**Online Supplement**

**Sample Description**

**Berlin Aging Study (BASE).** The full BASE sample comprised 516 residents of former West-Berlin districts who were randomly drawn from the obligatory city registry and stratified by age and gender (age: *M* = 84.92, *SD* = 8.66, range = 70–103; 50% women; for sample characteristics, see Lindenberger et al., 1999). Our report makes use of multi-wave data obtained from 228 participants who were born between 1889 and 1922. Testing took place at the participants’ place of residence (i.e., private household or institution) and was carried out in individual face-to-face sessions by trained research assistants. Sessions required an average of 90 minutes and, when necessary, were split into shorter units of assessment. Ethics approval for BASE was granted by the Berlin Medical Association. Initial testing took place in 1990/93 and was repeated up to 10 times over the following 13 years. With the exception of T2, each wave involved one multidisciplinary Intake Assessment followed by a series of discipline-specific Intensive Protocol sessions. More specifically, the Intake Assessment was designed to gain as much basic information as possible at an early stage. On average, it took 90 minutes to answer the 100 questions on a wide range of topics. The Intensive Protocol sessions were designed to assess the many facets of old age and aging using a variety of standardized instruments and measures selected by each of the four research units (Psychology, Internal Medicine, Psychiatry, Sociology). On average, each of the 14 Intensive Protocol sessions also lasted 90 minutes.

**Berlin Aging Study II (BASE-II).** The BASE-II sample included residents of the greater metropolitan area of Berlin. Potential participants were recruited via a participant pool at the Max Planck Institute for Human Development (Berlin) and via advertisements placed in local newspapers and the public transportation system. Our analyses make use of multi-wave data from 583 participants who were born between 1925 and 1948. Cognitive testing was carried out by trained interviewers in group-sessions of three to six participants. Ethics approval for BASE-II was granted by the ethics committee of the Charité-Universitätsmedizin Berlin and the ethics committee of the Max Planck Institute for Human Development, Berlin.

With our earlier-born BASE sample being born between 1889 and 1922, they had lived at least through the Second World War. As a consequence, the biographies of individuals in the earlier-born cohort are probably to a greater extent shaped by such major historical events. It is possible that those experiences had contributed to earlier-reported differences between the cohorts (e.g., substantially higher external control beliefs among the earlier-born cohort: Hueluer et al., 2016). We also note that we are not in a good position to empirically test East versus West differences: The earlier-born BASE sample was recruited only in the former West-Berlin, as recruitment took place before the fall of the Berlin Wall in 1989, and so cannot be compared with earlier-born older adults in East-Berlin. Finally, the large majority of participants in both samples (more than 95%) were of German citizenship since birth. It is thus an open question whether and how our findings generalize to more diverse samples.

**Quantification of Sample Selection.** We examined the extent of selectivity of each BASE subsample used here relative to the larger populations from which they were drawn, in a first step with the larger BASE and BASE-II samples and in a second step with the propensity-matched samples. Data obtained from participants aged 70+ in the nationally representative SOEP study in 1990 served as the reference for the BASE sample and SOEP data obtained from participants aged 60+ in 2010 served as the reference for the BASE-II sample. In particular, we first generated a variable indicating sample of origin (e.g., BASE vs. SOEP) and then estimated (separate) logistic regression analyses of education and household income on sample of origin (see Sassenroth, Kroh, & Wagner, 2013). Results for the larger BASE and BASE-II samples revealed that higher education (BASE: OR = 1.10, 95% Confidence Interval =1.04–1.16, *p* < .01; BASE-II: OR = 1.20, CI = 1.18–1.22, *p* < .001) and higher income (BASE: OR = 5.48, CI = 4.18–7.18, *p* < .001; BASE-II: OR = 2.09, CI = 1.88–2.31, *p* < .001) were each associated with participation in the BASE studies. This suggests that both BASE samples represent positive selections of the larger populations they were each drawn from (for an in-depth analyses of selection effects in BASE, see Lindenberger et al., 1999). Most important for our research question is that the earlier-born BASE cohort sample was considerably more select on income than the later-born BASE-II cohort sample. On education, the later-born BASE-II cohort sample was minimally more select than the earlier-born BASE cohort sample.

In our design that compares case-matched controls, the propensity-matched BASE and BASE-II samples remained positively select on education (BASE-I: OR = 1.16, CI = 1.09–1.24, *p* < .001; BASE-II: OR = 1.28, CI = 1.24–1.31, *p* < .001). The overlapping confidence intervals indicate that the matched samples do not differ in the degree by which they were positively select on education. In a similar vein, the matched BASE and BASE-II samples remained positively select on income, although the extent of selection was reduced (BASE: OR = 4.88, CI = 3.49–6.83, *p* < .001; BASE-II: OR = 2.19, CI = 1.84–2.60, *p* < .001). These selectivity analyses suggest that our case-matched control design represents a fair, if not conservative test of cohort differences because either the amount of positive selection was comparable across the matched BASE and BASE-II or matched BASE participants were even more positively select than the matched BASE-II participants.

References

Hueluer, G., Drewelies, J., Eibich, P., Duezel, S., Demuth, I., Ghisletta, P., Steinhagen–Thiessen, E., Wagner, G. G., Lindenberger, U., & Gerstorf, D. (2016). Cohort differences in psychosocial function over 20 years: Current older adults feel less lonely and less dependent on external circumstances. *Gerontology*, *62*(3), 354–361. <https://doi.org/10.1159/000438991>

Lindenberger, U., Gilberg, R., Litte, T. D., Nuthmann, R., Pötter, U., & Baltes, P. B. (1999). Sample selectivity and generalizability of the results of the Berlin Aging Study. In P. B. Baltes & P. B. Baltes (Eds.), *The Berlin Aging Study: Aging from 70 to 100* (pp. 56-82). New York, NY: Cambridge University Press.

Sassenroth, D., Kroh, M., & Wagner, G. G. (2013). *Selectivity processes in and weights for the Berlin Aging Study II (Base-II) (December 2013)*. SOEPpaper No. 608, Available at SSRN: [https://ssrn.com/abstract=2365707](https://ssrn.com/abstract%3D2365707) or [http://dx.doi.org/10.2139/ssrn.2365707](https://dx.doi.org/10.2139/ssrn.2365707)

**Statistical Power Considerations**

We are aware that statistical power in multilevel models (MLMs) is a complex issue. Among other reasons, MLMs are often more complex than simpler regression models because multiple sources of variance are involved. To evaluate statistical power, several factors need to be taken into account (e.g., Browne et al., 2009; Brysbaert & Stevens, 2018; Kreft & de Leeuw, 1998; Moerbeek et al., 2008; Snjiders & Bosker, 1993, 2012; Spybrook et al., 2011). First, various estimators for MLM parameter estimates (especially the standard errors) exist, including ordinary least squares, iterative generalized least squares, and restricted iterative least squares. In the current study, we have used maximum likelihood procedures that are known to provide overall the highest power, while maintaining type-I error rates at their nominal level. Second, it is an established finding that statistical power is generally greater for fixed effects than for random effects. Third, it is also known that statistical power is generally greater for fixed effects at more micro-levels (e.g., level-1) than for fixed effects at more macro-levels (e.g., level-2). For the former the total number of observations counts (though these need to be offset if the intra-class correlation is high), for the latter the number of level-2 units is very important. Following this logic, cross-level interactions (between a level-1 and a level-2 predictor) fare somewhere in between. Their statistical power depends on the combination of number of level-2 units and number of observations within the level-2 units (i.e., the number of level-1 units). Fourth, the statistical power for cross-level interactions depends also on the variance detected at level-2. This is so because the cross-level interaction is basically a level-1 predictor that attempts to explain variance at level-2. Fifth, statistical power is directly related to the effect size of the parameter of focus (i.e., the statistical power is larger if the effect size is larger).

These five points of consideration are all “classic” findings about power, and thus were all either mathematically derived or (via simulation work) found in the context of the MLM that contained linear fixed and linear random effects. The MLM we report in our study is a model with some nonlinear fixed effects, notably those about within-person change, and linear random effects. However, work on power in nonlinear MLMs is practically nonexistent. To the best of our knowledge, there exist no dedicated articles or chapters discussing this issue with the same focus and detail as for fully linear MLMs. However, several authors report that the basic principles listed above about power in fully linear MLMs ought to apply to nonlinear MLMs, at least those with linear random effects as implemented in the current study (e.g., Davidian & Giltinan, 1995; Pinheiro & Bates, 2000; Vonesh & Chinchilli, 1996). This is so because the estimation methods applied to MLMs with nonlinear fixed effects but linear random effects are basically linearizing the estimation of the fixed effects, so as to be able to derive characteristics (adjustment statistics, standard errors, etc.) as is done in classical fully linear MLMs.

With these general considerations in mind, we now move more specifically to our study. The effects that most interested us in our application to the BASE and BASE-II data are the cohort effects on the intercept and on the within-person age-related change parameters. Our major finding is that while we found the former to be substantial in size, the latter was not different from zero. One key question then is that perhaps our design, or our analysis, did not offer the power needed to pick up even a small, but significant, effect that would speak in favor of differential within-person change across the two cohorts. However, as noted, we conducted several sets of robustness analyses and still found no evidence for cohort effects on between-person age differences of within-person age changes. What we did not report (for brevity) in the article is that the variance in change detected in the fully linear MLM that we carried out as one such follow-up robustness analysis was incredibly lower (0.45) than it was in either nonlinear MLM (51.90 in Model 1 and 52.74 in Model 2). Thus, our nonlinear MLM offered much greater power for a cross-level interaction, such as the cohort effect on the random slope. Moreover, studies focusing on the statistical power to detect cross-level interactions suggest that with 150 level-2 units, five level-1 units per level-2 unit are sufficient to detect even relatively small effects. Our analyses are based on around 811 individuals (i.e., level-2 units), each with up to 11 and six assessments (i.e., level-1 units) in BASE, respectively, BASE-II. Granted, not all were observed during the entire study, but by adding the study and analytical aspects that favor statistical power, we believe that our approach maximized power to detect both between-person and within-person age gradients under the influence of cohort effects. At the same time, we fully acknowledge that more work is needed to corroborate and extend the current findings with results obtained from further studies and that a conservative interpretation is warranted with the in part large Standard Errors we have obtained for several of our parameter estimates.

References

Browne, W., Golalizadeh Lahi, M., & Parker, R. (2009). *A guide to sample size calculations for random effect models via simulation and the MLPowSim Software Package*. University of Bristol.

Brysbaert, M., & Stevens, M. (2018). Power Analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 9. <https://doi.org/10.5334/joc.10>

Davidian, M., & Giltinan, D. M. (1995). *Nonlinear models for repeated measurement data*. New York, NY: Chapman & Hall.

Kreft, I. G. G., & de Leeuw, J. (1998). *Introducing Multilevel Modeling* (1 edition). London, UK: Sage.

Moerbeek, M., Van Breukelen, G. J. P., & Berger, M. P. F. (2008). Optimal designs for multilevel studies. In Leeuw, J. & Meijer, E. (Eds)*Handbook of multilevel analysis* (pp. 177–205). New York, NY: Springer. <https://doi.org/10.1007/978-0-387-73186-5_4>

Pinheiro, J., & Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. New York, NY: Springer. <https://doi.org/10.1007/b98882>

Snijders, T. A., & Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational Statistics*, *18*(3), 237–259. <https://doi.org/10.2307/1165134>

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London, UK: Sage.

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A., Raudenbush, S., & To, A. (2011). Optimal design plus empirical evidence: Documentation for the “Optimal Design” software. *Survey Research Center of the Institute of Social Research at University of Michigan*, *5*, 2012.

Vonesh, E., & Chinchilli, V.M. (1997). Linear and Nonlinear Models for the Analysis of Repeated Measurements (1st ed.). New York, NY: CRC Press. <https://doi.org/10.1201/9781482293272>

**Statistical Analysis SAS Codes**

**Table 2**

TITLE1 'Exponential model with cohort effect on AMOUNT of exp. decline' ;

TITLE2 'Model 1 Without sex and educ (N=2008)' ;

**PROC** **NLMIXED** DATA=one MAXITER=**1000** ;

 traject = a + bbase2\*base2 + (d1+dbase2\*base2)\*(EXP(e\*wpA)) + bA\*bpA + bAbase2\*base2\*bpA +

 r\_lin\*rlin ;

 MODEL ds ~ NORMAL(traject, v\_e) ;

 RANDOM a d1 ~ NORMAL([m\_a,m\_d], [v\_a,c\_ad,v\_d])

 SUBJECT = id ;

 PARMS m\_a=**50** bA=**1** r\_lin=**.1** m\_d=**10** e=**.05** bbase2=**12** bAbase2=**0.1** dbase2=**.01**

 v\_a=**140** v\_d=**20**

 v\_e=**20** c\_ad=**0**;

**RUN** ;

TITLE1 'Exponential model with cohort effect on AMOUNT of exp. decline' ;

TITLE2 'Model 2 With sex and educ (N=1903)' ;

**PROC** **NLMIXED** DATA=one MAXITER=**1000** ;

 traject = a + bwomen\*women + beduc\*educc + bbase2\*base2 +

 bwomenBase2\*women\*base2 + beducBase2\*educc\*base2 +

 (d1+dwomen\*women)\*(EXP(e\*wpA)) +

 (d1+deduc\*educc)\*(EXP(e\*wpA)) +

 (d1+dbase2\*base2)\*(EXP(e\*wpA)) +

 (d1+dwomenbase2\*women\*base2)\*(EXP(e\*wpA)) +

 (d1+deducbase2\*educc\*base2)\*(EXP(e\*wpA)) +

 bA\*bpA + bAbase2\*base2\*bpA +

 r\_lin\*rlin ;

 MODEL ds ~ NORMAL(traject, v\_e) ;

 RANDOM a d1 ~ NORMAL([m\_a,m\_d], [v\_a,c\_ad,v\_d])

 SUBJECT = id ;

 PARMS m\_a=**40** bA=-**1** r\_lin=**1** m\_d=-**10** e=**.05** bwomen=**0.1** beduc=**.1**

 dwomen=**.1** deduc=**.1** bbase2=**12** bwomenBase2=**.1** beducBase2=**.1**

 bAbase2=**0.1** dbase2=**.01** dwomenbase2=**.1** deducbase2=**.1**

 v\_a=**100** v\_d=**20**

 v\_e=**20** c\_ad=**2** ;

**RUN** ;

**Table 3**

DM 'OUTPUT; CLEAR;';

TITLE1 'Exponential model with cohort effect on AGE OF ONSET of exp. decline' ;

TITLE2 'Model 3 Without sex and educ (N=2008)' ;

**PROC** **NLMIXED** DATA=one MAXITER=**1000** ;

 traject = a + bbase2\*base2 + d1\*(EXP(e\*(wpA+sbase2\*base2))) + bA\*bpA + bAbase2\*base2\*bpA +

 r\_lin\*rlin ;

 MODEL ds ~ NORMAL(traject, v\_e) ;

 RANDOM a d1 ~ NORMAL([m\_a,m\_d], [v\_a,c\_ad,v\_d])

 SUBJECT = id ;

 PARMS m\_a=**50** bA=**1** r\_lin=**.1** m\_d=-**10** e=**.05** bbase2=**12** bAbase2=-**1** sbase2=-**1**

 v\_a=**140** v\_d=**40**

 v\_e=**20** c\_ad=**0** ;

**RUN** ;

TITLE1 'Exponential model with cohort effect on AGE OF ONSET of exp. decline' ;

TITLE2 'Model 4 With sex and educ (N=1903)' ;

**PROC** **NLMIXED** DATA=one MAXITER=**1000** ;

 traject = a + bwomen\*women + beduc\*educc + bbase2\*base2 +

 bwomenBase2\*women\*base2 + beducBase2\*educc\*base2 +

 d1\*(EXP(e\*(wpA+swomen\*women))) +

 d1\*(EXP(e\*(wpA+seduc\*educc))) +

 d1\*(EXP(e\*(wpA+sbase2\*base2))) +

 d1\*(EXP(e\*(wpA+swomenbase2\*women\*base2))) +

 d1\*(EXP(e\*(wpA+seducbase2\*educc\*base2))) +

 bA\*bpA + bAbase2\*base2\*bpA +

 r\_lin\*rlin ;

 MODEL ds ~ NORMAL(traject, v\_e) ;

 RANDOM a d1 ~ NORMAL([m\_a,m\_d], [v\_a,c\_ad,v\_d])

 SUBJECT = id ;

 PARMS m\_a=**50** bA=-**1** r\_lin=**1** m\_d=-**10** e=**.05** bwomen=**1** beduc=**1**

 swomen=-**1** seduc=**1** bbase2=**12** bwomenBase2=**.1** beducBase2=-**.1**

 bAbase2=-**0.1** sbase2=-**1** swomenbase2=-**1** seducbase2=-**1**

 v\_a=**100** v\_d=**20**

 v\_e=**20** c\_ad=**2** ;

**RUN** ;

**Figure 2**

**DATA** data\_base;

 DO wpA= -**7.6** TO **8.6** BY **0.1**;

 id = wpA;

 ds = **40.4100** + (**12.2077** \* **0**) + (-**12.4005** + -**1.4055** \* **0**) \* EXP(**0.08951** \* wpA) +

 (-**0.3633** \* **0.0012550**) + (-**0.05774** \* **0** \* **0.0012550**) + (**0.8955** \* **2.5194223**) ;

 color = **1111111111**;

 OUTPUT; END;

**RUN**;

**DATA** data\_base2;

 DO wpA= -**7.6** TO **8.6** BY **0.1**;

 id = wpA;

 ds = **40.4100** + (**12.2077** \* **1**) + (-**12.4005** + -**1.4055** \* **1**) \* EXP(**0.08951** \* wpA) +

 (-**0.3633** \* **0.0012550**) + (-**0.05774** \* **1** \* **0.0012550**) + (**0.8955** \* **2.5194223**) ;

 color = **2222222222**;

 OUTPUT; END;

**RUN**;

**PROC** **MEANS** DATA=data\_base;

**RUN**;

**PROC** **MEANS** DATA=data\_base2;

**RUN**;

\*merging data files;

**DATA** gradient\_basebase2A;

 MERGE data\_base data\_base2;

 BY color;

 age=wpA + **77.8040388**;

 ager=ROUND(wpA + **77.8040388**);

 id=color;

**RUN**;

**PROC** **SORT** DATA=both;

 BY id;

**RUN**;

**PROC** **SORT** DATA=gradient\_basebase2A;

 BY id;

**RUN**;

**DATA** gradient\_basebase2B;

 MERGE gradient\_basebase2A both;

 BY id;

**RUN**;

\*plot;

GOPTIONS RESET=ALL;

GOPTIONS /\*FTEXT = triplex\*/ CTEXT=black HTEXT=**2** FTITLE=SWISSX HTITLE=**2**;

**RUN**;

TITLE1 ' ';

TITLE2 ' ';

**RUN**;

**PROC** **GPLOT** DATA=gradient\_basebase2B;

 SYMBOL1 REPEAT=**228** I=join V=dot H=**0.5** W=**0.5** C=black;

 SYMBOL2 REPEAT=**583** I=join V=dot H=**0.5** W=**0.5** C=red;

 SYMBOL3 REPEAT=**1** I=join V=none H=**.5** W=**7** C=black;

 SYMBOL4 REPEAT=**1** I=join V=none H=**.5** W=**7** C=red;

 AXIS1

 ORDER = (**0** to **100** by **10**)

 LABEL = (A=**90** 'Digit Symbol')

 MINOR = none

 OFFSET = (**2**)

 LENGTH = **78**;

 AXIS2

 ORDER = (**68** to **96** by **2**)

 LABEL = ('Chronological Age')

 MINOR = none

 OFFSET = (**2**)

 OFFSET= (**1** PCT, **1** PCT)

 LENGTH = **150**;

 PLOT ds \* age = id /NOLEGEND VAXIS=AXIS1 HAXIS=AXIS2 NOFRAME;

**RUN**;

**QUIT**;

Table S1 *Exponential Growth Models with Cohort Differences in the Amount of Exponential Decline,* ***Retests not included***

|  |  |  |  |
| --- | --- | --- | --- |
|  | Model 1 |  | Model 2 |
|  | Parameter estimate | Standard error |  | Parameter estimate | Standard error |
| Fixed effects |  |  |  |  |  |
| Intercept *γ00* | 34.03\* | 2.19 |  | 38.71\* | 4.42 |
| Between-person age *γ01* | –0.42\* | 0.17 |  | –0.52\* | 0.16 |
| Baseline amount of exponential decline *10* | –4.43\* | 2.15 |  | –1.39 | 0.90 |
| Exponential rate of decline ** | 0.10\* | 0.04 |  | 0.08\* | 0.04 |
| Covariates |  |  |  |  |  |
| Women *05* | – | – |  | 0.31 | 2.30 |
| Education *06* | – | – |  | 1.56\* | 0.52 |
| Women x amount of exponential decline *12* | – | – |  | 1.30 | 2.13 |
| Education x amount of exponential decline *13* | – | – |  | –0.14 | 0.51 |
| Cohort |  |  |  |  |  |
| Later-born BASE-II intercept *γ03* | 15.54\* | 2.16 |  | 10.48\* | 2.99 |
| Later-born BASE-II intercept x women *07* | – | – |  | 1.21 | 3.78 |
| Later-born BASE-II intercept x education *08* | – | – |  | –1.29 | 0.78 |
| Later-born BASE-II intercept x between-person age *γ04* | 0.05 | 0.21 |  | 0.08 | 0.21 |
| Later-born BASE-II x amount of exponential decline *11* | –3.94 | 2.06 |  | –3.10 | 2.83 |
| Later-born BASE-II x women x amount of exponential decline *14* | – | – |  | –0.49 | 3.66 |
| Later-born BASE-II x education x amount of exponential decline *15* | – | – |  | 0.33 | 0.78 |
| Random effects |  |  |  |  |  |
| Variance intercept *τ20* | 102.62\* | 30.72 |  | 99.01\* | 37.64 |
| Variance amount of exponential decline *τ20* | 46.80 | 40.77 |  | 2.64 | 2.59 |
| Cov. Intercept–amount of exponential decline *τ01* | –34.10 | 34.82 |  | –8.99 | 10.02 |
| Residual variance σ2 | 15.70\* | 0.72 |  | 15.62\* | 0.75 |
| *–2LL* | 13,336 |  | 12,574 |
| *BIC* | 13,410 |  | 12,700 |

*Note*. The baseline amount of exponential decline refers to men with average education level in the BASE cohort. Model 1: 2,008 observations. Model 2: 1,903 observations. \* *p* < .05.

Table S2 *Exponential Growth Models with Cohort Differences in the Age of onset of Exponential Decline,* ***Retests not included***

|  |  |  |  |
| --- | --- | --- | --- |
|  | Model 3 |  | Model 4 |
|  | Parameter estimate | Standard error |  | Parameter estimate | Standard error |
| Fixed effects |  |  |  |  |  |
| Intercept *γ00* | 34.29\* | 2.56 |  | 44.34\* | 6.61 |
| Between-person age *γ01* | –0.43\* | 0.17 |  | –0.44\* | 0.16 |
| Baseline amount of exponential decline *10* | –4.69 | 2.49 |  | –2.55 | 1.35 |
| Exponential rate of decline ** | 0.11\* | 0.05 |  | 0.05\* | 0.027 |
| Covariates |  |  |  |  |  |
| Women *05* | – | – |  | –0.19 | 1.17 |
| Education *06* | – | – |  | 1.34\* | 0.37 |
| Women x onset of exponential decline ** | – | – |  | –2.90 | 2.33 |
| Education x onset of exponential decline ** | – | – |  | 0.90 | 2.22 |
| Cohort |  |  |  |  |  |
| Later-born BASE-II intercept *γ03* | 12.39\* | 1.03 |  | 7.45\* | 0.86 |
| Later-born BASE-II intercept x women *07* | – | – |  | 2.32 | n.e. |
| Later-born BASE-II intercept x education *08* | – | – |  | –1.06\* | 0.37 |
| Later-born BASE-II intercept x between-person age *γ04* | 0.07 | 0.21 |  | 0.01 | 0.21 |
| Later-born BASE-II x age of onset of exponential decline ** | 1.33 | 1.22 |  | –2.19 | n.e. |
| Later-born BASE-II x women x age of onset of exponential decline ** | – | – |  | –2.12 | n.e. |
| Later-born BASE-II x education x age of onset of exponential decline ** | – | – |  | –2.64 | 2.97 |
| Random effects |  |  |  |  |  |
| Variance intercept *u0i* | 116.98\* | 46.05 |  | 98.27\* | 30.33 |
| Variance amount of exponential decline *u1i* | 33.98 | 37.86 |  | 5.47 | 3.39 |
| Cov. intercept–amount of exponential decline  | –33.85 | 40.84 |  | –15.52 | 10.52 |
| Error variance  | 15.58\* | 0.72 |  | 16.61\* | 1.11 |
| *–2LL* | 13,347 |  | 12,592 |
| *BIC* | 13,421 |  | 12,718 |

*Note*. n.e. = could not be estimated. The baseline amount of exponential decline refers to men with average education level in the BASE cohort. Model 3: 2,008 observations. Model 4: 1,903 observations. \* *p* < .01.

Tables S1 and S2 report from follow-up analyses that omit the retest effect in each of the four models we have tested. As can be obtained, findings are substantively identical to those reported in the main text: Differences in levels of cognitive functioning between the two cohorts are large, but there is no evidence for cohort differences in the rate of decline or in the onset of decline. The differences by education were also retained.

Three additional points are of note. First, for our most complex model 4, we were faced with model convergence problems such that not all standard errors could be estimated. Second, omitting the retest effects from any of the models resulted in substantial loss of fit relative to the main models that included retest effects (e.g., –2LL in Model 3: 13,347 vs. 13,305). Third, when retest effects were not modeled, the intercept difference between the two cohorts appeared even larger and was presumably overestimated.