ANNUAL REVIEWS

Annual Review of Sociology

Genealogical Microdata and Their Significance for Social Science

Xi Song[1] and Cameron D. Campbell[2]

[1]Department of Sociology, University of Chicago, Chicago, Illinois 60637; email: xisong@uchicago.edu

[2]Division of Social Science, Hong Kong University of Science and Technology, Hong Kong, China; email: camcam@ust.hk

Annu. Rev. Sociol. 2017. 43:75–99

First published as a Review in Advance on May 24, 2017

The *Annual Review of Sociology* is online at soc.annualreviews.org

https://doi.org/10.1146/annurev-soc-073014-112157

Keywords

genealogy, microdata, family, multigenerational research

Abstract

Despite long-standing recognition of the importance of family background in shaping life outcomes, only recently have empirical studies in demography, stratification, and other areas begun to consider the influence of kin other than parents. These new studies reflect the increasing availability of genealogical microdata that provide information about ancestors and kin over three or more generations. These data sets, including family genealogies, linked vital registration records, population registers, longitudinal surveys, and other sources, are valuable resources for social research on family, population, and stratification in a multigenerational perspective. This article reviews relevant recent studies, introduces and presents examples of the most important sources of genealogical microdata, identifies key methodological issues in the construction and analysis of genealogical data, and suggests directions for future research.

INTRODUCTION

The call by Mare (2011) to embrace a multigenerational perspective in studying such topics as stratification, population dynamics, and family organization identified novel directions for research on these topics. For inequality, already the subject of a very large literature (McCall & Percheski 2010, Neckerman & Torche 2007) that includes examinations of inequality over the very long term (Piketty 2014), new research has begun to consider multigenerational processes. Such studies extend the long-standing interest in parental influences on child outcomes in mobility studies by also considering the influences of grandparents and other more distant kin. A few depart completely from such traditional approaches by considering the overall influence of kin group membership or by considering outcomes such as long-term reproductive success instead of socioeconomic attainment. Theoretical and simulation-based investigations of the demographic dynamics of households, descent lines, and kin networks have recently been complemented by empirical studies that exploit multigenerational microdata.

The increasing availability of genealogical microdata has made these studies of multigenerational processes possible. By genealogical microdata, we mean (*a*) data that link individuals not only to their parents but also to more distant ancestors and living kin and (*b*) data that provide additional details on socioeconomic characteristics, household living arrangements, and/or demographic behavior. Such data enable stratification researchers to go beyond the traditional two-generation social mobility paradigm that focuses on parent-child associations, exploring micro/macro interactions between familial, institutional, and population changes over multiple generations. They allow family researchers to examine the role of kin networks outside the household in shaping social and demographic outcomes. They also allow demographic researchers to follow descent lines over multiple generations, identifying the social and economic factors that influence descent line growth, decline, or extinction. Some of these data sets are large, describing hundreds of thousands of individuals in past and present populations in Asia, North America, and Europe. Meanwhile, analysis of contemporary health and mortality data linked to genealogical microdata not only may facilitate traditional genetic studies but also may help advance understanding of gene-environment interactions.

Genealogical microdata constructed from diverse sources are now available for a variety of geographic settings and time periods. The most basic data include only a few key social and demographic characteristics for each individual. The richest microdata include longitudinal information about individuals over the courses of their lives and aggregate information about their communities and geographic areas. Historical data sets have been compiled from family genealogies, population registers, administrative certificates, church records, and census microdata. Contemporary data sets have been constructed from longitudinal survey data, linked administrative data, and in some cases, cross-sectional surveys that collect information about ancestry and kinship.

Our review briefly summarizes the major issues and debates in the study of multigenerational social, economic, and family processes; provides an overview of relevant genealogical microdata; discusses methodological issues; and concludes with suggestions for future research directions. By putting particular emphasis on sources of genealogical microdata, we seek to complement other recent overviews of multigenerational social and demographic processes (e.g., Mare 2011, 2014, 2015; Pfeffer 2014), which focus on laying out a research agenda, suggesting possible underlying mechanisms, and identifying methodological challenges. In contrast, our discussion of major issues and debates emphasizes recent empirical studies on multigenerational mobility and population processes that make use of the novel sources of genealogical microdata that we discuss. The following overview highlights data sets that are available for use or have distinctive properties that make them especially useful for specific questions. In discussing methodological issues, we consider not only methods for analyzing multigenerational data but also ways in which insights from their analysis may shape how we interpret results from other data sets.

MULTIGENERATIONAL FAMILY, SOCIAL, AND POPULATION PROCESSES

Studies of multigenerational family, social, and demographic processes have proliferated in recent years. Whereas interest in effects of grandparents on outcomes has been long-standing, recognition of the potential importance of other kin is more recent, as is the availability of suitable genealogical microdata. To help clarify the most important applications of the genealogical microdata presented below, we first introduce major issues and debates in the study of multigenerational processes. We begin with effects of characteristics of kin on social and economic attainment. We then present a set of studies that examine aspects of social mobility. We conclude with studies of the population dynamics of descent lines.

Kin Effects on Socioeconomic Attainment

Even though traditional studies of intergenerational mobility consider the effects of only parents' characteristics on attainment, the possibility that grandparents' characteristics may also have direct influences has been recognized for some time (Hodge 1966). Nevertheless, the implicit assumption in traditional approaches is that any influence of grandparents or more distant ancestors and kin is not direct but instead works entirely via their influence on intervening generations. If that assumption is correct, it is sufficient to consider parental characteristics and ignore other kin in studying the determinants of outcomes. Especially in societies outside Northwest Europe and North America, where grandparents play a more central role in the family, there are a variety of reasons to expect the characteristics of grandparents to matter; for example, grandparents may assist with childcare, educate grandchildren and shape their aspirations, or make bequests and inter vivos transfers that bypass parents. As Bengtson (2001) points out, understanding the role of grandparents in shaping life chances is, if anything, continuing to increase in importance because of population aging, changes in patterns of coresidence, and the growing diversity of kin relationships.

Empirical studies of grandparents' effects on socioeconomic attainment have been conducted for a variety of settings and time periods in East Asia, North America, and Europe. A number of studies test for evidence of direct effects of the characteristics of grandparents or even great-grandparents, independent of the characteristics of parents. One of the earliest studies was by Warren & Hauser (1997), who found no evidence of grandparents' effects in their analysis of data from the Wisconsin Longitudinal Study (WLS). This finding appeared to confirm the norm of noninterference suggested in the study of American grandparents by Cherlin & Furstenberg (1986). Many subsequent studies use contemporary, often nationally representative longitudinal data from the United Kingdom (Chan & Boliver 2013), the United States (Pfeffer & Killewald 2015, Wightman & Danziger 2014), Germany (Hertel & Groh-Samberg 2014), the Netherlands (Bol & Kalmijn 2016, Knigge 2016), mainland China (Zeng & Xie 2014), Denmark (Møllegaard & Jæger 2015), Finland (Erola & Moisio 2007), Sweden (Chaparro & Koupil 2014, Hällsten 2014, Modin et al. 2013), and Taiwan (Chiang & Park 2015). Other studies use historical data (Dribe & Helgertz 2017, Ferrie et al. 2016). The results of these studies are mixed. Associations vary across time, place, and dimension of socioeconomic status, indicating that grandparents' and great-grandparents' effects are heterogeneous and conditioned by the role of family organization.

Recent studies move beyond measurement of associations between grandparents and grandchildren to identify possible causal mechanisms that account for those associations. Sharkey &

Elwert (2011) address time-varying confounders involved in estimating grandparents' effects and show that grandparents' neighborhood experiences have lingering effects on grandchildren's educational outcomes. Solon (2014) presents a behavioral model in the tradition of Becker (1981) that makes a distinction between family endowment, such as genetic and cultural traits that a child inherited from parents, and family investment in a child's human capital development. The model, in contrast to reduced-form association models, takes into account family consumption behaviors and budget constraints in the multigenerational transmission of status. Overall, results on the causal effects of grandparents' characteristics are still limited. Existing descriptive analyses may be confounded by factors such as grandparents' ages (Silverstein & Marenco 2001), survival status (Zeng & Xie 2014), and geographic and relational closeness with grandchildren (King & Elder 1997), as well as family structure and living arrangement across generations (Dunifon & Kowaleski-Jones 2007, Song 2016). A closer look at the grandparents' effects also suggests that the roles of grandfathers and grandmothers and of paternal and maternal grandparents may differ (Kolk 2014, Modin et al. 2013, Olivetti et al. 2016). Grandmothers are often more involved in raising their grandchildren than are grandfathers. In patriarchal societies, paternal grandfathers may influence outcomes in other ways, even if they are not involved in childcare (Dong et al. 2016). Grandparents' effects on child well-being and survival are a major topic of study in the evolutionary demography literature and are already the subject of comprehensive literature reviews (Sear & Mace 2008).

More recent research considers kin other than grandparents, as well as overall differences in attainment chances according to kin group membership. Especially in regions outside North America and Northwest Europe, where the extended family was a key unit of social organization and people lived in large, complex households, distant kin were important, and clan or lineage membership played a salient role in structuring opportunity. In an analysis of Northeast China in the eighteenth and nineteenth centuries, Campbell & Lee (2003, 2008, 2011) found associations of attainment chances with characteristics of uncles, siblings, and cousins as well as lineage membership—a finding that confirms the importance of a broader network of kin, as suggested in a long line of anthropological studies of China (Cohen 1990, Freedman 1966). Even in North America, there is evidence that kin network membership influenced attainment chances: Jæger (2012) analyzed data from the WLS and found that even though associations of attainment chances with measured characteristics of specific individual kin are weak, there is nevertheless evidence of shared advantages or disadvantages in attainment chances across siblings and cousins.

Long-Term Social Immobility and Mobility

Stratification processes may be path dependent, in the sense that the characteristics of very distant ancestors may directly influence the outcomes of their descendants for many generations (Mare 2011, pp. 7–14). In the face of such processes, regression to the mean may take place slowly or not at all, leading to longer intergenerational persistence of status than expected from models based on parent-child associations. By contrast, in the first-order Markovian processes considered in the previous section, mathematical models and simulations demonstrate that anything other than total or near total immobility will eliminate influences from the distant past as generations succeed one another (Bartholomew 1967, Preston & Campbell 1993). As a result, the social and economic attainment of descendants of high- and low-status individuals will become indistinguishable relatively quickly.

Mechanisms underlying such path dependence may include, but are by no means limited to, the influences of great-grandparents and more distant ancestors that bypass intervening generations, trajectories of cumulative advantage or disadvantage that emerged in prior generations, and the influences of remote ancestors who experienced extreme success or adversity (Mare 2011). These effects may complement or substitute for those of parents and work through institutional mechanisms embedded in families, the law, educational institutions, and financial transactions (Coall & Hertwig 2010).

Empirical results from analysis of genealogical microdata that have the requisite generational span confirm the possibility that regression to the mean may be slower than implied by Markovian models based on parent-child associations. Campbell & Lee (2011) examined persistence in the relative status of lineages from the mid-eighteenth century to the end of the twentieth in genealogical microdata composed of historical household registers linked to a contemporary retrospective survey of a subset of contemporary descendants. They show that the rankings of descent groups according to educational attainment and attainment of official position in the late twentieth century was strongly correlated with rankings during the eighteenth and nineteenth centuries.

Results of indirect approaches based on other data also raise the possibility of persistence in family status well beyond that predicted by modeling or simulation based on two- or three-generation associations. When direct measures of socioeconomic statuses are unavailable, names may serve as a signal of ancestral economic success, ethnic origin, and migration history (Goldstein & Stecklov 2016, Olivetti & Paserman 2015). On the basis of the changes in the distribution of rare surnames initially associated with high status, Clark (2014) claims that mobility rates over several hundred or even a thousand years in medieval England, Qing dynasty China, Sweden, the United States, and many other places around the world were much lower than suggested by traditional studies of two-generation microdata. We discuss methodological issues in aggregate approaches below (see the section titled Methods and Models in the Absence of Microdata).

The long-term social mobility of families also interacts with families' marriage, reproduction, and mortality rates to create and sustain patterns of inequality across multiple generations. Whereas conventional mobility studies focus on the association of social status between two generations conditional on the existence of offspring, a more complete understanding of intergenerational influence treats successful reproduction as an integral part of the family transmission of status. In a multigenerational context, grandparents' influence on the composition of their grandchildren's generation is greater if they have more grandchildren, and less if they do not. These effects on composition may work through patterns of assortative mating, timing and levels of fertility, and the survival status of each generation (Breen & Ermisch 2016, Lawrence & Breen 2016, Mare & Maralani 2006, Matras 1961, Preston & Campbell 1993). Such mechanisms influence the composition of later generations via differential population growth of socioeconomic groups (Mare & Song 2014, Song & Mare 2017).

Descent Line Dynamics

Socioeconomic status in one generation influences not only the socioeconomic status of descendants but also their number. Clark (2007) showed that the descendants of socioeconomically advantaged men were overrepresented not only in the next generation but also several generations later. In other words, the descendants of high-status males account for a larger share of the population many generations later than the high-status males did in their own generation. Similar observations have long been the focus of formal demographers and population geneticists in their mathematical or simulation models that show the relationship between socioeconomic differentials in demographic behaviors and long-term population change (Bongaarts et al. 1987, Wachter et al. 1978, Wright 1929). Clark (2007) argues that this process could actually be a mechanism not only of demographic change but also of social change, if high-status males transmitted the behaviors, knowledge, and traits that made them successful. Cavalli-Sforza & Feldman (1981) examine more sophisticated formal models in which social change may be driven by various mechanisms of transmission, not only from parents to offspring but also between peers, between teachers and students, and by cultural and natural selection.

Empirical results from analysis of genealogical microdata highlight the potential for such data to help elucidate interactions among status, inheritance, reproduction, and survivorship in shaping population change. To unpack the micro-level mechanisms of such multigenerational processes, Song et al. (2015) examined social status influences on the number of patrilineal descendants in subsequent generations. Using data from historical China, they found that patrilineages founded by high-status males had higher growth rates for the next 150 years. Status effects had more to do with patrilines minimizing their chances of having no sons at all than with maximizing their numbers of sons in each generation. From an evolutionary genetics perspective, Goodman & Koupil (2009, 2010), relying on data from both patrilineal and matrilineal descent lines, have demonstrated similarities in their population dynamics and uncovered complex relationships among marriage, birth order, early-life biological characteristics, fertility, and socioeconomic status.

GENEALOGICAL MICRODATA

Here we introduce historical and contemporary genealogical microdata used for studies like the ones discussed in the section above. Such multigenerational data follow families for at least three and preferably more generations. They allow for the measurement of the characteristics of kin as well as aggregate characteristics of the kin group. Typically, these data include demographic behaviors such as fertility, mortality, marriage, and migration, as well as basic measures of socioeconomic status such as occupation. Certain sources also include living arrangements. The best sources include measures of individual characteristics at multiple points in time, as well as refined measures of socioeconomic status of income or wealth. Historical sources of genealogical microdata include family genealogies, household registers, censuses, and family reconstitutions based on church and vital records. Contemporary sources include a small number of longitudinal surveys that originally collected data on parents and grandparents and that have been in place long enough that they are now in their second or even third generation of respondents. We summarize major genealogical data infrastructure projects in **Table 1** and compare their characteristics in **Table 2**.

Family Genealogies

Genealogies were one of the first sources to be used in empirical studies of population dynamics and demographic behavior in Europe (Beeton et al. 1900, Dyke & Morrill 1980, Freeman 1935, Hollingsworth 1969, Knodel 2002, Wilson & Doering 1926, Winston 1932). They have also been used extensively as sources for the study of East Asian historical demography (Harrell & Pullum 1995; Liu 1978, 1981, 1985, 1995a,b; Telford 1990, 1992). Family genealogies typically present lineage members' kinship and pedigree information as lists, tables, charts, or written narratives. The amount of information contained in genealogies differs markedly across families. The most basic genealogies, generally unsuitable for analysis, provide little more than the names of family members and their pedigrees. For example, many Chinese family genealogies list only the male patrilineal descendants of a lineage founder. The most useful but least common genealogies include basic demographic information for lineage members, such as dates of birth, death, and marriage; the name(s) of their spouse(s); and in some cases their socioeconomic information.

Even the most detailed genealogies may be selective in terms of the families for whom they are available and the family members they record (Campbell & Lee 2002, Harrell 1987, Post et al. 1997, Willigan & Lynch 1982, Zhao 2001). Traditional genealogies were compiled retrospectively by surviving descendants and are therefore more likely to be available for families that were

Table 1 Major genealogical data infrastructure projects

Project abbreviation	Project name	Location	Access	Link or citation
Family genealogies				
CMGPD-IL	China Multi-Generational Panel Dataset–Imperial Lineage	Beijing and Shenyang, China	Proprietary	Lee et al. (1993)
UPDB	Utah Population Database	United States	Application	**http://healthcare.utah.edu/huntsmancancerinstitute/research/updb/access.php**
Synthetic genealogies				
Population registers				
CMGPD-LN	China Multi-Generational Panel Dataset–Liaoning	Northeast China	Download	**http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/27063**
CMGPD-SC	China Multi-Generational Panel Dataset–Shuangcheng	Northeast China	Download	**http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/35292**
KMGPD-TS	Korean Multi-Generational Panel Dataset–Tansung	Tansung, Korea	Application	**http://repository.ust.hk/ir/Record/1783.1-74215**
CTHRD	Colonial Taiwan Household Registration Database	Taiwan, China	Application	**http://www.demography.sinica.edu.tw/EN/en_background.htm**
NAC-SN	Japanese Ninbetsu-Aratame-Cho Population Register Database–Shimomoriya and Niita	Shimomoriya and Niita, Japan	Proprietary	**http://www.fl.reitaku-u.ac.jp/pfhp/index-e.html**
Parish registers and vital records				
LINKS/GENLIAS	LINKing System for historical family reconstruction	Netherlands	Application	**https://socialhistory.org/en/hsn/linking-system-historical-family-reconstruction-links**
HPR	Historical Population Registers	Norway	Application	**http://www.rhd.uit.no/nhdc/hpr.html**
UBCoS Multigen	Uppsala Birth Cohort Multigenerational Study	Sweden	Application	**http://www.chess.su.se/ubcosmg/**
SEDD	Scanian Economic Demographic Database	Sweden	Application	**http://www.ed.lu.se/databases/sedd**
POPLINK	Swedish POPLINK Demographic Database	Sweden	Application	**http://www.cedar.umu.se/english/order-data/**
TRA	French 3000 Family Study	France	Application	**http://tra.web.ined.fr/rubriques/donnees/uk_apply_data.htm**
PRDH	*Le Programme de Recherche en Démographie Historique*	Canada	Download	**https://www.genealogie.umontreal.ca/en/acces**
BALSAC	BALSAC population database	Canada	Application	**http://balsac.uqac.ca/english/acces-aux-donnees-2/acces-fin-de-recherche/**

(*Continued*)

Table 1 *(Continued)*

Project abbreviation	Project name	Location	Access	Link or citation
Census records				
IPUMS Linked	Integrated Public Use Microdata Series Linked Representative Samples, 1850–1930	United States	Download	**https://usa.ipums.org/usa/linked_data_samples.shtml**
Contemporary longitudinal data				
Longitudinal surveys				
PSID	US Panel Study of Income Dynamics	United States	Download	**https://psidonline.isr.umich.edu**
WLS	Wisconsin Longitudinal Study	Wisconsin, United States	Download	**http://www.ssc.wisc.edu/wlsresearch/**
NLSY79	National Longitudinal Study of Youth 79	United States	Download	**http://www.bls.gov/nls/nlsy79.htm**
HRS	Health and Retirement Study	United States	Download	**http://hrsonline.isr.umich.edu/**
FFCWS	Fragile Families & Child Wellbeing Study	United States	Download	**http://www.fragilefamilies.princeton.edu/**
ALSPAC	Avon Longitudinal Study of Parents and Children	United Kingdom	Application	**http://www.bristol.ac.uk/alspac/**
BHPS	British Household Panel Survey	United Kingdom	Download	**https://www.iser.essex.ac.uk/bhps**
SOEP	German Socio-Economic Panel	Germany	Download	**https://www.diw.de/en/soep**
CFPS	China Family Panel Studies	China	Download	**http://www.isss.edu.cn/cfps/EN/**
DLSY	Danish Longitudinal Survey of Youth	Denmark	Download	**http://dlsy.sfi.dk**
Retrospective surveys				
SSEE	Social Stratification in Eastern Europe after 1989	Bulgaria, Czech Republic, Hungary, Poland, Russian Federation, Slovakia	Download	**https://www.library.ucla.edu/social-science-data-archive/sseehome**
LHSCCC	Life Histories and Social Change in Contemporary China	China	Download	**https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/M889V1**

demographically successful. They are more likely to record recent ancestors; prominent ancestors; and ancestors who lived longer, married, and had more children. For example, Campbell & Lee (2002) showed that in a sample of Chinese family genealogies linked to household registers, high-status families and individuals, long-lived individuals, and individuals with more offspring are overrepresented.

Table 2 Characteristics of major genealogical datasets

	Period covered	Number of generations	Number of individuals	Variables								
				Birthdate	Death date	Both sexes	Occupation	Education	Income	Wealth or landholding	Migration	Disability or health outcomes
Family genealogies												
CMGPD-IL	1616–1936	16	83,256	Y	Y	N	Y	N	N	N	N	N
UPDB	Late 1700–present	7	>8 million	Y	Y	Y	Y	Y	Y	N	N	Y
Synthetic genealogies												
Population registers												
CMGPD-LN	1749–1909	6	266,091	Y	Y	Y	Y	Y	N	N	Y	Y
CMGPD-SC	1866–1913	6	107,551	Y	Y	Y	Y	Y	N	Y	Y	Y
KMGPD-TS	1678–1888	6	136,690	Y	Y	Y	Y	Y	N	N	Y	Y
CTHRD	1906–1945	4	103,151	Y	Y	Y	Y	N	N	Y	Y	Y
NAC-SN	1716–1870	6	6,257	Y	Y	Y	Y	N	N	Y	Y	Y
Parish registers and vital records												
LINKS/GENLIAS	1811–1920s	5	11 million	Y	Y	Y	Y	N	N	N	Y	N
HPR	1801–1964	3	9.7 million	Y	Y	Y	Y	Y	Y	N	Y	N
UBCoS Multigen	1915–2009	5	14,192	Y	Y	Y	Y	Y	Y	N	N	Y
SEDD	1646–2011	4	150,000	Y	Y	Y	Y	N	Y	Y	Y	Y
POPLINK	1680–1950	15	350,000	Y	Y	Y	Y	Y	N	N	Y	Y
TRA	1800–1939	7	70,000	N	Y	Y	Y	N	N	Y	Y	N
PRDH	1621–1849	9	440,000	Y	Y	Y	Y	Y	N	N	Y	N
BALSAC	1621–1965	12	5 million	Y	Y	Y	Y	Y	N	N	Y	N
Census records												
IPUMS Linked	1850–1930	2	500,000	Y	N	Y	Y	N	N	N	N	N

(Continued)

Table 2 (*Continued*)

					Variables								
		Period covered	Number of generations	Number of individuals	Birthdate	Death date	Both sexes	Occupation	Education	Income	Wealth or landholding	Migration	Disability or health outcomes
Contemporary longitudinal data													
Longitudinal surveys													
	PSID	1968–present	4	18,000	Y	Y	Y	Y	Y	Y	Y	Y	Y
	WLS	1957–present	3	10,317	Y	Y	Y	Y	Y	Y	Y	N	Y
	NLSY79	1979–present	3	12,686	Y	Y	Y	Y	Y	Y	Y	Y	Y
	HRS	1992–present	3	12,652	Y	Y	Y	Y	Y	Y	Y	Y	Y
	FFCWS	1998–present	3	5,000	Y	Y	Y	Y	Y	Y	Y	Y	Y
	ALSPAC	1991–present	3	14,000	Y	Y	Y	Y	Y	Y	N	N	Y
	BHPS	1991–present	3	10,300	Y	Y	Y	Y	Y	Y	Y	Y	Y
	SOEP	1984–present	3	30,000	Y	Y	Y	Y	Y	Y	Y	Y	Y
	CFPS	2010–present	3	30,000	Y	Y	Y	Y	Y	Y	Y	Y	Y
	DLSY	1968–2004	3	3,151	Y	Y	Y	Y	Y	Y	Y	Y	Y
Retrospective surveys													
	SSOASA	1991–1994	3	9,086	Y	Y	Y	Y	Y	Y	Y	Y	Y
	SSEE	1993–1994	3	28,500	Y	Y	Y	Y	Y	Y	Y	Y	Y
	LHSCCC	1996	3	6,090	Y	Y	Y	Y	Y	Y	Y	Y	Y

The number of individuals for longitudinal data refers to the rough estimate of the sample size in wave 1. Some statistics are obtained from the European Historical Population Sample Network website. The information may be slightly different from that provided on project websites.

Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; BHPS, British Household Panel Survey; CFPS, China Family Panel Studies; CMGPD-IL, China Multi-Generational Panel Dataset–Imperial Lineage; CMGPD-LN, China Multi-Generational Panel Dataset–Liaoning; CMGPD-SC, China Multi-Generational Panel Dataset–Shuangcheng; CTHRD, Colonial Taiwan Household Registration Database; DLSY, Danish Longitudinal Survey of Youth; FFCWS, Fragile Families & Child Wellbeing Study; HPR, Historical Population Register; HRS, Health and Retirement Study; IPUMS, Integrated Public Use Microdata Series; KMGPD-TS, Korea Multi-Generational Panel Dataset–Tansung; LHSCCC, Life Histories and Social Change in Contemporary China; NAC-SN, Japanese Ninbetsu-Aratame-Cho Population Register Database–Shimomoriya and Niita; NLSY79, National Longitudinal Survey of Youth 79; PRDH, Le Program de Recherche en Démographie Historique; PSID, US Panel Study of Income Dynamics; SEDD, Scanian Economic Demographic Database; SOEP, German Socio-Economic Panel; SSEE, Social Stratification in Eastern Europe after 1989; SSOASA, Survey of Socioeconomic Opportunity and Achievement in South Africa; UBCoS Multigen, Uppsala Birth Cohort Multigenerational Study; UPDB, Utah Population Database; WLS, Wisconsin Longitudinal Study.

There are nevertheless examples of genealogies that are suitable for use in multigenerational studies. For China, the China Multi-Generational Panel Dataset–Imperial Lineage (CMGPD-IL) records 120,000 individuals over 16 generations in the Qing (1644–1911) imperial lineage. The data record the patrilineal descendants of Takeshi, grandfather of the founder of the Qing dynasty, and his four brothers and male cousins, as well as the daughters of those patrilineal male descendants and all of their male descendants. In contrast with most Chinese genealogies, the CMGPD-IL was used for administration and had a government office dedicated to its upkeep. The Office of the Imperial Lineage registered imperial lineage members, supervised lineage activities, and maintained the lineage genealogy. At the beginning of the Qing dynasty, members of the lineage were a highly unrepresentative elite, most of whom held official positions or noble titles. As the lineage grew, however, more and more members experienced downward mobility. By the nineteenth century, many of them may not have differed much from other residents of the cities in which they lived. The CMGPD-IL has already been used in demographic studies that establish its reliability and completeness (Lee et al. 1993, Wang et al. 1995). Song et al. (2015) used it to study multigenerational processes. Korean lineage genealogies have properties that make them promising sources of multigenerational analysis. Although they are retrospective and therefore share some of the same limitations of Chinese family genealogies, there are notable differences in that many Korean genealogies provide information on daughters, spouses, adoptees, and in-laws, as well as relatively detailed information on socioeconomic status (Kim 2016, Kim & Park 2010, Park & Lee 2008).

Genealogies are also available for Europe and North America. One of the largest and most detailed is the Utah Population Database (UPDB). Its core consists of family genealogies that were compiled by members of the Church of Jesus Christ of Latter-day Saints. The individuals recorded in these genealogies have been linked to contemporary data from the state of Utah and other sources, as well as historical sources such as censuses. At present, the UPDB contains information on 8 million individuals from the late eighteenth century to the present, most of whom lived in Utah at some point. Linked sources include census data, birth and death certificates, marriage and divorce records, driver licenses, voter registrations, cancer registry records, medical records, and Medicare claims. The data are a rich source of information for demographic, genetic, biomedical, and epidemiological research, and access is available by application. They have been widely used for analyzing mortality and morbidity risks over individuals' life courses and the clustering of demographic behaviors and outcomes within family pedigrees (Bean et al. 1990, Goldgar et al. 1994, Hanson et al. 2015, Jennings et al. 2012, Slattery & Kerber 1993). The UPDB data have also been used extensively in genetic studies, especially those related to cancer (e.g., Smith et al. 1999, Taylor et al. 2011, Teerlink et al. 2012). As for Europe, Knodel (2002) used German genealogies to study fertility decline in Germany in the eighteenth and nineteenth centuries, and Hollingsworth (1964) used genealogies to study the demography of the British peerage from the seventeenth century to the twentieth century.

Synthetic Genealogies

There are many examples of genealogical microdata in which pedigrees have been synthesized by nominative linkage of parish and vital records, household registers, and most recently, censuses. One advantage of these databases relative to traditional genealogies is that because they rely on contemporaneous administrative data, they are prospective and should be less affected by survivorship and recall bias. Some of them are publicly available or accessible via application (see Ruggles 2014). In Europe and Québec, family reconstitution based on parish registers and vital records that link children to their parents has produced multigenerational data. Especially in East

Asia, iterative linkage of children to their parents in databases of population registers has allowed for reconstruction of descent lines. In North America and some other locations, recent linkage of census data has produced new data sets that follow families for three or more generations.

Parish registers and vital records. Research based on genealogies synthesized from parish and vital records dates back at least to Louis Henry (1968), a founder of historical demography, who devised the method that later became known as family reconstitution. This method relies on parish registers of baptisms, burials, and marriages that recorded the names of individuals and their parents to reconstruct the histories of individuals who belong to the same families (Hammel 1993, Norton 1980, Schofield & Wrigley 1979, Willigan & Lynch 1982, Wrigley 1997). Unlike many genealogies that recorded only patrilineal lines, family reconstitution can in principle trace both patrilineal and matrilineal lineages. Family reconstitution data were for some time the only available source for estimating fertility, mortality, and marriage rates in historical European populations, and the strengths and weaknesses are accordingly well known. The most important limitation of family reconstitutions as a source of genealogical microdata is that the ones used in early studies covered only a single parish or a sample of parishes. Individuals who moved in or out could not be traced to their origin or destination. Because mobility in historical Europe was high, only a small share of families remained in the same parish for more than three generations. Until recently, therefore, reconstitution over multiple generations was practical only for communities with highly sedentary populations.

Recently, however, there have been ambitious efforts to synthesize genealogical microdata at the national or regional level by linking parish and vital records with other administrative data. For the Netherlands, the LINKS project seeks to reconstruct families in the Netherlands from the nineteenth century to the early twentieth century by linkage of birth, marriage, and death certificates. It builds on the earlier GENLIAS database of linked marriage certificates (Oosten & Mandemakers 2007). These data have already been used for a study of grandparents' and great-grandparents' effects on attainment (Knigge 2016). The Historical Population Register (HPR) for Norway project will produce a database of linked census and church records to cover the population of that country from the early nineteenth century to 1964, when the current registration system was established (Thorvaldsen et al. 2015). For Scotland, an effort to link records for the Isle of Skye (Reid et al. 2002) has been under way for some time, and an even more ambitious project to link records for the entire country is in progress.[1]

A number of sources of genealogical microdata are now available for Sweden. For contemporary Sweden, the Multi-Generation Register maintained by Statistics Sweden records everyone born in Sweden since 1932 and anyone who was registered in Sweden after 1961. POPLINK was assembled from digitized Swedish parish records from the eighteenth century to the twentieth century in the Swedish Demographic Database at Umeå University. The database comprises information from approximately 350,000 individuals spanning up to 15 generations of families over nearly 300 years (Engberg et al. 2016). The current data include variables that are similar to those in many parish databases, but POPLINK can also be linked to modern Swedish registry resources and clinical data (Wisselgren et al. 2014). The Uppsala Birth Cohort Multigenerational Study (UBCoS Multigen) tracked all individuals born at the Uppsala University Hospital between 1915 and 1929 and traced their descendants to the present day. The sample includes males and females from the original birth cohort, along with their mothers, offspring, grandchildren, great-grandchildren, and great-great-grandchildren, and has already been used in the

[1] **http://www.lscs.ac.uk/projects/digitising-scotland/**.

studies of multigenerational processes (Goodman et al. 2012). Two other multigenerational population databases have been constructed in Sweden by linkage of parish, vital, and administrative records. The Scanian Economic Demographic Database (SEDD) contains more than 104,000 linked individual records from five parishes in Scania, a southern county of Sweden, during the period between 1813 and 1968. In addition to basic demographic and household data, the SEDD includes detailed information on socioeconomic status, including occupation, landholding, types of residence, income, height, poll-tax registration, and military history (Bengtsson et al. 2014).

The French TRA project (*L'Enquête TRA*) has produced a nationally representative sample of 3,000 French couples who married between 1803 and 1832 and their descendants (Bourdieu et al. 2014, Kesztenbaum 2008). To facilitate linkage of records of the same individuals held in different archives all over France, the sample included only those whose last name began with the letters TRA. One of the most notable features of the TRA database is that it includes 87,000 bequest records of 75,000 individuals. The wealth information was recorded at the time of death and is therefore an invaluable source for the study of multigenerational wealth dynamics (Arrondel & Grange 2003, Bonneuil & Rosental 1999, Bourdieu et al. 2008). Genealogical microdata rarely contain such detail on wealth. The TRA also includes demographic information drawn from vital registers in the nineteenth and twentieth centuries.

In North America, the PRDH (*Le Programme de Recherche en Démographie Historique*) at the Université de Montréal has reconstituted a large share of the population of Québec from the original settlement in the seventeenth century to the nineteenth century using parish registers and other administrative records (Légaré et al. 1972). The data set now consists of more than 2,350,000 individual records of births, marriages, deaths, family and household relationships, socioeconomic characteristics, places of residence, and origins of immigrants. From these data, the PRDH has reconstructed patrilineal and matrilineal pedigrees. A related project, the BALSAC population database, includes records for more than 5 million individuals in Québec from the seventeenth century to the late twentieth century. It combines the PRDH data with more recent baptism, marriage, and death records. These databases have been used extensively in studies of consanguinity and genetic studies more broadly. They provide firsthand evidence of the historical legacy of demographic behavior in the past for contemporary population composition (Bideau et al. 1997, Tremblay et al. 2008).

Population registers. For East Asia, historical population registers are a relatively new source of genealogical microdata. They represent an alternative to the family genealogies described above that were initially the mainstay of historical demographic research for East Asia. In contrast with most family genealogies, population registers were prospective. Extinct lineages were recorded alongside ones that flourished, so that there is no survivorship bias. Also in contrast with family genealogies, East Asian population registers record not only demographic behavior but also social and economic status. They typically record occupation and various hereditary ascribed statuses. In some cases, they record landholding or other wealth. Because Dong et al. (2015) already provide a comprehensive introduction to available household register databases for historical Korea, Japan, Northeast China, and Taiwan, we focus here only on the features specifically relevant to the study of multigenerational demographic and stratification processes.

The East Asian household register database with the greatest generational depth is the China Multi-Generational Panel Dataset–Liaoning (CMGPD-LN). It follows families in a selection of rural communities in what is now the province of Liaoning from 1749 to 1909. Many families are followed for as many as seven generations. For individuals born at the beginning of the twentieth century, approximately 60% can be linked to a patrilineal great-great-great-great-grandfather and his other male descendants, and more than 80% can be linked to a great-great-great-grandfather

(Lee et al. 2010, pp. 17–18). The data have already been used to study effects of characteristics of distant kin on socioeconomic attainment and demographic behavior (Campbell & Lee 2003, 2008). Automated record linkage has resolved the 1.3 million records of 260,000 individuals into more than 1,000 patrilineages. This allows for the patrilineage to be treated as a unit of analysis (Campbell & Lee 2011) and the social, economic, and demographic determinants of growth and extinction to be studied systematically (Song et al. 2015).

Other East Asian databases may be used to study grandparents' and, in some cases, great-grandparents' effects. The China Multi-Generational Panel Dataset–Shuangcheng (CMGPD-SC) follows 100,000 residents of recently settled farming communities in Northeast China between 1866 and 1913 and is also available for download at Interuniversity Consortium for Political and Social Research. Notably, it includes information on landholding at several points in time, allowing for the study of wealth mobility over multiple generations (Chen et al. 2010). For Japan, a large number of household registers covering communities in the eighteenth and nineteenth centuries survive, and many have been or are being transcribed into databases. These registers, collectively termed the Japanese Ninbetsu-Aratame-Cho Population Register Database (NAC), record a rich variety of social and economic statuses, including heritable ascribed statuses and, for many communities, landholding or other wealth. Registers for two communities, Shimomoriya and Niita, have been the subject of numerous published studies. These follow the residents of these communities on a nearly annual basis from 1716 to 1870. One-third of individuals can be linked at least to their grandparents, and nearly one-quarter can be linked at least to great-grandparents. For Taiwan, the Colonial Taiwan Household Registration Database (CTHRD) was constructed from household registers compiled between 1906 and 1945 by Japanese colonial authorities (Wolf & Huang 1980). These data sets have been used extensively in demographic studies and link one-third of individuals to their grandparents. Korean household register databases for Tansung and Daegu, which constitute the Korean Multi-Generational Panel Dataset–Tansung (KMGPD-TS) have been publicly released, but the potential for multigenerational studies is limited because high turnover makes linkage beyond parents difficult (HRWG 2003). For the Tansung registers, only 12% of individuals can be linked to grandparents (Dong et al. 2015, p. 1077).

Census records. Nominative linkage of census records has begun to yield genealogical microdata that cover three generations. For the United States and the United Kingdom, Joseph Ferrie and his collaborators have created nationally representative data sets that include family members across three generations by linking generations across historical population censuses from 1850 onward (Ferrie 1997, 2005; Ferrie et al. 2016; Long & Ferrie 2013). Details on the data linkage procedures and discussions of representativeness and potential biases are available in a paper by Long & Ferrie (2013). Nominative linkage of census records is especially challenging because many people shared the same names, and names changed or were written differently in succeeding censuses. Thus, whereas the resulting data sets are large and nationally representative, assumptions made or restrictions imposed during nominative linkage may lead the resulting multigenerational samples to be selective or otherwise unrepresentative. For example, the cross-generation linkage requires that offspring live in the same households as their parents on the 1850 census form. Otherwise, adults in 1880 cannot be matched to their childhood families in 1850 (Hout & Guest 2013, Xie & Killewald 2013).

Linked censuses nevertheless may soon become the core of much richer genealogical microdata that include linked administrative and survey data. Erola & Moisio (2007) linked the 1950 Finnish population census to population panel data from 1970 to 2000. Recently, a collaborative research group under the US Census Bureau has been investigating the possibility of integrating decennial censuses with other data collected by the Census Bureau, including Current Population Surveys, Survey of Income and Program Participation, and American Community Surveys (Grusky et al.

2015, Johnson et al. 2015, Massey 2016, Mitnik et al. 2016, Wagner & Layne 2014). IPUMS (Integrated Public Use Microdata Series) plans to further combine census data with vital records and family genealogies compiled by Ancestry.com to expand the time coverage of the longitudinal sample.[2] The resulting microdata will provide an unprecedented laboratory for studying multigenerational processes from the earliest US census dated 1790 to the present day.

Contemporary Survey Data

Growing numbers of contemporary surveys provide multigenerational data. Some longitudinal surveys have a genealogical design according to which descendants of original respondents are traced. Some retrospective surveys collect information on the parents and grandparents of respondents. Below we discuss each and provide examples.

Longitudinal surveys. Long-term panel data that are statistically representative of contemporary populations will become increasingly available in the next few decades. In recent decades, some longitudinal surveys have begun to follow offspring and descendants of the original respondents. One of the best examples of the potential for longitudinal surveys to become sources of genealogical microdata is the US Panel Study of Income Dynamics (PSID), the longest ongoing longitudinal household survey in the world. The project began in 1968 and has since collected socioeconomic and demographic information on the original respondents and their family members, including descendants. The project now includes families linked across as many as five generations, although most of the fourth and fifth generations are still too young to complete schooling and obtain social statuses (Andreski et al. 2013).

Contemporary surveys can also achieve three-generation coverage or greater by combining proxy responses for either parents or children with interviews of the other. The WLS, which follows a cohort of 1957 high school graduates, provides information on a subset of families for three generations. Parents of some original respondents were interviewed, and in later waves, those respondents with at least one adult child were asked to report their offspring's education and occupational information in later waves. These data have already been used in studies of kin associations with socioeconomic attainment (Jæger 2012). The Avon Longitudinal Study of Parents and Children (ALSPAC) has adopted a similar research design by following children who were born to 14,000 British women between 1991 and 1992 and then enrolling grandparents, siblings, and children of the original subjects in more recent waves (Golding et al. 2001). Ongoing panel studies such as the British Household Panel Survey (BHPS 1997) (Chan & Boliver 2013), Chinese Family Panel Studies (CFPS 2010) (Xie & Zhang 2016), and German Socio-Economic Panel (SOEP 2000) (Hertel & Groh-Samberg 2014) also have been used to construct three-generation samples in which researchers combine the retrospective and prospective features of these studies.

The National Longitudinal Survey also offers three-generation coverage via a combination of proxy responses for characteristics of the original respondents' parents and interviews with their children. The National Longitudinal Survey of Youth 79 (NLSY79) consists of a nationally representative sample of 12,686 men and women who were born between 1957 and 1964; they were first interviewed in 1979 and then reinterviewed on a regular basis afterward. The first wave of interviews collected information about the respondents' family background, including parents' relationships, living situation, socioeconomic statuses, and their siblings' ages, fertility, and education. Since 1986, the children born to the female respondents in the NLSY79 have been

[2] More information can be found at **https://pop.umn.edu/research/research-areas/population-data-science**.

followed up in a new project referred to as the NLSY79 Children and Young Adults (Cent. Hum. Resour. Res. Ohio State Univ. 2001). A similar data set, the Danish Longitudinal Survey of Youth (DLSY), followed 3,151 individuals around age 14 from 1968 until 2004 and has conducted separate surveys with the parents and all the children of DLSY respondents (Møllegaard & Jæger 2015).

Other longitudinal surveys achieve coverage of three adult generations by asking older respondents not only about their parents but also about their adult children. The Health and Retirement Study (HRS) is one example. The HRS, a longitudinal survey of the middle-aged and elderly in the United States, has collected detailed information about respondents' physical health, cognitive functioning, economic status and behaviors, and family structure and marriage. Respondents were asked to provide basic social and demographic information about their parents, siblings, children, and occasionally their stepparents and their children's spouses or partners (Heeringa & Connor 1995). In a module focused on financial and other transfers with relatives, respondents who had made transfers to their grandchildren were asked for their grandchildren's information (Servais 2010). The HRS has also genotyped a subsample of respondents and estimated relatedness among these respondents by kinship coefficients (Faul et al. 2014). These data have become one of the major sources for social scientists to quantify gene-environment interactions in shaping social inequality across generations (e.g., Conley et al. 2014, Domingue et al. 2014). The Fragile Families & Child Wellbeing Study, which follows 5,000 children born between 1998 and 2000 in large American cities, is also a three-generation study in the sense that it collects information on the parents and grandparents of those children, although the members of the third generation are still in their teens. If the study continues, it will contain three generations of adults within a few years (Reichman et al. 2001).

One potential limitation of these longitudinal surveys is that most of them adopt a genealogical design that follows only the original respondents and their direct descendants, but they do not collect information on the ancestors of their descendants' spouses. The ideal data would include both matrilineal and patrilineal ancestry, which would allow for examination of more complex multigenerational processes (Song & Mare 2017). By contrast, the longitudinal surveys typically yield data for either the patrilineal or matrilineal ancestors of subjects in the later generations, but not for both sets of grandparents. For example, most grandchildren in the PSID can be linked only to the grandparents, paternal or maternal, who were in the original 1968 PSID sample. The PSID addresses this partially by asking the partners of the original respondents for retrospective information about their parents. The PSID recently also added a Rosters and Transfers Module and a Childhood Retrospective Circumstances Study (PSID-CRCS) that asked a sample of respondents to provide basic social and demographic information for their kin.

Retrospective surveys. Cross-sectional surveys offer three-generation data for a number of settings where longitudinal data are not available. For example, Treiman and his collaborators conducted surveys in South Africa (the Survey of Socioeconomic Opportunity and Achievement in South Africa, or SSOASA), Eastern Europe[3] (Social Stratification in Eastern Europe after 1989, or SSEE), and China[4] (Life Histories and Social Change in Contemporary China, or LHSCCC) that collected retrospective data on the demographic and socioeconomic characteristics of the parents and grandparents of respondents during a period of rapid social transformation (Szelényi

[3]The complete data sets and documentation are publicly available for download from the Social Science Data Archive Dataverse (**http://dataverse.harvard.edu**).

[4]The data and documentation can be downloaded from the UCLA Social Science Data Archive: Data Portals (**http://data-archive.library.ucla.edu/lhsccs/chinaweb.html**).

& Treiman 1994, Treiman et al. 1996, Treiman & Walder 1996). Such retrospective data suffer limitations similar to those of the retrospective family genealogies discussed above. In particular, they suffer from retrospective sampling bias, which arises from the omission of childless populations in the parent and grandparent generations and the overrepresentation in the contemporary population of parents and grandparents who have higher-than-average fertility (Preston 1976, Song & Mare 2015). Recall error is also a potential problem because respondents may not actually know key details about their parents and grandparents.

Kinship data from surveys. Whereas the aforementioned data provide a longitudinal perspective on multigenerational processes, data that include records of living siblings, cousins, second cousins, and more distant kin offer a cross-sectional perspective. The demographic behavior and socioeconomic outcomes of siblings, cousins, and more distant kin may be influenced by their common background and their interactions with one another, so kin group membership may be an important source of variation between individuals, above and beyond the effects of common parentage or household residence (Hällsten 2014, Solon 1999). Compared with three-generation data that are used to measure direct and indirect associations between grandparents and grandchildren in a regression framework, cousin data provide an estimate of grandparents' gross effects working through both observed and unobserved variables via genetic, behavioral, social, cultural, and institutional pathways. When both cross-sectional data on kinship and longitudinal data on ancestry are available, we can compare grandparents' effects estimated from the grandparent-grandchildren regression analysis and the cousin analysis and show the proportion of grandparents' total effects that can be explained by the observed characteristics of the grandparents. By including both individuals and their kin as respondents, these cross-sectional surveys are also less prone to recall errors that arise when individuals report on their ancestors. Finally, the cross-sectional data on kinship networks may be more readily available than three-generation data, as we need to ascertain relationships only among individuals and do not need exact socioeconomic measures of grandparents.

Several recent studies have measured associations between kin using novel data and methods. Hällsten (2014) used the Multi-Generation Registers created by Statistics Sweden to identify first and second cousin pairs. He reported positive albeit weak correlations in academic performance, cognitive ability, educational outcomes, and occupational prestige among first cousins and second cousins. His approach is an extension of the sibling correlation method widely employed in the studies on family influences on individuals' social outcomes (e.g., Black et al. 2005, Bouchard & McGue 1981, Solon et al. 2000). Campbell & Lee (2011) used the CMGPD-LN and found that associations in the marriage and social attainment of patrilineage members were stronger than those of residents of the same village, but for other demographic outcomes, village of residence was more important than patrilineage membership.

METHODS AND MODELS IN THE ABSENCE OF MICRODATA

Given that genealogical microdata are still scarce, some researchers have used aggregated data to analyze long-term stratification processes. One example that we have already noted is the surname data employed by Gregory Clark and his colleagues based on rare surname frequencies and distributions in Qing dynasty China, medieval England, modern Sweden and the United States, and other places. They use the rate of decline from one generation to the next in the share of the population accounted for by people with rare surnames initially associated with high status to produce an indirect estimate of parent-child correlations in attainment (Clark 2014, Clark & Cummins 2015, Clark et al. 2015). In a related vein, Olivetti & Paserman (2015) approximate individuals' socioeconomic status by their first names and examine the trend in intergenerational income mobility in the United States from 1850 to 1940. They argue that a careful research design can yield valid mobility estimates. Caution is warranted, however, in interpreting indirect estimates from aggregated data produced by Clark and his collaborators. Torche & Corvalan (2016) evaluate the surname-based method by comparing aggregate mobility measures with those based on individual-level data. On the basis of their analysis, they conclude that mobility estimates from aggregate data and microdata are not comparable. By partitioning individual mobility into within-group mobility and between-group mobility, they show that the surname mobility estimates are essentially equivalent to the between-group intergenerational persistence parameter. The indirect approach based on aggregated data, in other words, overlooks social mobility that happens within groups.

In the absence of high-quality microdata at the individual and household levels, microsimulations provide an alternative source for examining family over decades and generations. The microsimulation method, also known as agent-based simulation (Billari et al. 2006), relies on assumptions about predetermined individual demographic rates to simulate the occurrence of demographic events among individuals in a hypothetical population. Simulations are particularly useful for producing estimates of the kinship networks that could have been achieved in the past under different assumptions about the levels of mortality, fertility, marriage, and migration. A counterfactual approach based on varying assumptions helps assess the relative importance of different factors in shaping population and family dynamics. Unlike mathematical models that obtain only estimates of average numbers of kin (Goodman et al. 1974), simulation models yield estimates of the distributions of individuals by number and types of kin. Previous research has developed many microsimulation programs such as AMBUSH (Howell & Lehotay 1978), BACKFOR (LeBras & Wachter 1978), SOCSIM (Wachter et al. 1978), MOMSIM (Ruggles 1992), and CAMSIM (Smith 1987). Applications include the relationship between contraceptive use and population growth (Howell & Lehotay 1978), the number and configuration of kin (De Vos & Palloni 1989, Hammel et al. 1979, Reeves 1987, Smith & Oeppen 1993), household composition (Clarke 1986, Ruggles 1987, Wachter et al. 1978, Zhao 2000), and assortative mating (Murphy 2006).

CONCLUSION

Developing a comprehensive understanding of the implications of long-term trends in social mobility and demographic processes within family lines, and the interplay of these processes in shaping descent line growth and extinction and changes in population composition, is crucial to advancing our understanding of stratification processes and population dynamics. By reviewing issues and debates in the study of multigenerational demographic and social processes and introducing major sources of genealogical microdata that can be used in relevant analyses, we have showcased the tremendous potential for further empirical studies. In the past decade or two, we have seen the growing availability of a variety of large-scale genealogical microdata that may be analyzed to achieve a better understanding of the complex nature of multigenerational processes. These data sources range from traditional family genealogies, parish records, and population registers to modern longitudinal or retrospective surveys that follow families over three or more generations and are designed to analyze contemporary family and household organizations. Genealogical microdata allow researchers to expand the scope of studies on the demography, economics, and sociology of families over long timescales.

The increasing availability of genealogical microdata from different locations and time periods raises the prospect of systematic comparison of multigenerational processes. We suggest that the next step in the analysis of genealogical microdata and study of multigenerational processes will be to pool multiple data sets and examine how features of social context condition relationships. At present, almost all multigenerational studies make use of a single data set, and observed relationships presumably reflect features specific to that context. Especially for studies of grandparents' effects, similar analyses of different data sets yield different results. Reconciling these differences and distinguishing what is universal from what is a systematic product of context will require analyses of pooled data sets, where the contextual characteristics of the population recorded in each data set are accounted for.

As Ganzeboom et al. (1991) have argued, all the important theoretical and empirical advances in comparative intergenerational stratification research have been accompanied by the development of new methods of data collection, measurement, and analysis. Just as the proliferation of mobility surveys around the world a few decades ago led to tremendous advances in the comparative understanding of stratification processes based on parent-child transmission of status, we expect the explosion in the availability of genealogical microdata to inspire new analyses and comparisons that will lead to a comprehensive multigenerational perspective on population dynamics, stratification processes, and the interactions between the two.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Andreski P, Beaule A, Dascola M, Duffy D, Leissou E, et al. 2013. *PSID main interview user manual: release 2013*. Inst. Soc. Res., Univ. Mich.

Arrondel L, Grange C. 2003. The accumulation and transmission of wealth over a long period: example of a rural family from Loire-Atlantique in the 19th and early 20th centuries. *Hist. Fam.* 8(1):103–34

Bartholomew DJ. 1967. *Stochastic Models for Social Processes*. London: Wiley

Bean L, Mineau GP, Anderton D. 1990. *Fertility Change on the American Frontier: Adaptation and Innovation*. Berkeley: Univ. Calif. Press

Becker GS. 1981. *A Treatise on the Family*. Cambridge, MA: Harvard Univ. Press

Beeton M, Yule GU, Pearson K. 1900. Data for the problem of evolution in man. V. On the correlation between duration of life and the number of offspring. *Proc. R. Soc. Lond.* 67:159–79

Bengtson VL. 2001. Beyond the nuclear family: the increasing importance of multigenerational bonds. *J. Marriage Fam.* 63:1–16

Bengtsson T, Dribe M, Quaranta L, Svensson P. 2014. *The Scanian Economic Demographic Database*, *Version 4.0 (Machine-Readable Database)*. Lund: Lund Univ., Cent. Econ. Demogr.

Bideau A, Desjardins B, Brignoli HP. 1997. *Infant and Child Mortality in the Past*. Oxford, UK: Clarendon Press

Billari FC, Fent T, Prskawetz A, Scheffran J, eds. 2006. *Agent-Based Computational Modelling: Applications in Demography*, *Social*, *Economic and Environmental Sciences*. New York: Springer

Black SE, Devereux PJ, Salvanes KG. 2005. Why the apple doesn't fall far: understanding intergenerational transmission of human capital. *Am. Econ. Rev.* 95(1):437–39

Bol T, Kalmijn M. 2016. Grandparents' resources and grandchildren's schooling: Does grandparental involvement moderate the grandparent effect? *Soc. Sci. Res.* 55:155–70

Bongaarts J, Burch TK, Wachter KW, eds. 1987. *Family Demography: Methods and Their Application*. Oxford, UK: Clarendon Press

Bonneuil N, Rosental PA. 1999. Changing social mobility in nineteenth-century France. *Hist. Methods J. Quant. Interdiscip. Hist.* 32(2):53–73

Bouchard TJ, McGue M. 1981. Familial studies of intelligence: a review. *Science* 212(4498):1055–59

Bourdieu J, Menéndez M, Postel-Vinay G, Suwa-Eisenmann A. 2008. Where have (almost) all the wealthy gone? Spatial decomposition of wealth trends in France, 1820–1939. *Rev. Agric. Environ. Stud.* 87(2):5–25

Bourdieu J, Kesztenbaum J, Postel-Vinay G. 2014. The TRA Project, a historical matrix. *Population* 69(2):217–48

Breen R, Ermisch J. 2016. *Educational reproduction in Great Britain: a prospective approach*. Presented at Res. Comm. Soc. Stratif. Mobil. (RC28) ISA, Bern, Switz.

Campbell CD, Lee JZ. 2002. State views and local views of population: linking and comparing genealogies and household registers in Liaoning, 1749–1909. *Hist. Comput.* 14:9–29

Campbell CD, Lee JZ. 2003. Social mobility from a kinship perspective: rural Liaoning, 1789–1909. *Int. Rev. Soc. Hist.* 47:1–26

Campbell CD, Lee JZ. 2008. Kin networks, marriage, and social mobility in late imperial China. *Soc. Sci. Hist.* 32(2):175–214

Campbell CD, Lee JZ. 2011. Kinship and the long-term persistence of inequality in Liaoning, China, 1749–2005. *Chin. Sociol. Rev.* 44(1):71–103

Cavalli-Sforza LL, Feldman MM. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton Univ. Press

Cent. Hum. Resour. Res. Ohio State Univ. 2001. NLSY79 User's Guide: *A Guide to the 1979–2000 National Longitudinal Survey of Youth Data*. Washington, DC: US Dep. Labor

Chan TW, Boliver V. 2013. The grandparent effect in social mobility: evidence from British birth cohort studies. *Am. Sociol. Rev.* 78(4):662–78

Chaparro MP, Koupil I. 2014. The impact of parental educational trajectories on their adult offspring's overweight/obesity status: a study of three generations of Swedish men and women. *Soc. Sci. Med.* 120:199–207

Chen S, Lee J, Campbell CD. 2010. Wealth stratification and reproduction in Northeast China, 1866–1907. *Hist. Fam. Int. Q.* 15(4):386–412

Cherlin AJ, Furstenberg FF. 1986. *The New American Grandparent: A Place in the Family, a Life Apart*. New York: Basic Books

Chiang YL, Park H. 2015. Do grandparents matter? A multigenerational perspective on educational attainment in Taiwan. *Soc. Sci. Res.* 51:163–73

Clark G. 2007. *A Farewell to Alms: A Brief Economic History of the World*. Princeton, NJ: Princeton Univ. Press

Clark G. 2014. *The Son Also Rises: Surnames and the History of Social Mobility*. Princeton, NJ: Princeton Univ. Press

Clark G, Cummins N. 2015. Malthus to modernity: wealth, status, and fertility in England, 1500–1879. *J. Popul. Econ.* 28(1):3–29

Clark G, Cummins N, Yu H, Vidal DD. 2015. Surnames: a new source for the history of social mobility. *Explorations Econ. Hist.* 55(1):3–24

Clarke M. 1986. Demographic processes and household dynamics: a micro-simulation approach. In *Population Structure and Models: Developments in Spatial Demography*, ed. R Woods, P Rees, pp. 245–72. Boston/London: Allen & Unwin

Coall DA, Hertwig R. 2010. Grandparental investment: past, present, and future. *Behav. Brain Sci.* 33:1–59

Cohen M. 1990. Lineage organization in North China. *J. Asian Stud.* 49(3):509–34

Conley D, Siegal ML, Domingue BW, Harris KM, McQueen MB, Boardman JD. 2014. Testing the key assumption of heritability estimates based on genome-wide genetic relatedness. *J. Hum. Genet.* 59:342–45

De Vos S, Palloni A. 1989. Formal models and methods for the analysis of kinship and household organization. *Popul. Index* 55(2):174–98

Domingue BW, Fletcher J, Conley D, Boardman JD. 2014. Genetic and educational assortative mating among US adults. *PNAS* 111:7996–8000

Dong H, Campbell CD, Kurosu S, Yang W, Lee JZ. 2015. New sources for comparative social science: historical population panel data from East Asia. *Demography* 52(3):1061–88

Dong H, Manfredini M, Kurosu S, Yang W, Lee JZ. 2016. Kin and birth order effects on male child mortality: three East Asian populations, 1716–1945. *Evol. Hum. Behav.* 38:208–16

Dribe M, Helgertz J. 2017. The lasting impact of grandfathers: class, occupational status, and earnings over three generations in Sweden 1815–2011. *J. Econ. Hist.* 76:969–1000

Dunifon R, Kowaleski-Jones L. 2007. The influence of grandparents in single-mother families. *J. Marriage Fam.* 69(2):465–81

Dyke B, Morrill WT. 1980. *Genealogical Demography*. New York: Academic

Engberg E, Westberg A, Edvinsson S. 2016. A unique source for innovative longitudinal research: the POPLINK database. *Hist. Life Course Stud.* 3:20–31

Erola J, Moisio P. 2007. Social mobility over three generations in Finland, 1950–2000. *Eur. Sociol. Rev.* 23:169–83

Faul J, Smith J, Zhao W. 2014. *Health and retirement study: candidate gene and SNP data description*. Health Retire. Study, Univ. Mich., Ann Arbor, MI. **http://hrsonline.isr.umich.edu/sitedocs/genetics/candidategene/CandidateGeneSNPDataDescription**

Ferrie JP. 1997. Immigrants and natives in the U.S.: comparative economic performance in the United States, 1850–1860 and 1965–1980. *Res. Labor Econ.* 16:319–41

Ferrie JP. 2005. The end of American exceptionalism? Mobility in the U.S. since 1850. *J. Econ. Perspect.* 19:199–215

Ferrie JP, Massey C, Rothbaum J. 2016. *Do grandparents and great-grandparents matter? Multigenerational mobility in the US, 1910–2013*. NBER Work. Pap. No. 22635, Natl. Bur. Econ. Res.

Freeman BC. 1935. Fertility and longevity in married women dying after the end of the reproductive period. *Hum. Biol.* 7(3):392–418

Freedman M. 1966. *Chinese Lineage and Society: Fukien and Kwangtung*. Lond. School Econ. Monogr. Soc. Anthropol., No. 33. London: Athlone Press

Ganzeboom HBG, Treiman DJ, Ultee WC. 1991. Comparative intergenerational stratification research: three generations and beyond. *Annu. Rev. Sociol.* 17:277–302

Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. 1994. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J. Natl. Cancer Inst.* 86(21):1600–8

Golding J, Pembrey M, Jones R, Study Team ALSPAC. 2001. ALSPAC—the Avon Longitudinal Study of Parents and Children. *Paediatr. Perinat. Epidemiol.* 15(1):74–87

Goldstein JR, Stecklov G. 2016. From Patrick to John F.: ethnic names and occupational success in the last era of mass migration. *Am. Sociol. Rev.* 81(1):85–106

Goodman A, Koupil I. 2009. Social and biological determinants of reproductive success in Swedish males and females born 1915–1929. *Evol. Hum. Behav.* 30(5):329–41

Goodman A, Koupil I. 2010. The effect of school performance upon marriage and long-term reproductive success in 10,000 Swedish males and females born 1915–1929. *Evol. Hum. Behav.* 31(6):425–35

Goodman A, Koupil I, Lawson DW. 2012. Low fertility increases descendant socioeconomic position but reduces long-term fitness in a modern post-industrial society. *Proc. R. Soc. B* 279(1746):4342–51

Goodman LA, Keyfitz N, Pullum TW. 1974. Family formation and the frequency of various kinship relationships. *Theor. Popul. Biol.* 5:1–27

Grusky DB, Smeeding TM, Snipp CM. 2015. A new infrastructure for monitoring social mobility in the United States. *Ann. Am. Acad. Polit. Soc. Sci.* 657:63–82

Hammel EA. 1993. Incomplete histories in family reconstitution: a sensitivity test of alternative strategies with historical Croatian data. In *Old and New Methods in Historical Demography*, ed. DS Reher, R Schofield, pp. 125–44. New York: Oxford Univ. Press

Hammel EA, McDaniel CK, Wachter KW. 1979. Demographic consequences of incest tabus: a microsimulation analysis. *Science* 205(4410):972–77

Hanson HA, Smith KR, Zimmer Z. 2015. Reproductive history and later-life comorbidity trajectories: a Medicare-linked cohort study from the Utah Population Database. *Demography* 52(6):2021–49

Harrell S. 1987. On the holes in Chinese genealogies. *Late Imp. China* 8(2):53–79

Harrell S, Pullum T. 1995. Marriage, mortality, and the developmental cycle in three Xiaoshan lineages. In *Chinese Historical Microdemography*, ed. S Harrell, pp. 141–62. Berkeley: Univ. Calif. Press

Heeringa SG, Connor J. 1995. *Technical description of the Health and Retirement Study sample design*. HRS/AHEAD Doc. Rep. 002, Inst. Soc. Res., Univ. Mich.

Henry L. 1968. Historical demography. *Daedalus* 97(2):385–96

Hertel FR, Groh-Samberg O. 2014. Class mobility across three generations in the US and Germany. *Res. Soc. Stratif. Mobil.* 35:35–52

Hodge RW. 1966. Occupational mobility as a probability process. *Demography* 3(1):19–34

Hollingsworth TH. 1964. *The Demography of the British Peerage*. London: Popul. Investig. Comm., Lond. Sch. Econ.

Hollingsworth TH. 1969. *Historical Demography*. Ithaca, NY: Cornell Univ. Press

HRWG (Househ. Regist. Work. Group). 2003. *Research on Tansung household registers*. Seoul: Sungkyunkwan Univ. Press (in Korean)

Hout M, Guest AM. 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850: comment. *Am. Econ. Rev.* 103(5):2021–40

Howell N, Lehotay VA. 1978. AMBUSH: a computer program for stochastic microsimulation of small human populations. *Am. Anthropol.* 80(4):905–22

Hällsten M. 2014. Inequality across three and four generations in Egalitarian Sweden: 1st and 2nd cousin correlations in socio-economic outcomes. *Res. Soc. Stratif. Mobil.* 35:19–33

Jæger MM. 2012. The extended family and children's educational success. *Am. Sociol. Rev.* 77:903–22

Jennings JA, Sullivan AR, Hacker JD. 2012. Intergenerational transmission of reproductive behavior during the demographic transition. *J. Interdiscip. Hist.* 42(4):543–69

Johnson DS, Massey C, O'Hara A. 2015. The opportunities and challenges of using administrative data linkages to evaluate mobility. *Ann. Am. Acad. Polit. Soc. Sci.* 657(1):247–64

Kesztenbaum L. 2008. Places of life events as bequestable wealth: family territory and migration in France, 19th and 20th centuries. In *Kinship and Demographic Behavior in the Past*, ed. T Bengtsson, GP Mineau, pp. 155–84. Dordrecht, Neth.: Springer

Kim K. 2016. Family demography in traditional Chosŏn Korea: survival strategies of families. In *A Global History of Historical Demography: Half a Century of Interdisciplinarity*, ed. A Fauve-Chamoux, I Bolovan, S Solver, pp. 411–33. Bern, Switz.: Peterland

Kim K, Park H. 2010. Family succession through adoption in the Chosun Dynasty. *Hist. Fam.* 4:443–52

King V, Elder GH. 1997. The legacy of grandparenting: childhood experiences with grandparents and current involvement with grandchildren. *J. Marriage Fam.* 59(4):848–59

Knigge A. 2016. Beyond the parental generation: the influence of grandfathers and great-grandfathers on status attainment. *Demography* 53:1219–44

Knodel JE. 2002. *Demographic Behavior in the Past: A Study of Fourteen German Village Populations in the Eighteenth and Nineteenth Centuries*. Cambridge, UK: Cambridge Univ. Press

Kolk M. 2014. Understanding transmission of fertility across multiple generations—socialization or socio-economics? *Res. Soc. Stratif. Mobil.* 35:89–103

Lawrence M, Breen R. 2016. And their children after them? The effect of college on educational reproduction. *Am. J. Sociol.* 122(2):532–72

LeBras H, Wachter KW. 1978. Living forbears in stable populations. In *Statistical Studies of Historical Social Structure*, ed. KW Wachter, EA Hammel, P Laslett, pp. 163–88. New York: Academic

Lee JZ, Campbell CD, Chen S. 2010. *China Multi-Generational Panel Dataset, Liaoning (CMGPD-LN) 1749–1909. User Guide.* Ann Arbor, MI: Interuniv. Consort. Political Soc. Res.

Lee JZ, Campbell CD, Wang F. 1993. The last emperors: an introduction to the demography of the Qing (1644–1911) imperial lineage. In *New and Old Methods in Historical Demography*, ed. DS Reher, R Schofield, pp. 361–82. Oxford, UK: Oxford Univ. Press

Légaré J, Lavoie Y, Charbonneau H. 1972. The early Canadian population: problems in automatic record linkage. *Can. Hist. Rev.* 53(4):427–42

Liu TJ. 1978. Chinese genealogies as a source for the study of historical demography. In *Studies and Essays in Commemoration of the Golden Jubilee of the Academia Sinica*, pp. 849–70. Taipei: Academia Sinica

Liu TJ. 1981. The demographic dynamics of some clans in the Lower Yangtze area, ca. 1400–1900. *Acad. Econ. Pap.* 9(1):115–60

Liu TJ. 1985. The demography of two Chinese clans in Hsiao-shan, Chekiang, 1650–1850. In *Family and Population in East Asian History*, ed. S Hanley, AP Wolf, pp. 13–61. Stanford, CA: Stanford Univ. Press

Liu TJ. 1995a. Demographic constraint and family structure in traditional Chinese lineages, ca. 1200–1900. In *Chinese Historical Microdemography*, ed. S Harrell, pp. 121–40. Berkeley: Univ. Calif. Press

Liu TJ. 1995b. Historical demography of south China lineages. In *Chinese Historical Microdemography*, ed. S Harrell, pp. 94–120. Berkeley: Univ. Calif. Press

Long J, Ferrie JP. 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850. *Am. Econ. Rev.* 103(4):1109–37

Mare RD. 2011. A multigenerational view of inequality. *Demography* 48:1–23

Mare RD. 2014. Multigenerational aspects of social stratification: issues for further research. *Res. Soc. Stratif. Mobil.* 35:121–28

Mare RD. 2015. Measuring networks beyond the origin family. *Ann. Am. Acad. Polit. Soc. Sci.* 657(1):97–107

Mare RD, Maralani V. 2006. The intergenerational effects of changes in women's educational attainments. *Am. Sociol. Rev.* 71:542–64

Mare RD, Song X. 2014. *Social mobility in multiple generations*. Work. Pap. PWP-CCPR-2014-014, Calif. Cent. Popul. Res., Univ. Calif., Los Angeles

Massey C. 2016. *Playing with matches: an assessment of accuracy in linked historical data*. Work. Pap., CARRA No. 2016–05, US Census Bur., June 10. **https://www.census.gov/library/working-papers/2016/adrm/carra-wp-2016-05.html**

Matras J. 1961. Differential fertility, intergenerational occupational mobility, and change in the occupational distribution: some elementary interrelationships. *Popul. Stud.* 15:187–97

McCall L, Percheski C. 2010. Income inequality: new trends and research directions. *Annu. Rev. Sociol.* 36:329–47

Mitnik PA, Cumberworth E, Grusky DB. 2016. Social mobility in a high-inequality regime. *Ann. Am. Acad. Polit. Soc. Sci.* 663(1):140–84

Modin B, Erikson R, Vågerö D. 2013. Intergenerational continuity in school performance: Do grandparents matter? *Eur. Sociol. Rev.* 29(4):858–70

Møllegaard S, Jæger MM. 2015. The effect of grandparents' economic, cultural, and social capital on grandchildren's educational success. *Res. Soc. Stratif. Mobil.* 42:11–19

Murphy M. 2006. The role of assortative mating on population growth in contemporary developed societies. See Billari et al. 2006, pp. 61–84

Neckerman KM, Torche F. 2007. Inequality: causes and consequences. *Annu. Rev. Sociol.* 33:335–57

Norton SL. 1980. The vital question: Are reconstituted families representative of the general population? In *Genealogical Demography*, ed. B Dyke, WT Morrill, pp. 11–22. New York: Academic

Olivetti C, Paserman MD. 2015. In the name of the son (and the daughter): intergenerational mobility in the United States, 1850–1940. *Am. Econ. Rev.* 105(8):2695–724

Olivetti C, Paserman MD, Salisbury L. 2016. *Three-generation mobility in the United States, 1850–1940: the role of maternal and paternal grandparents*, NBER Work. Pap. No. w22094, Natl. Bur. Econ. Res., Cambridge, MA

Oosten M, Mandemakers K. 2007. *GENLIAS release 2007_03, linked marriage certificates: five Dutch provinces (as a whole)*. Amsterdam: Int. Inst. Soc. Hist.

Park H, Lee S. 2008. A survey of data sources for studies of family and population in Korean history. *Hist. Fam.* 13:258–67

Piketty T. 2014. *Capital in the 21st Century*. Cambridge, MA: Harvard Univ. Press

Pfeffer FT. 2014. Multigenerational approaches to social mobility: a multifaceted research agenda. *Res. Soc. Stratif. Mobil.* 1(35):1–12

Pfeffer FT, Killewald A. 2015. *How rigid is the wealth structure and why? Inter- and multigenerational associations in family wealth*. Work. Pap. 15–845, Popul. Stud. Cent., Univ. Mich.

Post W, Van Poppel F, Van Imhoff E, Kruse E. 1997. Reconstructing the extended kin-network in the Netherlands with genealogies data: methods, problems, and results. *Popul. Stud.* 51(3):263–78

Preston SH. 1976. Family sizes of children and family sizes of women. *Demography* 13(1):105–14

Preston SH, Campbell CD. 1993. Differential fertility and the distribution of traits: the case of IQ. *Am. J. Sociol.* 98(5):997–1019

Reeves JH. 1987. Projection of number of kin. See Bongaarts et al. 1987, pp. 228–48

Reichman N, Teitler J, Garfinkel I, McLanahan S. 2001. Fragile families: sample and design. *Children Youth Serv. Rev.* 23(4/5):303–26

Reid A, Davies R, Garrett E. 2002. Nineteenth-century Scottish demography from linked censuses and civil registers: a 'sets of related individuals' approach. *Hist. Comput.* 14(1/2):61–86

Ruggles S. 1987. *Prolonged Connections: The Rise of the Extended Family in Nineteenth-Century England and America*. Madison, WI: Univ. Wis. Press

Ruggles S. 1992. Migration, marriage, and mortality: correcting sources of biases in English family reconstitutions. *Popul. Stud.* 46:507–22

Ruggles S. 2014. Big microdata for population research. *Demography* 51:287–97

Schofield R, Wrigley EA. 1979. Infant and child mortality in England in the late Tudor and early Stuart periods. In *Health, Medicine and Mortality in the Sixteenth Century*, ed. C Webster, pp. 61–95. Cambridge, UK: Cambridge Univ. Press

Sear R, Mace R. 2008. Who keeps children alive? A review of the effects of kin on child survival. *Evol. Hum. Behav.* 29(1):1–18

Servais MA. 2010. *Overview of HRS public data files for cross-sectional and longitudinal analysis*. Surv. Res. Cent., Inst. Soc. Res., Univ. Mich.

Sharkey P, Elwert F. 2011. The legacy of disadvantage: multigenerational neighborhood effects on cognitive ability. *Am. J. Sociol.* 116(6):1934–81

Silverstein M, Marenco A. 2001. How Americans enact the grandparent role across the family life. *J. Fam. Issues* 22(4):493–522

Slattery ML, Kerber RA. 1993. A comprehensive evaluation of family history and breast cancer risk: the Utah Population Database. *JAMA* 270(13):1563–68

Smith JE. 1987. The computer simulation of kin sets and kin counts. See Bongaarts et al. 1987, pp. 249–66

Smith JE, Oeppen J. 1993. Estimating numbers of kin in historical England using demographic microsimulation. In *Old and New Methods in Historical Demography*, ed. DS Reher, R Schofield, pp. 280–317. New York: Oxford Univ. Press

Smith KR, West JA, Croyle RT, Botkin JR. 1999. Familial context of genetic testing for cancer susceptibility: moderating effect of siblings' test results on psychological distress one to two weeks after BRCA1 mutation testing. *Cancer Epidemiol. Biomark. Prev.* 8:385–92

Solon G. 1999. Intergenerational mobility in the labor market. *Handb. Labor Econ.* 3:1761–800

Solon G. 2014. Theoretical models of inequality transmission across multiple generations. *Res. Soc. Stratif. Mobil.* 35:13–18

Solon G, Page ME, Duncan GJ. 2000. Correlations between neighboring children in their subsequent educational attainment. *Rev. Econ. Stat.* 82(3):383–92

Song X. 2016. Diverging mobility trajectories: grandparent effects on educational attainment in one- and two-parent families. *Demography* 53:1905–32

Song X, Campbell CD, Lee JZ. 2015. Ancestry matters: patrilineage growth and extinction. *Am. Sociol. Rev.* 80(3):574–602

Song X, Mare RD. 2015. Prospective versus retrospective approaches to the study of intergenerational social mobility. *Sociol. Methods Res.* 44(4):555–84

Song X, Mare RD. 2017. Short-term and long-term educational mobility of families: a two-sex approach. *Demography* 54:145–73

Szelényi I, Treiman DJ. 1994. *Social stratification in Eastern Europe after 1989: general population survey—provisional codebook*. Work. Pap., Dep. Sociol., Univ. Calif., Los Angeles

Taylor DP, Stoddard GJ, Burt RW, Williams MS, Mitchell JA, et al. 2011. How well does family history predict who will get colorectal cancer? Implications for cancer screening and counseling. *Genet. Med.* 13(5):385–91

Teerlink CC, Albright FS, Lins L, Cannon-Albright LA. 2012. A comprehensive survey of cancer risks in extended families. *Genet. Med.* 14:107–14

Telford T. 1990. Patching the holes in Chinese genealogies: mortality in the lineage populations of Tongcheng county, 1300–1880. *Late Imp. China* 11(2):116–36

Telford T. 1992. Covariates of men's age at first marriage: the historical demography of Chinese lineages. *Popul. Stud.* 46(1):19–35

Thorvaldsen G, Andersen T, Sommerseth HL. 2015. Record linkage in the Historical Population Register for Norway. In *Population Reconstruction*, ed. G Bloothooft, P Christen, K Mandemakers, M Schraagen, pp. 155–71. Dordrecht, Neth.: Springer

Torche F, Corvalan A. 2016. Estimating intergenerational mobility with grouped data (a critique of Clark's The Son Also Rises). *Sociol. Methods Res.* In press

Treiman DJ, Moeno S, Schlemmer L. 1996. *Survey of socioeconomic opportunity and achievement in South Africa—codebook*. Work. Pap., Dep. Sociol., Univ. Calif., Los Angeles

Treiman DJ, Walder AG. 1996. *Life histories and social change in contemporary China*. Soc. Sci. Data Arch., Univ. Calif., Los Angeles. Updated Jan. 4, 2016

Tremblay M, Vèzina H, Desjardins B, Houde L. 2008. Distant kinship and founder effects in the Quebec population. In *Kinship and Demographic Behavior in the Past*, ed. T Bengtsson, GP Mineau, pp. 259–77. Dordrecht, Neth: Springer

Wachter KW, Hammel EA, Laslett P, eds. 1978. *Statistical Studies of Historical Social Structure*. New York: Academic

Wagner D, Layne M. 2014. *The person identification validation system* (*PVS*): *applying the Center for Administrative Records Research and Applications'* (*CARRA*) *record linkage software*. CARRA Work. Pap. No. 2014–01, US Census Bur., July 1. **https://www.census.gov/content/dam/Census/library/working-papers/2014/adrm/carra-wp-2014-01.pdf**

Wang F, Lee JZ, Campbell CD. 1995. Marital fertility control among the Qing nobility: implications for two types of preventive check. *Popul. Stud.* 49(3):383–400

Warren JR, Hauser RM. 1997. Social stratification across three generations: new evidence from the Wisconsin Longitudinal Study. *Am. Sociol. Rev.* 62(4):561–72

Wightman P, Danziger S. 2014. Multi-generational income disadvantage and the educational attainment of young adults. *Res. Soc. Stratif. Mobil.* 35:53–69

Willigan JD, Lynch KA. 1982. *Sources and Methods of Historical Demography*. New York: Academic

Wilson EB, Doering CR. 1926. The elder Peirce's. *PNAS* 12(7):424–32

Winston S. 1932. Birth control and the sex-ratio at birth. *Am. J. Sociol.* 38(2):225–31

Wisselgren MJ, Edvinsson S, Berggren M, Larsson M. 2014. Testing methods of record linkage on Swedish censuses. *Hist. Methods J. Quant. Interdiscip. Hist.* 47(3):138–51

Wolf AP, Huang CS. 1980. *Marriage and Adoption in China, 1845–1945*. Stanford, CA: Stanford Univ. Press

Wright S. 1929. The evolution of dominance. *Am. Nat.* 63(689):556–61

Wrigley EA. 1997. *English Population History from Family Reconstitution 1580–1837* (No. 32). Cambridge, UK: Cambridge Univ. Press

Xie Y, Killewald A. 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850: comment. *Am. Econ. Rev.* 103(5):2003–20

Xie Y, Zhang C. 2016. *Children of the Cultural Revolution: multigenerational effects on educational outcomes in China*. Presented at Res. Comm. Soc. Stratif. Mobil. (RC28) ISA, Bern, Switz.

Zhao Z. 2000. Coresidential patterns in historical China: a simulation study. *Popul. Dev. Rev.* 26(2):263–93

Zhao Z. 2001. Chinese genealogies as a source for demographic research: a further assessment of their reliability and biases. *Popul. Stud.* 55(2):181–93

Zeng Z, Xie Y. 2014. The effects of grandparents on children's schooling: evidence from rural China. *Demography* 51(2):599–617

Annual Review of Sociology

Volume 43, 2017

Contents

Policy

Indexes

Errata

An online log of corrections to *Annual Review of Sociology* articles may be found at http://www.annualreviews.org/errata/soc