

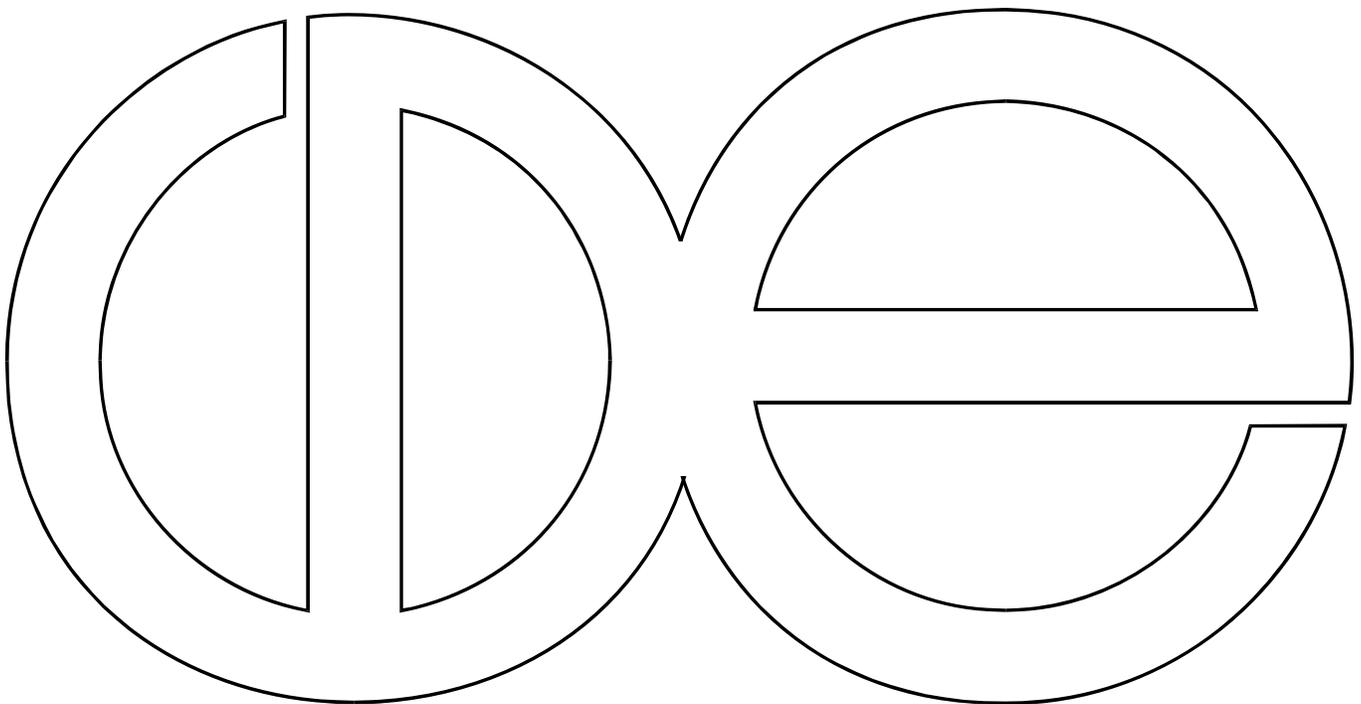
**Center for Demography and Ecology
University of Wisconsin-Madison**

**Recent Developments in Longitudinal Studies
of Aging in the United States**

Robert M. Hauser

David Weir

CDE Working Paper No. 2009-05



**Recent Developments in Longitudinal Studies
of Aging in the United States¹**

Robert M. Hauser

University of Wisconsin-Madison

David Weir

University of Michigan

September 1, 2009

¹ Prepared for the September 2009 Conference on Contemporary European and American Societies, Institute of European and American Studies, Academia Sinica, Taipei, Taiwan. Support for this research was provided by the National Institute on Aging (AG-9775 and AG-21079), the Vilas Estate Trust, and the Center for Demography of Health and Aging at the University of Wisconsin-Madison. The opinions expressed herein are those of the author. This paper partly reflects my collaborations and communications with Robert Willis and David Weir of the Health and Retirement Study and with Taissa S. Hauser, Carol Roan, and other staff and investigators of the Wisconsin Longitudinal Study. Parts of the text are drawn from Hauser and Willis (2005) and from a presentation by Weir to the RAND Summer Institute on the Demography of Aging, July 2009. Address correspondence to Robert M. Hauser, Department of Sociology, The University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706, USA, or E-MAIL to HAUSER@SSC.WISC.EDU

Two developments dominate the rapid evolution of large-scale, longitudinal studies of aging. First, there is now an international cohort of longitudinal studies of aging that one might better describe in terms of a network of investigators and samples than as a collection of discrete projects. Second – and equally importantly – in design and content they have increasingly become interdisciplinary and biosocial. In this essay, we focus on developments in the U.S. and do not attempt to cover international developments in a serious or comprehensive way. Also, the essay is mainly concerned with issues of survey design and method; we have made little attempt to review research findings.

In the first section of this paper, we briefly sketch the pace and social context of population aging. Next, we review some of the history and methodological features of large-scale, population-based longitudinal studies in the U.S. In the third section, following a brief description of the larger set of American studies, we focus on selected features of the design and content of two major, continuing longitudinal studies – the Health and Retirement Study and the Wisconsin Longitudinal Study.

Population Aging in the U.S. and around the World

In the decades following World War II, a mixture of high birth rates, improved survivorship, and subsequent declines in fertility has led to rapid and continuing population aging around the world. The U.S. is no exception. In 1950, only 8.3 percent of the population was aged 65 or older. By 1980, it was 11.2 percent; in 2010 it will be 13.0 percent; and by 2040 it is projected to be 20.4 percent. There is also rapid growth of the “oldest old,” those 80 years and older. In 1950, those 80 years old or older were just 13.3 percent of the older population. By 1980, they were 21.4 percent and by 2010 they will be 29.3 percent of the older population. And

it is projected that by 2040, more than one third of the older population, 35.5 percent, will be 80 years or older (National Institute on Aging and U.S. Census Bureau 2009: Table B-1, pp. 138-39). One highly visible signal of population aging in the U.S. is the forthcoming entry of the first of the baby boom cohorts into the “over 65” category, which will occur in 2011. Those cohorts have strained institutions and organizational structures throughout their life course, and those stresses will persist as the baby boomers age. The survivors of the “Jamesgu” and “Marys” who once overfilled kindergartens and elementary schools in the 1950s will place new demands on senior services – the Social Security, Medicare, and Medicaid systems; centers for independent and assisted living, nursing homes, and hospitals; and their family and community support systems.²

To be sure, many other nations – especially in the developing world – are on cp"even faster aging trajectory. For example, between 2008 and 2040, the U.S. is projected to experience a 107 percent increase in the population aged 65 and over, but that increase will place the U.S. in the middle rank among nations. At least 20 developing nations are expected to experience increases of 200 percent or more in their older populations in the same period (National Institute on Aging and U.S. Census Bureau 2009: 13).

A recent overview of population aging offered nine “challenges and opportunities” that show why population aging matters (National Institute on Aging and U.S. Department of State 2007: 3):

² In 1946 James and Mary were the most popular American baby names <<http://www.ssa.gov/OACT/babynames/index.html>>.

- “The overall population is aging. For the first time in history, and probably for the rest of human history, people age 65 and over will outnumber children under age 5.
- Life expectancy is increasing. Most countries, including developing countries, show a steady increase in longevity over time, which raises the question of how much further life expectancy will increase.
- The number of oldest old is rising. People age 85 and over are now the fastest growing portion of many national populations.
- Noncommunicable diseases are becoming a growing burden. Chronic noncommunicable diseases are now the major cause of death among older people in both more developed and less developed countries.
- Some populations will shrink in the next few decades. While world population is aging at an unprecedented rate, the total population in some countries is simultaneously declining.
- Family structures are changing. As people live longer and have fewer children, family structures are transformed, leaving older people with fewer options for care.
- Patterns of work and retirement are shifting. Shrinking ratios of workers to pensioners and people spending a larger portion of their lives in retirement increasingly strain existing health and pension systems.

- Social insurance systems are evolving. As social insurance expenditures escalate, an increasing number of countries are evaluating the sustainability of these systems.
- New economic challenges are emerging. Population aging will have dramatic effects on social entitlement programs, labor supply, trade, and savings around the globe and may demand new fiscal approaches to accommodate a changing world.”

Had it been published little more than a year later, these challenges probably would have been elaborated or supplemented to include major changes in economic activity that affect older populations, such as the meltdown in retirement savings and loss of jobs accompanying the current, worldwide economic recession.

Heterogeneity is a key feature of population aging in the U.S. One obvious aspect is variation in educational and economic standing among the elderly, which affect health and well-being and access to social institutions and supports for them. There are also important variations in population aging among major demographic groups. For example, Hispanics – the fastest growing ethnic population in the U.S. – are presently a young population. Only 4.9 percent of Hispanics were aged 65 or above in 2000, and of these only 19.2 percent were 80 years or above. The white population was far older; 13.6 percent were aged 65 or above, and of those 27.0 percent were 80 or above. The black population was intermediate in aging between whites and Hispanics in 2000; 8 percent were aged 65 or above, and of those, 23.4 percent were aged 80 or above. By 2040, the U.S. Census Bureau projects that 12.2 percent of Hispanics will be aged 65 or above, and 25.6 percent of those will be 80 years or older. That is, by 2040, the Hispanic

population will be almost as old as the white population was four decades earlier. However, the white population will continue to age, and by 2040 21.2 percent of whites are projected to be aged 65 or above, among whom 35.3 percent will be aged 80 or above. In 2040, the black population is projected to remain intermediate in aging between whites and Hispanics with 16.9 percent aged 65 or above and 29.9 percent of those aged 80 or above (U.S. Bureau of the Census 2008). As there will be large race-ethnic differences in the share of the population at higher ages, so the demands for supporting social, economic, and health resources will vary across those groups. One might expect, as well, that those differences in population composition will lead to race-ethnic differences in political priorities for support of the elderly.

Elderly immigrants are a related source of heterogeneity in aging in the U.S. In addition to the aging of working-age immigrants, the growth of the older, immigrant population is fed by a policy that permits immigration of elderly family members of legal migrants. Thus, many immigrant adults bring their parents to the U.S. There are now nearly 4.5 million immigrants who are more than 65 years old, who comprise 11.8 percent of the older population of the U.S. They are heavily concentrated in California, New York, and Florida, where more than half of older immigrants live. They are of very heterogeneous national origins. Only about twenty percent were born in Mexico or Cuba, and a larger share is from Europe or the former Soviet Union. Other countries contributing large numbers of older migrants are China, the Philippines, India, El Salvador, Vietnam, and Korea (Terrazas 2009).

Another important and potentially influential source of heterogeneity in aging in the U.S. is geographic. The extent of population aging varies widely across the country. Figure 1 shows the percentage of the population at ages 65 and above by state, based on three years of data from

the American Community Survey, 2005-2007. Utah, Colorado, Texas, and Georgia are relatively young, while Florida, a destination for many retirees, and several other states – North Dakota, Iowa, West Virginia, Pennsylvania, and Maine – have much larger shares of elderly. In the latter states, aging is more a consequence of out-migration by the young than of in-migration of the old. Variation in the location of persons aged 85 years and over follows much the same pattern.³

Geographic heterogeneity of population aging is even more striking in Figure 2, which shows the geographic distribution of the percentage of individuals aged 65 and above by county. Two relatively young states – Nevada and Arizona – have several counties in which the share of the elderly is quite large. Also, that the variation in the share of elderly is far larger among counties than among states. In Figure 1, the data classes cover the range from 6.7 to 16.9 percent, while in Figure 2 they cover the range from less than 3.8 percent (the large gray areas) to 31.5 percent. These and other aspects of heterogeneity in population aging make for complex problems in studying as in providing goods, services, and public and familial support for the older population of the United States.

Longitudinal Population Studies

Large scale, long-term longitudinal studies of human populations are a relatively new social invention. In the U.S., for example a handful of small long-lived life-course studies was undertaken early in the last century – notably Lewis Terman’s studies of highly able youth (Terman 1925; Burks et al. 1959; Terman and Oden 1959a; Terman and Oden 1959b; Oden 1968; Martin and Kubzansky 2005) and the studies of Oakland and Berkeley, California,

³ These data and maps were obtained from the U.S. Bureau of the Census (http://factfinder.census.gov/servlet/ACSSAFFPeople?_submenuId=people_3&_sse=on).

samples that are best known for the work of Glen Elder and John Clausen (Elder 1974; Clausen 1993). A 1991 compendium of American longitudinal studies consists mainly of studies that are recent, small, highly specialized, or short-lived (Young et al. 1991). A more recent collection of autobiographical retrospectives about American longitudinal studies exhibits similar limits (Phelps et al. 2002). Just two of eleven studies were national in scope and otherwise unselective – the National Longitudinal Studies of Labor Market Experience and the Panel Study of Income Dynamics – but both of them were initiated only in the late 1960s.

Up to this point, we have left the subject of this essay undefined, but we would suggest five criteria for a major longitudinal study:

1. The study must be based on one or more well-defined populations and be regional or national in scope.
2. There must be a large number of cases, at least several thousand, and the investigators must be able to retain them in the study.
3. The study must cover a wide window in time, at least a decade and preferably longer.
4. The study must measure key variables across multiple content domains.
5. The study must repeat key measurements or obtain retrospective histories of key variables.

In addition, we will add that for any such social-scientific study to be considered “major,” the data must be public – or at least readily accessible to all qualified researchers. There are two reasons for this. The first is that the cost of major longitudinal studies is very high, so it is cost-effective for them to be widely used. The second is that real scientific findings must be

reproducible by any qualified researcher (National Research Council 1985; Hauser 1987; National Research Council and Research Data in a Digital Age 2009). Some or all of the data from each of the studies mentioned below is in the public domain and available from a project or archival website.

As recently as the late 1950s, most social scientists in the U.S. simply did not believe that it would be possible to retain a large and representative sample across a long period of time, and – despite the ongoing example of the remarkable British birth cohort study of 1946 (Douglas and Blomfield 1958; Douglas 1964; Douglas et al. 1968; Wadsworth 1991) – no such studies were undertaken. Indeed, the U.S. has still not succeeded in creating a successful, large-scale longitudinal birth cohort study, though the British have repeated their success with new cohorts at 12-year intervals. The nascent National Children’s Study may at last fill that gap – with a sample of some 100,000 births – but there remain serious problems with its study design (National Research Council (U.S.). Panel to Review the National Children's Study Research Plan. et al. 2008).

William H. Sewell’s seven-year follow-up of 10,317 Wisconsin high graduates of 1957 (Little 1958; Little 1959) provided a “proof of concept” for long-term panel studies. In 1962, Sewell learned that the 1957 survey schedules and punch cards were sitting unused in the University administration building. Sewell had long been interested in the formation and consequences of youthful aspirations, but he had lacked access to an appropriate population for study. At that time, social scientists had little real evidence about the extent of social and economic mobility between generations in the United States. Only in 1962 was the first large national study of social mobility in America conducted (Blau and Duncan 1967). Researchers

could do little more than speculate about the processes of selection and socialization that accounted for social stability or social movement.

Sewell selected a random, one-third sample of the graduates, consisting of 10,317 cases, for further study. This was possible because the original questionnaires included the students' names and the names and addresses of their parents. Moreover, Sewell was able to locate IQ test scores for all of the students and tax records of parents' incomes for about 90 percent of them. Reasoning that, seven years after high school graduation, the graduates themselves would have little inclination to participate in a survey, Sewell chose instead to survey their parents. His theory was that parents were always interested in talking about their children. He designed a very brief survey instrument, which fitted on one side of a folded post-card. Two of the other sides were outgoing and incoming addresses, and the fourth contained a letter of introduction. The survey ascertained an educational history, occupation, marital status, military service, and – for women only – husband's occupation.

After five waves of mailing, diligent address searches, and a telephone follow-up of non-respondents, Sewell and his colleagues achieved an 87.2 percent response rate. This was unprecedented at the time, and it encouraged others to undertake large longitudinal studies of youth. Thus began the decennial series of national longitudinal studies of high school students – the National Longitudinal Study of the High School Class of 1972, High School and Beyond, and the National Educational Longitudinal Study of 1988 – and the massive, but less successful Project Talent (Flanagan et al. 1966; Lohnes et al. 1966; Cureton and American Institute of Education 1968; Carrel et al. 1975). By design, the national longitudinal studies – undertaken by the U.S. Department of Education – were each terminated with the completion of the cohort's schooling

or soon thereafter. Project Talent also ended in less than two decades because of heavy sample attrition, but there is now a realistic possibility that it will be resurrected as a study of aging. Preliminary work suggests that it will be possible to trace surviving participants from 1960 to the present time.

Given a focus on aging, social structure, and health, what are the major longitudinal studies in the U.S.? What makes them “major”? How do their design and content compare? What are their strengths and weaknesses? And what, in the national context, is an appropriate number and mix of such studies? Five basic survey designs may yield useful longitudinal data:

- One-time cross-sections
- Repeated cross-sections
- Short longitudinal
- Long longitudinal
- Repeated longitudinal

One-time cross-sectional surveys may yield true or synthetic longitudinal data. Retrospective reports yield the former, and the validity of such data depends heavily and contingently on the characteristics of the sample, the content and detail of material to be recalled, and the quality of the survey protocol. Often, retrospective data are no worse than contemporary data, for example, in the case of people’s reports of the educational attainments and occupations of their parents (Bielby et al. 1977). By synthetic data, we refer to the classification of survey responses by age, that is, to synthetic cohort analysis, which depends on the oft-violated theory that each successive cohort recapitulates the experience of its predecessors.

Provided that survey coverage, methods, and measures are comparable across time, repeated cross-section surveys can provide very valuable, but limited longitudinal data. Under these conditions, one may track the means of variables across time, and – given some explanatory covariates that do not change – one can track changes in the effects of those variables across time. For example, Hauser and his collaborators have used dozens of successive October Current Population Surveys to study trends and differentials in the progression of cohorts through the educational process (Hauser et al. 2004a; Hauser et al. 2004b; Frederick and Hauser 2008).⁴ The limitation of such studies is that – absent perfect retrospection – they cannot provide covariances among variables measured at different times. Thus, in the case of the October CPS, each survey includes very limited and possibly unreliable information about prior enrollment and grade retention, but no full history of educational enrollment and progress. Moreover, longitudinal analyses depend on the assumption, which is partly false, that social background characteristics like parents' educational attainments, occupations, or incomes and family structures do not change over time. That is, in looking at relationships of educational outcomes to social background within the same cohort but across occasions, one might well ask the extent to which observed changes in those relationships are attributable to changes in the dependence of education on background and to changes in the background variables themselves. In principle, this limitation would not hold in the case of presumably unchanging characteristics like race-ethnicity and gender, but reports of youth's reports of their race-ethnicity are known to vary across social settings (Harris and Sim 2002).

⁴ The Current Population Survey (CPS), carried out by the U.S. Bureau of the Census, is the monthly household survey from which the Bureau of Labor Statistics obtains a key economic indicator, the unemployment rate. Each October, the CPS includes a special supplement on school enrollment.

Short and long longitudinal surveys share some of the same analytic characteristics. Assuming that attrition is neither selective nor excessive, it is possible to do everything that one can do with repeated cross-sectional surveys – within the covered cohort(s). But one can do more, by analyzing the dynamics of relationships among all variables across time. Some longitudinal surveys are short by design because they are intended to cover a specific set of events or a specific interval of the life course. Such is the case for the several national longitudinal surveys of educational attainment in the U.S. that we have already mentioned. Another short-term longitudinal survey – whose value is largely unexploited – is the Current Population Survey. Its rotation group structure is designed so each sample household remains in the CPS for the same four months in two successive years. The main purpose of that design is to reduce sampling variance in trend measures, but it also permits studies of short-term (one-year) changes in social and economic characteristics. Another, successful short term study is the Early Childhood Longitudinal Study – Kindergarten Class, supported by the Department of Education to provide data on early school experiences. It followed a large cohort of children from kindergarten in the 1998-99 school year through the 8th grade.

MIDUS (Midlife in the United States) is a short-term longitudinal study of aging and the life course that may eventually turn into a long-term study (Brim et al. 2004). The first wave of MIDUS was based on a national telephone survey of American adults in 1995-96, which was developed with support from the MacArthur Foundation. In 2004-06, a second wave of MIDUS was carried out with major support from the National Institute on Aging (<http://midus.wisc.edu/>). It includes a large sample of twins and non-twin siblings and an oversample of older adults. It is very rich in variables related to health and aging, with a major emphasis on the measurement of

psychological variables, both cognitive and non-cognitive. Subsamples have provided comprehensive biomarkers, cortisol samples, and neurological assessments. However, even in the first wave, the MIDUS response rate was relatively low, leading to a substantial upward bias in the socioeconomic composition of the sample, and attrition was a serious problem between the first and second waves. Thus, researchers should be very careful in analyzing public data from MIDUS, especially when population representation is at issue.

The key factors distinguishing short- and long-term longitudinal studies are more operational than analytic. Can research support be sustained over a long period of time? Can the sample be retained over a long period of time? Can participants be traced, and will they continue to volunteer their time and information? Is the sample originally large enough to be useful after a substantial share of participants have died, migrated, or left the study for other reasons? We have already mentioned one very long-term, large-scale longitudinal study, the WLS, and there are now several successful examples in the U.S.⁵

Another positive exemplar is the Panel Study of Income Dynamics (PSID), which has followed a cohort of households since 1968, including descendants and spouses or partners of the original household members. The sample has grown so large that it has occasionally been necessary to trim its size as well as to refresh the base to include new immigrants. The PSID has become a rich resource for studies of aging and the life course. Data access and documentation for the PSID have improved over the years, so it no longer deserves the reputation it once had for

⁵ Kuo et al. (2001) provide extensive, but dated comparisons of the design and content of the Wisconsin Longitudinal Study, Midlife in the United States, the Health and Retirement Study, the National Survey of Families and Households, and American's Changing Lives. The report is available at <http://www.ssc.wisc.edu/cde/cdewp/2001-06.pdf>.

being especially difficult to use. The National Institute on Aging has provided support for use of the PSID in studies of pensions, retirement planning, and financial well-being over the life course. Since the PSID attempts to maintain representation of the U.S. population across time (Fitzgerald et al. 1998a; Fitzgerald et al. 1998b), one might think of it as having some of the advantages of independent, repeated cross-sectional studies as well as those of a longitudinal cohort study. Also, for many years the PSID has been paralleled by a German sibling, the German Socio-Economic Panel (GSOEP), and by the more recent Panel Study of Family Dynamics (PSFD) in Taiwan, thus encouraging comparative analysis.

American's Changing Lives (ACL) is another national longitudinal study with a strong emphasis on aging. Its four waves – 1986, 1989, 1994, and 2002 – covered a wide range of sociological psychological, mental, and physical health items. The ACL research program was

“designed to investigate the following: (1) the ways in which a wide range of activities and social relationships that people engage in are broadly "productive," (2) how individuals adapt to acute life events and chronic stresses that threaten the maintenance of health, effective functioning, and productive activity, and (3) sociocultural variations in the nature, meaning, determinants, and consequences of productive activity and relationships”

(<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/04690/>).

The National Survey of Families and Households is another multi-wave study that may be useful in research on aging in the United States. Its three waves – 1987-88, 1992-93, and 2001-02 provided a unique store of data on American Families because the study was designed to interview multiple members of each sample household. It began with 13,017 persons aged 19

or more, and it provides extensive self-reported accounts of relationships among spouses, children, and parents. Unfortunately, the NSFH has ended, and despite some effort its supporting agency, the National Institute of Child Health and Human Development, has not yet developed a comparable vehicle to monitor family structure and change in the U.S.

The National Social Life, Health, and Aging Project (NSHAP) is a new and soon-to-become longitudinal population study supported by the National Institute on Aging (<http://www.norc.org/nshap>). It began in 2005-06 with about 3000 interviews of a nationally representative sample of adults aged 57 to 85. The investigators, led by Linda Waite, are at the University of Chicago, and fieldwork is conducted by the National Opinion Research Center (NORC). The NORC website describes the study as

“the first population-based study of health and social factors on a national scale, aiming to understand the well-being of older, community-dwelling Americans by examining the interactions among physical health, illness, medication use, cognitive function, emotional health, sensory function, health behaviors, and social connectedness. ... Data collection elicited: 1) demographic characteristics; 2) social networks; 3) social and cultural activity; 4) physical and mental health including cognition; 5) well-being; 6) illness; 7) medications and alternative therapies; 8) history of sexual and intimate partnerships; and 9) patient-physician communication.”

In fact, what is most distinctive about NSHAP is its focus on the previously neglected subject of sexual behavior among the elderly. One might think of it as an extension of the earlier, highly controversial, and privately funded survey of American sexual behavior, also carried out by the

University of Chicago and NORC (Laumann 2000; Laumann and Michael 2001). The earlier study had failed to include older participants.

Repeated longitudinal panels – provided they are truly comparable – combine all of the benefits of the other four designs and are yet more difficult to carry off. Successful examples are rare. One major – and less than successful – repeated short term study is the Survey of Income and Program Participation (SIPP). This general household survey, carried out by the U.S. Bureau of the Census, was originally designed to measure short-term changes in economic standing and participation in social welfare programs. From 1984 through 1993, each annual panel lasted for three years, and interviews were carried out every four months (National Research Council 1993). After that, the panels were designed to last four years and were not refreshed annually. The major reasons for these changes were the burden of maintaining multiple panels in the field simultaneously and the difficulty of weighting and analyzing data created using that design. In principle, SIPP could provide current measures of poverty based on far more detailed information than that obtained in the current official source, the March Current Population Survey. However, attrition was so great after the redesign, especially among the poor and racial-ethnic minorities, that both cross-sectional and longitudinal validity was compromised, and SIPP is now being redesigned for a second time (National Research Council 2009).

The four national longitudinal studies of secondary school students carried out by the Department of Education follow the plan of repeated longitudinal designs. Much, but not all of the content in these large-scale studies has been comparable from one panel to the next. The major problem is that each panel has been abandoned at or shortly after most members of the

cohort reached adulthood. Thus, the particular interests and obligations of the Department of Education prevented these studies from spanning the life course.

In the late 1960s, a research group of economists at Ohio State University initiated a quartet of longitudinal studies, all based on households that had been retired from the Current Population Survey (<http://www.bls.gov/nls/home.htm>). They were targeted at specific population groups that were believed to have distinctive labor force experiences: Young men aged 14 to 24 when first interviewed in 1966 (discontinued in 1981 because of high attrition); young women aged 14 to 24 when first interviewed in 1968 (discontinued in 1993); mature women aged 30 to 44 when first interviewed in 1967 (discontinued in 1993); and older men aged 45 to 49 when first interviewed in 1966 (discontinued in 1990). Experience with these four studies led to the development of a new youth cohort survey, the National Longitudinal Study of Youth 1979 (NLSY79), which sampled all youth within each sample household who were 14 to 21 years old at the time of the initial survey. Unfortunately, the design of this youth cohort study was not strictly comparable to its predecessor, but it was also far more comprehensive. For example, the earlier study had attempted to collect and calibrate test score data from schools attended by the youthful participants, while the 1979 study administered all participants the same comprehensive test battery – the Armed Services Vocational Aptitude Battery (ASVAB) in 1980. However, even this level of standardization created analytic problems because of the differing ages at which participants took the assessment and because some participants had left school before taking it (Neal and Johnson 1996). The NLSY79 continues with bi-annual interviews now that the surviving participants are at midlife, 44 to 52 years old, and we hope that this sample will be followed to extinction. There is also an important spin-off of the NLSY79, a longitudinal study

of births to women in the sample. Children of the NLSY somewhat resembles a true birth cohort study, but a number of biases are introduced by dint of the fact that births occurred selectively over a period of many years. A third youth panel – now in its tenth annual wave – was begun with 12 to 17 year olds in 1997. While the design and content of the three successive youth panels are admirable in many ways, new investigators joined the project at each juncture, and they emphasized extensions of content and improvements in design over comparability from one period to the next.

One repeated longitudinal study, supported by the National Institute on Aging, is no longer in existence and both were of relatively short duration. There were two panels of the Longitudinal Study of Aging (LSOA I and II), both drawn from participants in the National Health Interview Survey. LSOA I covered 7,500 persons aged 70 and above and interviewed them in 1984, 1986, 1988, and 1990. According to the CDC (Centers for Disease Control) website (<http://www.cdc.gov/nchs/lsoa/lsoa1.htm>), it was intended to

- “make data on the oldest-old and on people moving into that age group available to the research community,
- describe the continuum from functionally independent living in the community, through dependence, including institutionalization, to death,
- measure change in the functional status and in the living arrangements of older people,
- provide mortality rates for demographic, social, economic, and health characteristics that are not available from the vital statistics system, and
- provide measures of health care use for individuals over time.”

In addition to the biennial interviews, LSOA matched Medicare data and the National Death Index, thus obtaining multiple cause of death records. LSOA II replicated the design and content of LSOA I with older participants in the 1994 National Health Interview Survey. Follow-up interviews of 9,400 participants were carried out in 1997-98 and in 1999-2000. According to the CDC (<http://www.cdc.gov/nchs/lsao/lsao2.htm>), “The LSOA II data, when used in conjunction with data from the original LSOA, enables researchers to determine whether the prevalence and incidence of functioning, pathology, and impairments in the elderly population have changed over 10 years and whether the change is due to differences in cohort characteristics or to technological and medical advancements. This provides researchers and policy planners with an opportunity to examine trends in ‘healthy aging,’ the determinants of these trends, the differences in these trends in two cohorts of older Americans, and their social and economic consequences.” For reasons that are not at all clear to me – perhaps the advent of the Health and Retirement Study and the closely related content and purposes of the National Long Term Care Survey – or the relatively short duration of the two LSOA panels, those data have not been heavily used.

The National Long Term Care Survey (NLTC) began with a sample of nearly 36,000 persons drawn from national Medicare enrollment files in 1982 (<http://www.nlts.aas.duke.edu/>). Supported by the National Institute on Aging, it was later augmented with some 20,000 new Medicare enrollees (aged 65) in four refresher samples at five-year intervals. Field work has been carried out by the U.S. Bureau of the Census, even though the NLTC was designed and run by a research group at Duke University. Response and retention rates have been exceptionally high, and the NLTC has been diligent in following individuals in institutions,

e.g., nursing homes, as well as those in the community. The study is perhaps best known for its finding that disability among American elderly has steadily declined. Because of problems with the management of the study, the National Institute on Aging ran a fresh competition for the NLCTS in 2009. The award was won by a research group at Johns Hopkins University. It has proposed to create an entirely new sample with fieldwork by WESTAT, a large, nonprofit survey firm, rather than continuing the existing panel after 2004. The main reason for abandoning the surviving panel members is that Duke University and the U.S. Bureau of the Census had established consent agreements with participants that required renewed consent from each participant – a most daunting task – if the study were ever moved to another research or survey organization.

The Health and Retirement Study – of which we will say more later – provides continuous longitudinal coverage of the U.S. population aged 50 and above. It is the “Cadillac” of longitudinal studies of aging in the U.S. It began with a sample of about 12,000 individuals – both spouses among married couples – who were aged 51 to 61 in 1992. Later, it added two older cohorts (AHEAD and CODA) in order to complete coverage of older Americans, thus growing to a steady state of about 20,000 participants. It has subsequently added younger refresher cohorts, and it over-samples minorities and Floridians. HRS has also established links with data from the Social Security Administration, the National Death Index, Medicare data, and employer records. Coverage was excellent in the initial wave, and sample retention has also been good. Data cover cognition, health, and a variety of attitudes and preferences, but the greatest emphasis is on careful and detailed economic measurement. Data on upbringing and on events before the age of 50 are retrospective or thin in HRS. Spousal reports provide important relational content,

but there are no data from siblings or children. There have been many special-purpose supplements and continuing investments in the content of the surveys and in the quality of measures.

This overview would not be complete without brief mention of two other types of studies, epidemiologic panel studies and historical studies. The former include the Alameda County (California) Health and Ways of Living Study, which “collected information for 6,928 respondents (including approximately 500 women aged 65 years and older) on chronic health conditions, health behaviors, social involvements, and psychological characteristics. The 1974 questionnaire was sent to 6,246 living subjects who had responded in 1965, and were able to be located. A total of 4,864 individuals responded in 1974. Questions were asked on marital and life satisfaction, parenting, physical activities, employment, and childhood experiences.

Demographic information on age, race, height, weight, education, income, and religion was also collected” (<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/06838>). The original survey covered a stratified random sample of Alameda County residents who were 21 or older, plus married residents older than 16. Third, fourth, and fifth waves of data were collected in 1994, 1995, and 1999, but by 1999 there were only 2,123 remaining participants (<http://www.icpsr.umich.edu/cocoon/NACDA/STUDY/04432.xml>). Even with allowance for mortality, panel attrition appears to have been very high across the 34 years of the panel.

Another important epidemiologic panel study was entitled Established Populations for Epidemiologic Studies of the Elderly. It was a multi-site study, covering “all noninstitutionalized persons 65 years of age and older in East Boston, Massachusetts, New Haven, Connecticut, Iowa and Washington Counties, Iowa, and five counties in north central North Carolina.” Its goals

were “to describe and identify predictors of mortality, hospitalization, and placement in long-term care facilities and to investigate risk factors for chronic diseases and loss of functioning” (<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/09915>). In addition to the baseline survey, there were six follow-ups as well as mortality close-outs. Among other things, data from this study were used to propose that allostatic load, a “cumulative measure of biological dysregulation” was implicated in mortality and accounted for about a third of socioeconomic differentials in mortality (Seeman et al. 2001; Seeman et al. 2004).

A team led by the eminent economic historian, Robert Fogel, labored for decades to create a longitudinal database of information about soldiers in the Union Army – who served in our 19th century civil war – that now includes about 40,000 men. The file now includes military records, pension records – which are rich in health data – and census records, and the study has so far yielded novel findings about trends in nutritional status, health, survival, and aging processes (Fogel 1993; Fogel 1994; Fogel and Costa 1997; Costa and Kahn 2003; Fogel 2004; Costa and Kahn 2009).

Recent Developments in the HRS and the WLS

We turn now to a selective account of recent developments in the Health and Retirement Study (HRS) and the Wisconsin Longitudinal Study (WLS) – the latter because Hauser leads and is most familiar with its strengths and weaknesses and the former because Weir leads it, and the HRS is the premier longitudinal study of aging in the U.S.⁶ Both studies have been supported by

⁶ Hauser also serves on a scientific advisory committee of the Health and Retirement Study.

the National Institute on Aging for almost two decades, but they have different histories and overlapping, but distinct content and uses.

Almost five decades ago, when a professor of education at the University of Wisconsin-Madison carried out a survey of high school seniors to assess the demand for postsecondary schooling, no one would have imagined that it would create the foundation for a major study of the life course and aging. The WLS began in large part because that survey provided a complete, identifiable list of the population. That sample was followed up successfully in 1964 and 1975 – and expanded to include randomly selected siblings in that wave. Only in 1992 to 1993, when the graduates were 53 and 54 years old, was the content of the surveys expanded to shift the focus of the study from education, careers, and family to health and aging. The fourth wave of the WLS was carried out in 2004-06, as the graduates began to reach retirement age, and the sample was then expanded to include the spouses or widows of graduates and of their siblings. A new wave of home interviews with graduates and siblings is now in preparation for fielding in the second quarter of 2010. This is a major departure from the methodology of past waves of the WLS; all earlier follow-ups have been by telephone and or mail. Face-to-face (CAPI and CASI) interviews will make it possible to overcome the likely decline in hearing of many participants, to facilitate more extensive and intensive cognitive measurements, and to permit direct anthropometric measurement.⁷

In contrast, the Health and Retirement Study was created in a deliberate effort to establish a longitudinal observatory for multi- and interdisciplinary studies of aging populations (Juster

⁷ For reviews of the past design and content of the WLS, see Sewell et al. (2004), Hauser (2005; 2009b), Hauser and Willis (2005), and Huang and Hauser (2009).

and Suzman 1995), but its design and coverage, also, has evolved across time. The HRS design began with a decision to represent the U.S. population in 1992 at ages 51 to 61 (and their spouses of whatever age). Because only 16.6 percent of households contained people in this age range, a large screening survey was undertaken by the Survey Research Center at the University of Michigan. The 1992 screening survey was also used to identify other cohorts that have become part of the HRS: the AHEAD sample (born in 1890 to 1923), the CODA sample (born in 1924 to 1930), and “war babies” (born in 1942 to 1947). Thus, the HRS currently represents all cohorts of Americans born between 18; 0 and 1947.

In 2004, a fresh screening sample was drawn to add a cohort of “Early Boomers” born in 1948 to 1953; here the household eligibility rate was even lower, about 12.4 percent. This effort generated a random sample of younger cohorts that will be used in the 2010 wave of the HRS, and those in older cohorts were selected for the National Social Life, Health and Aging Project, which is also supported by the NIA. Other older participants in the screening sample were asked to join a project on measurement of cognition that is being used to “reengineer” HRS cognitive measures using state-of-the-art ability tests and psychometric methods. Finally, a small number of cases were retained for use as a “test bed” sample for use by the HRS.

Whatever else may be problematic in the repertoire of aging surveys that are sponsored in whole or part by NIA, they all start with probability samples. This has great advantages. One can generalize to each of the populations that has been sampled, whether it is national or regional, of the whole adult population or of specific age cohorts.⁸ Moreover, if they are large enough, these

⁸ In Europe, surveys like the HRS are now underway, notably, ELSA (English Longitudinal Study of Ageing) and SHARE (Survey of Health, Ageing, and Retirement in Europe). Fortunately, each of these studies uses probability sampling, thus overcoming a tradition of quota sampling in several of the participating countries.

population samples may be used as frames for the study of special populations. For example, ADAMS (Aging, Demographics, and Memory Study) has identified for intensive study a stratified random sample of about 850 persons from the Health and Retirement Study, based on survey measures of cognitive impairment. These individuals have had in-home visits by a trained nurse and psychometric technician to obtain clinically valid measures of dementia.

Similarly, Marsha Seltzer and Jan Greenberg's study of nonnormative parenting outcomes within the Wisconsin Longitudinal Study has identified several hundred participants (high school graduates or their siblings) for intensive study if one or more of their children have a developmental disability or severe mental illness. In each of these cases, the base samples are large enough to generate as many or more cases for intensive study than one often finds in clinical research, but the subpopulations are also representative of all instances meeting the definition of the special subpopulation. Another example of this is simply that the baseline samples, both in the HRS and the WLS, are of populations large enough to yield a substantial sample, many years later, of participants at advanced ages.

Longitudinal observation of the same individuals is necessary, but not sufficient to provide a useful base for analyses of change in the process of aging. Thus, while studies of unique cohorts, like the WLS graduates, are of great scientific value, they do not provide a sound basis for studies of inter-period change, nor do they permit accumulation of age-specific data for scientifically significant subpopulations in different cohorts.⁹ Moreover, unique cohort studies

⁹ Because the WLS covers siblings as well as the 1957 graduates, it is possible to carry out some intercohort comparisons within the study population. However, such comparisons necessarily pertain to differently selected populations. For example, graduates must have completed high school, but their siblings need not have done so. Graduates may be singletons, but siblings cannot be singletons.

are likely to become obsolete as the world changes and as policy issues and scientific methods evolve. More generally, a commitment to the collection of data on a succession of cohorts permits researchers to study the impact of events—changes in laws and policies, new diseases and treatments, variations in economic conditions—that cannot be foreseen by the survey’s designers. This is one of the great strengths of the HRS.

Recent Developments in the HRS

It is difficult to underestimate the scope and influence of the Health and Retirement Study. From 1992 through 2008 more than 30,000 people were interviewed at least once, 155,000 interviews were completed, 7,500 workers retired, 10,000 participants died (leading to 9,000 interviews with survivors), 22,000 people were linked to their Social Security records, and 17,000 people were linked to their Medicare records. As of the end of 2008, the HRS had yielded more than 1,400 publications, and there has been a steady acceleration in the rate of publication. HRS publications reflect the work of more than 1,000 authors or co-authors and more than 10,000 registered data users.

Moreover, the success and productivity of the HRS has influenced the development of large-scale, longitudinal studies of aging in many other nations. These include the Mexican Health and Ageing Study (MHAS), English Longitudinal Study of Ageing (ELSA), Korean Longitudinal Study of Aging (KLoSA), Japanese Study of Aging and Retirement (JSTAR), China Aging and Retirement Longitudinal Study (CHARLS), Longitudinal Aging Study for India (LASI). In addition to these independent national projects, the Survey of Health, Ageing and Retirement in Europe (SHARE) is a consortium of HRS and ELSA harmonized studies (partly supported by the U.S. NIA). SHARE began with a set of 11 nations, but its third wave now

includes 16 nations (<http://www.share-project.org/>). There are elements of HRS comparability, also, in The Irish Longitudinal Study of Aging (TILDA) and the Canadian Longitudinal Study of Aging (CLSA), while the HRS has strongly influenced the design and content of other American aging studies – proudly including the WLS.

Measuring Wealth

The HRS has successfully estimated the implicit wealth in Social Security benefits through its link to SSA records, and it is now undertaking a most ambitious effort to improve the measurement of private pension benefits. In its early waves, the HRS asked participants about pensions on current and past jobs, including characteristics of plans and employer names. Then, it contacted employers to request Summary Plan Descriptions (SPD) without naming participants. Every six years, it coded plan rules from the SPD into a benefit calculation and then merged the reported characteristics with the plan rules to calculate benefits and wealth.

This did not work well. Participants did not know about key features of their pension coverage – even plan type and coverage. Most employers in the private sector did not respond to requests for their SPD, and the original benefit calculation program was inflexible and failed to provide some needed output. Thus, the fully informative set of merged respondents and plans was small and unrepresentative. Moreover, this data collection scheme missed changes in pension plans, which have occurred frequently since the beginning of the 1990s.

Thus, the HRS is now changing its pension-data protocol. First, it is improving respondent-level data by revising the survey pension sequence and using linked administrative data about individuals. The old pension sequence created problems because it was dependent on the participant's knowledge of plan type – defined benefit (DB) vs. defined contribution (DC) –

and many participants did not know their plan type. Thus, the new sequence does not condition on plan type.

Second, the new sequence is improving private firm-level data by building a longitudinal employer database starting with Form 5500 and using multiple sources of information about plan rules.¹⁰ There are fewer problems matching to government pension plans. Under the previous scheme, only about 20 percent of private employers responded to HRS requests for SPDs. However, under the new protocol, participants are asked to request the SPDs from their employers – which is a legal entitlement – and this increases coverage to 35 percent. Coverage of about 75 percent or more of employers is achieved using Form 5500 once the employer identification number (EIN) is known. However, there is a long reporting lag for the Forms 5500. Taking account of all three sources of plan information (and the possibility of multiple responses for the same firm), the HRS projects that overall coverage will be increased to about 65 percent.

This scheme is useful primarily for defined benefit plans, while are rapidly disappearing in both the public and private sectors of the economy. In the case of wealth in defined contribution programs, one alternative is to use the scheme of getting plan descriptions, reconstructing wages, contributions, and returns on investments, but there is also a much easier way to ascertain DC pension wealth: Simply ask directly. This has advantages because one can never truly learn all past contributions, rates of return, lump-sum withdrawals, transfers, and consolidations. Moreover, a trial comparison of account statements and self-reported DC balances was carried out by the HRS in 2004 and yielded a strong, linear relationship between

¹⁰ Form 5500 is a required annual report by firms that was developed by the Department of Labor, the Internal Revenue Service, and the Pension Benefit Guarantee Corporation in order to comply with requirements of ERISA, the Employee Retirement Income Security Act.

the two. Finally, the HRS is developing a new pension calculator program to take advantage of the improved primary data about DB plans. Thus, the HRS is taking on an expensive data augmentation task that will be useful to many data users but far too demanding and expensive for them to undertake independently.

Public Administrative Data Links

Both the HRS and the WLS track and link deaths of sample members to the National Death Index (NDI). Both studies use a combination of their own respondent tracking activities with the Social Security Death Index, which provides a timely listing of decedents names and places of death that are accessible through their Social Security Numbers. The additional, restricted NDI data – including multiple causes of death – are provided at low cost as a service of the National Center for Health Statistics (<http://www.cdc.gov/nchs/ndi.htm>). Without the NDI service it would be almost as burdensome to collect official cause of death data as private firm's pension data, for the several states (and some localities within states) have a variety of burdensome procedures that govern and restrict the release of death information.

Both the HRS and WLS have linked participants) survey and Medicare data, but the HRS is much further along in that endeavor. Both projects suffered in this process because of the uncertainties and delays associated with administration of HIPAA (the Health Insurance Portability and Accountability Act of 1996). At the time of the last (2004-06) WLS surveys, few of the graduates, but nearly half the siblings had reached age 65. The older siblings were consented by telephone, and after a very long effort, those data have been made available on a highly restricted basis to a group of WLS collaborators in the University of Wisconsin School of Medicine and Public Health. The HRS contracted with a private firm, Acumen, LLC, through a

Small Business Innovation Research grant from the National Institute on Aging. It has now obtained data on all types of claims from 1993 to 2005, and updates are in process. These include summary files by quarter and year and for intervals between HRS interviews. Fortunately, while the raw data are highly restricted, it may become possible to release public-use summary data.

Complementary Data Collection

The HRS is engaged in several complementary data collection activities. Consumption and time use are important components of economic models of household production.

The consumption of economic goods is very hard to measure. For example, it takes many hours to complete the Consumer Expenditure Survey (CEX) (<http://www.bls.gov/cecx/>), yet economic theory focuses heavily on consumption. Likewise, activities (time use) provide a valuable perspective on changes in people's lives with aging and disability. In 2001, the HRS initiated CAMS (Consumption and Activities Mail Survey) in 40 percent of participating households in order to assess consumption and time use. This survey has been repeated on the same households every two years, that is, in off-years when there is no regular HRS interview.

CAMS has been very effective, despite its limitations relative to the CEX. It covers about 95 percent of consumption as measured by the CEX. Analyses based on CAMS have laid to rest the "retirement consumption puzzle." That is, consumption changes little at retirement – despite reduced income – mostly due to a reduction in work-related expenses plus some substitution of home for market production, e.g., cooking at home rather than going to restaurants.

The first round of the Prescription Drug Study (PDS) was carried out in the Fall of 2005 in anticipation of the new Part D of the Medicare program. It collected detailed data on medication use, along with information about attitudes toward and knowledge about the Part D

Medicare program. The PDS was repeated in the Fall of 2007. Briefly, Part D uses private-sector insurers to offer subsidized drug coverage, but because there are myriad plans to choose from – differing from state to state – choosing whether to join Part D and, if so, what plan to join, is a daunting cognitive task. Indeed, some thought that the complexity of Part D choices would undermine the program, while others assumed (incorrectly) that just about any senior could use an interactive government website to make a well-informed choice. This choice process is complicated further by the fact that the cost of participation in Part D increases with the length of time between eligibility and enrollment. That is, eligible individuals must choose between joining the program early at low cost – but when they may not need the coverage – and joining it later at substantially higher cost.

Prescription Drug Use

The PDS found that HRS participants were far more concerned about Part D during than after the enrollment period. Perceived difficulty with the program was greater among those with lower levels of cognitive functioning during the enrollment period than among those with higher cognitive functioning, but after the enrollment period, that differential was erased. What happened is that seniors used their social networks to help them make acceptable decision. More people with lower levels of cognitive ability than with higher levels of cognitive ability sought help in the Part D decision from another person. Thus, differential use of social networks eliminated the impact of cognitive inequality on this difficult decision process. Moreover, those who choose not to enroll in Part D actually used fewer prescription drugs – at least during the observation period. In that sense, people’s choices appear to have been sound, and HRS participants had high levels of satisfaction with the Part D choices that they made. Finally, there

appeared to be no change in prescription drug use, either between the 2005 PDS and the 2007 PDS or between the core HRS surveys in 2004 and 2006.

Internet Surveys

In collaboration with investigators of the RAND Corporation and the American Life Panel (ALP), the HRS has been exploring the use of web-based survey protocols as a substitute mode for core telephone and face-to-face interviews. One key research issue is to determine the content and timing for which internet surveys should be preferred to traditional survey modes. If successful – and to the extent web-based survey data are comparable to those obtained by traditional modes – web-based surveys can yield the same content at much lower cost. Another possibility is that web-based surveys cannot substitute for core survey operations, but can be used for new, supplementary content or as a substitute for supplementary mail surveys. A second key issue that complicates the first is differential access to the internet. In the HRS, internet access varies directly with educational attainment and inversely with age. Moreover, to a substantial degree, the effect of educational attainment in this context reflects differences in cognitive functioning (Freese and Rivas 2005; Freese et al. 2006b). Thus, participants with internet access are a highly selected subset of all participants.

The Financial Crisis

Given the extensive economic data obtained in the HRS, it would appear to be a valuable resource for information about the effects of the 2008-09 financial crisis on the incomes and assets of older Americans. But the timing was off. The 2008 round of HRS interviews was almost complete before the crash, while the next regular round in 2010 was almost two years away. For this reason, an internet survey of HRS participants was carried out from April to June

2009 in collaboration with the RAND American Life Panel, using instruments developed in part by John McArdle and Robert Willis to provide in-depth measures of financial knowledge (the cognitive economics survey).

Briefly, 62 percent of HRS participants had stocks, 401ks, IRAs (Individual Retirement Accounts), mutual funds, or other stock-based instruments. They typically lost about 20 percent of equity in IRAs and 401ks, 25 percent in mutual funds, and 30 percent in single company stocks. The vast majority of HRS participants, 87 percent, owned their own homes, and 28 percent of those reported a decline in the value of their homes. Forty-seven percent had mortgages, but relatively few experienced serious problems on that account. About 7 percent of mortgagees owed more than their homes were worth; 4 percent had fallen behind in payments; and only 0.3 percent went into foreclosure. Overall, the HRS participants were affected by the financial crisis, but the problems were more of perception than reality. The crisis had some impact on psychological well-being, but there were no shifts in core economic or social behaviors.

Representation in the Core HRS Survey

A fundamental concern – in the WLS as in the HRS – is representation. There are reasons to worry both about selection bias in initial recruitment and participation and about bias in attrition across time. In the case of the WLS, there was essentially no nonresponse at baseline (in 1957, where students were “captive” participants), but there have been a modest share of refusals as each new (or older) cohort has entered the HRS. On the other hand, some initial biases can be ignored. For example, the HRS began as a survey of community-dwelling individuals and

couples, but because it follows people across time into nursing homes, it now fully represents the population in nursing homes.¹¹

A RAND study (Kapteyn et al. 2006) has examined attrition in the HRS from 1992 to 2006. It found some statistically significant predictors, but these had trivial effects on the distributions of key outcomes in 2006, primarily because HRS often succeeded in re-interviewing individuals who had missed a wave of the study. The main factor affecting non-response in the HRS is cognition, which negatively affects willingness and ability to participate in an interview. Similar findings have been reported for the WLS (Hauser 2005; Freese 2006). In the HRS, one indicator of this differential is the acceleration with age of proxy interviews taken because of cognitive limitation of participants. By taking proxy interviews, HRS maintains higher coverage of low-functioning individuals than ELSA, where there is a sharp gradient in the probability of continued participation with cognitive functioning at the immediately prior wave. Although one cannot obtain all of the data from a proxy respondent that might be obtained from a fully functional target respondent, it is possible to ascertain some valuable data, including assessments of the impact of cognitive disability on the family.

An important implication of the cognitive response differential is that efforts to measure the prevalence of cognitive impairment require careful attention to survey participation rates by level of cognitive functioning. In this regard, it is advantageous for longitudinal studies to track all current or prior participants.

¹¹ The percentages of the population in nursing homes as measured in the HRS virtually matches the percentages in the American Community Survey of the U.S. Bureau of the Census at ages 65 to 74, 75 to 84, and at and above age 85.

Enhancements of the Face-to-Face Interview

In 2006, the face-to-face HRS interview was enhanced in a random half of the sample with an array of biomarkers. Physical performance measures included assessments of grip strength, lung capacity (peak flow), a timed walk, and a balance exercise. Anthropometric measures were height, weight, and waist circumference. Blood pressure was assessed. Blood spots were taken and assayed for HbA1c, cholesterol, and CRP. Finally, DNA samples were obtained. In addition, a psychosocial questionnaire was administered for the first time. These measures deliberately overlapped with those in other studies: in-home biomarkers as in NSHAP, psychosocial measures as in MIDUS and WLS, and social supports as in ELSA and WLS. The full content of the HRS psychosocial instruments is shown in Figure 3.¹²

As suggested by Weinstein and Willis (2001) the addition of biomarkers to the HRS protocol potentially has enormous scientific value. The sample is nationally representative. There are longitudinal observations, both before and after the biomarkers have been measured. There are potential comparisons, both domestic and international. Since HRS spans a broad age range – everything over 50 years – there are many possible choices for scientific and policy attention. At younger ages, one very important issue is the metabolic syndrome – defined as three or more of hyperglycemia (>109 mg/dL), high blood pressure (>130/85 mm Hg), central obesity (waist circumference >102 cm in men and >88 cm in women), low HDL cholesterol (<40 mg/dL (1.04 mmol/L) in men and <50 mg/dL (1.29 mmol/L) in women), and elevated triglycerides (>150 mg/dL). Along with a sedentary lifestyle, it is strongly associated with diabetes mellitus,

¹² There is a great deal of overlap between this array of measures and those covered in the WLS since 1993-94.

coronary heart disease, and lipodystrophy. Prevalence of the metabolic syndrome is high. In the third National Health and Nutrition Examination Survey, 1988-1994, the rate was under 15 percent at ages 30 to 39, 22 percent at ages 40 to 49, 35 percent at ages 50 to 59, and about 45 percent at ages 60 to 69 (Ford et al. 2002). Moreover, the prevalence is increasing. Between 1988-1994 and 1999-2000 the rate had increased from 36.3 percent to 40.3 percent among 40 to 59 year-old men and from 30.5 percent to 33.8 percent among 40 to 59 year-old women. Moreover, the rate among 20 to 39 year-old women almost doubled, from 10.8 percent to 19.1 percent (Ford et al. 2004). Thus, it is most important to follow trends and correlates of the metabolic syndrome among baby boomers and following cohorts.

In older cohorts, frailty is a major health issue whose prevalence and correlates can be followed in the HRS. A standard definition of frailty was defined by Fried et al. (2001) as three or more of unintentional weight loss in the prior year, lack of grip strength, poor endurance and low energy, slow walking gait, and low physical activity level. Across an eight year follow-up period, mortality, hospitalizations, and disabilities were much higher among frail than among non-frail or intermediate classes. This has strong policy implications because the prevalence of frailty is an indicator of the need for long-term care services.

However, there are also challenges in the collection, storage, dissemination, and protection of such data, thus requiring a strong scientific rationale for each such measure and for the collective array. It is far more expensive – in interview time, training costs, equipment, and post-field processing – to collect biomarkers and biospecimens than to ask more survey questions. Moreover, although prior experience suggests that almost all survey participants will agree to contribute biomarkers and biospecimens, at least in face-to-face situations, it is not clear

how such potentially or actually invasive procedures may affect attrition in the long run. The HRS limited its biomarker collection to measures that could be obtained in the home by survey interviewers. It also chose to spread the biomarker data collection over two waves: half in 2006 and the other half-sample in 2008.

Fortunately, HRS staff had some earlier experience with the collection of biomarkers before introducing this protocol to the full sample. These included the ADAMS (dementia) supplement, a diabetes supplement, and face-to-face interviews. The ADAMS study measured the prevalence of dementia in a nationally representative sample drawn from the HRS. This required two stages of consent, both at the University of Michigan (home of the HRS) and Duke University (home of ADAMS). The protocol included a buccal DNA sample, and there was nearly 100 percent compliance with this request. While the two-stage consent process reduced the response rate – to just 60 percent – the study clearly demonstrated the feasibility of in-home requests for biomarkers.

The diabetes study was a mail survey of 2500 self-reported diabetics in the HRS. Here there was a two-stage data collection process. The first stage was a questionnaire, which achieved an 80 percent response rate. The second stage was a self-administered fingerstick blood spot that was assayed for HbA1c, which achieved a 65 percent response rate among stage one participants. Thus the net response rate was 52 percent. This was not much lower than the rate in ADAMS and far lower in cost, and the data obtained were of high quality. However, there was a substantial race-ethnic differential in participation. HRS investigators again drew the conclusion that two-stage consent or data collection processes were problematic. Moreover, there were

larger race-ethnic differentials in participation when the biomarkers were collected by mail than by administration in home interviews.

In 2004, the Social Security administration provided supplemental support for in-person interviews in order to improve consent rates in selected population groups. The NIA topped this off in order to make the special sample fully representative. The biomarkers obtained in this study were height, weight, and physical performance. This study confirmed the feasibility of obtaining these measures in household interviews on a large scale. In general, there was good cooperation with the physical performance protocol, but there was some selectivity in non-response. This suggested a need to ask self-rating questions in addition to the performance measures. In addition, this round of interviews was a successful pilot test of the mailback psychosocial questionnaire, for an 80 percent response rate was achieved. HRS investigators concluded that the extra cost and respondent burden of face-to-face interviews was justifiable in light of the demonstration that new content could be collected that would improve interdisciplinary research opportunities.

Comparisons of self-reported and measured height and weight in the face-to-face survey led to important findings. Briefly, the correlation between measured and self-reported weight was very high ($r = 0.98$), but there was a downward bias of just over 3 pounds in the self-reports. Self-reported and measured height were also highly correlated ($r = 0.94$). Again, self-reports were slightly larger than the direct measurements, but the bias grew larger as height increased. Errors in self-reported height and weight were correlated with age, gender, and years of educational attainment, and errors in height were also associated with memory capacity. While the errors appeared to be small, they are very consequential. For example, the guideline for a

diagnosis of obesity is BMI \geq 30, which is very near the mean of BMI. Thus, a relatively small error in BMI leads to a big change in estimated obesity. Based on self-reports 29.5 percent were obese, but based on measured height and weight, 38.2 percent were obese. A 4 percent increase in estimated BMI implied a 29 percent increase in obesity. The lesson here is that researchers and policy-makers should be very careful in their use of arbitrary cutpoints.

Obtaining reliable and valid assays has been a serious problem. In the case of blood spots most biological measures are expressed in quantities per unit volume, yet the blood spots are dry. This is an intrinsic problem. One example of the difficulty is that estimates of CRT (C-Reactive Protein), an indicator of inflammation, vary systematically between studies that have used different laboratories. Holding BMI constant, estimates are always larger in NHANES than in NSHAP and larger in NSHAP than in the HRS. However, this does not appear to be a problem in the case of HbA1c, where estimates for diabetics and non-diabetics are consistent among the three studies.

Institutional requirements for informed consent to obtain biospecimens are far more demanding than those for ordinary survey interviews, even though the entire protocol falls well within the definition of “minimal risk” in the Common Rule (45CFR46), the regulation that governs research with human participants: “the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.” Finally, there are serious issues in maintaining confidentiality because information obtained from biomarkers may be sensitive, e.g., pertain to serious health disorders or the potential for such disorders, or because it may be potentially identifiable, as in the case of even a

minimal set of genetic information (Homer et al. 2008; O'Brien 2009). Thus, even where the data collection protocol itself presents minimal risk, the collection of biomarkers and biospecimens creates new challenges and strains between the protection of confidentiality and the practice of sharing data (National Research Council 1985; Hauser 1987; National Research Council and Research Data in a Digital Age 2009).

Other New Content in the HRS

In 2008, the HRS introduced new content about women's health, including recall of hysterectomy, age at hysterectomy, menopausal status, and age at natural menopause. This is an important set of data to obtain for entering HRS cohorts, for about half of women will have experienced menopause by age 50.¹³ One check on the quality of these data is that the cumulative percentages of women in the HRS who have experienced menopause by specific ages agrees among women who reported retrospectively to agree with contemporaneous reports.

Another important new content area is child health. After a positive experience with these items in the supplementary internet survey, HRS has now included questions about child health in the regular interval cycle. Similar questions have been asked as part of the PSID, ELSA, and the WLS. They include a rating of general health in childhood, experience with a list of childhood diseases, and a report of length absences from school because of illness or injury. The rationale for asking such questions is the increasing scientific interest in the interplay between health and cognitive, social, and economic success across the life course (Palloni 2006).

¹³ Extensive reports about women's health issues, including dates of menarche, menopause, hysterectomy, oophorectomy, use and discontinuation of hormone replacement therapy, and menopausal symptoms were collected by the WLS in 1993-94 and in 2004-06.

A final innovation in the HRS in 2010 will be a set of adaptive cognitive measurements, developed in collaboration with John McArdle of the University of Southern California. Previous HRS cognition measures have focused on age-related cognitive decline, especially loss of memory and dementia. They have not previously included key intellectual capacities such as fluid reasoning, cognitive speed, and executive function. These gaps in the assessments are being addressed through the development of adaptive testing sequences. The essential idea behind adaptive testing is to ask people items that are appropriate to their ability level,¹⁴ not to waste time on items that are either much too difficult or much too easy for them.¹⁵ Thus, in the absence of prior knowledge about an individual's ability, the first question would be of medium ability, and subsequent questions would be easier or more difficult, depending on the response to the first question. Using this branching scheme, it is possible to measure an ability accurately with a small number of items, and it offers a better experience to participants because they are less likely to get a lot of items that are hard (unpleasant) or easy (boring). The process can be made even more efficient if there is prior knowledge about the participant's ability; that is, the first question asked would be near their expected level of performance, rather than near the middle of the overall distribution of performance. However, in order to build an adaptive test, it is necessary to know the relative difficulty of the items in advance, that is, to pre-norm them, and for this reason the HRS assessments are based on the Woodcock-Johnson tests.

¹⁴ "Ability" is used here in the psychometric sense, referring to whatever is being measured, not to the generic psychological construct of cognitive ability.

¹⁵ In the U.S. the Graduate Record Examinations (GRE) are now administered adaptively.

In the development process, during 2004 and 2006, selected participants were asked a subset of a sample of three minutes of questions using number series. In this test, a prototypic item would be something like “1, 3, 5, ...” what comes next? There were 45 possible questions, and each participant was asked only five or six items to assess this ability. The adaptive cognitive testing worked very well, and there is some reason to consider it in the measurement of other psychological characteristics, such as personality, well-being, and depression, provided standard items in those series can be pre-normed.

Recent Developments in the WLS

Recording Interviews

In development of the 2004-2006 round of the Wisconsin Longitudinal Study, project staff decided that it would be desirable to record all of the telephone interviews. The original reason for our investigation of recording technology was that two of the collaborators, Nora Cate Schaeffer and Douglas Maynard, wanted to obtain high quality recordings of about 1000 randomly selected interviews that could be used for intensive analysis of respondent-interviewer interaction in an older population (Gathman et al. 2008). A second reason, which applied to parts of all interviews, was that some of the more attractive protocols for cognitive assessment by telephone could not be administered and coded reliably unless the responses were recorded. Furthermore, recordings could be used to validate appropriate administration of the assessments.

We soon learned that almost all respondents would agree to full recordings and to their retention for research purposes. This was confirmed in pretests, and a consent protocol was developed. All respondents were asked at the beginning of the interview whether it could be recorded. If they agreed, the interview was recorded, and they were asked at the end of the

interview if the recording could be retained for research purposes. If the respondent declined to be recorded at the beginning of the interview, the recording equipment was turned off, but the respondent was asked again, at the beginning of the cognitive assessments, to give permission for just that portion of the interview to be recorded.

The WLS recording technology was developed to meet four main criteria:

(1) Recordings should be digital and stored as *.WAV or *.MP3 files, thus permitting random access for research purposes. Thus, standard or digital audio taping was eliminated as a possibility.

(2) Control of the recording process should take place automatically through the CATI software (CASES) and not require separate activation or adjustment by interviewers.

(3) The recordings should be as high in audio quality as possible, given the limited bandwidth of telephone lines.

(4) Because audio files are potentially identifiable, as well as valuable for research purposes – including some that may lie in the distant future – file storage procedures should be both secure and redundant.

The cost of the recording software and equipment was modest, less than \$250 per workstation. In order to assure security of the recordings, they have been placed on a large dedicated server that is independent of other data servers and made accessible only to specific WLS project staff.

Aside from the continuing value of the recordings in research, they also proved useful in the process of instrument development. For example, it was efficient for researchers to listen to each instance of a pretest telephone module, e.g., a family roster or employment history, in order to detect and solve problems in the logic and content of the instrument and to identify problems

in interviewing that could then be addressed in training sessions. The recordings were also very useful in post-field editing of survey responses. However, this had one disadvantage; recordings created an opportunity and an incentive for endless editing.

Because of the demonstrated feasibility and utility of recordings of telephone interviews, we are also planning to record the full home interviews with graduates and siblings in 2010-2011. This round of the study will be carried out by computer assisted personal interviewing (CAPI), and it is easy to set up a microphone connected directly to the interviewer's PC to record the session. In addition to the several advantages of recording interviews in 2004-06, the new recordings will serve two other purposes. First, as part of the protocol we will ask the participant to describe everything that is going on in a picture handed to them. This is the well known "cookie theft" prompt (Goodglass and Kaplan 1983). Scoring is based on the full text of the reply, so it is critical to record this sample of extended speech. Second, the recordings are a strong form of social control. Telephone interviews were carried out from a call center, where supervisory staff could monitor all activities. Face-to-face home interviews are conducted in isolation, but recordings provide evidence that an interview actually took place.

Coding Photographic Images of Participants

High school yearbooks from 1957, the senior year of the WLS graduates, were collected over the past several years, at first mainly from public sources – high schools and public libraries – beginning with larger schools in larger places. This effort yielded approximately 7500 photos, from which a well-designed sample of 3000 was drawn and coded. A second round of collection and scanning was carried out during the first half of 2007 in conjunction with the distribution of respondent reports from the 2004-2006 round of the WLS. Thus, the WLS collection of scanned

yearbooks now covers a large majority share of the original sample of 10,317 graduates. The remaining photos were coded during 2008, yielding a total of 8600 photographs rated on facial mass and facial attractiveness.

Each set of facial appearance codes was assigned independently by multiple coders using a detailed, computer-assisted protocol. The yearbook photograph is the only information about the target individual presented to the coder. In both cases, the target picture was presented in combination with a visual scale that is anchored by gender-specific photographs of 1957 Wisconsin high school seniors who were not in the WLS sample. For example, Figures 4 and 5 show displays similar to those used to code men's and women's facial mass.¹⁶

In the case of facial mass, six graduate students independently coded facial mass following these instructions.

1. Click on this internet site: (Suppressed)

You will be prompted for a name and password. This is highly confidential information, and you are responsible for safeguarding it. Do not share this information with anyone for any reason!

Username: *****

Password: *****

2. In each session, you will code approximately 300 photos, 150 for boys and 150 for girls. At this time, choose whether you want to code girls or boys first. In subsequent coding sessions, please alternate which gender you choose first.

3. You will be asked for your name, sex and birthday. Please respond to these questions

¹⁶ The individuals shown in the photographs are members of Wisconsin's class of 1957, but they are not in the WLS, nor were these photos used in the actual rating protocols.

consistently at each coding session. Follow the format in this illustration:

Name: Jane Doe

Sex: female

Birthday: October 22, 1970

4. Next, you will see a series of photos aligned at the top of the page with several triangle markers placed beneath them. These photos are a guide to coding the yearbook photos that appear at the bottom of the page. Also at the bottom of the page is a question: How heavy is this person? To answer this question, follow these instructions:

A. In 10 seconds or less, form an initial reaction about where you think this person fits on the scale. Click once on that point.

B. After you choose a particular point, the following question will appear: Does the red symbol indicate your choice? Do not answer this question immediately, but rather proceed to step C.

C. Examine the person's neck. How wide is this person's neck relative to the other photos? Do you see evidence of a "double-chin" or other fat deposits around the neck?

D. Examine the person's cheeks. Are this person's cheeks "puffy," full or sunken? How does this compare to other photos in the scale?

E. After you examine the person's neck and cheeks, decide whether your initial reaction still seems appropriate. If so, click "yes."

F. If not, then (1) answer "no," (2) click on the new point that seems most appropriate and (3) immediately answer "yes."

G. Repeat steps A-F for all photos.

H. Once finished, choose the other gender and code those photos following these instructions (Reither 2005).

Very high reliability (Cronbach's $\alpha = 0.91$) was obtained in the ratings of facial mass with six independent codings of each photograph (Reither 2005). In the second round of rating, each rater also recoded several hundred previously coded photos in order to calibrate the two rounds of coding.

A similar protocol was used to code facial attractiveness, based on 12 independent ratings of each photo by contemporaries of the graduates. In pilot research, six ratings were obtained from college students, and six ratings were obtained from contemporaries of the graduates (Meland 2002). Both older and younger coders were used because standards of facial attractiveness might have changed. However, the only substantial difference in the ratings by older and younger coders was that the older coders were more reliable, so older coders have been used consistently in all production coding. Again, we have obtained repeated ratings of several hundred previously coded photos in order to calibrate the two rounds of coding. As in the ratings of facial mass, the ratings of facial attractiveness were highly reliable (Cronbach's $\alpha = 0.87$). There is little reason to think that external information may have prejudiced judgments of the coders. However, in the WLS data, it is possible to control contextual clues affecting the ratings using measures of the social background and academic performance of the participants; in fact, despite the nominal ignorance of coders, who saw only the cropped portraits, there were modest effects of the socioeconomic status and IQ of the participants on the ratings that they received.¹⁷

¹⁷ As in the case of recordings of WLS interviews, the scanned photographs are stored securely and entirely separately from any other WLS data.

The ratings have already been used in several studies. Freese, Meland, and Irwin (2006a) found that expressions of positive emotion in the photos were unrelated to later life marital and health outcomes; this failed to confirm an earlier psychological study based on a small sample of Pomona College graduates. Reither, Hauser, and Swallen (2009) showed that the ratings of facial mass are an acceptable proxy for the Body Mass Index (BMI) and are related to morbidity and mortality by ages 53 and 54. Hauser (2009a) confirmed this finding with respect to mortality through 2008 in the full set of 8600 coded senior photos, and he also showed that greater facial attractiveness was associated with lower mortality. Sicinski (2009) has found that facial attractiveness is positively associated with the economic prospects of spouses (for women, but not for men) and that it is strongly related to earnings early, but not later in workers careers. Finally using the initial release of codes for 3000 photos, Jokela (2009) obtained findings suggesting that facial attractiveness was modestly, but non-monotonically related to “reproductive success” (number of children) among WLS graduates.

Because of the demonstrated scientific utility of the adolescent photographs, we have decided to obtain new photographs of WLS participants during the forthcoming home interviews. This will enable us to assess the long-term stability of facial features among WLS participants and, also, to assess other aspects of physical appearance. Our decision was prompted by the recent use of photographs in a Danish study to assess the progress of aging (Christensen et al. 2008).

*Studying the Recently Bereaved*¹⁸

Advances in medical technology enable chronically ill older adults to extend the length, though not necessarily the quality, of their lives (National Institutes of Health 2005). Most dying older adults have limited mobility, impaired cognitive functioning, and discomfort. If dying patients have not made formal plans for their end-of-life medical care or have not informed others of their treatment preferences, they may have little control over the care they receive. Difficult decisions about stopping or prolonging treatment typically fall upon family members who may be distressed, or who may disagree about appropriate care (Kramer et al. 2006). When patients have not articulated their preferences, health care providers may prolong futile treatments, which may cause physical and emotional distress to the patient, and financial and emotional strain for family members.

Recognizing the financial and emotional strains associated with problematic or unwanted end-of-life care, practitioners encourage advance care planning. This practice encompasses completing a living will, appointing a durable power of attorney for health care (DPAHC), and discussing treatment preferences with family, physicians, and caregivers (Lipkin 2006). Despite widespread endorsement of such practices by organizations like the American Medical Association (1996), no large-scale, community-based studies have assessed whether and how specific aspects of advance care planning protect bereaved family members from the psychological, and interpersonal strains that may accompany a “bad death” (Powis et al. 2004).

¹⁸ This section of the paper is the work of Professor Deborah Carr, Department of Sociology, Rutgers University.

The financial costs associated with illness and death also may create considerable stress for bereaved family members, compounding the psychological stress of the dying process. This distress may be exacerbated by financial uncertainty, probate costs, and difficulties in settling estates -- especially in cases where the deceased did not offer guidance through financial planning mechanisms, such as a legal will, life insurance designations, or pension plan options (Holden and Brand 2003). Given that older adults wish to protect their family members from psychological and financial distress (Fried and Bradley 2003), documenting whether and how financial end-of-life planning benefits surviving kin may provide greater incentive for planning, and help practitioners identify which practices are most helpful (and which cause unintended difficulties) for the dying and their kin.

WLS investigators plan to examine the impact of both medical and financial end-of-life planning on the psychological and financial well-being survivors of a sample of recent decedents. Because detailed data on the decedents' end-of-life planning were obtained prior to the time of death, the WLS has the unique opportunity to examine prospectively the influence of end-of-life planning. The WLS also will investigate three possible pathways linking end-of-life planning with survivor well-being: the quality of care received by the decedent, the nature of the dying process, and conflicts surrounding end-of-life care -- where assessments are made retrospectively by bereaved family members. This will include an evaluation of the extent to which the observed effects of end-of-life planning are due to selection; that is, the pre-existing traits and resources of those who plan, such as relationship quality, income, or education. Parallel measures will be obtained from two bereaved significant others who are deemed most knowledgeable about the death. The aim is to interview the decedent's DPAHC, where

applicable. Subjective, retrospective proxy assessments of end-of-life experiences are known to be unreliable, partly reflecting the distressed state of the survivor and variable levels of involvement in end-of-life care (Morrison et al. 2000). Obtaining two reports from next-of-kin will permit more accurate assessment of the impact of end-of-life planning, particularly with respect to the quality of and conflicts surrounding end-of-life care, and quality of dying. Multiple reports will also make it possible to explore whether survivors' discrepant perceptions of the dying process are associated with conflict during the dying process and survivor distress after the loss.

Since 1975 parallel interviews have been conducted with one randomly selected sibling of the graduates at roughly comparable time periods; most siblings are within eight years of age of the graduates and are now in their mid-60s to mid-70s. In 2004, parallel interviews also were conducted with the spouses of both the 1957 graduates and siblings. Thus, the WLS provides a full record of family relations, health, socioeconomic status, income, assets, and cognitive ability over the life course. In 2004, the WLS also collected extensive data on medical end-of-life planning (i.e., a living will, a durable power of attorney for health care (DPAHC), or held discussions about end-of-life treatment preferences). Respondents also were asked about five financial planning instruments that provide for asset transfers upon death: a legal will, revocable trust, life insurance, pension plans, and beneficiary and executor designations.

With these issues in mind, deaths to WLS graduates and siblings since the most recent wave of data collection (2004-06) are being tracked, and the WLS will interview two significant others per decedent. Significant others will be selected based on their closeness to the decedent, and knowledge of the death; the vast majority will be spouses, adult children, and siblings of the

decedent. Two survivors per decedent will be interviewed by telephone, including the person named as DPAHC. The majority of a projected 1,350 interviews with bereaved persons will occur nine months following the death, although there will also be a comparison sample of about 300 longer-bereaved persons who lost their spouse in the two year period (2007-09) prior to the start of data collection. With these data, the project will investigate the extent to which both the individual and combined components of the decedent's financial and health-related end-of-life planning affect the surviving kin's psychological well-being (e.g., depressive symptoms, anxiety, alcohol use, grief symptoms, spiritual well-being), and financial well-being (e.g., changes in assets, perceived financial strain, direct costs of end-of-life care). The project will evaluate the extent to which the relationship between end-of-life planning and survivors' well-being reflects three possible pathways: family and professional conflict surrounding end-of-life care; the locus and quality of care received by decedent; and quality of decedent's dying experience (i.e., mediation analysis). Finally, the project will assess the extent to which the observed associations reflect pre-existing characteristics of the decedent and survivor(s), such as socioeconomic resources, relationship quality, or depressed affect, which may affect both end-of-life planning and end-of-life/bereavement outcomes.

Genetic Analyses and Data in the WLS

In the WLS, the collection of biomarkers is similar to, but more limited than in the HRS. In addition to the proxy measure of youthful BMI, height and weight have been collected by self-report since 1993-94. The forthcoming round of home interviews will include other anthropometric measures and a set of performance measures that largely overlaps that in the HRS: grip strength, peak flow (puff test), walking gait, and a chair rise. However, the WLS does

not plan to collect blood spots. The collection of DNA in the WLS is scarcely a unique undertaking; many social surveys – of young as well as older samples – have done so. However, the rationale and operational procedures for doing so may yet be of general interest.

Research examining psychological consequences of family roles and transitions has not adequately addressed the possibility that these effects could reflect, to some extent, genetic endowments that are associated both with psychological functioning and specific configurations of family roles. If such genetic influences exist, the relationships between family roles and mental health could be partly spurious, reflecting selection of individuals based on genetic predispositions rather than causal effects of family roles. For example, using the WLS sample of twins, full non-twin siblings, half-, step-, and adopted siblings, Pudrovská (2008) found that the association between parenthood and psychological well-being at midlife is partly explained by genetic influences. We plan to use DNA assays to examine specific polymorphisms that may be implicated in the relationships between family roles and mental health in mid- and later life. Recent research suggests that at least two polymorphisms may be of particular importance in this respect: 5-HTTLPR and DRD4.

The 5-HTTLPR s-containing genotype has been associated with anxiety-related personality traits, affective disorders, and suicide (Collier et al. 1996; Lesch et al. 1996; Bondy et al. 2000; Willeit et al. 2003), although the evidence is mixed (Willis-Owen et al. 2005; Middeldorp et al. 2007). In addition to psychological well-being, 5-HTTLPR was shown to be related to reproductive behaviors. Research among rhesus macaques indicates that males who are homozygous for the diallelic 5-HTTLPR polymorphism have higher fecundity during early and late adulthood than their heterozygous counterparts who reproduce mainly at an intermediate age

(Krawczak et al. 2005). Suggestive findings were obtained in human studies as well. Hamer (2002) reports a statistically significant association between 5-HTTLPR and frequency of sexual activity among middle-aged men. In addition, the 5-HTTLPR variant was shown to be associated with the number of sex partners, although this association was moderated by religious attendance (Halpern et al. 2006). Similarly, DRD4 was shown to be associated both with psychological traits, such as novelty seeking and neuroticism (Benjamin et al. 2000; Tochigi et al. 2006), and reproductive behaviors, such as age at first sexual intercourse and the number of sexual partners (Miller et al. 1999; Hamer 2002; Guo and Tong 2006). Thus, in future research we plan to examine the extent to which the psychological impact of family roles and transitions in later life is explained by variants of 5-HTTLPR, DRD4, and other polymorphisms.

Plug and Vijverberg have used the identification of children of WLS participants as biological or adoptive to improve the estimation of the effects of parental characteristics on educational success and to estimate its heritability (Plug and Vijverberg 2003; Plug 2004; Plug and Vijverberg 2005).

As noted above, the WLS research design has been expanded to include siblings (including a modest number of twin pairs) and cousins. Moreover, the biological relationships among graduates, their (nominal) siblings, and the offspring of graduates and siblings have been identified by survey reports, e.g., as biological, step, or adopted.

Sibling Resemblance

Several studies have used WLS data to focus on similarities and differences among graduates and siblings with respect to cognition, educational attainment, and economic success and on the extent to which failure to control common family factors may lead to bias in

estimated effects of educational attainment on economic success. These studies invite further investigation with genetically informed designs.

Just as the Wisconsin Model was inspired by Blau and Duncan's (1967) classic work, our development of sibling data in the WLS was inspired by the creative efforts of Duncan (1968; Duncan et al. 1972), Jencks (1972), and their collaborators.¹⁹ They had gathered and synthesized diverse data on the socioeconomic resemblance of siblings; the advantage of the WLS was that we could obtain such data within a single, well-defined population.²⁰

During the 1975 interview, we obtained a roster of each respondent's living siblings. The roster included age, gender, and highest level of schooling. This has informed our analyses of the influence of age, gender, birth order, sibship size, and child spacing on ability, educational attainment, and occupational achievement. A strength of these data is that they provide detailed information both on families and on samples of persons. A second advantage is that they include detailed information on the structural characteristics of sibships. By using the data on families, on sibling pairs, and on our respondents, we have been able to overcome two problems: Data on families are inherently confounded with temporal changes in the larger society, while data on samples of persons risk confounding family structure with other characteristics of the family (Sewell and Hauser 1977). On the other hand, this design fails to distinguish among effects of genetic makeup, common family environments, or those of distal social environments that are shared by siblings, e.g., neighborhoods or schools.

¹⁹ In fact, one insightful passage in Blau and Duncan (1967:320-28) anticipated many strengths and problems of models of sibling resemblance.

²⁰ Since the mid-1970s, well designed studies of sibling resemblance in youth have proliferated. These include the National Longitudinal Study of Youth and the Adolescent Health Survey.

In each sibling roster, we identified a randomly selected sister or brother by name and ascertained that person's occupation, current address, and the name of the last secondary school (if any) attended by that person in the state of Wisconsin. We randomly chose a highly stratified probability sample of selected siblings for interview during 1977, using much the same instruments and methods as the 1975 graduate survey. We also ascertained data on the mental ability of the siblings from the Wisconsin State Testing Service. Interviews were completed with 90 percent of the selected siblings, and test scores were obtained for 80 percent of them. The 1977 sample contains approximately 1,500 same-sex pairs of siblings and 500 cross-gender pairs of siblings. In the 1992 to 1994 round of the WLS, we expanded the sibling sample and, again, used essentially the same instruments as for the graduates.

With these data, we have analyzed the effects of the family on measured mental ability, on educational, occupational, and economic achievements, and on selected indicators of health. The resemblance of siblings raised together is, of course, a fundamental indicator of the force with which the family functions to create and maintain systems of social differentiation and inequality. Sibling resemblance captures the effects of social and economic background, of family structure, and of other commonalities of social and psychological functioning of the family. It is possible to give sibling resemblance an explicit interpretation to the extent that shared familial characteristics have been measured.

The effect of family socioeconomic background on educational and socioeconomic attainment may lead to bias in simple regressions of occupational status or other outcomes of schooling on educational attainment. This is usually called "family bias." Hauser (1984) developed a structural equation model of sibling resemblance in occupational attainment that

took into account possible family bias. He then tested this model on data on sibling pairs from the WLS sample and Olneck's (1976) Kalamazoo sample. Family bias in the effect of schooling on occupational status was less than others have claimed, and estimates of family bias are very sensitive to specifications of response variability in schooling. He also provided some useful methods for cross-population comparisons of structural equation models.

Hauser and Mossel (1985; 1988) used these models to estimate the regression of occupational status on schooling in a sample of 518 pairs of brothers from the WLS sibling sample, but their models also corrected for survey response error. They found that the regression of occupational status on educational attainment was relatively insensitive to both response error and to the specification of family factors. Family membership accounted for about half of the variance in occupational status, but there was little evidence that failure to control family background led to upward bias in estimates of the effect of schooling on occupational status.²¹

Hauser and Sewell (1986) further analyzed the influence of family effects in models of education, occupational status, and earnings for fraternal pairs from the WLS sibling sample. They modeled the effects of measured and unmeasured family background, mental ability, and educational attainment on occupational status and earnings. Their models were estimated from incomplete data (for brother pairs who may or may not have been interviewed in 1977), with corrections for measurement error, and they permitted direct comparisons of within- and between-family regressions. There was no evidence that effects of family background lead to bias in the effects of mental ability on schooling or in the effects of schooling on occupational status or earnings. Family background did have large effects on ability and schooling but smaller

²¹ Hauser and Mossel (1988) provided a more extensive methodological treatment of these models, including detailed documentation of estimation procedures.

effects on socioeconomic achievement. Hauser and Sewell carried out parallel analyses of Olneck's (1977) data on Kalamazoo brothers and concluded that, after correcting for attenuation, the Kalamazoo data provided no more evidence of family bias than the Wisconsin data.

Hauser (1988) extended the methodological work on sibling resemblance by comparing two linear models educational and occupational resemblance among pairs of Wisconsin brothers.²² Each model decomposes the regression of occupational status on schooling into a between-family regression and a pair of within-family regressions, one for each sibling. In the first model, within-family regressions are written in unique within-family factors for each brother, and in the second model, the within-family regressions are in total educational and occupational variables. The two models are equivalent when within-family regressions are the same for each member of the pair. Otherwise, they are not equivalent; the second model implies two, distinct between-family regressions. Using the WLS sample of sibling pairs, Hauser showed that, under certain conditions, either model may exhibit symptoms of near-underidentification or even be underidentified. In the Wisconsin data, because family bias was minimal, this problem was more likely to occur when using the second model than the first.

Hauser and Wong (1989) further extended the work on sibling resemblance and intersibling effects using data from a Nebraska sibling sample originally studied by Benin and Johnson (1984). The reanalysis showed no unusually high level of resemblance between brothers in the data, but there was a strikingly low similarity between older sisters and younger brothers. Hauser and Wong then reanalyzed Olneck's (1976) Kalamazoo data and found clear evidence of reciprocal effects of brother's levels of education, net of common family background and the

²² Gillespie (1991) and Hauser (1991) debated the value of these models.

effect of each brother's mental ability on his schooling. The unconstrained effect of older brother's schooling on younger brother's schooling was larger than the reciprocal effect, but a model of equal effects fit almost as well. This left open two questions, whether reciprocal influence occurs in other types of sibling pairs and whether there is a predominant influence of older on younger siblings. Lee (1989) used the structural model developed by Hauser and Wong to examine sibling resemblance and cross-sibling effects on educational attainment in WLS sibling pairs. However, that analysis was inconclusive because of problems with missing data.

Although studies of social stratification position the family of origin as a key agent, few studies have explored family effects across careers. Hauser, Sheridan and Warren (Warren et al. 2002) used WLS data on sibling pairs to explore the relationship between social background, cognitive ability, educational attainment, and occupational standing across the life course. They extended the model of sibling resemblance estimated by Hauser and Sewell (1986) in three important ways. They indexed occupational standing in two different ways, using typical educational levels and typical income levels (Hauser and Warren 1997). They included pairs of sisters and of sisters and brothers as well as brother pairs. They also modeled occupational status at four times: first job, job in 1970, job in 1975 or 1977, and job in 1992 or 1994.

Hauser, Sheridan and Warren found that women are at least at parity with men in occupational education, but they fall far below men in occupational income. Women and men experience growth in occupational education, both in the transition from their families of orientation to the labor market and throughout their careers. However, women typically enter occupations with low levels of income, early in their careers, and subsequent growth never leads to parity with their fathers, peers, or brothers. Occupational levels of education are more

persistent across generations and within the career than occupational levels of income or earnings.

Across families and regardless of gender, Hauser, Sheridan, and Warren (1999) found that educational attainment levels are determined largely by cognitive ability and to a lesser degree by social background; family levels of occupational standing are determined largely by family education levels. Within families, cognitive ability also affects occupational standing primarily through schooling. Occupational inequalities and the effects of educational attainment on those inequalities both tended to decline across the life course. Finally, if there is a “central” variable in the attainment process, it is educational attainment, not cognitive ability. Cognitive ability, like social background, affects occupational careers primarily through the length of schooling; net of schooling, the effects of cognitive ability are very small.²³

Depression

MacLean and Hauser (2000; 2001) looked at the resemblance of brothers and sisters of the WLS in respect to a 20-item version of the CES-D that was administered in the 1993-94 surveys (Radloff 1977; Robins et al. 1981). The gist of their findings was that the factorial components of depression were more a state than a trait, that is, that there was very little resemblance between siblings (no more than 15 percent of shared reliable variance) and that the only socioeconomic characteristic substantially (and negatively) correlated with depression was wealth.

There is much more to be done with these data, for the CES-D items were repeated in 2004-06, and we also have repeated measurements of a lifetime depression history designed by

²³ Compare Herrnstein and Murray's (1994) excessive claims about the importance of cognitive ability in social stratification (also, see Hauser et al. 2000).

Ronald Kessler. Of course, there are now several candidate alleles predisposing to depression, as well as evidence of gene-environment interaction (Caspi and Moffitt 2003; Caspi et al. 2003; Moffitt et al. 2005), but the generality of this finding has been challenged in a recent meta-analysis (Risch et al. 2009).

Personality

In 1993-94 and 2004-06, the WLS surveys of graduates and siblings have measured the basic dimensions of personality known as the “Big Five”: extraversion (vs. social inhibition); agreeableness (vs. antagonism); conscientiousness (vs. lack of direction); neuroticism (vs. emotional stability); and openness to experience (Digman 1990). Although the Five-Factor Model has its detractors (Eysenck 1992), it is arguably the most widely used taxonomy of personality traits. We obtained self-reports for 29 items in Johns’ BFI-54 (1990; 1991) to measure personality variables for graduates and siblings in the last two rounds of the WLS. Internal consistency reliabilities were 0.69 (A), 0.64 (C), 0.76 (E), 0.78 (N), and 0.62 (O) in among graduates in 1993, and they were virtually the same in 2004.

We have found that from age 54 to age 65, agreeableness increased by .04 standard deviations and neuroticism, conscientiousness, openness and extroversion decreased by .19, .15, .07, and .02 standard deviations respectively. These findings are consistent with previous studies, except the change in conscientiousness. Contrary to theoretical expectation, conscientiousness decreased considerably. This unexpected result may reflect the age of our sample. Between test and retest, over 50% of the sample retired from full- or part-time jobs. Job separation may lower levels of conscientiousness, as previous studies (mostly based on students and young adults) have shown a linkage between conscientiousness and the work domain (Neyer and Asendorpf

2001). The unexpected decrease in conscientiousness also may reflect selective mortality. By looking separately at personality measures in 1993 and 1994 in relation to subsequent mortality, other sources of survey attrition, and continuing participation in the WLS, we can learn more about this.

We also examined trait differences and changes by sex, IQ, and socioeconomic status of origin family. Consistent with previous studies, we found that women score higher on agreeableness (.43 s.d.), neuroticism (.22 s.d.), extroversion (.11 s.d.) and conscientiousness (.07 s.d.). In addition, IQ and family SES are both positively correlated with openness and negatively correlated with neuroticism. With regard to temporal change, a strong pattern of parallel development emerged: Men and women, groups of different levels of IQ and family SES show almost identical changes in the Big Five. As a result, cross-sectional group differences in personality traits remain constant over time.

A prevailing view in psychological research is that personality is close to immutable, especially among older adults. For example, McCrae and Costa (2003, Chs. 6-7, esp. pp. 124) write that in their longitudinal analyses, correcting for attenuation, “if we had perfect measures we would find near-perfect stability of personality traits.” At first glance, this would appear to be the case, also, in our repeated measurements of personality. Simple corrections for attenuation (using Cronbach’s α), yield inter-temporal correlations of 0.91 (A), 0.95 (C), 0.98 (E), 0.90 (N), and 1.18 (O)—the last of which is plainly out of bounds. This anomaly has led us to develop less naïve models of the measurement of personality among 4078 WLS graduates who answered all items on both occasions. We have developed preliminary structural equation models of the measurement and persistence of personality that are based on the assumption that responses to

the BFI-54 items are ordered categorical representations of normally distributed variables. If we assume that each item indicates one factor on each occasion, and all errors of indicators are uncorrelated, we estimate inter-temporal correlations of 0.92 (A), 0.86 (C), 0.90 (E), 0.92 (N), and 0.93 (O). These are just slightly lower than the disattenuated correlations. However, this model fit poorly, so we introduced more realistic assumptions. Our best fitting model to date yields an acceptable Root Mean Square Error (RMSEA = 0.048). It incorporates three methodological effects: (a) correlation across occasions of the errors of identically worded items, (b) correlations within occasions of the errors in adjacent items, and (c) a common factor on each occasion for negatively phrased items. In this model, the estimated inter-temporal correlations of the Big Five are substantially reduced: 0.83 (A), 0.80 (C), 0.83 (E) 0.83 (N), and 0.77 (O).²⁴ That is, part of the apparent persistence in indexes of personality variables is attributable to individual tendencies to respond in the same way to the same items across occasions and similarly to negatively phrased items. None of the personality correlations is “small” or even “moderate,” but they leave room for intra-personal change across time. A correlation of ~0.9 explains 81% of variance, while a correlation of ~0.8 explains only 64% of variance. In recent months, we also have tested the implications of different measurement assumptions (interval vs. ordinal measurement) on stability of personality, on tendencies to respond positively as well as negatively, and on tendencies to use the same response category for adjacent items.

Behavioral genetic studies indicate that human personality traits are under a substantial degree of genetic influence (Plomin et al. 1994; Bouchard 1997). Yet, despite recent developments in molecular genetics, only modest progress has been made in identifying specific

²⁴ If we assume interval measurement in the BFI items, we obtain a better fitting model and generally lower estimates of the inter-temporal correlations: 0.73 (A), 0.77 (C), 0.80 (E) 0.76 (N), and 0.79 (O).

genetic variants that demonstrate robust and consistent associations with personality traits (Munafo et al. 2003; Munafo et al. 2008).

Genetic markers available in the WLS have received attention in personality research. MAOA genetic variation was associated with agreeableness and conscientiousness among men (Rosenberg et al. 2006). Males carrying the 3-VNTR MAOA gene variant (209 bp) had significantly lower values on openness (Samochowiec et al. 2004), whereas high function alleles (3.5 repeat and longer) were associated with higher neuroticism among men but not women (Eley et al. 2003). In contrast, in a Japanese sample, MAOA was not associated with Big Five traits (Tochigi et al. 2006b). The dopamine receptor genes also were implicated in personality. DRD2 and DRD4 were shown to be associated with neuroticism (Lee et al. 2005; Wacker et al. 2005; Tochigi et al. 2006a), and DRD4 was related to extraversion (Benjamin et al. 1996; Eichhammer et al. 2005).

The 5-HTTLPR polymorphism was extensively studied with respect to neuroticism; yet, findings remain mixed and inconclusive. Some studies indicate that the 5-HTTLPR short allele is associated with higher neuroticism (Lesch et al. 1996; Greenberg et al. 2000; Jang et al. 2001; Sen et al. 2004; Dragan and Oniszczenko 2006), whereas others found no association between 5-HTTLPR and neuroticism (Stoltenberg et al. 2002; Samochowiec et al. 2004; Willis-Owen et al. 2005; Middeldorp et al. 2007). One study suggested that individuals with the short variant of the allele (ss or sl) were significantly lower on neuroticism compared with individuals with the long allele variant, a pattern opposite to that of other reports (Brummett et al. 2003). In addition to neuroticism, 5-HTTLPR may be associated with openness (Stoltenberg et al. 2002),

agreeableness (Greenberg et al. 2000; Jang et al. 2001; Sen et al. 2004), and extraversion (Gillihan et al. 2007), although the robustness of these associations needs to be assessed further.

One reason for the inconsistency and non-replicability of previous findings is the small magnitude of effect sizes (Munafo et al. 2008). Personality is a complex trait and is, therefore, influenced by many genes, each with a relatively small effect (Greenberg et al. 2000; Savitz and Ramesar 2004). A putatively weak relationship between genotype and phenotype should only be detectable through the use of large samples (Savitz and Ramesar 2004). Yet, most studies relied on samples that were too small to reliably detect the effects of specific genes (e.g., Brummett et al., 2003; Dragan & Oniszczenko, 2006; Gillihan et al., 2007; Tochigi et al., 2006a, b). Large-scale studies, such as the WLS, offer the best means of achieving sufficient power to detect associations of small magnitude. Moreover, the WLS can overcome other limitations of existing research. Many studies of genetic influences on personality were based on clinical or non-representative samples, e.g., depressed patients (Brummett et al. 2003), families with alcoholism (Stoltenberg et al. 2002), families with high blood pressure (Sen et al. 2004), and twin studies (Jang et al. 2001; Eley et al. 2003; Middeldorp et al. 2007). In contrast, the WLS is broadly representative of white non-Hispanic young-old men and women. Finally, research has largely focused on young or early midlife adults (Dragan and Oniszczenko 2006; Rosenberg et al. 2006; Gillihan et al. 2007). Thus, little is known about genetic influences on personality in a general population of older adults, and findings from the WLS can provide important insights in this respect.

Like other complex behavioral outcomes, personality traits result from the interplay of genetic variations with environmental influences (Loehlin 1992; National Research Council

2006). The gene-environment interaction approach assumes that environmental factors, such as life events, influence personality, and genes affect individual susceptibility to environmental challenges (Caspi et al. 2003; Moffitt et al. 2005; Rutter et al. 2006). The WLS measures a wide array of life transitions; thus, we may examine whether certain genotypes increase the likelihood of personality change following stressful events, such as widowhood.

Cognition and Decision-Making

Over the past 20 years, both the government and employers have shifted more of the responsibility associated with planning and managing the retirement years towards individuals. The capacity of individuals to make good choices about their investments, medical care, insurance, and other domains of increased uncertainty and personal responsibility may be an important determinant of both financial and health-related well-being in old age. Moreover, technological advances—ranging from the expansion of the Internet to the development of life-extending medical technologies—also offer older adults an unprecedented number of options about health insurance, choice of health care providers, and types of medical treatments. The contexts in which people make these choices may play a significant role in shaping the quality of those choices, and ultimately, in the quality of life for older Americans.

One of the key baseline measures in the WLS is a standard IQ test, the Henmon-Nelson Test of Mental Ability (Henmon and Holt 1931; Henmon and Nelson 1946; Henmon and Nelson 1954), which was administered to about 70 percent of graduates in their freshman year of high school and to all of them in the junior year. Coverage of selected siblings is about 80 percent. In 1993-94, the surveys administered a single cognitive assessment, the abstract reasoning (similarities) module of the WAIS (Wechsler 1958; Wechsler 1981). In 2004-05, we repeated

this assessment, but also added a digit recall and ordering task, an immediate and delayed word recall task, and timed tests of verbal fluency. We plan to repeat these assessments and add several more intensive assessments of cognition in the forthcoming round of home interviews. While these assessments are neither optimal nor comparable across time, we believe that they provide a feasible basis for assessments of levels and changes in cognition across the life course within the WLS samples.

In this connection, we are collaborating with the economist, David Laibson, and his colleagues about several lines of work with the WLS. We plan collaborate in analyzing our data for many of the nearly 400 SNPs that Laibson and colleagues are using in analyses of the Age, Gene/Environment Susceptibility-Reykjavik Study (AGES-RS) (Benjamin et al. 2006). That is, we will cross-validate preliminary findings from AGES-RS. We also plan to add decision-making tasks designed by Laibson and colleagues to the forthcoming wave of WLS surveys. Most participants will undertake a hypothetical set of tasks pertaining to small, immediate payoffs vs. later payoffs with known probabilities, while a highly stratified subsample of graduates and their siblings will play the same games for modest, but real monetary stakes. While the AGES-RS is very rich in biomedical measures, the WLS has far more evidence about actual decision-making across the life-course, beginning with the decisions about post-secondary school activities that were at issue when the study began (Sewell et al. 2004).

Genetic Data in the WLS

Until recently, the WLS had obtained no direct biomedical measurements. In 1993-94 and 2004-05, we asked a large number of questions about health. We had ascertained weight and

height contemporaneously and retrospectively, and we have used high school yearbook photographs to obtain a proxy measure of BMI in adolescence (Reither et al. 2009).

Early in 2007, we collected approximately 4500 DNA samples from WLS graduates by mail using Oragene/DNA saliva kits. These cover just under 70 percent of survivors who participated fully (by mail and telephone) in the 2004 graduate surveys. In 2008, we collected 2300 additional samples from a sample of biological siblings of the graduates. Briefly, participants are sent an Oragene DNA/Genotek kit containing a vial somewhat like an oversized contact lens case (http://www.dnagenotek.com/products_oragene.htm). They are instructed to place saliva in the lower part of the container, then screw on the top, shake it, and return it in a sealed bag. When the vial is sealed, a preservative is released. The average DNA yield is much larger than that in more widely used collection schemes, e.g., a median of 2 μg for a buccal swab vs. 110 μg for the Oragene kit.

While response was acceptable for a mail-out protocol, we plan to make a second attempt to collect DNA and consent forms from non-respondents during the forthcoming home interviews. Collection of DNA samples has raised serious security issues, and we have made every effort to eliminate any possibility of personal re-identification of the samples. We have a completely different set of identification numbers for the DNA samples from those used elsewhere in the study, and no identifiable information is directly linked to the samples. The key link between DNA and other WLS data is very tightly controlled. While we expect to share the DNA data – like all other WLS data – with qualified investigators, our current plan is for statistical analysis to take place only within the secure data enclave of the Center for Demography of Health and Aging at the University of Wisconsin-Madison.

Our DNA samples can be kept indefinitely at room temperature, so they would be available for repeated assays for many years to come without any further action on our part, other than to store them at room temperature. Initial laboratory work has confirmed the quality and volume of the DNA samples collected to date, but because there is a possibility of isolating other useful biological indicators from the saliva, we have extracted and frozen the samples in a laboratory with a secure backup power source. Initial laboratory assessments have established that most of our DNA samples are large enough to permit 10 to 12 independent sets of assays as new scientific opportunities arise. With funding now in hand, we have completed an initial round of assays to identify 95 alleles, including the APOE gene (a genetic marker for late onset Alzheimer's disease) and 5-HTTLPR, 5HTR1A, 5HTR2A, 5HTR2C, MAOA, TPH, DAT1, DRD2, and DRD4 (markers for psychological characteristics, in particular, personality traits and indicators of psychological distress, e.g., depressive symptoms, alcohol use, hostility, and anxiety).

Concluding Remarks

Large-scale longitudinal studies of aging have many strengths. They tell us “the big picture,” by offering true descriptions of population characteristics. They make it possible to model longitudinal processes. They often provide relational data by directly or indirectly collecting the characteristics of individuals related to primary participants. For example, the HRS typically samples both members of a married couple, while the WLS includes the siblings (and sometimes, spouses) of graduates. These studies also provide insights into multiple social and economic contexts, e.g., as defined by gender, race-ethnicity, geographic location, or type of employment. Such studies can also tell us something about “the small picture,” unique situations.

For example, most research on parents with disabled adult children is based on clinical cases or “snowball samples,” while the size of the WLS has made it possible to identify hundreds of such cases in a representative sample (Greenberg et al. 1993; Seltzer et al. 1995; Greenberg et al. 1999; Seltzer et al. 2001; Seltzer et al. 2004; Taylor et al. 2008). Similar observations apply to the ADAMS spin-off of HRS, which differs from most studies of dementia because it pertains to a representative sample of such persons.

A final strength of large-scale longitudinal studies is the possibility of serendipitous research opportunities and findings. The WLS studies of parents with disabled adult children began with the observation that several participants had identified such children and commented about them in the course of the 1993-94 interviews.

Another such opportunity followed – by several years – an introductory talk that Hauser gave to a group of cancer specialists. An anesthesiologist was present, and he later proposed that the WLS could be used to study whether experience with various types of surgeries and anesthetics was associated with later cognitive decline, including Alzheimer’s Disease. Thus, the forthcoming round of the WLS will include a brief series of questions to identify surgeries with general anesthesia. Those who have had such an experience will be asked whether they are willing to participate in an additional telephone interview about their surgery. A sample of such cases will be re-contacted in order to identify the time and place of the surgery and to obtain consent for a record search. Then, a research nurse will obtain the original hospital record of the event. That set of data will then be available to investigate the consequences of various surgeries and modes of anesthesia.

There are also many challenges to success in large-scale longitudinal studies. In the past, most such studies have been of short duration. They have often been short of bio-indicators – though that situation is improving quickly. Such studies are very costly, and loss of population coverage is always a threat to validity. They are also burdensome to participants; it is a remarkable fact that, despite overall declines in willingness to participate in social surveys, well-designed and executed studies have been able to maintain participation over long periods of time. The exemplars in the U.S. include the HRS and WLS and, also, the 1979 National Longitudinal Study of Youth. Finally, as discussed above in relation to HRS biomarker collection, the protection of research participants is a continuing challenge.

Ideally, we should like to see a system of continuous longitudinal studies in the U.S. that would cover the entire population from birth to death. This is not to suggest that there should be one central source of such studies, but that universities and government entities should collaborate in moving toward that goal. Each such study should be interdisciplinary in design, measurement, and analysis. The several studies should overlap in content and design, but not be so similar that the same gaps or problems could occur in every study. There should be a serious effort to establish comparability in the measurement of key variables and, of course, sharing of data to the greatest extent possible within the limits imposed by protection of privacy and confidentiality of participants.

Cited References

- American Medical Association, Council on Scientific Affairs. 1996. "Good Care of the Dying Patient." *Journal of the American Medical Association* 275:474-78.
- Benin, Mary Holland and David R. Johnson. 1984. "Sibling Similarities in Educational Attainment: A Comparison of Like-Sex and Cross-Sex Sibling Pairs." *Sociology of Education* 57:11-21.
- Benjamin, Daniel J., Christopher F. Chabris, Edward Glaeser, Vilmundur Gudnason, Tamara B. Harris, David Laibson, Lenore Launer, and Shaun Purcell. 2006. "Genoeconomics." in *National Research Council Workshop on Collecting and Utilizing Biological Indicators and Genetic Information in Social Science Surveys*. Washington, DC.
- Benjamin, J., L. Li, C. Patterson, B. D. Greenberg, D. L. Murphy, and D. H. Hamer. 1996. "Population and Familial Association between the D4 Dopamine Receptor Gene and Measures of Novelty Seeking." *Nat Genet* 12:81-4.
- Bielby, William T., Robert M. Hauser, and David L. Featherman. 1977. "Response Errors of Black and Nonblack Males in Models of the Intergenerational Transmission of Socioeconomic Status." *American Journal of Sociology* 82:1242-88.
- Blau, Peter M. and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: John Wiley and Sons.

- Bondy, B., A. Erfurth, S. de Jonge, M. Kruger, and H. Meyer. 2000. "Possible Association of the Short Allele of the Serotonin Transporter Promoter Gene Polymorphism (5-HTTLPR) with Violent Suicide." *Mol Psychiatry* 5:193-5.
- Bouchard, T. J. 1997. *The Genetics of Personality*. Boca Raton, FL: CCR Press.
- Brim, Orville Gilbert, Carol D. Ryff, and Ronald C. Kessler. 2004. *How Healthy Are We? : A National Study of Well-Being at Midlife*. Chicago :: University of Chicago Press.
- Brummett, B. H., I. C. Siegler, D. R. McQuoid, I. K. Svenson, D. A. Marchuk, and D. C. Steffens. 2003. "Associations among the Neo Personality Inventory, Revised and the Serotonin Transporter Gene-Linked Polymorphic Region in Elders: Effects of Depression and Gender." *Psychiatr Genet* 13:13-8.
- Burks, Barbara Stoddard, Dortha Williams Jensen, and Lewis M. Terman. 1959. *Follow-up Studies of a Thousand Gifted Children*, Edited by L. M. Terman. Stanford, California: Stanford University Press.
- Carrel, Kathleen S., Clifford A. Potts, Emily A. Campbell, Education National Institute of, and Research American Institutes for. 1975. *Project Talent's Nonrespondent Follow-up Survey the 10th Grade Special Sample : Interim Report*. Palo Alto, Calif: American Institutes for Research.
- Caspi, A. and T. E. Moffitt. 2003. "Gene–Environment Interactions in Psychiatry: Joining Forces with Neuroscience." *J. Neurosci* 23:1125-1132.
- Caspi, A., K. Sugden, T. E. Moffitt, A. Taylor, I. W. Craig, H. Harrington, J. McClay, J. Mill, J. Martin, A. Braithwaite, and R. Poulton. 2003. "Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene." *Science* 301:386-9.

- Christensen, Kaare, Lise Bathum, and Lene Christiansen. 2008. "Biological Indicators and Genetic Information in Danish Twin and Oldest-Old Surveys." in *Biosocial Surveys*, edited by M. Weinstein, J. W. Vaupel, and K. W. Wachter. Washington, DC: National Academies Press.
- Clausen, John A. 1993. *American Lives: Looking Back at the Children of the Great Depression*. New York: The Free Press.
- Collier, D. A., G. Stober, T. Li, A. Heils, M. Catalano, D. Di Bella, M. J. Arranz, R. M. Murray, H. P. Vallada, D. Bengel, C. R. Muller, G. W. Roberts, E. Smeraldi, G. Kirov, P. Sham, and K. P. Lesch. 1996. "A Novel Functional Polymorphism within the Promoter of the Serotonin Transporter Gene: Possible Role in Susceptibility to Affective Disorders." *Mol Psychiatry* 1:453-60.
- Costa, D and M Kahn. 2009. *Heroes and Cowards: The Social Face of War*: Princeton Univ Pr.
- Costa, DL and ME Kahn. 2003. "Cowards and Heroes: Group Loyalty in the American Civil War*." *Quarterly Journal of Economics* 118:519-548.
- Cureton, Edward Eugene and American Institutes for Research. 1968. *A Factor Analysis of Project Talent Tests and Four Other Test Batteries*. Palo Alto: American Institutes for Research.
- Digman, J. M. 1990. "Personality Structure: Emergence of the Five-Factor Model." *Annual Review of Psychology* 41:417-440.
- Douglas, James William Bruce. 1964. *The Home and the School: A Study of Ability and Attainment in the Primary School*. London: Macgibbon & Kee.

- Douglas, James William Bruce and J. M. Blomfield. 1958. *Children under Five: The Results of a National Survey Made by a Joint Committee of the Institute of Child Health (University of London), the Society of Medical Officers of Health, and the Population Investigation Committee*. London: Allen & Unwin.
- Douglas, James William Bruce, J. M. Ross, and Howard Russell Simpson. 1968. *All Our Future: A Longitudinal Study of Secondary Education*. London: P. Davies.
- Dragan, W. L. and W. Oniszczenko. 2006. "Association of a Functional Polymorphism in the Serotonin Transporter Gene with Personality Traits in Females in a Polish Population." *Neuropsychobiology* 54:45-50.
- Duncan, Otis Dudley. 1968. "Ability and Achievement." *Eugenics Quarterly* 15:1-11.
- Duncan, Otis Dudley, David L. Featherman, and Beverly Duncan. 1972. *Socioeconomic Background and Achievement*. New York: Seminar Press.
- Eichhammer, P., P. G. Sand, P. Stoertebecker, B. Langguth, M. Zowe, and G. Hajak. 2005. "Variation at the Drd4 Promoter Modulates Extraversion in Caucasians." *Mol Psychiatry* 10:520-2.
- Elder, Glen H., Jr. 1974. *Children of the Great Depression*. Chicago: University of Chicago Press.
- Eley, T. C., E. Tahir, A. Angleitner, K. Harriss, J. McClay, R. Plomin, R. Riemann, F. Spinath, and I. Craig. 2003. "Association Analysis of Maa and Comt with Neuroticism Assessed by Peers." *Am J Med Genet B Neuropsychiatr Genet* 120:90-6.
- Eysenck, H. J. 1992. "Four Ways Five Factors Are Not Basic." *Personality and individual differences* 13:667-673.

- Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998a. "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources* 33:251-299.
- . 1998b. "An Analysis of the Impact of Sample Attrition on the Second Generation of Respondents in the Michigan Panel Study of Income Dynamics." *Journal of Human Resources* 33:300-344.
- Flanagan, John Clemans, William W. Cooley, P. R. Lohnes, L. F. Schoenfeldt, R. W. Holdeman, J. Combs, and et al. 1966. *Project Talent: One-Year Follow-up Studies*. Pittsburgh, PA: American Institutes for Research, University of Pittsburgh.
- Fogel, RW. 1993. "New Sources and New Techniques for the Study of Secular Trends in Nutritional Status, Health, Mortality, and the Process of Aging." NBER.
- . 1994. "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy." *The American Economic Review*:369-395.
- . 2004. "Changes in the Process of Aging During the Twentieth Century: Findings and Procedures of the Early Indicators Project." *Population and Development Review*:19-47.
- Fogel, RW and DL Costa. 1997. "A Theory of Technophysio Evolution, with Some Implications for Forecasting Population, Health Care Costs, and Pension Costs." *Demography*:49-66.
- Ford, ES, WH Giles, and WH Dietz. 2002. "Prevalence of the Metabolic Syndrome among Us Adults Findings from the Third National Health and Nutrition Examination Survey." Pp. 356-359, vol. 287: Am Med Assoc.
- Ford, ES, WH Giles, and AH Mokdad. 2004. "Increasing Prevalence of the Metabolic Syndrome among Us Adults a Table Elsewhere in This Issue Shows Conventional and Systeme

- International (Si) Units and Conversion Factors for Many Substances." Pp. 2444-2449, vol. 27: Am Diabetes Assoc.
- Frederick, Carl B. and Robert M. Hauser. 2008. "Have We Put an End to Social Promotion? Changes in School Progress among Children Aged 6 to 17 from 1972 to 2005." *Demography* 45:719-740.
- Freese, Jeremy. 2006. "Cognitive Skills and Survey Nonresponse: Evidence from Two Longitudinal Studies." Center for Demography and Ecology, University of Wisconsin-Madison, Madison, Wisconsin.
- Freese, Jeremy, Sheri Meland, and William Irwin. 2006a. "Expressions of Positive Emotion in Photographs, Personality, and Later-Life Marital and Health Outcomes." *Journal of Research and Personality*.
- Freese, Jeremy and Salvador Rivas. 2005. "Cognition, Personality, and Individual Response to Technological Change: The Case of the Internet Adoption."
- Freese, Jeremy, Salvador Rivas, and Eszter Hargittai. 2006b. "Cognitive Ability and Internet Use among Older Adults: The Digital Divide in the Twenty-First Century." *Poetics* 34:236-249.
- Fried, LP, CM Tangen, J Walston, AB Newman, C Hirsch, J Gottdiener, T Seeman, R Tracy, WJ Kop, and G Burke. 2001. "Frailty in Older Adults Evidence for a Phenotype." *Journals of Gerontology Series A: Biological and Medical Sciences* 56:146-157.
- Fried, TR and EH Bradley. 2003. "What Matters to Seriously Ill Older Persons Making End-of-Life Treatment Decisions?: A Qualitative Study." *Journal of Palliative Medicine* 6:237-244.

- Gathman, ECH, DW Maynard, and NC Schaeffer. 2008. "The Respondents Are All above Average: Compliment Sequences in a Survey Interview." *Research on Language & Social Interaction* 41:271-301.
- Gillespie, Michael W. 1991. "Conceptual Issues in Models of Sibling Resemblance: Comment on Hauser and Mossel (1985) and Hauser (1988)." *American Journal of Sociology* 97:196-206.
- Gillihan, S. J., M. J. Farah, G. M. Sankoorikal, J. Breland, and E. S. Brodtkin. 2007. "Association between Serotonin Transporter Genotype and Extraversion." *Psychiatr Genet* 17:351-4.
- Goodglass, H and E Kaplan. 1983. *Boston Naming Test*: Lea & Febiger Philadelphia.
- Greenberg, B. D., Q. Li, F. R. Lucas, S. Hu, L. A. Sirota, J. Benjamin, K. P. Lesch, D. Hamer, and D. L. Murphy. 2000. "Association between the Serotonin Transporter Promoter Polymorphism and Personality Traits in a Primarily Female Population Sample." *Am J Med Genet* 96:202-16.
- Greenberg, J. S., M. M. Seltzer, and J. R. Greenley. 1993. "Aging Parents of Adults with Disabilities: The Gratifications and Frustrations of Later-Life Caregiving." *Gerontologist* 33:542-50.
- Greenberg, J. S., M. M. Seltzer, G. I. Orsmond, and M. W. Krauss. 1999. "Siblings of Adults with Mental Illness or Mental Retardation: Current Involvement and Expectation of Future Caregiving." *Psychiatr Serv* 50:1214-9.
- Guo, G. and Y. Tong. 2006. "Age at First Sexual Intercourse, Genes, and Social Context: Evidence from Twins and the Dopamine D4 Receptor Gene." *Demography* 43:747-769.

- Halpern, C. T., C. E. Kaestle, G. Guo, and D. D. Hallfors. 2006. "Gene-Environment Contributions to Young Adult Sexual Partnering." *Arch Sex Behav*.
- Hamer, D.H. 2002. "Genetics of Sexual Behavior." Pp. 257-272 in *Molecular Genetics and the Human Personality*, edited by J. Benjamin, R. P. Ebstein, and R. H. Belmaker. Washington, DC: American Psychiatric Publishing.
- Harris, DR and JJ Sim. 2002. "Who Is Multiracial? Assessing the Complexity of Lived Race." *American Sociological Review*:614-627.
- Hauser, Robert M. 1984. "Some Cross-Population Comparisons of Family Bias in the Effects of Schooling on Occupational Status." *Social Science Research* 13:159-87.
- . 1987. "Sharing Data: It's Time for Asa Journals to Follow the Folkways of a Scientific Sociology." *American Sociological Review* 52.
- . 1988. "A Note on Two Models of Sibling Resemblance." *American Journal of Sociology* 93:1401-23.
- . 1991. "Reply to Michael Gillespie's Comment on Models of Sibling Resemblance." *American Journal of Sociology* 97:206-209.
- . 2005. "Survey Response in the Long Run: The Wisconsin Longitudinal Study." *Field Methods* 17:3-29.
- . 2009a. "Beauty, Brains, and Bulges: Life Course Consequences of Adolescent Attractiveness, Ability, and Adiposity." in *Research Committee on Social Stratification, International Sociological Association*. Beijing, China.

- . 2009b. "The Wisconsin Longitudinal Study: Designing a Study of the Life Course." Pp. 29-50 in *The Craft of Life Course Research*, edited by G. H. Elder and J. Z. Giele. New York: The Guilford Press.
- Hauser, Robert M. and Peter A. Mossel. 1985. "Fraternal Resemblance in Educational Attainment and Occupational Status." *American Journal of Sociology* 91:650-73.
- . 1988. "Some Structural Equation Models of Sibling Resemblance in Educational Attainment and Occupational Status." Pp. 108-37 in *Structural Modeling by Example: Applications in Educational, Sociological, and Behavioral Research*, edited by P. Cuttance and R. Ecob. Cambridge: Cambridge University Press.
- Hauser, Robert M., Devah I. Pager, and Solon J. Simmons. 2004a. "Race-Ethnicity, Social Background, and Grade Retention." Pp. 97-114 in *Can Unlike Students Learn Together? Grade Retention, Tracking, and Grouping*, edited by H. J. Walberg, A. J. Reynolds, and M. C. Wang. Greenwich, Connecticut: Information Age Publishing.
- Hauser, Robert M. and William H. Sewell. 1986. "Family Effects in Simple Models of Education, Occupational Status, and Earnings: Findings from the Wisconsin and Kalamazoo Studies." *Journal of Labor Economics* 4:S83-S115.
- Hauser, Robert M., Jennifer T. Sheridan, and John Robert Warren. 1999. "Socioeconomic Achievements of Siblings in the Life Course: New Findings from the Wisconsin Longitudinal Study." *Research on Aging* 21:338-378.
- Hauser, Robert M., Solon J. Simmons, and Devah I. Pager. 2004b. "High School Dropout, Race-Ethnicity, and Social Background from the 1970s to the 1990s." Pp. 85-106 in *Dropouts*

- in America: Confronting the Graduation Rate Crisis*, edited by G. Orfield. Cambridge, Massachusetts: Harvard Educational Publishing Group.
- Hauser, Robert M. and John Robert Warren. 1997. "Socioeconomic Indexes for Occupations: A Review, Update, and Critique." Pp. 177-298 in *Sociological Methodology 1997*, edited by A. E. Raftery. Cambridge: Basil Blackwell.
- Hauser, Robert M., John Robert Warren, Min-Hsiung Huang, and Wendy Y. Carter. 2000. "Occupational Status, Education, and Social Mobility in the Meritocracy." Pp. 179-229 in *Meritocracy and Economic Inequality*, edited by K. Arrow, S. Bowles, and S. Durlauf. Princeton: Princeton University Press.
- Hauser, Robert M. and Robert J. Willis. 2005. "Survey Design and Methodology in the Health and Retirement Study and the Wisconsin Longitudinal Study." Pp. 209-235 in *Aging, Health, and Public Policy: Demographic and Economic Perspectives*, edited by L. J. Waite. New York: Population Council.
- Hauser, Robert M. and Raymond Sin-Kwok Wong. 1989. "Sibling Resemblance and Inter-Sibling Effects in Educational Attainment." *Sociology of Education* 62:149-71.
- Henmon, V. A. C. and Frank Oscar Holt. 1931. *A Report on the Administration of Scholastic Aptitude Tests to 34,000 High School Seniors in Wisconsin in 1929 and 1930 Prepared for the Committee on Cooperation, Wisconsin Secondary Schools and Colleges*. Madison: Bureau of Guidance and Records of the University of Wisconsin.
- Henmon, V. A. C. and M. J. Nelson. 1946. *Henmon-Nelson Tests of Mental Ability, High School Examination - Grades 7 to 12 - Forms a, B, and C. Teacher's Manual*. Boston: Houghton-Mifflin Company.

- . 1954. *The Henmon-Nelson Tests of Mental Ability. Manual for Administration*. Boston: Houghton-Mifflin Company.
- Herrnstein, Richard J. and Charles Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: The Free Press.
- Holden, K and J Brand. 2003. "Income Change and Distribution Upon Widowhood: Comparison of Britain, Us, and Germany." *Pensions: Challenges and Reform, Aldershot, Ashgate*:211-225.
- Homer, N, S Szelinger, M Redman, D Duggan, W Tembe, J Muehling, JV Pearson, DA Stephan, SF Nelson, and DW Craig. 2008. "Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-Density Snp Genotyping Microarrays." *PLoS Genetics* 4.
- Huang, Min-Hsiung and Taissa S. Hauser. 2009. "Tracking Persons from High School through Adult Life: Lessons from the Wisconsin Longitudinal Study." in *IEAS Conference on Contemporary European and American Societies*. Taipei, Taiwan.
- Jang, K. L., S. Hu, W. J. Livesley, A. Angleitner, R. Riemann, J. Ando, Y. Ono, P. A. Vernon, and D. H. Hamer. 2001. "Covariance Structure of Neuroticism and Agreeableness: A Twin and Molecular Genetic Analysis of the Role of the Serotonin Transporter Gene." *J Pers Soc Psychol* 81:295-304.
- Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.

- John, Oliver. 1990. "The Big Five Factor Taxonomy: Dimensions of Personality in the Natural Language and Questionnaires." Pp. 66-100 in *Handbook of Personality Theory and Research*, edited by L. A. Pervin. New York: Guilford.
- . 1991. "Big Five Inventory (Bfi-54)." University of California, Berkeley: Institute of Personality Assessment and Research.
- Jokela, M. 2009. "Physical Attractiveness and Reproductive Success in Humans: Evidence from the Late 20th Century United States." *Evolution and Human Behavior*.
- Kapteyn, A, PC Michaud, JP Smith, and A Van Soest. 2006. "Effects of Attrition and Non-Response in the Health and Retirement Study." Institute for the Study of Labor (IZA), Bonn, Germany.
- Kramer, B. J., A. Boelk, and C. Auer. 2006. "Family Conflict at the End-of-Life: Lessons Learned in a Model Program for Vulnerable Older Adults." *Journal of Palliative Medicine* 9:791-801.
- Krawczak, M., A. Trefilov, J. Berard, F. Bercovitch, M. Kessler, U. Saueremann, P. Croucher, P. Nurnberg, A. Widdig, and J. Schmidtke. 2005. "Male Reproductive Timing in Rhesus Macaques Is Influenced by the 5httlpr Promoter Polymorphism of the Serotonin Transporter Gene." *Biol Reprod* 72:1109-13.
- Kuo, Hsiang-Hui Daphne, Nadine F. Marks, Hyun-Joon Park, Taissa S. Hauser, and Robert M. Hauser. 2001. "Surveys of the Life Course and Aging: Some Comparisons." Center for Demography and Ecology The University of Wisconsin-Madison, Madison, Wisconsin.
- Laumann, Edward O. 2000. *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.

- Laumann, Edward O. and Robert T. Michael. 2001. *Sex, Love, and Health in America: Private Choices and Public Policies*. Chicago: University of Chicago Press.
- Lee, I. H., C. C. Cheng, Y. K. Yang, T. L. Yeh, P. S. Chen, and N. T. Chiu. 2005. "Correlation between Striatal Dopamine D2 Receptor Density and Neuroticism in Community Volunteers." *Psychiatry Res* 138:259-64.
- Lee, Mehng-Lurn. 1989. "Some Structural Models of Family and Inter-Sibling Effects on Educational Attainment: Findings from Wisconsin Data." Department of Sociology, University of Wisconsin-Madison.
- Lesch, K. P., D. Bengel, A. Heils, S. Z. Sabol, B. D. Greenberg, S. Petri, J. Benjamin, C. R. Muller, D. H. Hamer, and D. L. Murphy. 1996. "Association of Anxiety-Related Traits with a Polymorphism in the Serotonin Transporter Gene Regulatory Region." *Science* 274:1527-31.
- Lipkin, KM. 2006. "Brief Report: Identifying a Proxy for Health Care as Part of Routine Medical Inquiry." *Journal of General Internal Medicine* 21:1188-1191.
- Little, J. Kenneth. 1958. *A State-Wide Inquiry into Decisions of Youth About Education Beyond High School--Follow-up Studies*. Madison, WI: University of Wisconsin School of Education.
- . 1959. *Explorations into the College Plans and Experiences of High School Graduates: A State-Wide Inquiry*. Madison, WI: University of Wisconsin School of Education.
- Loehlin, J. C. 1992. *Genes and Environment in Personality Development*. Newburg Park, CA: Sage Publications.

- Lohnes, Paul R., Pittsburgh University of, Office Project Talent, Pittsburgh University of, and Education School of. 1966. *Measuring Adolescent Personality Project Talent Five-Year Follow-up Studies, Interim Report 1*. Pittsburgh: University of Pittsburgh.
- MacLean, Alair and Robert M. Hauser. 2000. "Socioeconomic Status and Depression among Adult Siblings." Center for Demography and Ecology, The University of Wisconsin-Madison, Madison, Wisconsin.
- . 2001. "Measuring Depression among Adult Siblings." Center for Demography and Ecology, The University of Wisconsin-Madison.
- Martin, Laurie T. and Laura D. Kubzansky. 2005. "Childhood Cognitive Performance and Risk of Mortality: A Prospective Cohort Study of Gifted Individuals." *Am. J. Epidemiol.* 162:887-890.
- McCrae, Robert R. and Paul T. Costa. 2003. *Personality in Adulthood: A Five-Factor Theory Perspective*. New York: Guilford Press.
- Meland, Sheri A. 2002. "Objectivity in Perceived Attractiveness: Development of a New Methodology for Rating Facial Physical Attractiveness." Unpublished MA thesis, Department of Sociology, University of Wisconsin-Madison.
- Middeldorp, C. M., E. J. de Geus, A. L. Beem, N. Lakenberg, J. J. Hottenga, P. E. Slagboom, and D. I. Boomsma. 2007. "Family Based Association Analyses between the Serotonin Transporter Gene Polymorphism (5-Httlpr) and Neuroticism, Anxiety and Depression." *Behav Genet* 37:294-301.

- Miller, W. B., D. J. Pasta, J. MacMurray, C. Chiu, H. Wu, and D. E. Comings. 1999. "Dopamine Receptor Genes Are Associated with Age at First Sexual Intercourse." *J Biosoc Sci* 31:43-54.
- Moffitt, T. E., A. Caspi, and M. Rutter. 2005. "Strategy for Investigating Interactions between Measured Genes and Measured Environments." *Journal of the American Medical Association* 62:473-481.
- Morrison, RS, AL Siu, RM Leipzig, CK Cassel, and DE Meier. 2000. "The Hard Task of Improving the Quality of Care at the End of Life." Pp. 743-747, vol. 160: Am Med Assoc.
- Munafo, M. R., T. G. Clark, L. R. Moore, E. Payne, R. Walton, and J. Flint. 2003. "Genetic Polymorphisms and Personality in Healthy Adults: A Systematic Review and Meta-Analysis." *Mol Psychiatry* 8:471-84.
- Munafo, M. R., B. Yalcin, S. A. Willis-Owen, and J. Flint. 2008. "Association of the Dopamine D4 Receptor (Drd4) Gene and Approach-Related Personality Traits: Meta-Analysis and New Data." *Biol Psychiatry* 63:197-206.
- National Institute on Aging and U.S. Census Bureau. 2009. "An Aging World: 2008." National Institutes of Health, U.S. Department of Commerce, Washington, DC.
- National Institute on Aging and U.S. Department of State. 2007. "Why Population Aging Matters: A Global Perspective." U.S. Department of Health and Human Services, U.S. Department of State, Washington, DC.
- National Institutes of Health. 2005. "State-of-the-Science Conference Statement on Improving End-of-Life Care." National Institutes of Health, Bethesda, Maryland.

National Research Council. 2006. *Genes, Behavior, and the Social Environment: Moving Beyond the Nature/Nurture Debate*. Washington, DC: National Academy Press.

National Research Council (U.S.), Panel to Review the National Children's Study Research Plan, National Research Council (U.S.), Committee on National Statistics, National Research Council (U.S.), Board on Children Youth and Families, and National Research Council (U.S.), Board on Population Health and Public Health Practice. 2008. *The National Children's Study Research Plan: A Review*. Washington, D.C.: National Academies Press.

National Research Council, Committee on Ensuring the Utility and Integrity of and Committee on Science Research Data in a Digital Age, Engineering, and Public Policy. 2009. *Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age*. Washington, D.C.: National Academy Press.

National Research Council, Committee on National Statistics. 1985. *Sharing Research Data*, Edited by S. E. Fienberg, M. E. Martin, and M. L. Straf. Washington, D.C.: National Academy Press.

National Research Council, Panel on the Census Bureau's Reengineered Survey of Income and Program Participation, Committee on National Statistics, Commission on Behavioral and Social Sciences and Education. 2009. *Reengineering the Survey of Income and Program Participations*, Edited by C. Citro and J. K. Scholz. Washington, D.C.: National Academy Press.

National Research Council, Panel to Evaluate the Survey of Income and Program Participation, Committee on National Statistics, Commission on Behavioral and Social Sciences and

- Education. 1993. *The Future of the Survey of Income and Program Participation*, Edited by C. Citro and G. Kalton. Washington, D.C.: National Academy Press.
- Neal, Derek A. and William R. Johnson. 1996. "The Role of Pre-Market Factors in Black-White Wage Differences." *Journal of Political Economy* 104:869-895.
- Neyer, Franz J. and Jens B. Asendorpf. 2001. "Personality-Relationship Transaction in Young Adulthood." *Journal of Personality and Social Psychology* 81:1190-1204.
- O'Brien, SJ. 2009. "Stewardship of Human Biospecimens, DNA, Genotype, and Clinical Data in the Gwas Era." *Annual Review of Genomics and Human Genetics* 10:9.1-9.17.
- Oden, Melita H. 1968. "The Fulfillment of Promise: 40-Year Follow-up of the Terman Gifted Group." *Genetic Psychology Monographs* 77:3-93.
- Olneck, Michael R. 1976. "The Determinants of Educational Attainment and Adult Status among Brothers: The Kalamazoo Study." Harvard Graduate School of Education.
- . 1977. "On the Use of Sibling Data to Estimate the Effects of Family Background, Cognitive Skills, and Schooling: Results from the Kalamazoo Brothers Study." Pp. 125-62 in *Kinometrics: Determinants of Socioeconomic Success within and between Families*, edited by P. Taubman. Amsterdam: North-Holland.
- Palloni, Alberto. 2006. "Reproducing Inequalities: Luck, Wallets, and the Enduring Effects of Childhood Health." *Demography* 43:587-615.
- Phelps, Erin, Frank F. Furstenberg, and Anne Colby. 2002. *Looking at Lives: American Longitudinal Studies of the Twentieth Century*. New York: Russell Sage Foundation.
- Plomin, R., M. J. Owen, and P. McGuffin. 1994. "The Genetic Basis of Complex Human Behaviors." *Science* 264:1733-9.

- Plug, E. and W. Vijverberg. 2003. "Schooling, Family Background, and Adoption: Is It Nature or Is It Nurture?" *Journal of Political Economy* 111:611-641.
- . 2005. "Does Family Income Matter for Schooling Outcomes? Using Adoptees as a Natural Experiment." *Economic Journal* 115:879-906.
- Plug, Erik. 2004. "Estimating the Effect of Mother's Schooling on Children's Schooling Using a Sample of Adoptees." *American economic review* 94:358-368.
- Powis, J, E Etchells, DK Martin, SK MacRae, and PA Singer. 2004. "Can A " Good Death" Be Made Better?: A Preliminary Evaluation of a Patient-Centred Quality Improvement Strategy for Severely Ill in-Patients." *BMC Palliative Care* 3:2.
- Pudrovska, T. 2008. "Psychological Implications of Motherhood and Fatherhood in Midlife: Evidence from Sibling Models." *Journal of Marriage and the Family* 70:168-81.
- Radloff, L. 1977. "The Ces-D Scale: A Self-Report Depression Scale for Research in the General Population." *Applied Psychological Measurement* 1:385-401.
- Reither, Eric. 2005. "Why Are Our Waistlines Expanding? Age-Period-Cohort Analyses of the Obesity Epidemic and a Critical Examination of Mass Preparation Theory." Department of Sociology, University of Wisconsin-Madison.
- Reither, Eric N., Robert M. Hauser, and Karen E. Swallen. 2009. "Predicting Adult Health and Mortality from Adolescent Facial Characteristics in Yearbook Photographs." *Demography* 46:27-41.
- Risch, N, R Herrell, T Lehner, KY Liang, L Eaves, J Hoh, A Griem, M Kovacs, J Ott, and KR Merikangas. 2009. "Interaction between the Serotonin Transporter Gene (5-Httlpr), Stressful Life Events, and Risk of Depression: A Meta-Analysis." *JAMA* 301:2462.

- Robins, L. N., J. E. Helzer, J. Croughan, and K. S. Ratcliff. 1981. "National Institute of Mental Health Diagnostic Interview Schedule. Its History, Characteristics, and Validity." *Arch Gen Psychiatry* 38:381-9.
- Rosenberg, S., A. R. Templeton, P. D. Feigin, D. Lancet, J. S. Beckmann, S. Selig, D. H. Hamer, and K. Skorecki. 2006. "The Association of DNA Sequence Variation at the Maa0 Genetic Locus with Quantitative Behavioural Traits in Normal Males." *Hum Genet* 120:447-59.
- Rutter, M., T. E. Moffitt, and A. Caspi. 2006. "Gene–Environment Interplay and Psychopathology: Multiple Varieties but Real Effects." *Journal of Child Psychology and Psychiatry* 47:226-261.
- Samochowiec, J., S. Syrek, P. Michal, A. Ryzewska-Wodecka, A. Samochowiec, J. Horodnicki, M. Zakrzewska, and J. Kucharska-Mazur. 2004. "Polymorphisms in the Serotonin Transporter and Monoamine Oxidase a Genes and Their Relationship to Personality Traits Measured by the Temperament and Character Inventory and Neo Five-Factor Inventory in Healthy Volunteers." *Neuropsychobiology* 50:174-81.
- Savitz, J. B. and R. S. Ramesar. 2004. "Genetic Variants Implicated in Personality: A Review of the More Promising Candidates." *Am J Med Genet B Neuropsychiatr Genet* 131:20-32.
- Seeman, T. E., B. S. McEwen, J. W. Rowe, and B. H. Singer. 2001. "Allostatic Load as a Marker of Cumulative Biological Risk: Macarthur Studies of Successful Aging." *Proc Natl Acad Sci U S A* 98:4770-5.
- Seeman, TE, E Crimmins, MH Huang, B Singer, A Bucur, T Gruenewald, LF Berkman, and DB Reuben. 2004. "Cumulative Biological Risk and Socio-Economic Differences in

- Mortality: MacArthur Studies of Successful Aging." *Social Science & Medicine* 58:1985-1997.
- Seltzer, M. M., J. S. Greenberg, F. J. Floyd, Y. Pettee, and J. Hong. 2001. "Life Course Impacts of Parenting a Child with a Disability." *American Journal on Mental Retardation* 106:265-286.
- Seltzer, M. M., J. S. Greenberg, and M. W. Krauss. 1995. "A Comparison of Coping Strategies of Aging Mothers of Adults with Mental Illness or Mental Retardation." *Psychol Aging* 10:64-75.
- Seltzer, Marsha Mailick, Jan S. Greenberg, Frank J. Floyd, and Jinkuk Hong. 2004. "Accommodative Coping and Well-Being of Midlife Parents of Children with Mental Health Problems or Developmental Disabilities." *American Journal of Orthopsychiatry* 74:187-195.
- Sen, S., S. Villafuerte, R. Nesse, S. F. Stoltenberg, J. Hopcian, L. Gleiberman, A. Weder, and M. Burmeister. 2004. "Serotonin Transporter and Gabaa Alpha 6 Receptor Variants Are Associated with Neuroticism." *Biol Psychiatry* 55:244-9.
- Sewell, William H. and Robert M. Hauser. 1977. "On the Effects of Families and Family Structure on Achievements." Pp. 255-83 in *Kinometrics: The Determinants of Educational Attainment, Mental Ability, and Occupational Success within and between Families*, edited by P. J. Taubman. Amsterdam: North Holland.
- Sewell, William H., Robert M. Hauser, Kristen W. Springer, and Taissa S. Hauser. 2004. "As We Age: The Wisconsin Longitudinal Study, 1957-2001." Pp. 3-111 in *Research in Social Stratification and Mobility*, vol. 20, edited by K. Leicht. London: Elsevier.

- Sicinski, Kamil. 2009. "Essays on Beauty and Economic Success." Doctoral Thesis, Economics, University of Wisconsin, Madison.
- Stoltenberg, S. F., G. R. Twitchell, G. L. Hanna, E. H. Cook, H. E. Fitzgerald, R. A. Zucker, and K. Y. Little. 2002. "Serotonin Transporter Promoter Polymorphism, Peripheral Indexes of Serotonin Function, and Personality Measures in Families with Alcoholism." *Am J Med Genet* 114:230-4.
- Taylor, JL, JS Greenberg, MM Seltzer, and FJ Floyd. 2008. "Siblings of Adults with Mild Intellectual Deficits or Mental Illness: Differential Life Course Outcomes." *Journal of family psychology: JFP: journal of the Division of Family Psychology of the American Psychological Association (Division 43)* 22:905.
- Terman, Lewis M. 1925. *Mental and Physical Traits of a Thousand Gifted Children*. Stanford, California: Stanford University Press.
- Terman, Lewis M. and Melita H. Oden. 1959a. *The Gifted Child Grows Up: Twenty-Five Years' Follow-up of a Superior Group*, Edited by L. M. Terman. Stanford, California: Stanford University Press.
- . 1959b. *The Gifted Group at Mid-Life: Thirty-Five Years' Follow-up of the Superior Child*, Edited by L. M. Terman. Stanford, California: Stanford University Press.
- Terrazas, Aaron. 2009, "Older Immigrants in the United States", Retrieved August 31, 2009, 2009 (<http://www.migrationinformation.org/USfocus/display.cfm?ID=727#top>).
- Tochigi, M., H. Hibino, T. Otowa, C. Kato, T. Marui, T. Ohtani, T. Umekage, N. Kato, and T. Sasaki. 2006a. "Association between Dopamine D4 Receptor (Drd4) Exon Iii Polymorphism and Neuroticism in the Japanese Population." *Neurosci Lett* 398:333-6.

- Tochigi, M., T. Otowa, H. Hibino, C. Kato, T. Otani, T. Umekage, T. Utsumi, N. Kato, and T. Sasaki. 2006b. "Combined Analysis of Association between Personality Traits and Three Functional Polymorphisms in the Tyrosine Hydroxylase, Monoamine Oxidase a, and Catechol-O-Methyltransferase Genes." *Neurosci Res* 54:180-5.
- U.S. Bureau of the Census. 2008, "National Population Projections, Released 2008 (Based on Census 2000)" *U.S. Population Projections*,
<http://www.census.gov/population/www/projections/methodstatement.html>,
<http://www.census.gov/population/www/projections/downloadablefiles.html>).
- Wacker, J., M. Reuter, J. Hennig, and G. Stemmler. 2005. "Sexually Dimorphic Link between Dopamine D2 Receptor Gene and Neuroticism-Anxiety." *Neuroreport* 16:611-4.
- Wadsworth, Michael E. J. 1991. *The Imprint of Time: Childhood, History, and Adult Life*. Oxford: Clarendon Press.
- Warren, John Robert, Jennifer T. Sheridan, and Robert M. Hauser. 2002. "Occupational Stratification across the Life Course: Evidence from the Wisconsin Longitudinal Study." *American Sociological Review* 67:432-455.
- Wechsler, David. 1958. *The Measurement and Appraisal of Adult Intelligence*. London: Balliere Tindall and Cox Ltd.
- . 1981. *Wais-R Manual: Wechsler Adult Intelligence Scale-Revised*. San Antonio: The Psychological Corporation.
- Weinstein, Maxine and Robert J. Willis. 2001. "Stretching Social Surveys to Include Bioindicators: Possibilities for the Health and Retirement Study, Experience from the Taiwan Study of the Elderly." Pp. 250-275 in *Cells and Surveys: Should Biological*

Measures Be Included in Social Science Research?, edited by C. E. Finch, J. W. Vaupel, and K. G. Kinsella. Washington, D.C.: National Academy Press.

Willeit, M., N. Praschak-Rieder, A. Neumeister, P. Zill, F. Leisch, J. Stastny, E. Hilger, N.

Thierry, A. Konstantinidis, D. Winkler, K. Fuchs, W. Sieghart, H. Aschauer, M.

Ackenheil, B. Bondy, and S. Kasper. 2003. "A Polymorphism (5-Httlpr) in the Serotonin Transporter Promoter Gene Is Associated with Dsm-Iv Depression Subtypes in Seasonal Affective Disorder." *Mol Psychiatry* 8:942-6.

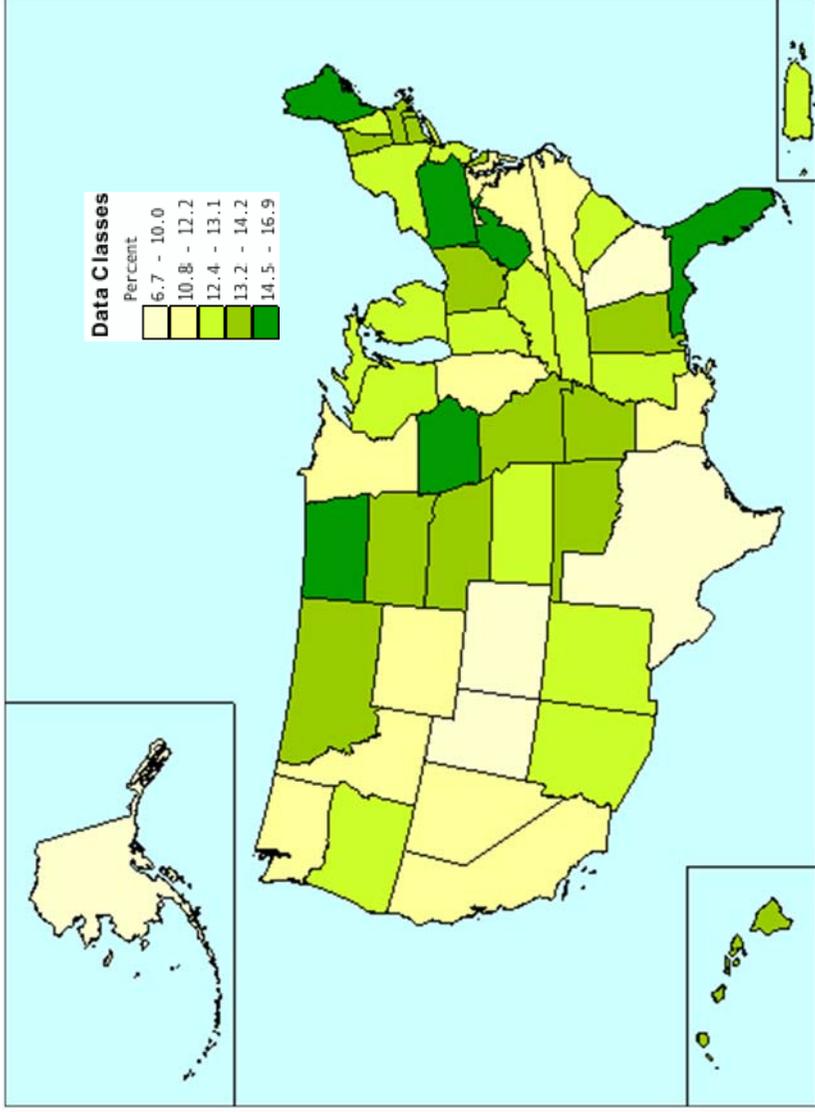
Willis-Owen, S. A., M. G. Turri, M. R. Munafo, P. G. Surtees, N. W. Wainwright, R. D. Brixey,

and J. Flint. 2005. "The Serotonin Transporter Length Polymorphism, Neuroticism, and Depression: A Comprehensive Assessment of Association." *Biol Psychiatry* 58:451-6.

Young, Copeland H., Kristen L. Savola, and Erin Phelps. 1991. *Inventory of Longitudinal*

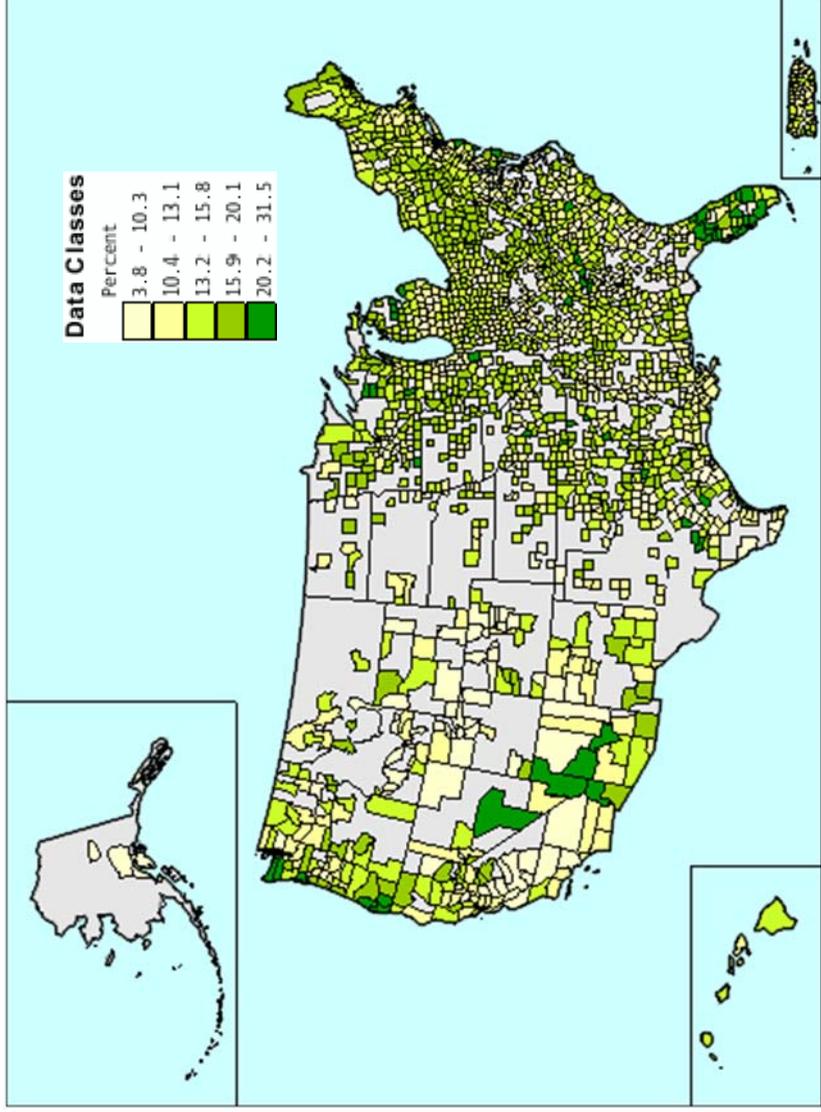
Studies in the Social Sciences. Newbury Park, Calif. :: Sage Publications.

Figure 1. Percentage of the Total Population Who Are 65 Years and Over by State: 2007



Source: 2005-07 American Community Survey 3-Year Estimates from <http://factfinder.census.gov>.

Figure 2. Percentage of the Total Population Who Are 65 Years and Over by County: 2007



Source: 2005-07 American Community Survey 3-Year Estimates from <http://factfinder.census.gov>.

Figure 3. HRS Psychosocial Content

Well-being

- Life satisfaction
- Domain satisfaction*
- Positive affect
- Negative affect
- Purpose in life
- Personal growth
- Self acceptance
- Financial strain

Lifestyle

- Activities in life*
- Neighborhood evaluation
- Religiosity
- Discrimination
- Lifetime traumas
- Early life experiences*
- Stressful life events
- Ongoing stress

Social Relationships

- Spouse/child/kin/friends
- Positive support
- Negative support
- Closeness
- Loneliness
- Early parental relationships*
- Friend contact
- Child contact

Personality

- Extraversion
- Neuroticism
- Openness
- Agreeableness
- Conscientiousness
- Cynical hostility
- Anxiety
- Anger

Work

- Work stress
- Work discrimination
- Work satisfaction
- Capacity to work
- Effort-reward balance
- Work support
- Work/family priorities
- Work/life balance

Self-related Beliefs

- Personal mastery
- Perceived constraints
- Hopelessness
- Subjective age*
- Perceptions of aging*
- Subjective social status
- Optimism
- Pessimism

* Indicates items added in 2008

Figure 4. Photo Rating Scale for Men's Facial Mass

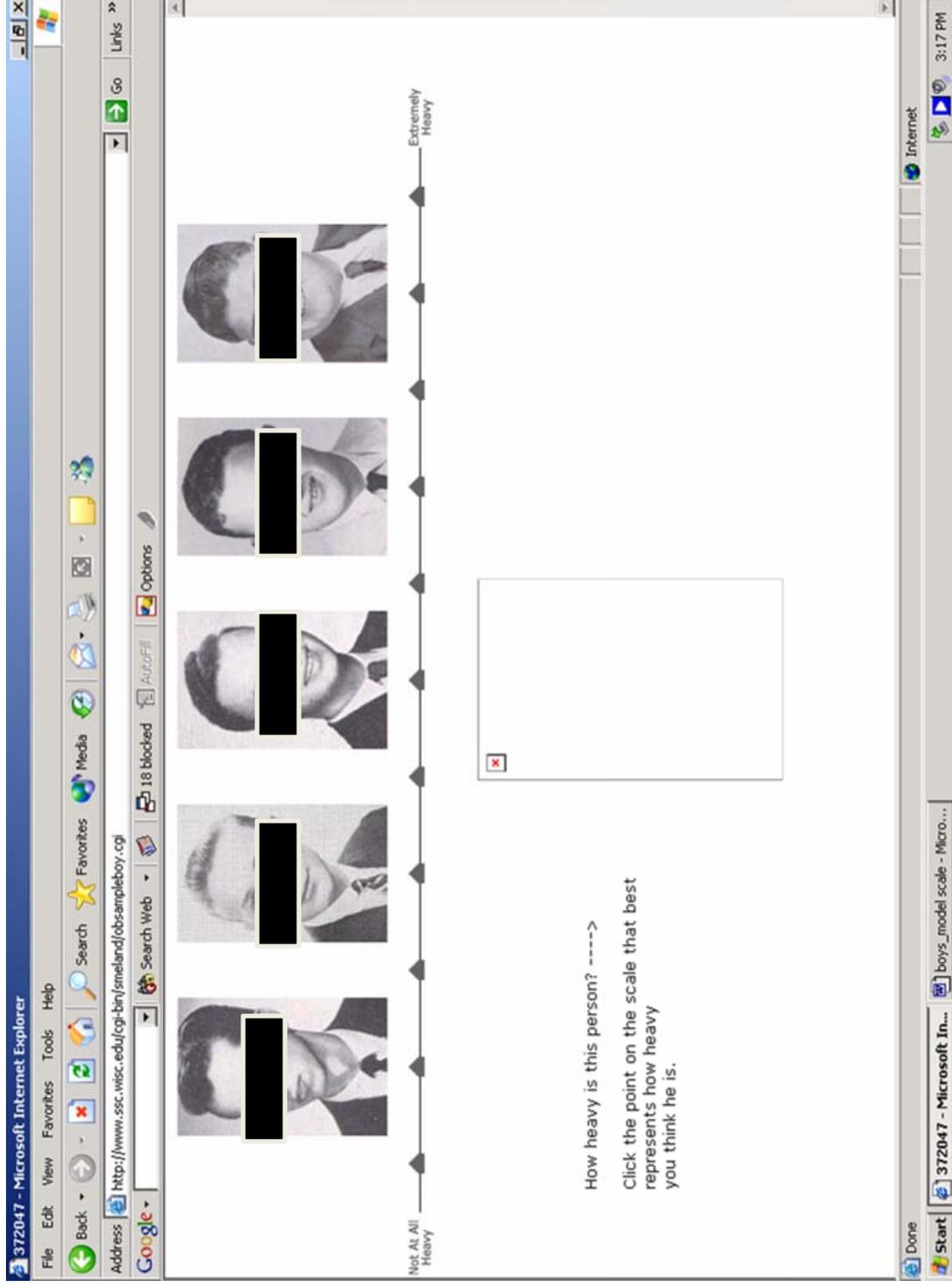
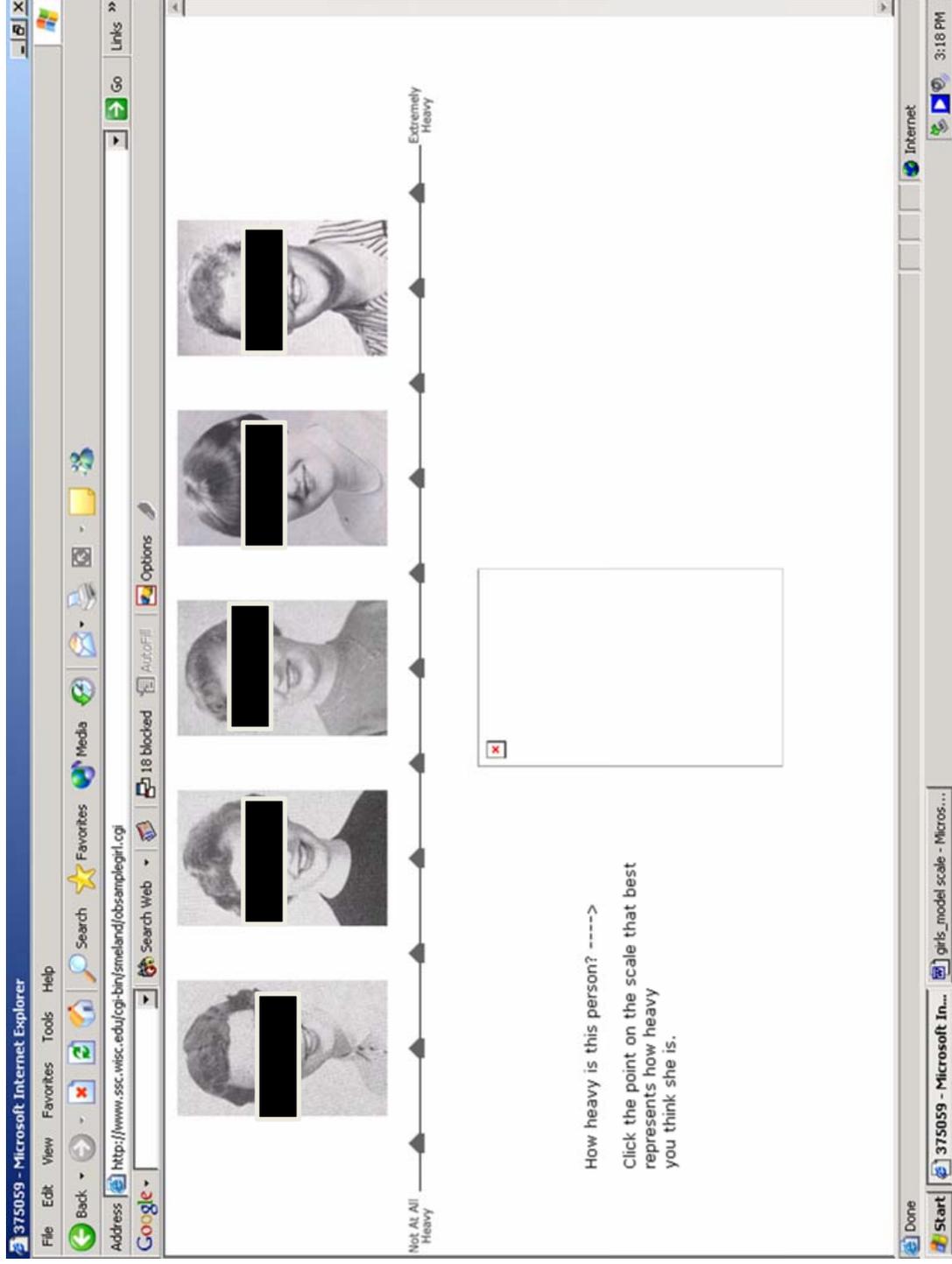


Figure 5. Photo Rating Scale for Women's Facial Mass



Center for Demography and Ecology
University of Wisconsin
1180 Observatory Drive Rm. 4412
Madison, WI 53706-1393
U.S.A.
608/262-2182
FAX 608/262-8400
comments to: hauser@ssc.wisc.edu
requests to: cdepubs@ssc.wisc.edu