SOM for Measures Section

Classification of Included Nations

In the current study, we included Mexico, which is considered an upper-middle-income and developing nation and the United States, Australia, Germany, and South Korea, which are considered high-income nations. The distinction follows the United Nations (Millennium Development Indicators: World and Regional Groupings, 2003; UN, 2020) that uses the per capita gross national income (GNI) established by the World Bank to classify countries by their level of development. In particular, upper-middle-income countries are those with between $3,996 and $12,375 per capita GNI and high-income countries are those with incomes of more than $12,375 per capita GNI (UN, 2020). We note that although the countries included are classified as upper-middle-income and high-income nations, given that our goal is to compare the findings pertaining to the US of increasing deaths of despair and declining mental and physical health across nations, we deliberately chose countries from other parts of the world that are comparable in rankings to the US.

Selection of Measures

We strongly believe that measurement invariance across nations and cultures has always been and continues to be one of the major challenges for cross-cultural comparisons. We acknowledge that integrated data sources exist that allow cross-sectional comparisons of the same measure across many different nations (e.g., Gallup poll, World Values Survey, etc.). However, such an integrated data source is not available for longitudinal data, thereby precluding any direct comparison of repeated within-person assessments across nations within the same data set. We can thus only operate with a work around of conducting parallel analyses of independent data sets that contain at face value similar constructs. To reduce the number of unknowns, we have thus decided to (a) focus on domains that have been demonstrated — in the limited cross-national research that exists so far — to be pertinent for middle-aged adults (midlife peaks in

distress and well-being, Blanchflower & Oswald, 2020); onset of decline for cognition and physical health: Hughes et al., 2018; Chen & Sloan, 2015; Infurna et al., 2020) and (b) select measures that have long been recognized and broadly used in cross-national comparisons and for which measurement and cross-national equivalence has been suggested in other studies (e.g., Hu & Lee, 2011; Jain et al., 2016; Jebb et al., 2020; Miyamoto et al., 2018; Stephens et al., 2014).

Data Collection Time Periods

Each of the five panel surveys included data collected in similar time periods, from the early 2000s to 2018. Although the frequency of data collection varied, the fact that data were collected during the same time period further enables us to compare historical trends on mental and physical health across nations. This is especially the case for persons born in the 1940s, 1950s, and 1960s and across all five countries. Furthermore, the HRS and SOEP contain data from the 1990s onwards. This inclusion of additional data provides us with the opportunity to more completely explore historical trends for persons born in the 1930s (i.e., 1930 to 1939; see Figures 1 to 5). To further allow for the comparability of our findings across cohorts and nations, in our statistical models we center age for all models and all nations at age 55 — an age period that is contained in each of the cohorts shown in Figures S1 to S5 and falls in the middle of the age range that is typically considered midlife.

Figures S1-S5 graphically illustrates the number of observations provided for each birth year bin (i.e., 1930-1939, 1940-1949, 1950-1959, 1960-60) at age of assessment for each study. This further shows the degree of overlap in the assessments across the birth years and how the most amount of overlap is observed for the birth years in the 1940s, 1950s, and 1960s, with age 55 being encompassed in each birth year bin.

Taken together, data for the studies included have been collected in roughly similar windows of historical time and are directly comparable across a broad age range of midlife for people born across multiple decades at around the middle of the last century.

Measures

United States: HRS. Eight items from the Center for Epidemiologic Studies-Depression (CES-D) scale (Radloff, 1977) were used to assess *depressive symptoms*. Items asked participants whether they had (1) or had not (0) experienced the following symptoms "much of the time during the past week": feeling depressed, everything was an effort, restless sleep, was not happy, felt lonely, did not enjoy life, felt sad, and could not get going. The sum across items is taken as an indicator of the number of depressive symptoms an individual experienced frequently. The shorter scale with different response format (yes or no vs. several response categories ranging from *rarely or none of the time* to *most or all of the time*) used in the HRS has demonstrated highly similar construct and external validity as the standard CES-D (see Kohout, Berkman, Evans, & Cornoni-Huntley, 1993; Steffick, 2000). Wave-to-wave correlations ranged from .53 to .60. The intraclass correlation was 0.527 indicating that 53% of the total variance in depressive symptoms was between-person variance and 47% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Functional limitations were measured using a composite sum index of the number of everyday activities participants reported having any difficulty completing, including walking several blocks, climbing one flight of stairs, pushing or pulling large objects, lifting or carrying 10 lb (4.53 kg) of weight, and picking up a dime. Higher scores represent greater functional limitations or poorer physical functioning (Rodgers & Miller, 1997). Although abbreviated versions of standard activities of daily living (ADL) and instrumental activities of daily living (IADL) questionnaires were used, the HRS's measures of functional limitations are comparable with the standard scales (see Fonda & Herzog, 2004; Rodgers & Miller, 1997). Wave-to-wave correlations ranged from .68 to .75. The intraclass correlation was 0.676, indicating that 68% of the total variance in functional limitations was between-person variance and 32% was within-

person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Health conditions were assessed with a sum index of the number self-reported physician-diagnosed medical conditions, including high blood pressure, diabetes, cancer or malignant tumor, lung disease, heart condition, stroke, psychiatric problems, and arthritis. A sum score was created and higher scores indicate reporting more health conditions. Wave-to-wave correlations ranged from .89 to .92. The intraclass correlation was 0.784, indicating that 78% of the total variance in health conditions was between-person variance and 22% was within-person variance. The data thus appeared to contain both more between-person differences than within-person variation over time.

Self-rated health was assessed using a single-item asking participants to rate his or her health on a five-point scale ranging from 1 (excellent) to 5 (poor). The item was reverse scored so that higher scores indicate better self-reported health. Wave-to-wave correlations ranged from .66 to .71. The intraclass correlation was 0.664, indicating that 66% of the total variance in self-rated health was between-person variance and 34% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Episodic memory was measured using a unit-weight composite of performances on the immediate and delayed free-recall tests (see Ofstedal, Fisher, & Herzog, 2005). The immediate recall test was typically given during the first interview quarter and asked participants to recall as many nouns as possible from a list of 10 nouns selected from four lists. For the delayed recall test, interviewers asked participants after a period of about five minutes again to recall as many nouns as possible out of the original word list. We used the percentage of words correctly remembered from both tests, ranging from 0 to 20, with higher scores representing more words remembered or better memory. In the 1992 and 1994 assessments, the word lists included 20

words, instead of 10 words. Because the data were skewed and scores were much lower for these assessments, we did not include them in our analyses. Furthermore, our statistical models accounted for practice effects by including occasion-specific parameters (see Ghisletta et al., 2014). Wave-to-wave correlations ranged from .39 to .63. The intraclass correlation was 0.499, indicating that 50% of the total variance in episodic memory was between-person variance and 50% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Australia: HILDA. Participants' reported on their *life satisfaction* annually, answering the question "How satisfied are you with your life, all things considered?" using a 0 (*totally unsatisfied*) to 10 (*totally satisfied*) rating scale. This item has been used widely in psychological research (see Fujita & Diener, 2005; Lucas et al., 2003). Wave-to-wave correlations ranged from .54 to .65. The intraclass correlation was 0.548, indicating that 55% of the total variance in life satisfaction was between-person variance and 45% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Mental health was measured using the mental health subscale of the SF-36 (Gerstorf, Windsor, Hoppmann, & Butterworth, 2013; Ware et al., 1994). This measure consists of five items that asked whether participants "been a nervous person", "felt so down in the dumps nothing could cheer you up", "felt calm and peaceful", "Felt down", and "been a happy person" during the past 4 weeks. Items were answered on a scale from 1 (all of the time) to 6 (none of the time). Following standard scoring procedures (see Ware et al., 1994), mental health was standardized using the Australian normed population averages and standard deviations, with higher scores indicating better mental health. Wave-to-wave correlations ranged from .60 to .72. The intraclass correlation was 0.620, indicating that 62% of the total variance in mental health was between-person variance and 38% was within-person variance. The data thus appeared to

contain both substantial amounts of between-person differences and within-person variation over time.

Vitality is a subscale of the SF-36 that consists of four items that asked whether participants were "full of life", "have a lot of energy", "felt worn out", or "felt tired" during the past 4 weeks. Items were answered on a scale from 1 (all of the time) to 6 (none of the time). Following standard scoring procedures (see Ware et al., 1994), vitality was standardized using the Australian normed population averages and standard deviations, with higher scores indicating better vitality. Wave-to-wave correlations ranged from .63 to .74. The intraclass correlation was 0.647, indicating that 65% of the total variance in vitality was between-person variance and 35% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

General health is a subscale of the SF-36 that consists of 5 items, answered on a scale from 1 to 5. Specific items asked whether participants "got sick a little easier than other people," "were as healthy as anybody they knew," "expected their health to get worse," "overall health," and "health rated as compared to a year ago." Following standard scoring procedures (see Ware et al., 1994), general health was standardized using the Australian normed population averages and standard deviations, with higher scores indicating better general health. Wave-to-wave correlations ranged from .73 to .83. The intraclass correlation was 0.745, indicating that 75% of the total variance in general health was between-person variance and 25% was within-person variance. The data thus appeared to contain both more between-person differences than within-person variation over time.

Physical functioning is a subscale of the SF-36 that consists of 10 items asking participants whether, during the past 4 weeks, their health limited them across various activities, answered on a scale, "yes, limited a lot," "yes, limited a little," and "no, not limited at all." Specific items asked whether participants' health limited them in "vigorous activities" and

"moderate activities," and difficulty with the ability to "lift, carry groceries," "climb several flights of stairs," "climb one flight of stairs," "bend, kneel," "walk a mile," "walk several blocks," "walk one block," "bathe, dress." Following standard score procedures (see Ware et al., 1994), physical functioning was standardized using the Australian normed population averages and standard deviations, with higher scores indicating better physical functioning. Wave-to-wave correlations ranged from .61 to .80. The intraclass correlation was 0.665, indicating that 67% of the total variance in physical functioning was between-person variance and 33% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Germany: SOEP. We use data on life satisfaction that was provided annually from 1992 to 2018. Participants' reported on their *life satisfaction* annually, answering the question "How satisfied are you with your life, all things considered?" using a 0 (*totally unsatisfied*) to 10 (*totally satisfied*) rating scale. This item has been used widely in psychological research (see Fujita & Diener, 2005; Gerstorf, Hueluer, Wagner, Kunzmann, & Ram, 2018; Lucas et al., 2003). Wave-to-wave correlations ranged from .53 to .65. The intraclass correlation was 0.509, indicating that 51% of the total variance in life satisfaction was between-person variance and 49% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Self-rated health was assessed annually between 1992 to 2018. using a single-item asking participants to rate his or her health on a five-point scale ranging from 1 (excellent) to 5 (poor). The item was reverse scored so that higher scores indicate better self-reported health. Wave-to-wave correlations ranged from .61 to .68. The intraclass correlation was 0.556, indicating that 56% of the total variance in self-rated health was between-person variance and 44% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

The SF-12 was included beginning in 2002 and given to participants biennially through 2018. The *mental health* component of the SF-12 includes items pertaining to vitality, social functioning, role emotional, and mental health, with specific items including the extent to which participants in the last 4 weeks felt "run-down, melancholy", "well-balanced", "used energy", "accomplished less due to emotional problems", "less careful due to emotional problems", and "limited socially due to health last 4 weeks". Wave-to-wave correlations ranged from .49 to .56. The intraclass correlation was 0.487, indicating that 49% of the total variance in mental health was between-person variance and 51% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

The *physical health* component of the SF-12 includes items pertaining to physical functioning, role physical, bodily pain, and general health. Specific items included "current self-rated health", "health affecting ascending stairs", "state of health affects tiring tasks", "strong physical pain in last 4 weeks", "accomplished less due to physical problems", and "limited socially due to health last 4 weeks". Wave-to-wave correlations ranged from .68 to .72. The intraclass correlation was 0.611, indicating that 61% of the total variance in physical health was between-person variance and 39% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

South Korea: KLoSA. *Life satisfaction* was assessed at each biennial assessment using a single item that asked participants how satisfied they are with their overall quality with life using a scale from 0 (dissatisfied) to 100 (satisfied). Wave-to-wave correlations ranged from .47 to .57. The intraclass correlation was 0.456, indicating that 46% of the total variance in life satisfaction was between-person variance and 54% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Depressive symptoms were comprised of seven items and asked participants how often they reported the following symptoms in the past week using a scale from 1 (rarely or none of the time [less than one day]), 2 (some or a little of time [1~2 days]), 3 (occasionally or a moderate amount of time [3~4 days]), and 4 most of all of the time [5~7 days]). Specific items included "I felt depressed", "I felt that everything I did was an effort", "I was happy (reverse scored)", "my sleep was restless", "I enjoyed life (reverse scored)", "I felt lonely", and "I could not get "going". A sum score was created with higher values indicative of reporting more depressive symptoms. Wave-to-wave correlations ranged from .41 to .48. The intraclass correlation was 0.320 indicating that 32% of the total variance in depressive symptoms was between-person variance and 68% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Self-rated health was assessed using a single-item asking participants to rate his or her health on a five-point scale ranging from 1 (very good) to 5 (very bad). The item was reverse scored so that higher scores indicate better self-reported health. Wave-to-wave correlations ranged from .55 to .60. The intraclass correlation was 0.471, indicating that 47% of the total variance in self-rated health was between-person variance and 53% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Health conditions were assessed with a sum index of the number self-reported physician-diagnosed medical conditions, including high blood pressure, diabetes or high blood sugar or not, cancer or malignant tumor, chronic lung disease, liver disease, heart-related disease, cerebrovascular disease, emotional nervous or psychiatric problems, arthritis and prostatic disease. A sum score was created and higher scores indicates reporting more health conditions. Wave-to-wave correlations ranged from .91 to .96. The intraclass correlation was 0.822, indicating that 82% of the total variance in health conditions was between-person variance and

18% was within-person variance. The data thus appeared to contain both more between-person differences than within-person variation over time.

Functional limitations were assessed by combining an activities of daily living (ADL) and instrumental activities of daily living (IADL) scales. For each item, a 1 was assigned if participants needed to be helped partially or fully in order to complete each task, otherwise, given a 0. The ADL component consisted of seven items and inquired whether participants needed help with dressing, washing face and hair and brushing teeth, bathing and showering, eating prepared food, getting in/out of bed and walking across a room, using the toilet, and continence. The IADL component consisted of 10 items and inquired whether participants need help with grooming, household chores, preparing meals, doing laundry, going out to nearby places, using transportation, shopping at a shop, managing money, making and taking a call and taking medications. The two components were summed with higher scores indicating greater functional limitations. Wave-to-wave correlations ranged from .60 to .64. The intraclass correlation was 0.648, indicating that 65% of the total variance in functional limitations was between-person variance and 35% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

The Korea version of the Mini-Mental State Examination (MMSE) was used to assess overall *cognition* and cognitive tasks like thinking and learning. The MMSE index is the number of correctly answered questions with the high score being 30. Scores below 17 indicate the participant is suspected of dementia, scores in the range of 18-23 indicate the respondent has a decline in cognitive function, and a score of 24 or more suggests the respondent is considered normal. The measure consists of measures pertaining to cognitive ability regarding time (date-year-month-day), place (city), memory test (memorizing three words), attention and calculation, memory test, use of belongings, sentence pronunciation, and following direction. Furthermore, our statistical models accounted for practice effects by including occasion-specific parameters

(see Ghisletta et al., 2014). Wave-to-wave correlations ranged from .69 to .75. The intraclass correlation was 0.508, indicating that 51% of the total variance in cognition was between-person variance and 49% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Mexico: MHAS. *Depressive symptoms* were measured with a modified version of the CES-D that includes 9 items, asking participants whether they reported experiencing certain feelings the majority of the time during the week prior to the interview (coded as 1) or not (coded as 0). Specific items included "depressed", "everything was an effort", "restless sleep", "happiness (reverse scored)", "loneliness", "enjoy life (reverse scored)", "sadness", "felt tired", and "had a lot of energy (reverse scored)". A sum score was created with higher values indicative of reporting more depressive symptoms. Wave-to-wave correlations ranged from .42 to .51. The intraclass correlation was 0.614, indicating that 61% of the total variance in depressive symptoms was between-person variance and 39% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Self-rated health was assessed using a single-item asking participants to rate his or her health on a five-point scale ranging from 1 (excellent) to 5 (poor). The item was reverse scored so that higher scores indicate better self-reported health. Wave-to-wave correlations ranged from .38 to .43. The intraclass correlation was 0.419, indicating that 42% of the total variance in self-rated health was between-person variance and 58% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Functional limitations were assessed from 14 items that cover both ADL and IADL indices. The stem for each item asked participants "Because of a health problem, do you have difficulty with …". Participants responded to each item with a yes or no. The specific items

were: walking several blocks, running or jogging one kilometer, difficulty walking one block, sitting for about two hours, getting up from a chair after sitting for long periods, climbing several flights of stairs without resting, climbing one flight of stairs without resting, difficulty with stooping, kneeling, or crouching, reaching or extending your arms above shoulder level, pulling or pushing large objects like a living-room chair, lifting or carrying objects weighing over 5 kg, like a heavy bag of groceries, picking up a 1 peso coin from the table, difficulty with dressing, including putting on shoes and socks, and does anyone ever help you get dressed. A sum score was created with higher values indicative of greater functional limitations. Wave-to-wave correlations ranged from .48 to .61. The intraclass correlation was 0.459, indicating that 46% of the total variance in functional limitations was between-person variance and 54% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Health conditions were assessed with a sum index of the number self-reported physician or medical-personnel diagnosed medical conditions, including hypertension, diabetes, cancer or malignant tumor, lung disease, heart condition, stroke, and arthritis. A sum score was created and higher scores indicates reporting more health conditions. Wave-to-wave correlations ranged from .53 to .68. The intraclass correlation was 0.548, indicating that 55% of the total variance in health conditions was between-person variance and 45% was within-person variance. The data thus appeared to contain both substantial amounts of between-person differences and within-person variation over time.

Selection of Statistical Model of Analyses

We acknowledge that alternative modeling approaches would have been feasible with the data at hand. In the end, multiple reasons had led us to select the growth curve models over such alternatives (Grimm, Ram, & Estabrook, 2017). To begin with, our major objective has been to track how key indicators of physical and mental health develop as people move through their

midlife years (i.e., within-person change) and how this may differ across historical times (i.e., between-person differences). With this rationale in mind, we selected the use of growth curve models as an analytic approach that represents good practice and is well-established in the broader developmental and adult development literature. Doing so puts us in the position that results obtained from our analyses can directly be compared with a myriad of other empirical reports in a straightforward manner. Second, the type of model we opted for represents a parsimonious way to model levels at a given age and time-related trends and do so in ways that are immediately comparable across the five national data sets – an inclusion of non-linear (e.g., quadratic) trends in some of the studies but not others would have hampered direct comparability. Third, we acknowledge that certain modeling techniques such as latent basis models may at times selectively outperform the growth curve models used here. A major disadvantage of latent basis models is that they cannot accommodate time-unstructured data. Yet, the latter (growth curve models) have repeatedly been demonstrated to provide robust alternatives and offer the genuine advantage that incorporation of predictor variables (here, birth year as our variable of prime interest) is possible in straightforward ways that are directly comparable across models; something that is unfortunately not always a given for analytic alternatives such as a latent basis model. Methodologically, we did test different model/covariance structures such as auto-regressive versus unstructured and the approach of unstructured we utilize here provided the most parsimonious way to examine our research questions.

Full information maximum likelihood. As background information, the approach utilized has repeatedly been shown to indeed provide estimates of average within-person change in the overall sample whether or not an individual stayed in the sample over time (i.e., missing at random [MAR] assumption; see Gerstorf et al., 2007). In addition, we note that Full Information Maximum Likelihood estimation procedures (FIML, as implemented here) and alternative

approaches such as Multiple Imputation are both working on missing-at-random (MAR) assumptions – with the attrition-related variables informing the parameter estimation. We think the relevant question is about whether the MAR assumptions are justifiable – and whether all the attrition-relevant variables have been tested and/or included. To begin with, FIML approaches have indeed been found effective and robust even in the presence of moderate violations of MAR assumptions as long as attrition-informative variables are included in the estimation (cf. Lövdén, Ghisletta, & Lindenberger, 2004). This is so because in longitudinal studies, the data are correlated across measurement occasions. Thus, even in the presence of (relatively minor) selectivity effects in change, the initial level is to some degree predictive of the missing value after dropout --- that is, the level data carry information about observed, and unobserved, change. Accordingly, with likelihood analyses under MAR we can, at least to some extent, account for selectivity effects in change in addition to capturing the largest selectivity effects (i.e., level). Therefore, analyses under the MAR assumption offer a relatively robust approach to the present research questions (see also Rabbitt et al., 2004). The attrition-informative variables included in our models were age, gender, and education.

SOM for Results Section

In the following, we elaborate on results arising from our analyses that go beyond those describing similarities and differences in cohort effects within and across nations. We also discuss findings pertaining to gender differences within each nation.

Physical health. Age trends were observed in the US across each of the three physical health outcomes. On average, increasing age was associated with more health conditions across cohorts. Focusing on functional limitations and self-rated health, those born in the 1930s, on average, exhibited declining functional limitations and slight improvements in self-rated health, whereas those born in the 1950s and 1960s exhibited an opposite trend in that increasing age was associated with increasing functional limitations and declining self-rated health. Those born in

the 1940s, on average, showed relative stability across age for both functional limitations and self-rated health.

For middle-aged adults in Australia, only level differences were observed across cohorts with the within-person age trends being similar over time.

Focusing on Germany, the age trend for self-rated health mimicked that of the US in that those born in the 1930s, on average, reported slight increases in self-rated health over age and those in the 1940s exhibited relative stability. For cohorts of persons born in the 1950s and 1960s, age trends revealed declines over chronological age.

Focusing on South Korea, health conditions showed an age trend of general increases over chronological age across cohorts, whereas functional limitations showed declines over chronological age and no age trend was observed for self-rated health.

Focusing on Mexico, health conditions and functional limitations showed slight age trends of increases over chronological cage, whereas self-rated health showed declines, but later-born cohorts the change was not as strong.

Mental health. Focusing on the US, age trends were observed for both depressive symptoms and episodic memory. Later-born cohorts in the US, on average, exhibited higher levels of depressive symptoms, but relative stability across age. This contrasts with episodic memory in that each cohort showed declines over chronological age and the rate differed across the cohorts shown.

Focusing on Australia, of the three outcomes of mental health, only life satisfaction showed differences in age trends across cohorts. Later-born cohorts, on average, reported stability in life satisfaction across chronological age, whereas earlier-born cohorts reported increases over chronological age. The trajectories for mental health and vitality show each outcome being lowest in the 40s and early 50s, with gradual increases over chronological age.

Focusing on Germany, age trends were observed for both life satisfaction and mental health. Earlier-born cohorts in Germany reported steady increases in life satisfaction, with this leveling off for later born cohorts. General increases over age were observed across cohorts for mental health.

Focusing on South Korea, age trends were observed for both depressive symptoms and general cognitive status (MMSE). There were slight increases in depressive symptoms over chronological age. The rate of decline in general cognitive status differed across cohorts, with later-born cohorts exhibiting less steep declines across chronological age.

Focusing on Mexico, later-born cohorts exhibited declining trends of depressive symptoms over chronological age, as compared to earlier-born cohorts that showed relative stability.

Summary of age trend findings. Across the nations included, we observed a non-convergence of age trends. Such non-convergence across the various outcomes examined is a typical finding (for review, see Sliwinski, Hoffman, & Hofer, 2010) and reflects the fact that in long-term longitudinal studies that when, for example, 55-year olds in a given study age 5 years, they will not necessarily show the same patterns as those who started participating in the study as 60-year olds — the latter typically start off at higher levels of functioning (here, fewer functional limitations). This is the very reason why it is so important to separate the age and selection trend from the within-person change effects/trends that we are primarily interested in.

Gender. Across the nations, there were both main effects of gender and to some extent gender moderating cohort effects. On average, women in the US, Australia, and Mexico reported poorer mental health and physical health. Regarding cognition, women, on average in the US exhibited better memory, whereas in South Korea women exhibited poorer cognition. In Germany, the only main effect of gender was women reporting better mental health.

Focusing on gender moderating the impact of cohort, interestingly, there were differences between the US and Australia with South Korea. Later-born cohorts of women in the US, on

average, reported lower self-rated health and more depressive symptoms and in Australia, later-born cohorts of women reported poorer vitality. Conversely, in South Korea, later-born cohorts of women reported better self-rated health and fewer depressive symptoms and exhibited better cognition. In Germany and Mexico, there were no evidence to suggest for gender moderating cohort differences across each of the outcomes examined.

Findings pertaining to race in the US sample. In the second model for the US sample, we included race as a covariate, which was coded 1 = White and 0 = non-White (African American, Hispanic, or Other). We found that there was a main effect for race, such that individuals who were white, on average, were more likely to report fewer health conditions, fewer functional limitations and better self-rated health, and exhibit fewer depressive symptoms and better memory. There was only one race by cohort interaction that was statistically significant and it was for depressive symptoms; later-born cohorts of individuals who were white, on average, exhibited fewer depressive symptoms.

References

Chen, Y., Sloan, F.A. (2015) Explaining disability trends in the U.S. elderly and near-elderly population. *Health Services Research, 50,* 1528-1549.

Gerstorf, D., Hueluer, G., Wagner, G. G., Kunzmann, U., & Ram, N. (2018). Terminal change across facets of affective experience and domain satisfaction: Commonalities, differences, and bittersweet emotions at the end of life. *Developmental Psychology*, *54,* 2382-2402. doi: 10.1037/dev0000599

Gerstorf, D., Lövdén, M., Röcke, C., Smith, J., & Lindenberger, U. (2007). Well-being affects changes in perceptual speed in advanced old age: Longitudinal evidence for a dynamic link. *Developmental Psychology, 43,* 705-718.

Ghisletta, P., Bäckman, L., Bertram, L., Brandmaier, A., Gerstorf, D., Liu, L., & Lindenberger, U. (2014). The Val/Met polymorphism of the Brain-Derived Neurotrophic Factor (BDNF) gene predicts decline in perceptual speed in older adults. *Psychology and Aging, 29,* 384-392. doi: 10.1037/a0035201

Hu, P., & Lee, J. (2012). Harmonization of Cross-National Studies of Aging to the Health and Retirement Study: Chronic Medical Conditions. Santa Monica, CA: RAND Corporation.

Hughes, M. L., Agrigoroaei, S., Jeon, M., Bruzzese, M., & Lachman, M. E. (2018). Change in cognitive performance from midlife into old age: Findings from the Midlife in the United States (MIDUS) study. *Journal of the International Neuropsychological Society, 24*(8), 805-820.

Infurna, F. J., Gerstorf. D., & Lachman, M. E. (2020). Midlife in the 2020s: Opportunities and challenges. *American Psychologist, 75,* 470-485. https://doi.org/10.1037/amp0000591

Jain, U., Min, J., & Lee, J. (2016). Harmonization of cross-national studies of aging to the health and retirement study – user guide: family transfer – informal care. *CESR-Schaeffer Working Paper Series, Paper no: 2016-008.*

Jebb, A. T., Morrison, M., Tay, L., & Diener, E. (2020) Subjective Well-Being Around the World: Trends and Predictors Across the Life Span, *Psychological Science, 31 (3),* 293-305. https://doi.org/10.1177/0956797619898826

Lövdén, M., Rönnlund, Wahlin, A., Bäckman, L., Nyberg, & Nilsson, L-G. (2004). The extent of stability and change in episodic and semantic memory in old age: Demographic and predictors of level and change. *The Journals of Gerontology, Series B, 59,* P130-P134.

McArdle, J. J., Fisher, G. G., & Kadlec, K. M. (2007). Latent variable analyses of age trends from the Health and Retirement Study, 1992–2004. *Psychology and Aging*, 22, 525–545.

Millennium Development Indicators: World and Regional Groupings (2003). Retrieved December 8, 2020, from https://millenniumindicators.un.org/unsd/mi/worldbank.htm

Miyamoto, Y., Yoo, J., Levine, C. S., Park, J., Boylan, J. M. … Ryff, C. D. (2018). Culture and social hierarchy: Self- and other-oriented correlates of socioeconomic status across cultures. *Journal of Personality and Social Psychology, 115 (3),* 427-445. http://dx.doi.org/10.1037/pspi0000133

Ofstedal, M. B., Fisher, G. G., & Herzog, A. R. (2005). *Documentation of cognitive functioning measures in the Health and Retirement Study* (HRS/AHEAD Documentation Report DR-006). Ann Arbor, MI: University of Michigan.

Rabbitt, P., Diggle, P., Holland, F., & McInnes, L. (2004). Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. *The Journals of Gerontology, Series B, 59,* P84-P97.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385-401.

Rodgers, W. L., & Miller, B. (1997). A comparative analysis of ADL questions in surveys of older people. *The Journals of Gerontology: Psychological Sciences, 52*B, 21–36.

Sliwinski, M., Hoffman, L., & Hofer, S. M. (2010). Evaluating convergence of within-person change and between-person age differences in age-heterogenous longitudinal studies. *Research in Human Development, 7,* 45-60.

Stephens, N. M., Markus, H. R., & Phillips, L. T. (2014). Social class culture cycles: How three gateway contexts shape selves and fuel inequality. *Annual Review of Psychology, 65(1),* 611-634. http://10.1146/annurev-psych-010213-115143

United Nations (2020). *World Economic Situation and Prospects.* Retrieved from https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/WESP2020_Annex.pdf

Ware, J. E., Kosinski, M., & Keller, S. D. (1994). *SF-36 physical and mental health summary scales: A user's manual*. Boston, MA: The Health Institute, New England Medical Center.

Table S1

Descriptive Information and Measures of Mental and Physical Health for Each of the Datasets Used in the Current Study

	United States (HRS)	Australia (HILDA)	Germany (SOEP)	South Korea (KLoSA)	Mexico (MHAS)
Data Collection Years	1992 to 2018 (biennially)	2001 to 2018 (annually)	1992 to 2018 (annually) 2002 to 2018 (biennially)	2006 to 2016 (biennially)	2001, 2003, 2013, 2015, 2018
Sample Size	*N* = 28,219	*N* = 10,836	*N* = 27,822 *N* = 22,610	*N* = 6,402	*N* = 16,564
Age Range	40 – 65	40 – 65	40 – 65	45 – 65	40 – 65
Years of Education	*M* = 12.61 (*SD* = 3.21), range: 0 to 17	*M* = 13.32 (*SD* = 2.46), range: 11 to 20	*M* = 11.98 (*SD* = 2.74), range: 7 to 18 *M* = 12.24 (*SD* = 2.75), range: 7 to 18	*M* = 9.34 (*SD* = 4.36), range: 0 to 20	*M* = 5.99 (*SD* = 4.80), range: 0 to 22
Gender (% Women)	56%	52%	49%	55%	57%
Race/Ethnicity (% White)	68%				
Year of birth	*M* = 1948 (*SD* = 10.62), range: 1930 to 1969	*M* = 1955 (*SD* = 8.77), range: 1936 to 1969	*M* = 1953 (*SD* = 10.56), range: 1927 to 1969 *M* = 1956 (*SD* = 8.98), range: 1937 to 1969	*M* = 1951 (*SD* = 6.20), range: 1938 to 1961	*M* = 1953 (*SD* = 9.20), range: 1936 to 1969
Mental Health Measures	Depressive symptoms (CES-D), Episodic memory (Immediate and Delayed Recall)	Life satisfaction (single-item), Mental health (SF-36), Vitality (SF-36)	Life satisfaction (single-item), Mental health (SF-12)	Life satisfaction (single-item), Depressive symptoms (CES-D), General cognition (MMSE)	Depressive symptoms (CES-D)

Physical Health Measures	Self-rated health (single-item), Functional limitations (ADL, IADL), Number of health conditions	General health (SF-36), Physical functioning (SF-36)	Self-rated health (single-item), Physical health (SF-12)	Self-rated health (single-item), Functional limitations (ADL, IADL), Number of health conditions	Self-rated health (single-item), Functional limitations (ADL, IADL), Number of health conditions

Note. HRS = Health and Retirement Study. HILDA = Household, Income, and Labour Dynamics of Australia. SOEP = German Socioeconomic Panel Study. KLoSA = Korean Longitudinal Study of Aging. MHAS = Mexico Health and Aging Study. CES-D = Center for Epidemiologic Studies-Depression. MMSE = Mini-Mental State Examination. ADL = Activities of Daily Living. IADL = Instrumental Activities of Daily Living. For the SOEP, data collection for life satisfaction and self-rated health was annually from 1992 to 2018 and for mental and physical health biennially from 2002 to 2018. For the SOEP, the second set of descriptive information for demographics refers to the sample that provided data on mental and physical health biennially from 2002 to 2018.

Table S2

Growth models Examining Cohort differences Across Physical Health in United States, Australia, and Germany

	United States: HRS						**Australia: HILDA**				**Germany: SOEP**			
	Health conditions		Functional limitations		Self-rated health		Physical Functioning		General Health		Physical Health		Self- Rated health	
Parameters	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed Effects														
Intercept	1.71*	0.02	1.18*	0.03	3.01*	0.03	75.54*	0.79	63.03*	0.80	46.38*	0.15	2.83*	0.02
Time	0.08*	0.003	0.03*	0.005	-0.05*	0.004	-0.93*	0.18	-0.61*	0.16	-0.67*	0.06	-0.03*	0.004
Time^2	0.0003*	0.00008	0.002	0.0001	-0.00009	0.0001	-0.01*	0.002	0.0009	0.002	-0.01	0.005	0.0008*	0.00006
Age	0.06*	0.004	-0.03	0.006	0.02*	0.005	-0.28	0.14	0.07	0.14	-0.39*	0.02	0.02*	0.003
Age × Time	0.0005	0.0005	0.002	0.0008	0.002^	0.0007	0.04	0.03	0.03	0.03	0.007	0.008	0.001	0.0006
Cohort														
Year of Birth	-0.007*	0.001	-0.05*	0.005	0.04*	0.004	0.71*	0.14	0.53*	0.15	-0.006	0.02	0.04*	0.003
Year of Birth^2	--	--	0.001*	0.0002	-0.002*	0.0001	-0.027*	0.007	-0.03*	0.007	--	--	-0.001*	0.00008
Year of Birth × Time	0.0004*	0.0001	0.003*	0.0007	0.001*	0.0006	0.06	0.03	0.04	0.03	-0.009	0.005	-0.0005	0.0005
Year of Birth^2 x Time	--	--	-0.0002*	0.00003	-0.00006	0.00002	-0.002	0.001	-0.002	0.001	--	--	0.00003	0.00002
Year of Birth × Age	-0.0002	0.0002	0.005*	0.0007	-0.003	0.0006	-0.02	0.02	-0.03	0.03	-0.0004	0.001	-0.003*	0.0003
Year of Birth^2 x Age	--	--	-0.00007	0.00002	0.00004*	0.00002	-0.0004	0.0005	-0.0008	0.0005	--	--	0.0001	0.000008
Year of Birth × Age × Time	0.00008*	0.00002	-0.0002	0.00009	0.00004	0.00008	-0.003	0.003	-0.001	0.003	-0.0004	0.001	0.00006	0.00005
Year of Birth^2 x Age x Time	--	--	-0.00005	0.000003	-0.00015	0.000003	0.00001	0.00007	-0.00008	0.00006	--	--	0.000008	0.000001
Random effects														
Var intercept	1.79*	0.02	1.42*	0.01	0.84*	0.009	337.68*	5.37	369.16*	5.57	65.69*	0.88	0.47*	0.005
Var. time	0.009*	0.0001	0.006*	0.0002	0.002*	0.00007	1.31*	0.04	1.22*	0.03	0.83*	0.03	0.002*	0.001
Cov. Intercept, time	0.06	0.001	0.02*	0.001	0.005*	0.0007	6.59*	0.39	3.21*	0.35	3.71*	0.15	0.003*	0.0004
Residual variance	0.19*	0.001	0.50*	0.002	0.39*	0.002	151.33*	0.81	102.11*	0.55	31.79*	0.20	0.35*	0.001

Note. ^ = p = .01, * p < .01. Intraclass correlation (ICC). ICC for HRS Outcomes: Health Conditions = 0.784; Functional Limitations = 0.676; Self-rated Health = 0.664. ICC for Australia Outcomes: Physical functioning = 0.665; General Health = 0.745. ICC for Germany Outcomes: Physical Health = 0.611; Self-rated health = 0.556.

Table S3

Growth models Examining Cohort Differences Across Physical Health in South Korea, and Mexico

	South Korea: KLoSA						Mexico: MHAS					
	Functional limitations		Self-rated health		Health conditions		Functional limitations		Health conditions		Self-rated health	
Parameters	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed Effects												
Intercept	0.60*	0.07	3.25*	0.03	0.54*	0.04	3.02*	0.05	0.91*	0.01	2.28*	0.01
Time	0.11*	0.02	-0.04*	0.01	0.04*	0.009	0.14*	0.009	0.04*	0.002	-0.01*	0.002
Time^2	0.003*	0.0007	-0.001	0.0004	-0.0003	0.0003	0.005*	0.0007	0.001*	0.0002	-0.00008	0.0003
Age	-0.08*	0.01	-0.01	0.006	0.08*	0.01	0.03*	0.009	0.01*	0.002	-0.007^	0.003
Age × Time	-0.02*	0.05	0.004	0.002	0.005	0.002	-0.003	0.001	-0.00004	0.0004	-0.009^	0.0004
Cohort												
Year of Birth	-0.12*	0.03	0.03*	0.006	0.03	0.02	-0.05*	0.004	-0.01*	0.001	0.01*	0.001
Year of Birth^2	0.007*	0.003	--	--	-0.004^	0.002	--	--	--	--	--	--
Year of Birth × Time	-0.03*	0.007	0.004	0.002	0.004	0.003	-0.004	0.008	-0.0003	0.0002	-0.00002	0.0002
Year of Birth^2 x Time	0.001	0.0007	--	--	-0.0006	0.0003	--	--	--	--	--	--
Year of Birth × Age	0.009^	0.004	0.0008	0.0004	-0.007*	0.002	-0.001^	0.0007	-0.0006	0.0002	0.0005*	0.0002
Year of Birth^2 x Age	-0.00002	0.002	--	--	-0.00004	0.0001	--	--	--	--	--	--
Year of Birth × Age × Time	0.002	0.001	0.00003	0.0001	-0.0009	0.0004	0.0002	0.0001	0.00003	0.00003	0.00005	0.00003
Year of Birth^2 x Age x Time	0.000006	0.00004	--	--	0.00003	0.00002	--	--	--	--	--	--
Random effects												
Var intercept	1.69*	0.04	0.29*	0.008	0.83*	0.02	5.29*	0.09	0.49*	0.007	0.29*	0.006
Var. time	0.01*	0.0008	0.003*	0.0002	0.006*	0.0001	0.19*	0.009	0.01*	0.0007	0.002*	0.0006
Cov. Intercept, time	0.06*	0.004	-0.004*	0.0009	0.02*	0.001	0.02*	0.002	0.002*	0.0008	0.003*	0.00009
Residual variance	0.72*	0.009	0.35*	0.004	0.05*	0.0007	4.62*	0.05	0.31*	0.003	0.39*	0.004

Note. ^ = p = .01, * p < .01. Intraclass correlation (ICC). ICC for South Korea Outcomes: Functional limitations = 0.648; Self-rated health = 0.471; Health conditions = 0.822. ICC for Mexico Outcomes: Functional limitations = 0.459; Health conditions = 0.548; Self-rated health = 0.419.

Table S4

Growth models Examining Cohort differences Across Mental Health in United States, Australia, and Germany

	United States: HRS				**Australia: HILDA**						**Germany: SOEP**			
	Depressive Symptoms		Episodic Memory		Vitality		Life Satisfaction		Mental Health		Life Satisfaction		Mental Health	
Parameters	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed Effects														
Intercept	1.40*	0.031	52.85*	0.58	53.83*	0.71	7.87*	0.03	73.19*	0.61	6.31*	0.04	50.04*	0.26
Time	0.03*	0.004	-1.00*	0.09	-0.14	0.17	0.001	0.006	-0.02	0.15	0.02*	0.007	-0.07	0.10
Time^2	-0.002*	0.0002	-0.007	0.003	-0.004	0.002	0.001*	0.0002	-0.002	0.002	0.002*	0.0001	-0.02*	0.006
Age	0.03*	0.006	-0.97*	0.09	0.32*	0.13	0.03*	0.005	0.47*	0.11	0.10*	0.006	0.36*	0.05
Age × Time	-0.0008	0.0007	-0.04*	0.02	0.002	0.03	0.002	0.001	0.02	0.03	-0.005*	0.001	0.03	0.02
Cohort														
Year of Birth	0.02*	0.0002	-0.09*	0.07	0.33*	0.13	-0.01*	0.003	0.25^	0.11	0.07*	0.005	0.18*	0.05
Year of Birth^2	--	--	--		-0.02*	0.007	--	--	-0.02*	0.006	-0.001*	0.0001	-0.005	0.002
Year of Birth × Time	-0.003*	0.0003	0.02*	0.003	0.03	0.03	0.001	0.0007	0.03	0.03	-0.008*	0.0009	0.007	0.02
Year of Birth^2 x Time	--	--	--	--	-0.003	0.002	--	--	-0.002	0.001	0.0003*	0.00003	0.002	0.001
Year of Birth × Age	-0.0008^	0.0008	-0.02*	0.002	-0.02	0.01	-0.001*	0.0003	-0.02	0.01	-0.004*	0.0006	0.003	0.005
Year of Birth^2 x Age	--	--	--	--	-0.0005	0.0005	--	--	-0.0007	0.0004	0.00002	0.00001	-0.0006	0.0001
Year of Birth × Age × Time	-0.0003*	0.00004	-0.001*	0.0004	-0.00009	0.003	0.00006	0.00003	-0.0003	0.003	0.0008*	0.00009	0.007*	0.002
Year of Birth^2 x Age x Time	--	--	--	--	-0.0001	0.00006	--	--	-0.00005	0.00005	-0.0007*	0.00002	-0.0002*	0.00005
Random effects														
Var intercept	2.39*	0.03	137.28*	1.70	275.58*	4.32	1.34*	0.02	198.08*	3.16	1.73*	0.02	59.61*	1.04
Var. time	0.004*	0.0002	0.17*	0.02	2.92*	0.29	0.007*	0.0002	0.73*	0.03	0.008*	0.0002	5.13*	0.47
Cov. Intercept, time	0.06*	0.002	0.86*	0.15	0.93*	0.03	0.005*	0.002	1.88*	0.22	0.009*	0.002	0.14^	0.06
Residual variance	1.89*	0.009	138.34*	0.78	132.60*	0.70	1.01*	0.005	108.39*	0.58	1.42*	0.005	45.92*	0.31

Note. ^ = p = .01, * p < .01. Intraclass correlation (ICC). ICC for US Outcomes: Depressive symptoms = 0.527; Episodic memory = 0.499. ICC for Australia Outcomes: Vitality = 0.647; Life satisfaction = 0.548; Mental health = 0.620. ICC for Germany Outcomes: Life satisfaction = 0.509; Mental health = 0.487.

Table S5

Growth models Examining Cohort Differences Across Mental Health in South Korea, and Mexico

	South Korea: KLoSA						Mexico: MHAS	
	Life satisfaction		Depressive symptoms		MMSE		Depressive symptoms	
Parameters	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed Effects								
Intercept	61.18*	0.64	11.53*	0.14	27.41*	0.12	3.42*	0.04
Time	-0.55	0.26	0.18*	0.06	-0.23*	0.05	0.009	0.007
Time^2	0.05*	0.01	-0.007*	0.002	-0.01*	0.005	-0.0006	0.0009
Age	0.34	0.15	0.12*	0.04	-0.14^	0.06	-0.004	0.008
Age × Time	0.13	0.06	-0.01	0.007	0.006	0.012	0.0007	0.001
Cohort								
Year of Birth	0.59*	0.13	-0.02	0.05	0.012	0.05	-0.02*	0.003
Year of Birth^2	--	--	0.01*	.005	--	--	--	--
Year of Birth × Time	0.11	0.05	-0.01	0.02	0.006	0.01	-0.0005	0.0007
Year of Birth^2 x Time	--	--	0.0006	0.001	--	--	--	--
Year of Birth × Age	0.14	0.009	0.01	0.007	0.007*	0.002	-0.003	0.0005
Year of Birth^2 x Age	--	--	-0.0007	0.0003	--	--	--	--
Year of Birth × Age × Time	0.003	0.002	0.0009	0.003	-0.0007	0.0005	-0.0002*	0.0001
Year of Birth^2 x Age x Time	--	--	0.00006	0.00009	--	--	--	--
Random effects								
Var intercept	137.73*	3.71	3.87*	0.12	5.46*	0.14	3.05*	0.06
Var. time	1.29*	0.10	0.06*	0.004	0.04*	0.003	0.03*	0.005
Cov. Intercept, time	-3.37*	0.44	0.11*	0.02	0.06*	0.01	0.006*	0.0009
Residual variance	159.60*	1.92	6.63*	0.08	4.49*	0.06	3.81*	0.04

Note. ^ = p = .01, * p < .01. Intraclass correlation (ICC). ICC for South Korea Outcomes: Life satisfaction = 0.456; Depressive symptoms = 0.320; MMSE = 0.508. ICC for Mexico Outcome: Depressive symptoms = 0.614.

Table S6

Growth Models Examining Cohort and Education Differences Across Physical Health in United States, Australia, and Germany

	United States: HRS						**Australia: HILDA**				**Germany: SOEP**			
	Health conditions		Functional limitations		Self-rated health		Physical Functioning		General Health		Physical Health		Self-Rated Health	
Parameters	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed Effects														
Intercept	1.76*	0.02	1.15*	0.03	3.01*	0.03	74.30*	0.83	62.34*	0.84	46.07*	0.15	2.92*	0.02
Time	0.08*	0.003	0.03*	0.004	-0.05*	0.004	-0.82*	0.21	-0.64*	0.18	-0.67*	0.06	-0.03*	0.004
Time^2	0.0003*	0.00008	0.002*	0.0001	-0.00005	0.0001	-0.01*	0.002	0.0009	0.002	-0.009	0.005	0.0008*	0.00006
Age	0.05*	0.004	-0.02*	0.006	0.007	0.005	-0.35^	0.15	-0.01	0.15	-0.33*	0.03	0.01*	0.004
Age × Time	-0.00004	0.0005	0.002*	0.0008	0.003*	0.0007	0.02	0.04	0.04	0.04	0.01	0.008	0.002*	0.0007
Cohort														
Year of Birth	-0.008*	0.001	-0.03*	0.005	0.02*	0.004	0.61*	0.16	0.42*	0.16	-0.0006	0.02	0.03*	0.003
Year of Birth^2	--	--	0.0008*	0.0001	-0.0008*	0.0001	-0.03*	0.008	-0.03*	0.008	--	--	-0.0009*	0.0001
Year of Birth × Time	0.0003	0.0002	0.004*	0.0007	0.002*	0.0006	0.03	0.04	0.04	0.03	-0.008	0.005	0.0004	0.0005
Year of Birth^2 x Time	--	--	-0.0002*	0.00003	-.00001	0.00002	-0.0008	0.002	-0.004*	0.002	--	--	-0.0004	0.0002
Year of Birth × Age	0.00005	0.0002	0.004*	0.0007	-0.001^	0.0006	-0.02	0.02	-0.03	0.01	-0.003	0.001	-0.002*	0.0003
Year of Birth^2 x Age	--	--	-0.00004	0.00002	-0.00005	0.00002	-0.0005	0.0005	-0.001	0.0006	--	--	-0.00004	0.00008
Year of Birth × Age × Time	0.00007*	0.00003	-0.0003*	0.00009	0.00003	0.00007	-0.0006	0.003	-0.004	0.003	-0.001*	0.0003	0.000008	0.00005
Year of Birth^2 x Age x Time	--	--	-0.00004	0.000003	-0.00002	0.00002	-0.00002	0.00008	-0.00008	0.00007	--	--	0.000004	0.000001
Correlates														
Education	-0.09*	0.007	-0.11*	0.01	0.12*	0.008	2.04*	0.34	1.29*	0.35	0.73*	0.05	0.05*	0.009
Education x Time	-0.003*	0.0008	0.002	0.001	0.0009	0.001	-0.05	0.07	0.004	0.07	0.03	0.02	0.002	0.001
Education x Year of Birth	0.002*	0.0004	0.0008	0.002	0.002	0.001	-0.09	0.06	-0.05	0.06	0.002	0.006	0.0004	0.001
Education x Year of Birth^2	--	--	0.00003	0.00005	-0.0001*	0.00004	0.005	0.003	0.005	0.003	--	--	0.00002	0.00004
Education x Age	0.004	0.001	0.003	0.002	-0.002	0.001	-0.05	0.06	-0.02	0.06	-0.02	0.009	-0.001	0.0013
Education x Age x Year of Birth	-0.0001	0.0002	0.000007	0.0002	-0.0001	0.0002	0.009	0.007	0.008	0.006	0.001	0.002	0.0001	0.0001
Education x Age x Year of Birth^2	--	--	0.000005	0.000006	-0.00007	0.00005	-0.0001	0.0002	-0.00001	0.0002	--	--	-0.00006	0.00003
Education x Year of Birth x Time	0.000004	0.00005	-0.0007*	0.0002	-0.00009	0.0002	0.02	0.01	-0.004	0.01	-0.002	0.002	-0.00004	0.00002
Education x Year of Birth^2 x Time	--	--	--	--	0.000001	0.000006	-0.0009	0.0007	0.0007	0.0006	--	--	0.00002^	0.00006
Education x Age x Time	0.0001	0.0002	-0.0004	0.0003	-0.0003	0.0002	0.02	0.01	-0.005	0.01	-0.008^	0.003	-0.0005	0.0002
Education x Age x Time x Year of Birth	0.000008	0.000007	0.00006	0.00003	0.00001	0.00001	-0.001	0.001	0.001	0.001	0.00006	0.0001	0.00003	0.00001
Woman	0.21*	0.02	0.43*	0.02	0.03	0.02	-2.14*	0.56	2.40*	0.58	0.27	0.17	0.05^	0.02
Woman x Time	-0.003	0.001	-0.0008	0.002	0.009*	0.001	-0.07	0.03	0.03	0.03	-0.05	0.03	-0.003*	0.001
Woman x Year of Birth	-0.0002	0.004	-0.008^	0.003	0.004	0.003	-0.02	0.09	-0.17	0.09	0.005	0.01	-0.004	0.002
Woman x Year of Birth^2	0.000009	0.0002	0.0002	0.0001	-0.0004*	0.0001	-0.0001	0.005	0.005	0.005	--	--	0.0001	0.00008
White	-0.26*	0.03	-0.35*	0.03	0.33*	0.02	--	--	--	--	--	--	--	--
White x Time	-0.008*	0.002	-0.01*	0.002	-0.01*	0.002	--	--	--	--	--	--	--	--
White x Year of Birth	-0.005	0.004	-0.006	0.004	0.003	0.003	--	--	--	--	--	--	--	--
White x Year of Birth^2	0.0004^	0.0002	0.0004	0.0002	-0.0003*	0.0001	--	--	--	--	--	--	--	--
Random Effects														
Var intercept	1.73*	0.02	1.23*	0.01	0.71*	0.008	321.75*	5.14	361.30*	5.47	65.69*	0.88	0.45*	0.005
Var. time	0.009*	0.0001	0.005*	0.0001	0.002*	0.00008	1.30*	0.04	1.21*	0.04	0.83*	0.03	0.002*	0.00005
Cov. Intercept, time	0.06*	0.001	0.04*	0.001	0.007*	0.0006	6.17*	0.37	3.10	0.34	3.71*	0.15	0.003*	0.0004
Residual variance	0.19*	0.001	0.52*	0.003	0.39*	0.002	151.33*	0.81	102.10	0.55	31.79*	0.20	0.35*	0.001

Note. ^ = p = .01, * p < .01.

Table S7

Growth Models Examining Cohort and Education Differences Across Physical Health in South Korea, and Mexico

	South Korea: KLoSA						Mexico: MHAS					
	Functional limitations		Self-rated health		Health conditions		Functional limitations		Health conditions		Self-rated health	
Parameters	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed Effects												
Intercept	0.46*	0.07	3.39*	0.03	0.55*	.04	2.22*	0.08	0.79*	0.02	2.63*	0.02
Time	0.08*	0.02	-0.05*	0.01	0.06*	0.008	0.10*	0.01	0.03*	0.004	-0.007^	0.003
Time^2	0.003*	0.0006	-0.001	0.0005	-0.0003	0.0001	0.002*	0.0010	0.0006*	0.0002	0.0008	0.0002
Age	-0.05*	0.02	-0.01	0.007	0.06*	0.0009	0.05*	0.02	0.01*	0.002	-0.02	0.003
Age × Time	-0.01*	0.005	0.005	0.002	0.0006	0.002	0.002	0.003	0.00007	0.0006	-0.001*	0.0006
Cohort												
Year of Birth	-0.06*	0.01	0.02*	0.006	-0.0008	0.008	-0.03*	0.008	-0.01*	0.002	0.004	0.001
Year of Birth^2	--	--	--	--	--	--	--	--	--	--	--	--
Year of Birth × Time	-0.01*	0.004	0.006	0.002	-0.002	0.001	0.0007	0.001	0.0002	0.0005	-0.0004	0.0003
Year of Birth^2 x Time	--	--	--	--	--	--	--	--	--	--	--	--
Year of Birth × Age	0.0008	0.001	0.001*	0.0005	-0.002*	0.0006	-0.0002	0.0001	0.0002	0.0003	0.0005	0.0003
Year of Birth^2 x Age	--	--	--	--	--	--	--	--	--	--	--	--
Year of Birth × Age × Time	0.0003	0.0002	-0.00002	0.0001	0.000005	0.0001	-0.0002	0.0002	0.0001	0.00006	0.00005	0.00003
Year of Birth^2 x Age x Time	--	--	--	--	--	--	--	--	--	--	--	--
Correlates												
Education	-0.03	0.01	0.06*	0.007	-0.01	0.009	-0.13*	0.01	-0.02*	0.002	0.05*	0.002
Education x Time	-0.006	0.004	-0.005	0.003	0.003	0.002	-0.006*	0.002	-0.0003	0.0004	0.00005	0.0004
Education x Year of Birth	0.004	0.003	-0.002	0.001	-0.004	0.002	0.002	0.001	0.0005	0.0003	-0.0003	0.0003
Education x Year of Birth^2	--	--	--	--	--	--	--	--	--	--	--	--
Education x Age	0.003	0.003	-0.004	0.002	-0.005	0.002	0.003	0.002	0.001	0.0006	-0.001^	0.0005
Education x Age x Year of Birth	0.004	0.003	0.00009	0.0001	0.00004	0.001	-0.0001	0.0002	0.00001	0.00004	0.00005	0.00003
Education x Age x Year of Birth^2	--	--	--	--	--	--	--	--	--	--	--	--
Education x Year of Birth x Time	0.001	0.0008	0.0007	0.0006	-0.0008	0.0004	0.0003	0.0002	0.00003	0.00007	-0.00002	0.00007
Education x Year of Birth^2 x Time	--	--	--	--	--	--	--	--	--	--	--	--
Education x Age x Time	0.002	0.001	0.0007	0.0007	-0.0009	0.0005	0.0005	0.0004	-0.00006	0.0001	-0.00001	0.0001
Education x Age x Time x Year of Birth	0.00002	0.00005	-0.00005	0.00003	0.00006	0.00002	-0.00005	0.00003	0.00002	0.000008	--	--
Woman	-0.32*	0.04	-0.11*	0.02	0.06	0.02	1.25*	0.06	0.31*	0.02	-0.14*	0.02
Woman x Time	-0.005	0.005	0.01*	0.003	0.0003	0.002	0.001	0.009	-0.002	0.002	0.007*	0.002
Woman x Year of Birth	0.002	0.006	0.008*	0.003	-0.01	0.003	0.004	0.008	-0.003	0.002	-0.0007	0.002
Woman x Year of Birth^2	--	--	--	--	--	--	-0.00001	0.0007	-0.0003	0.0002	0.00009	0.0002
Random effects												
Var intercept	1.67*	0.04	0.25*	0007	0.76*	0.01	4.65*	0.09	0.48*	0.008	0.22*	0.006
Var. time	0.01*	0.0008	0.003*	0.0002	0.006*	0.0001	0.19*	0.001	0.01*	0.0001	0.002*	0.0001
Cov. Intercept, time	0.06*	0.004	-0.002*	0.0009	0.03*	0.001	0.02*	0.009	0.002*	0.0008	0.0003*	0.0006
Residual variance	0.72*	0.009	0.35*	0.004	0.06*	0.0008	4.64*	0.06	0.31*	0.004	0.39*	0.005

Note. ^ = p = .01, * p < .01.

Table S8

Growth Models Examining Cohort and Education Differences Across Mental Health in United States, Australia, and Germany

	United States: HRS				**Australia: HILDA**						**Germany: SOEP**			
	Depressive Symptoms		Episodic Memory		Vitality		Life Satisfaction		Mental Health		Life Satisfaction		Mental health	
Parameters	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed Effects														
Intercept	1.39*	0.05	51.79*	0.43	58.23*	0.46	7.89*	0.03	72.82*	0.62	6.37*	0.04	49.66*	0.16
Time	0.05*	0.007	-0.83*	0.06	-0.14	0.19	-0.008	0.006	-0.04	0.15	0.03*	0.007	-0.30*	0.07
Time^2	-0.002*	0.0002	0.007^	0.003	-0.004	0.002	0.001*	0.0002	-0.002	0.002	0.002*	0.0001	-0.02*	0.006
Age	0.04*	0.009	-0.57*	0.05	0.23	0.13	0.02*	0.005	0.38*	0.11	0.098	0.007	0.40*	0.02
Age × Time	-0.006*	0.001	0.06*	0.01	0.0004	0.04	0.003	0.001	0.02	0.03	-0.006*	0.001	0.07*	0.01
Cohort														
Year of Birth	0.04*	0.007	-0.07*	0.02	0.25	0.14	-0.01*	0.003	0.17	0.11	0.06*	0.005	0.19*	0.02
Year of Birth^2	-0.0009*	0.0002	--	--	-0.023*	0.007	--	--	-0.02*	0.006	-0.001*	0.0002	--	--
Year of Birth × Time	-0.006*	0.001	0.03*	0.003	0.03	0.04	0.001	0.0007	0.03	0.03	-0.008*	0.0009	0.05*	0.007
Year of Birth^2 x Time	0.0001*	0.00004	--	--	-0.003	0.001	--	--	-0.001	0.001	0.0004*	0.00003	--	--
Year of Birth × Age	-0.004	0.001	0.007*	0.003	-0.03	0.01	-0.001*	0.0003	-0.02	0.01	-0.004*	0.0006	-0.003	0.002
Year of Birth^2 x Age	0.0001*	0.00003	--	--	-0.0004	0.0005	--	--	-0.0008	0.0004	0.00003	0.00001	--	--
Year of Birth × Age × Time	-0.0003^	0.0001	-0.002*	0.0004	-0.0007	0.003	0.00004	0.00004	-0.0005	0.003	0.0008	0.0009	0.0006	0.0004
Year of Birth^2 x Age x Time	0.00001*	0.000004	--	--	-0.0001	0.00007	--	--	-0.00005	0.00005	-0.0007*	0.0002	--	--
Correlates														
Education	-0.13*	0.01	2.07*	0.08	1.06*	0.31	-0.01	0.01	0.71*	0.15	-0.0006	0.02	0.23*	0.05
Education x Time	-0.02	0.002	0.02	0.01	0.01	0.06	0.006*	0.002	0.02	0.03	0.01*	0.002	0.008	0.02
Education x Year of Birth	-0.007*	0.002	-0.03*	0.004	-0.04	0.06	0.002	0.001	-0.009	0.02	0.01*	0.002	-0.0005	0.006
Education x Year of Birth^2	0.0004*	0.00008	--	--	0.002	0.003	--	--	--	--	-0.0003*	0.00007	--	--
Education x Age	-0.005	0.003	-0.03	0.02	0.02	0.06	0.005*	0.002	0.03	0.02	0.01*	0.002	0.01	0.009
Education x Age x Year of Birth	0.001*	0.0003	0.0007	0.0007	0.006	0.006	-0.0001	0.0001	-0.001	0.001	-0.0004	0.0002	-0.0005	0.0005
Education x Age x Year of Birth^2	-0.0002	0.000009	--	--	-0.0002	0.0002	--	--	--	--	-0.00006	0.000005	--	--
Education x Year of Birth x Time	-0.0002	0.0003	-0.001	0.0008	0.002	0.01	-0.0005	0.0002	0.0005	0.003	-0.0008^	0.0003	0.0003	0.002
Education x Year of Birth^2 x Time	0.00002	0.00001	--	--	0.0002	0.0006	--	--	--	--	0.00001	0.00001	--	--
Education x Age x Time	-0.0002	0.0004	-0.002	0.003	0.0007	0.01	-0.0008	0.0004	-0.002	0.004	-0.001*	0.0004	0.002	0.004
Education x Age x Time x Year of Birth	0.0001*	0.00004	-0.00002	0.0001	0.0004	0.001	0.000007	0.00001	0.0002	0.0002	0.00004	0.00002	-0.00005	0.0001
Woman	0.33*	0.03	6.76*	0.26	-1.74*	0.51	0.11*	0.04	-1.27*	0.44	-0.04	0.04	1.43*	0.19
Woman x Time	0.002	0.003	0.06^	0.03	0.01	0.03	-0.003	0.003	-0.01	0.03	-0.003	0.002	-0.007	0.03
Woman x Year of Birth	-0.01	0.004	-0.11^	0.04	-0.21*	0.08	0.0006	0.003	-0.11	0.07	-0.008	0.004	-0.008	0.01
Woman x Year of Birth^2	0.006*	0.002	-0.001	0.002	0.005	0.004	--	--	0.004	0.004	0.0003	0.0002	--	--
White	-0.31*	0.04	5.98*	0.32	--	--	--	--	--	--	--	--	--	--
White x Time	0.003	0.005	0.04	0.03	--	--	--	--	--	--	--	--	--	--
White x Year of Birth	-0.015*	0.006	0.04	0.05	--	--	--	--	--	--	--	--	--	--
White x Year of Birth^2	0.0007	0.0002	-0.002	0.002	--	--	--	--	--	--	--	--	--	--
Random Effects														
Var intercept	2.14*	0.02	94.93*	1.32	269.60*	4.23	1.34*	0.02	194.72*	3.11	1.67*	0.02	58.57*	1.08
Var. time	0.004*	0.0002	0.17*	0.02	0.92*	0.03	0.007*	0.002	0.73*	0.03	0.008*	0.0002	4.20*	0.30
Cov. Intercept, time	0.06*	0.002	0.83*	0.13	2.81*	0.28	0.004^	0.002	1.81*	0.22	0.008*	0.001	5.08*	0.48
Residual variance	1.89*	0.01	138.51*	0.79	132.63*	0.70	1.01*	0.005	108.40*	0.58	1.41*	0.005	45.73*	0.32

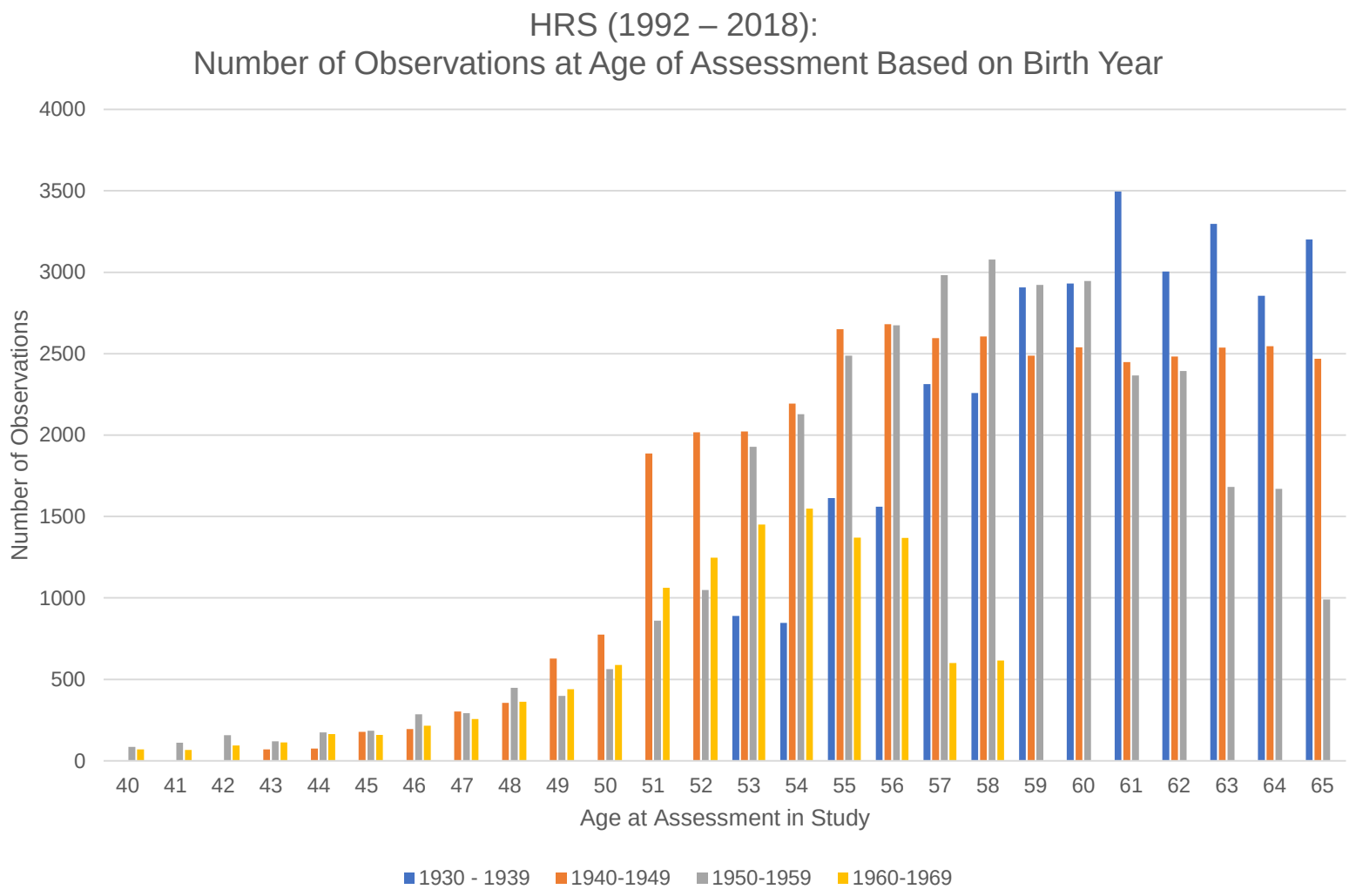
Note. ^ = p = .01, * p < .01.

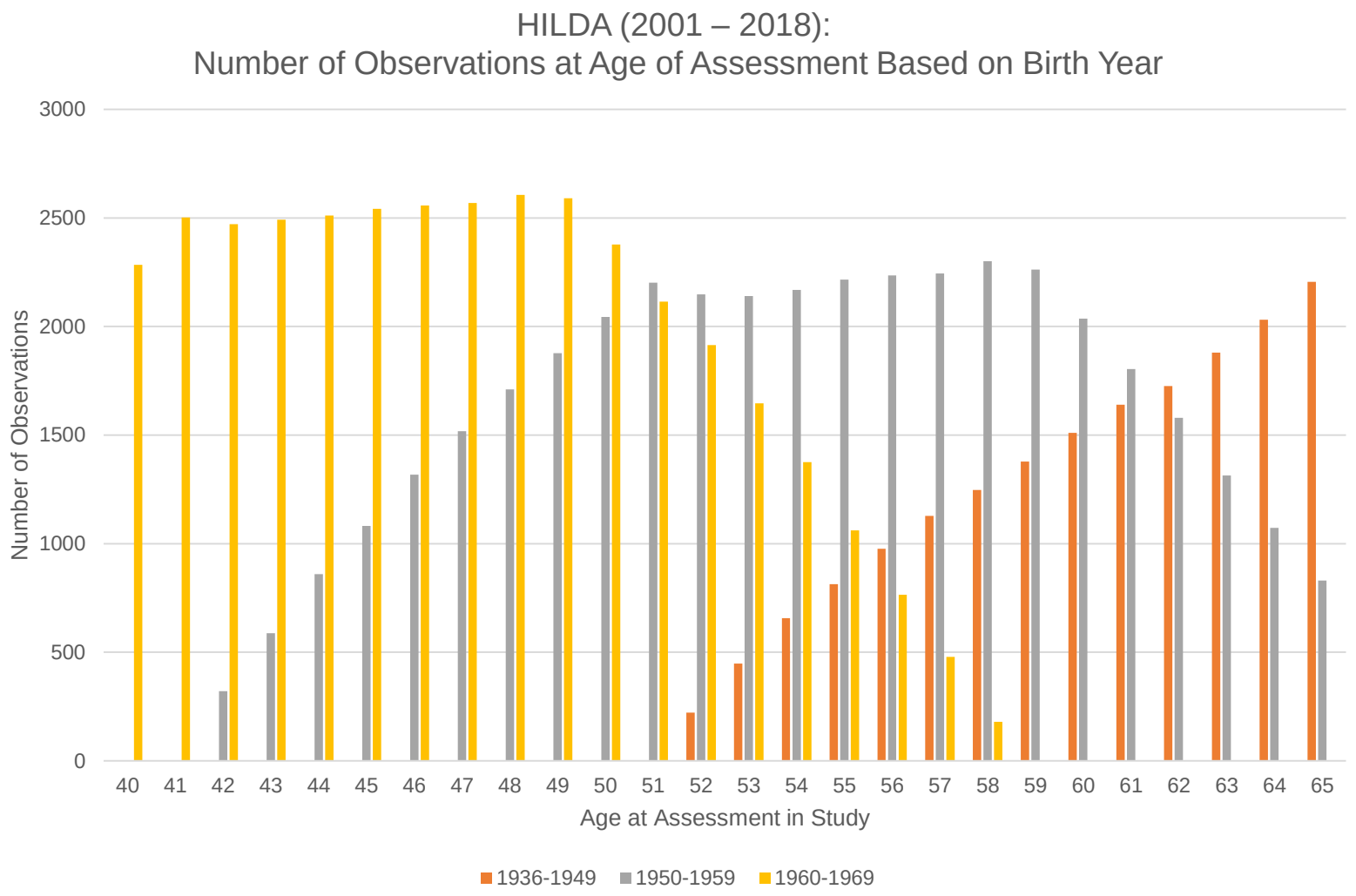
Table S9

Growth Models Examining Cohort and Education Differences Across Mental Health in South Korea, and Mexico

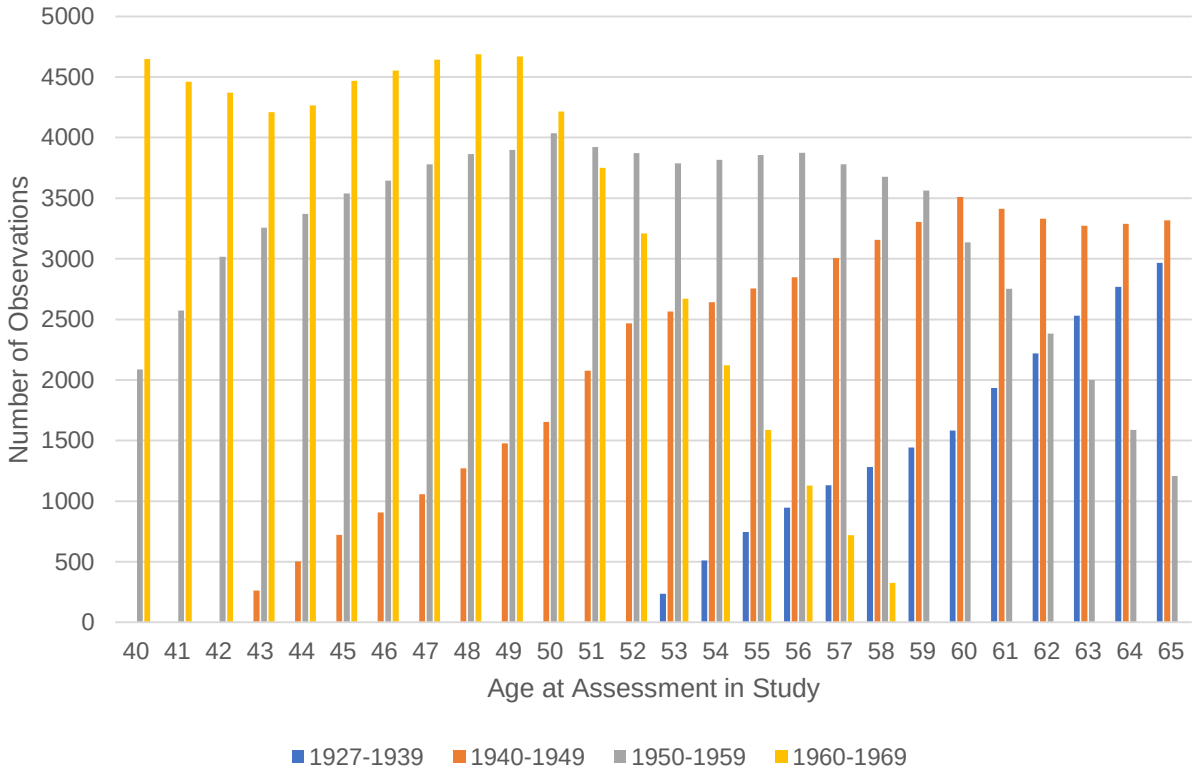
	South Korea: KLoSA						Mexico: MHAS	
	Life satisfaction		Depressive symptoms		MMSE		Depressive symptoms	
Parameters	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed Effects								
Intercept	64.46*	0.68	11.03*	0.13	27.97*	0.12	2.57*	0.06
Time	-0.93*	0.28	0.23*	0.06	-0.21*	0.06	-0.01	0.01
Time^2	0.05*	0.009	-0.006*	0.002	-0.02*	0.005	-0.003*	0.0007
Age	0.21	0.15	0.18*	0.02	-0.14^	0.06	0.02	0.01
Age × Time	0.18*	0.06	-0.02	0.01	-0.003	0.013	0.003	0.003
Cohort								
Year of Birth	0.20	0.14	0.13*	0.03	-0.06	0.04	-0.002	0.006
Year of Birth^2	--	--	--	--	--	--	--	--
Year of Birth × Time	0.17*	0.05	-0.01	0.01	0.002	0.01	0.0001	0.001
Year of Birth^2 x Time	--	--	--	--	--	--	--	--
Year of Birth × Age	0.02	0.01	-0.006*	0.002	0.0005^	0.002	-0.001	0.001
Year of Birth^2 x Age	--	--	--	--	--	--	--	--
Year of Birth × Age × Time	0.001	0.002	0.0004	0.0005	0.0001	0.0005	-0.0002	0.0002
Year of Birth^2 x Age x Time	--	--	--	--	--	--		
Correlates								
Education	1.46*	0.15	-0.15*	0.03	0.25*	0.028	-0.13*	0.008
Education x Time	-0.15^	0.06	0.03	0.01	0.02	0.01	-0.002	0.001
Education x Year of Birth	-0.09*	0.03	0.01	0.001	-0.01	0.005	0.002*	0.0009
Education x Year of Birth^2	--	--	--	--	--	--	--	--
Education x Age	-0.12*	0.03	0.02	0.007	-0.007	0.006	0.003	0.002
Education x Age x Year of Birth	0.002	0.002	-0.0002	0.0004	-0.0004	0.004	-0.00007	0.0001
Education x Age x Year of Birth^2	--	--	--	--	--	--	--	--
Education x Year of Birth x Time	0.02	0.01	-0.003	0.002	-0.003	0.002	0.00009	0.0002
Education x Year of Birth^2 x Time	--	--	--	--	--	--	--	--
Education x Age x Time	0.03	0.01	-0.004	0.003	-0.005	0.003	0.0002	0.00002
Education x Age x Time x Year of Birth	-0.0007	0.0006	0.0002	0.0001	0.0003*	0.0001	0.00001	0.00003
Woman	1.05*	0.38	0.16	0.07	-0.18*	0.07	0.95*	0.05
Woman x Time	-0.03	0.07	-0.01	0.01	0.005	0.01	-0.005	0.007
Woman x Year of Birth	0.14	0.06	-0.03*	0.01	0.04*	0.01	0.005	0.006
White x Year of Birth^2	--	--	--	--	--	--	--	--
Random effects								
Var intercept	121.42*	3.40	3.73*	0.11	4.43*	0.12	2.52*	0.06
Var. time	1.25*	0.10	0.06*	0.004	0.04*	0.004	0.03*	0.005
Cov. Intercept, time	-2.61*	0.42	0.12*	0.01	0.10*	0.01	0.006*	0.0009
Residual variance	159.51*	1.91	6.62*	0.08	4.50*	0.06	3.83*	0.04

Note. ^ = p = .01, * p < .01.

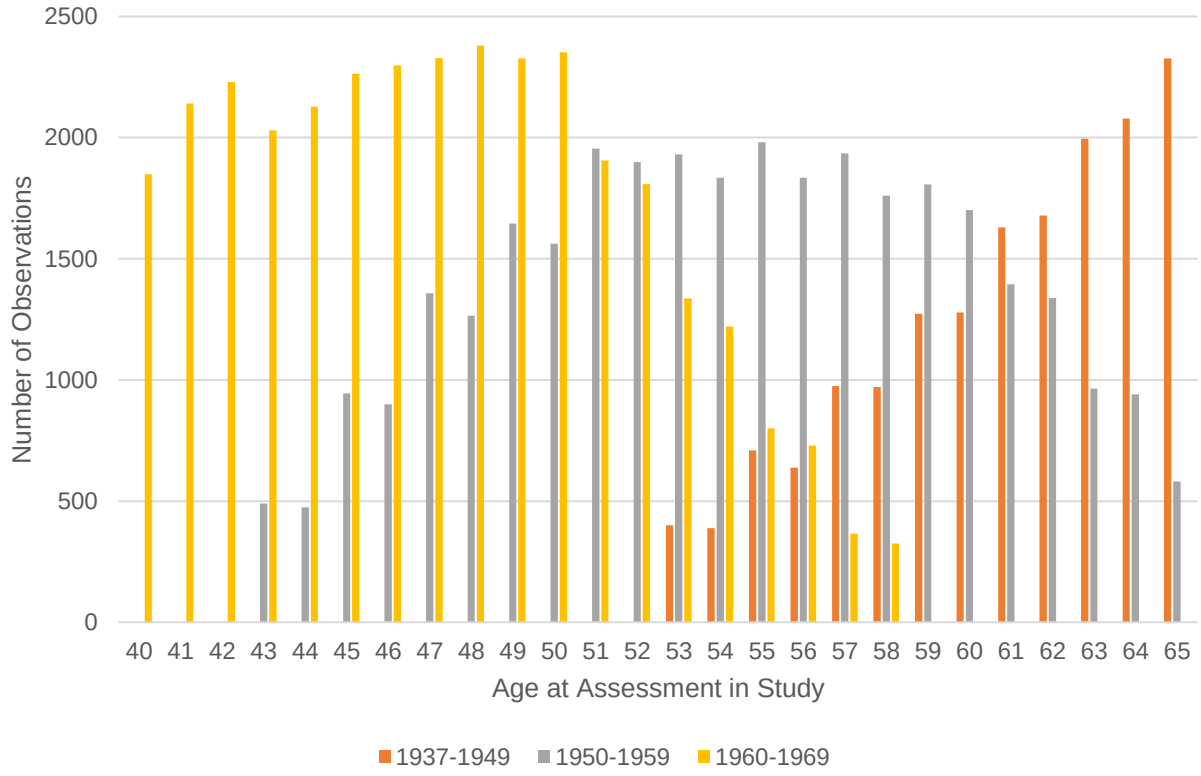

HRS (1992 – 2018):
Number of Observations at Age of Assessment Based on Birth Year
Number of Observations
4000
3500
3000
2500
2000
1500
1000
500
0
40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65
Age at Assessment in Study
1930 - 1939
1940-1949
1950-1959
1960-1969

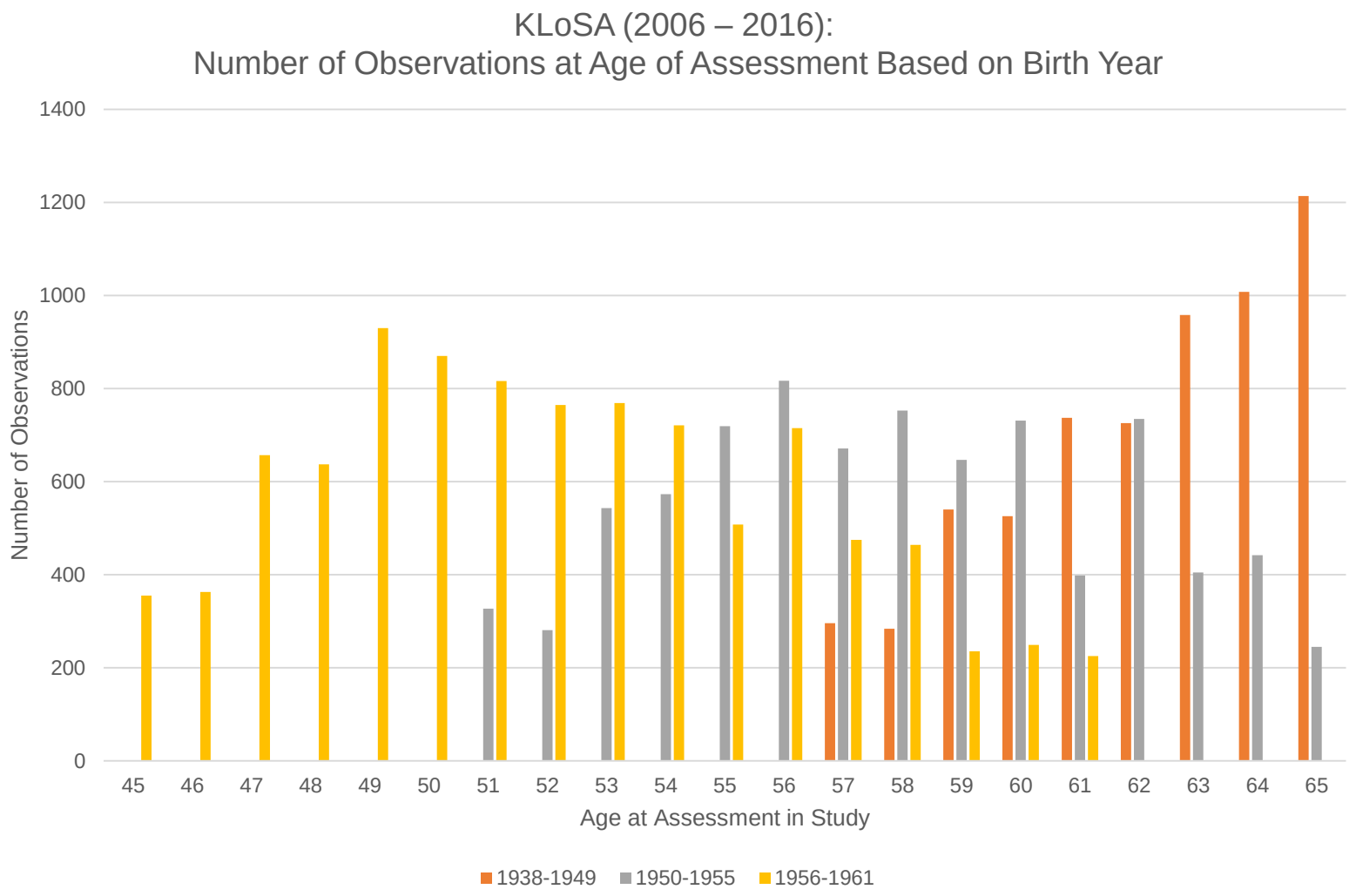

HILDA (2001 – 2018):
Number of Observations at Age of Assessment Based on Birth Year
Number of Observations
3000
2500
2000
1500
1000
500
0
40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65
Age at Assessment in Study
1936-1949
1950-1959
1960-1969

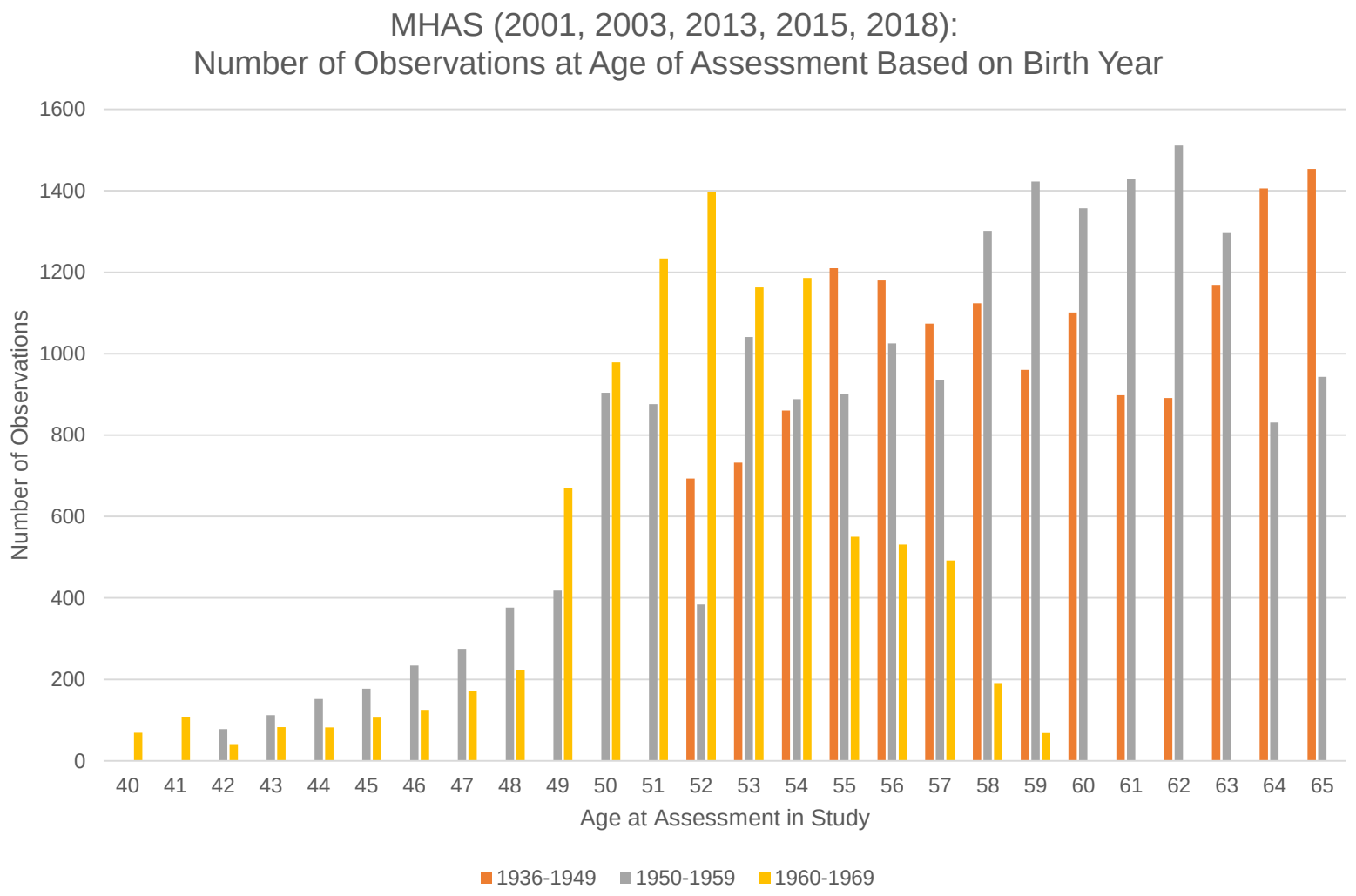
SOEP (1992 – 2018):
Number of Observations at Age of Assessment Based on Birth Years



SOEP (2002 – 2018):
Number of Observations at Age of Assessment Based on Birth Year




KLoSA (2006 – 2016):
Number of Observations at Age of Assessment Based on Birth Year
Number of Observations
1400
1200
1000
800
600
400
200
0
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
Age at Assessment in Study
1938-1949
1950-1955
1956-1961


MHAS (2001, 2003, 2013, 2015, 2018):
Number of Observations at Age of Assessment Based on Birth Year
Number of Observations
1600
1400
1200
1000
800
600
400
200
0
40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65
Age at Assessment in Study
1936-1949
1950-1959
1960-1969