDR. doi:10.4054/MPIDR-WP-2022-005.

Cochran, W.G. (1977). *Sampling techniques* (3d ed). Wiley.

Colasurdo, A. and Omenti, R. (2024). Using online genealogical data for demographic research: An empirical examination of the FamiLinx database. *Demographic Research* 51(41): 1299–1350. doi:10.4054/DemRes.2024.51.41.

De Bel, V., Bokányi, E., Hank, K., and Leopold, T. (2024). A parallel kinship universe? Using Dutch kinship network data to replicate Kolk et al.'s (2023) demographic account of kinship networks in Sweden. OSF preprint. doi:10.31235/osf.io/3k6nq.

Dudel, C. and Klüsener, S. (2021). Male–female fertility differentials across 17 high-income countries: Insights from a new data resource. *European Journal of Population* 37(2): 417–441. doi:10.1007/s10680-020-09575-9.

Elliott, M. R. and Valliant, R. (2017). Inference for nonprobability samples. *Statistical Science* 32(2): 249–264. doi:10.1214/16-STS598.

Erens, B., Burkill, S., Couper, M.P., Conrad, F., Clifton, S., Tanton, C., Phelps, A., Datta, J., Mercer, C.H., Sonnenberg, P., Prah, P., Mitchell, K.R., Wellings, K., Johnson, A.M., and Copas, A.J. (2014). Nonprobability web surveys to measure sexual behaviors and attitudes in the general population: A comparison with a probability sample interview survey. *Journal of Medical Internet Research* 16(12): e276. doi:10.2196/jmir.3382.

Feng, K., Song, X., and Caswell, H. (2025). The present and future dementia burden in China: Kinship-based projections and global comparison. *Demography* 62(6): 1917–1944. doi:10.1215/00703370-12345648.

Fingerman, K.L., Huo, M., and Birditt, K.S. (2020). Mothers, fathers, daughters, and sons: Gender differences in adults' intergenerational ties. *Journal of Family Issues* 41(9): 1597–1625. doi:10.1177/0192513X19894369.

Freese, J. and Jin, O. (2025). Online nonprobability samples. *Annual Review of Sociology* 51: 109–128. doi:10.1146/annurev-soc-090524-043117.

Gil-Clavel, S. and Mulder, C.H. (2024). Does Twitter data mirror the European North–South family ties divide? A comparative analysis of tweets about family. *Population Research and Policy Review* 43(4): 48. doi:10.1007/s11113-024-09891-6.

Gittelman, S.H., Thomas, R.K., Lavrakas, P.J., and Lange, V. (2015). Quota controls in survey research: A test of accuracy and intersource reliability in online samples. *Journal of Advertising Research* 55(4): 368–379. doi:10.2501/JAR-2015-020.

Hank, K. (2007). Proximity and contacts between older parents and their children: A European comparison. *Journal of Marriage and Family* 69(1): 157–173. doi:10.1111/j.1741-3737.2006.00351.x.

Hays, R.D., Liu, H., and Kapteyn, A. (2015). Use of internet panels to conduct surveys. *Behavior Research Methods* 47(3): 685–690. doi:10.3758/s13428-015-0617-9.

Herbell, K. and Zauszniewski, J.A. (2018). Facebook or Twitter? Effective recruitment strategies for family caregivers. *Applied Nursing Research* 41: 1–4. doi:10.1016/j.apnr.2018.02.004.

Hünteler, B., Hank, K., Alburez-Guiterrez, D., and Leopold, T. (2026). A typology of younger adults’ nuclear and extended family relations in the United States. *Journal of Marriage and Family* 88(1): 284–298. doi:10.1111/jomf.70014.

Jerit, J. and Barabas, J. (2023). Are nonprobability surveys fit for purpose? *Public Opinion Quarterly* 87(3): 816–840. doi:10.1093/poq/nfad037.

Kaplanis, J., Gordon, A., Shor, T., Weissbrod, O., Geiger, D., Wahl, M., Gershovits, M., Markus, B., Sheikh, M., Gymrek, M., Bhatia, G., MacArthur, D.G., Price, A.L., and Erlich, Y. (2018). Quantitative analysis of population-scale family trees with millions of relatives. *Science* 360(6385): 171–175. doi:10.1126/science.aam9309.

Kennedy, C., Mercer, A., Keeter, S., Hatley, N., McGeeney, K., and Gimenez, A. (2016). *Evaluating online nonprobability surveys. Vendor choice matters; widespread errors found for estimates based on blacks and Hispanics.* (Numbers, Facts and Trends Shaping the World). Pew Research Center. https://www.pewresearch.org/methods/2016/05/02/evaluating-online-nonprobability-surveys/.

Kocar, S. and Baffour, B. (2023). Comparing and improving the accuracy of nonprobability samples: Profiling Australian surveys. *Methods*, *Data*, *Analyses* 17(2): 36. doi:10.12758/mda.2023.04.

Kohut, A., Keeter, S., Doherty, C., Dimock, M. and Christian, L. (2012). Assessing the representativeness of public opinion surveys. Pew Research Center. https://law.osu.edu/electionlaw/litigation/documents/Veasey7885.pdf.

Kolk, M., Andersson, L., Pettersson, E., and Drefahl, S. (2023). The Swedish kinship universe: A demographic account of the number of children, parents, siblings, grandchildren, grandparents, aunts/uncles, nieces/nephews, and cousins using national population registers. *Demography* 60(5): 1359–1385. doi:10.1215/00703370-10955240.

Kuhnt, A.-K. and Steinbach, A. (2014). Diversität von Familie in Deutschland. In: Steinbach, A., Hennig, M., and Arránz Becker, O. (eds.). *Familie im Fokus der Wissenschaft.* Wiesbaden: Springer Fachmedien: 41–70. doi:10.1007/978-3-658-02895-4_3.

Lavrakas, P.J., Pennay, D., Neiger, D., and Phillips, B. (2022). Comparing probability-based surveys and nonprobability online panel surveys in Australia: A total survey error perspective. *Survey Research Methods* 16(2): 241–266. doi:10.18148/srm/2022.v16i2.7907.

Lee, S. and Valliant, R. (2009). Estimation for volunteer panel web surveys using propensity score adjustment and calibration adjustment. *Sociological Methods and Research* 37(3): 319–343. doi:10.1177/0049124108329643.

Lehdonvirta, V., Oksanen, A., Räsänen, P., and Blank, G. (2021). Social media, web, and panel surveys: Using non-probability samples in social and policy research. *Policy and Internet* 13(1): 134–155. doi:10.1002/poi3.238.

Leopold, T., Becker, C.C., Buyukkececi, Z., Çineli, B., and Raab, M. (2024). KINMATRIX: A new data resource for studies of families and kinship. *Demographic Research* 51(25): 789–808. doi:10.4054/DemRes.2024.51.25.

Leopold, T., Becker, C.C., Buyukkececi, Z., Çineli, B., and Raab, M. (2025a). *KINMATRIX – Uncovering the kinship matrix* (Version 1.0.0) [Data set]. GESIS. doi:10.4232/1.14380.

Leopold, T., Raab, M., Becker, C.C., Buyukkececi, Z., and Çineli, B. (2025b). Mapping modern kinship networks: First results from the KINMATRIX survey. *Journal of Marriage and Family* 87(2): 478–504. doi:10.1111/jomf.13049.

Leopold, T. and Kalmijn, M. (2026). Kinship in Black and White families: The strength of horizontal ties. *American Journal of Sociology* (forthcoming).

Levy, P.S. and Lemeshow, S. (2008). *Sampling of populations: Methods and applications* (1st ed.). Wiley. doi:10.1002/9780470374597.

Lindner, J.R., Murphy, T.H., and Briers, G.E. (2001). Handling nonresponse in social research. *Journal of Agricultural Education* 42(4): 43–53. doi:10.5032/jae.2001.04043.

MacInnis, B., Krosnick, J.A., Ho, A.S., and Cho, M.-J. (2018). The accuracy of measurements with probability and nonprobability survey samples: Replication and extension. *Public Opinion Quarterly* 82(4): 707–744. doi:10.1093/poq/nfy038.

Mercer, A.W., Kreuter, F., Keeter, S., and Stuart, E.A. (2017). Theory and practice in nonprobability surveys. *Public Opinion Quarterly* 81(S1): 250–271. doi:10.1093/poq/nfw060.

Mocho, H., Martins, C., Dos Santos, R., and Nunes, C. (2024). Parental involvement and stress in children's quality of life: A longitudinal study with Portuguese parents during the COVID-19 pandemic period. *Children* 11(4): 440. doi:10.3390/children11040440.

Mondal, M., Punuru, A., Cheng, T.-W.S., Vargas, K., Gundry, C., Driggs, N.S., Schill, N., Carlson, N., Bedwell, J., Lorenc, J.Q., Ghosh, I., Li, Y., Fulda, N., and Page, X. (2023). A tale of two cultures: Comparing interpersonal information disclosure norms on Twitter. *Proceedings of the ACM on Human-Computer Interaction* 7(CSCW2): 1–40. doi:10.1145/3610045.

Mönkediek, B. and Bras, H. (2014). Strong and weak family ties revisited: Reconsidering European family structures from a network perspective. *The History of the Family* 19(2): 235–259. doi:10.1080/1081602X.2014.897246.

Murray-Watters, A., Zins, S., Silber, H., Gummer, T., and Lechner, C.M. (2023). River sampling – a fishing expedition: A non-probability case study. *Methods, Data, Analyses* 17(1): 3–28. doi:10.12758/mda.2022.05.

Nandram, B., Choi, J.W., and Liu, Y. (2021). Integration of nonprobability and probability samples via survey weights. *International Journal of Statistics and Probability* 10(6): 5. doi:10.5539/ijsp.v10n6p5.

Paolacci, G. and Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science* 23(3): 184–188. doi:10.1177/0963721414531598.

Pasek, J. (2016). When will nonprobability surveys mirror probability surveys? Considering types of inference and weighting strategies as criteria for correspondence. *International Journal of Public Opinion Research* 28(2): 269–291. doi:10.1093/ijpor/edv016.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., and Damer, E. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods* 54(4): 1643–1662. doi:10.3758/s13428-021-01694-3.

Pesando, L.M., Dorélien, A., St-Denis, X., and Santos, A. (2023). Demography as a field: Where we came from and where we are headed. *Canadian Studies in Population* 50(4). doi:10.1007/s42650-023-00076-8.

Pruchno, R.A., Brill, J.E., Shands, Y., Gordon, J.R., Genderson, M.W., Rose, M., and Cartwright, F. (2008). Convenience samples and caregiving research: How generalizable are the findings? *The Gerontologist* 48(6): 820–827. doi:10.1093/geront/48.6.820.

Reher, D.S. (1998). Family ties in Western Europe: Persistent contrasts. *Population and Development Review* 24(2): 203–234. doi:10.2307/2807972.

Reher, D.S. (2004). The demographic transition revisited as a global process. *Population, Space and Place* 10(1): 19–41. doi:10.1002/psp.313.

Rohr, B., Silber, H., and Felderer, B. (2025). Comparing the accuracy of univariate, bivariate, and multivariate estimates across probability and non-probability surveys with population benchmarks. *Sociological Methodology* 55(1): 121–154. doi:10.1177/00811750241280963.

Rosenbaum, P.R. and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41–55. doi:10.1093/biomet/70.1.41.

Salvatore, C. (2023). Inference with non-probability samples and survey data integration: A science mapping study. *METRON* 81(1): 83–107. doi:10.1007/s40300-023-00243-6.

Sivak, E. and Smirnov, I. (2018). *Gender bias in sharenting: Both men and women mention sons more often than daughters on social media*. arXiv preprint. doi:10.48550/arXiv.1803.10329.

Smith, M.B., Blakemore, J.K., Ho, J.R., and Grifo, J.A. (2021). Making it (net)work: A social network analysis of “fertility” in Twitter before and during the COVID-19 pandemic. *F&S Reports* 2(4): 472–478. doi:10.1016/j.xfre.2021.08.005.

Smith, T.W. (1995). Trends in non-response rates. *International Journal of Public Opinion Research* 7(2): 157–171. doi:10.1093/ijpor/7.2.157.

Stelter, R. and Alburez-Gutierrez, D. (2022). Representativeness is crucial for inferring demographic processes from online genealogies: Evidence from lifespan dynamics. *Proceedings of the National Academy of Sciences* 119(10): e2120455119. doi:10.1073/pnas.2120455119.

Sytkina, K., and Leopold, T. (2026). Parental separation and the quality of ties to extended kin. *Demography* (forthcoming).

Tofallis, C. (2015). A better measure of relative prediction accuracy for model selection and model estimation. *Journal of the Operational Research Society* 66(8): 1352–1362. doi:10.1057/jors.2014.103.

Tourangeau, R., Michael Brick, J., Lohr, S., and Li, J. (2017). Adaptive and responsive survey designs: A review and assessment. *Journal of the Royal Statistical Society Series A: Statistics in Society* 180(1): 203–223. doi:10.1111/rssa.12186.

United Nations (2024). World population prospects 2024: Summary of results. (UN DESA/POP/2024/TR/NO. 9). United Nations. https://population.un.org/wpp/assets/Files/WPP2024_Summary-of-Results.pdf.

Vehovar, V., Toepoel, V., and Steinmetz, S. (2016). Non-probability sampling. In: Wolf, C., Joye, D., Smith, T., and Fu, Y. (eds.). *The SAGE handbook of survey methodology*. Newcastle: SAGE: 329–345. doi:10.4135/9781473957893.n22.

Vicente, P. (2023). Sampling Twitter users for social science research: Evidence from a systematic review of the literature. *Quality and Quantity* 57(6): 5449–5489. doi:10.1007/s11135-023-01615-w.

Williams, I. and Alburez-Gutierrez, D. (2025). DemoKin: Estimate population kin distribution [Computer software]. GitHub. https://github.com/IvanWilli/DemoKin.

Williams, I., Alburez-Gutierrez, D., Song, X., and Caswell, H. (2021). DemoKin: An R package to implement demographic matrix kinship models. [Computer software]. CRAN. https://cran.r-project.org/package=DemoKin.

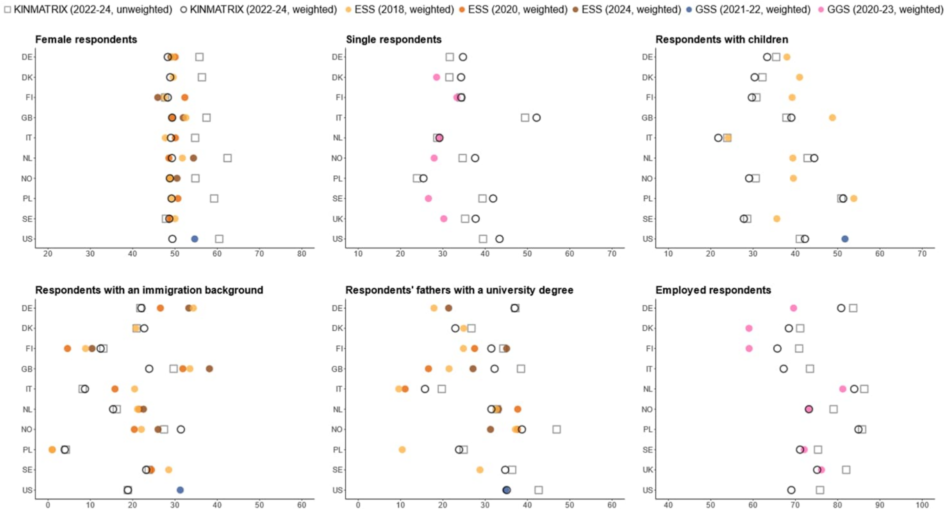
Wiśniowski, A., Sakshaug, J.W., Perez Ruiz, D.A., and Blom, A.G. (2020). Integrating probability and nonprobability samples for survey inference. *Journal of Survey Statistics and Methodology* 8(1): 120–147. doi:10.1093/jssam/smz051.

Yeager, D.S., Krosnick, J.A., Chang, L., Javitz, H.S., Levendusky, M.S., Simpser, A., and Wang, R. (2011). Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly* 75(4): 709–747. doi:10.1093/poq/nfr020.

Zack, E.S., Kennedy, J., and Long, J.S. (2019). Can nonprobability samples be used for social science research? A cautionary tale. *Survey Research Methods* 13(2): 215–227. doi:10.18148/SRM/2019.V13I2.7262.

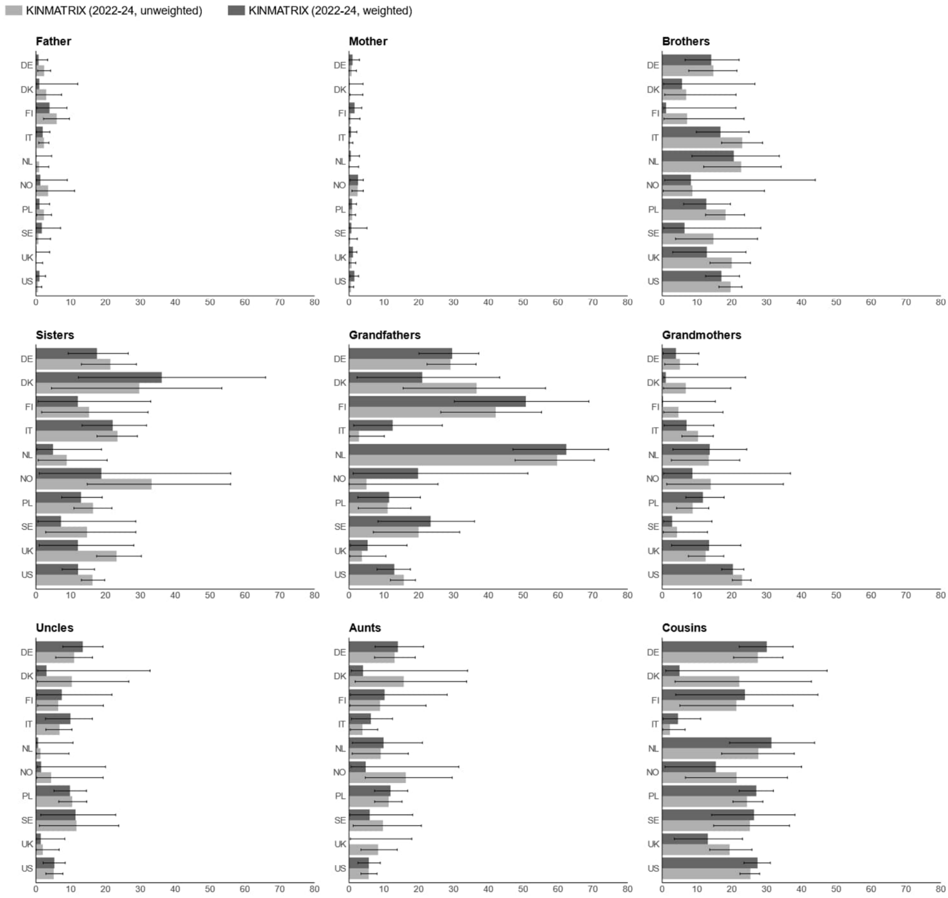
Appendix

Figure A-1: Sociodemographic characteristics (mean share in %), including small benchmark samples of n<200



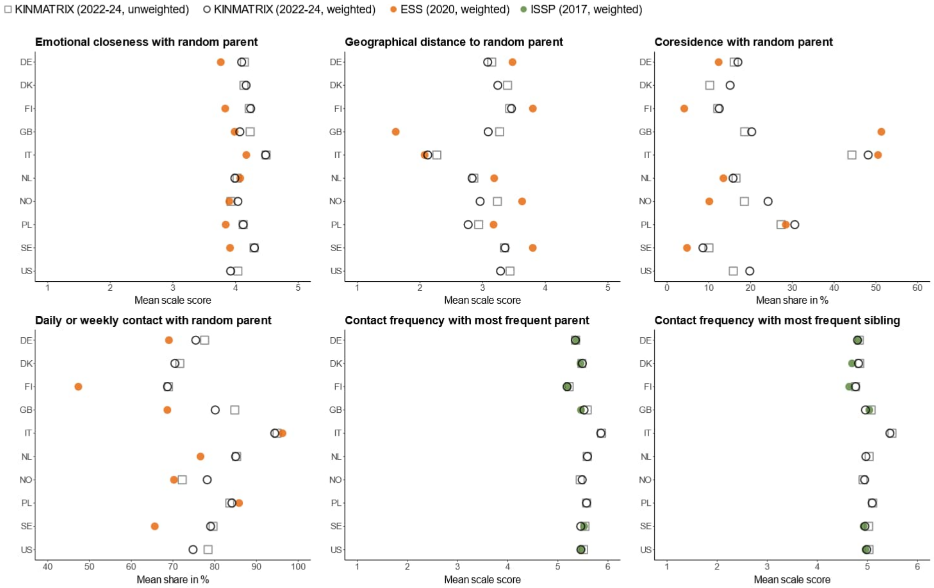
Note: DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States.
Source: Own calculation and illustration.

Figure A-2: sMAPE in % of the mean number of kin alive using DemoKin (2024) as reference



Note: The symmetric mean absolute percentage error (sMAPE) indicates the degree to which KINMATRIX estimates deviate from DemoKin estimates in absolute percentages (for the direction of deviations, see Figure 2). DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States.
Source: Own calculation and illustration.

Figure A-3: Family relationships, including small benchmark samples of n<200



Note: DE=Germany, DK=Denmark, FI=Finland, IT=Italy, NL=Netherlands, NO=Norway, PL=Poland, SE=Sweden, UK=United Kingdom/GB=Great Britain, US=United States.
Source: Own calculation and illustration.