|  |
| --- |
| eTable 1. Logistic regression of dementia, reporting average marginal effects. |
| Variable | Model 1 | Model 2 | Model 3 | Model 4 |
|  | AME | 95% CI | AME | 95% CI | AME | 95% CI | AME | 95% CI |
| Time (Unit: 10 years) | -0.004 | (-0.007, -0.002) | 0.000 | (-0.005, 0.005) | -0.013 | (-0.018, -0.008) | -0.005 | (-0.010, -0.001) |
| Person-waves | 180,162 | 180,162 | 180,162 | 180,162 |
| Model 1 adjusts for age, age squared, sex, race/ethnicity, proxy status, and proxy\*sex.Model 2 adjusts for age, age squared, sex, race/ethnicity, proxy status, proxy\*sex, interview number, and proxy\*interview number.Model 3 adjusts for age, age squared, sex, race/ethnicity, proxy status, proxy\*sex, interview number, proxy\*interview number, and total interviews.Model 4 adjusts for age, age squared, sex, race/ethnicity, proxy status, proxy\*sex, interview number, proxy\*interview number, total interviews, and education. |

|  |
| --- |
| eTable 2. Logistic regression of dementia, non-parametric year trend. |
| Variable | Model 1 | Model 2 | Model 3 | Model 4 |
|  | OR | 95% CI | OR | 95% CI | OR | 95% CI | OR | 95% CI |
| Year (ref: 1995-1999) |  |  |  |  |  |  |  |
| 2000-2004 | 0.96 | (0.91, 1.0) | 1.2 | (1.1, 1.4) | 1.1 | (0.98, 1.3) | 1.2 | (1.1, 1.4) |
| 2005-2009 | 0.98 | (0.90, 1.1) | 1.4 | (1.2, 1.6) | 1.2 | (1.0, 1.4) | 1.4 | (1.2, 1.6) |
| 2010-2014 | 0.87 | (0.80, 0.95) | 1.2 | (1.0, 1.5) | 0.78 | (0.65, 0.93) | 1.0 | (0.85, 1.2) |
| Person-waves | 180,162 | 180,162 | 180,162 | 180,162 |
| Model 1 adjusts for age, age squared, sex, race/ethnicity, proxy status, and proxy\*sex.Model 2 adjusts for age, age squared, sex, race/ethnicity, proxy status, proxy\*sex, interview number, and proxy\*interview number.Model 3 adjusts for age, age squared, sex, race/ethnicity, proxy status, proxy\*sex, interview number, proxy\*interview number, and total interviews.Model 4 adjusts for age, age squared, sex, race/ethnicity, proxy status, proxy\*sex, interview number, proxy\*interview number, total interviews, and education. |

|  |
| --- |
| eTable 3. Logistic regression of dementia among self-respondents. |
| Variable | Model 1 | Model 2 | Model 3 | Model 4 |
|  | OR | 95% CI | OR | 95% CI | OR | 95% CI | OR | 95% CI |
| Time (Unit: 10 years) | 0.77 | (0.71, 0.83) | 1.0 | (0.87, 1.2) | 0.73 | (0.64, 0.83) | 0.94 | (0.83, 1.1) |
| Person-waves | 166,113 | 166,113 | 166,113 | 166,113 |
| Model 1 adjusts for age, age squared, sex, and race/ethnicity.Model 2 adjusts for age, age squared, sex, race/ethnicity, and interview number.Model 3 adjusts for age, age squared, sex, race/ethnicity, interview number, and total interviews.Model 4 adjusts for age, age squared, sex, race/ethnicity, interview number, total interviews, and education. |
| eTable 4. Logistic regression of dementia among proxy-respondents. |
| Variable | Model 1 | Model 2 | Model 3 | Model 4 |
|  | OR | 95% CI | OR | 95% CI | OR | 95% CI | OR | 95% CI |
| Time (Unit: 10 years) | 1.3 | (1.2, 1.5) | 0.78 | (0.60, 1.0) | 0.61 | (0.48, 0.78) | 0.62 | (0.49, 0.78) |
| Person-waves | 14,049 | 14,049 | 14,049 | 14,049 |
| Model 1 adjusts for age, age squared, sex, and race/ethnicity.Model 2 adjusts for age, age squared, sex, race/ethnicity, and interview number.Model 3 adjusts for age, age squared, sex, race/ethnicity, interview number, and total interviews.Model 4 adjusts for age, age squared, sex, race/ethnicity, interview number, total interviews, and education. |

\* Mark Lee, Andrew Halpern-Manners, and John Robert Warren

\* Comment on Hale et al

\* Code for Data Cleaning and Analysis

\* To successfully run this code, three data files are required:

\* 1. The RAND longitudinal data file, version P (titled rndhrs\_p.dta)

\* 2. The HRS tracker file (titled trk2018tr\_r.dta)

\* 3. The Langa-Weir cognitive data file (titled cogfinalimp\_9516wide.dta)

\*

\* All data files are publicly available on the HRS website.

\* Load data

cd "" // Insert file path to folder where data are stored between quotation marks

set maxvar 15000

use rndhrs\_p.dta, clear

drop hhidpn // this is the numeric id

rename rahhidpn hhidpn // rahhidpn is the string id I need to merge with the cognition data

\* Select data

keep hhidpn r\*wtcrnh r\*agey\_e hacohort raracem rahispan ragender raeduc

\* Rename weight variables

gen rwt1992 = r1wtcrnh

gen rwt1993 = r2wtcrnh

replace rwt1993 = . if hacohort>=2

gen rwt1994 = r2wtcrnh

replace rwt1994 = . if hacohort<2

gen rwt1995 = r3wtcrnh

replace rwt1995 = . if hacohort>=2

gen rwt1996 = r3wtcrnh

replace rwt1996 = . if hacohort<2

gen rwt1998 = r4wtcrnh

gen rwt2000 = r5wtcrnh

gen rwt2002 = r6wtcrnh

gen rwt2004 = r7wtcrnh

gen rwt2006 = r8wtcrnh

gen rwt2008 = r9wtcrnh

gen rwt2010 = r10wtcrnh

gen rwt2012 = r11wtcrnh

gen rwt2014 = r12wtcrnh

drop r\*wtcrnh

\* Rename age variables

gen age1992 = r1agey\_e

gen age1993 = r2agey\_e

replace age1993 = . if hacohort>=2

gen age1994 = r2agey\_e

replace age1994 = . if hacohort<2

gen age1995 = r3agey\_e

replace age1995 = . if hacohort>=2

gen age1996 = r3agey\_e

replace age1996 = . if hacohort<2

gen age1998 = r4agey\_e

gen age2000 = r5agey\_e

gen age2002 = r6agey\_e

gen age2004 = r7agey\_e

gen age2006 = r8agey\_e

gen age2008 = r9agey\_e

gen age2010 = r10agey\_e

gen age2012 = r11agey\_e

gen age2014 = r12agey\_e

drop r\*agey\*

drop hacohort

\* Recode demographic variables

gen race\_eth=.

replace race\_eth=1 if raracem==1 & rahispan==0

replace race\_eth=2 if raracem==2 & rahispan==0

replace race\_eth=3 if rahispan==1

replace race\_eth=4 if raracem==3 & rahispan==0

label define racecat 1 "white" 2 "black" 3 "hispanic" 4 "other"

label values race\_eth racecat

drop raracem rahispan

recode ragender (1=1 "male")(2=0 "female")(.m=.), gen(male)

recode raeduc(1=1 "<HS")(2/3=2 "HS")(4/5=3 "Coll")(.m=.), gen(edu)

drop raeduc

\* Merge with tracker file

merge 1:1 hhidpn using trk2018tr\_r.dta

keep if \_merge==3

\* Rename IWMODE IWWAVE and IWYEAR variables

rename aiwmode iwmode1992

rename aiwwave iwwave1992

rename aiwyear iwyear1992

rename biwmode iwmode1993

rename biwwave iwwave1993

rename biwyear iwyear1993

rename ciwmode iwmode1994

rename ciwwave iwwave1994

rename ciwyear iwyear1994

rename diwmode iwmode1995

rename diwwave iwwave1995

rename diwyear iwyear1995

rename eiwmode iwmode1996

rename eiwwave iwwave1996

rename eiwyear iwyear1996

rename fiwmode iwmode1998

rename fiwwave iwwave1998

rename fiwyear iwyear1998

rename giwmode iwmode2000

rename giwwave iwwave2000

rename giwyear iwyear2000

rename hiwmode iwmode2002

rename hiwwave iwwave2002

rename hiwyear iwyear2002

rename jiwmode iwmode2004

rename jiwwave iwwave2004

rename jiwyear iwyear2004

rename kiwmode iwmode2006

rename kiwwave iwwave2006

rename kiwyear iwyear2006

rename liwmode iwmode2008

rename liwwave iwwave2008

rename liwyear iwyear2008

rename miwmode iwmode2010

rename miwwave iwwave2010

rename miwyear iwyear2010

rename niwmode iwmode2012

rename niwwave iwwave2012

rename niwyear iwyear2012

rename oiwmode iwmode2014

rename oiwwave iwwave2014

rename oiwyear iwyear2014

keep hhidpn stratum secu rwt\* age\* male race\_eth edu iwmode\* iwwave\* iwyear\*

recode iwyear1998(9998=1998)

\* Merge with Langa-Weir data

merge 1:1 hhidpn using cogfinalimp\_9516wide.dta

rename cogtot27\_imp\* cogscore\*

\* Interview number

\* The variable int\_yearYYYY records how many interviews the respondent had completed on or before wave YYYY

recode iwwave\*(0=.)

egen int\_year1995 = rownonmiss(iwwave1995)

egen int\_year1996 = rownonmiss(iwwave1995 iwwave1996)

egen int\_year1998 = rownonmiss(iwwave1995 iwwave1996 iwwave1998)

egen int\_year2000 = rownonmiss(iwwave1995 iwwave1996 iwwave1998 iwwave2000)

egen int\_year2002 = rownonmiss(iwwave1995 iwwave1996 iwwave1998 iwwave2000 iwwave2002)

egen int\_year2004 = rownonmiss(iwwave1995 iwwave1996 iwwave1998 iwwave2000 iwwave2002 iwwave2004)

egen int\_year2006 = rownonmiss(iwwave1995 iwwave1996 iwwave1998 iwwave2000 iwwave2002 iwwave2004 iwwave2006)

egen int\_year2008 = rownonmiss(iwwave1995 iwwave1996 iwwave1998 iwwave2000 iwwave2002 iwwave2004 iwwave2006 iwwave2008)

egen int\_year2010 = rownonmiss(iwwave1995 iwwave1996 iwwave1998 iwwave2000 iwwave2002 iwwave2004 iwwave2006 iwwave2008 iwwave2010)

egen int\_year2012 = rownonmiss(iwwave1995 iwwave1996 iwwave1998 iwwave2000 iwwave2002 iwwave2004 iwwave2006 iwwave2008 iwwave2010 iwwave2012)

egen int\_year2014 = rownonmiss(iwwave1995 iwwave1996 iwwave1998 iwwave2000 iwwave2002 iwwave2004 iwwave2006 iwwave2008 iwwave2010 iwwave2012 iwwave2014)

replace int\_year1995=. if iwwave1995==.

replace int\_year1996=. if iwwave1996==.

replace int\_year1998=. if iwwave1998==.

replace int\_year2000=. if iwwave2000==.

replace int\_year2002=. if iwwave2002==.

replace int\_year2004=. if iwwave2004==.

replace int\_year2006=. if iwwave2006==.

replace int\_year2008=. if iwwave2008==.

replace int\_year2010=. if iwwave2010==.

replace int\_year2012=. if iwwave2012==.

replace int\_year2014=. if iwwave2014==.

\* Total interviews completed

\* The variable totalint records how many interviews the respondent comppleted between 1995 and 2014.

egen totalint = rowmax(int\_year1995-int\_year2014)

keep hhidpn secu stratum rwt\* age\* male race\_eth edu iwwave\* iwmode\* iwyear\* cogfunction\* cogscore\* proxy\* int\_year\* totalint

drop cogfunction2016 cogscore2016 proxy2016

\* Reshape data

reshape long age iwwave iwmode iwyear rwt cogfunction cogscore proxy int\_year, i(hhidpn) j(year)

\* Time-dependent covariates

replace age=age-50

gen agesq = age\*age

gen year\_ctr=iwyear-2000

recode proxy(1/2=1)(5=0)

recode iwmode(1=0 "face to face")(2=1 "phone")(8=.), gen(phone)

recode int\_year(1=1 "1")(2=2 "2")(3/4=3 "3-4")(5/7=4 "5-7")(8/10=5 "8+"), gen(intnum)

\* total interviews

recode totalint(1=1 "1")(2=2 "2")(3/4=3 "3-4")(5/7=4 "5-7")(8/10=5 "8+"), gen(intever)

\* Dementia status

recode cogfunction(1/2=0)(3=1), gen(dem)

\* Restrict sample

\* We restrict the sample to relevant years and ages and drop individuals with missing data on covariates.

keep if iwwave==1

drop if year<1995

drop if year>2014

drop if dem==.

drop if age<0

drop if age==.

drop if race\_eth==.

drop if male==.

drop if edu==.

drop if phone==.

drop if proxy==.

drop if rwt==0 | rwt==.

\* Person identifier

sort hhidpn year

egen justone=tag(hhidpn)

\* Baseline Characteristics for table 1

gen realage = age+50

summ realage [aw=rwt] if justone==1

tab male [aw=rwt] if justone==1

tab race\_eth [aw=rwt] if justone==1

tab edu [aw=rwt] if justone==1

\* Recode age and year variables

gen year10 = year\_ctr/10

gen age10 = age/10

gen age10sq = age10\*age10

\* Declare survey design

svyset secu, strata(stratum) weight(rwt) vce(linearized) singleunit(missing)

\* Regressions with linear year trend

svy linearized : logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.proxy i.proxy#i.male

margins, dydx(year10) // marginal effects

svy linearized : logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.proxy i.proxy#i.male i.intnum i.proxy#i.intnum

margins, dydx(year10) // marginal effects

svy linearized : logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.proxy i.proxy#i.male i.intnum i.proxy#i.intnum i.intever

margins, dydx(year10) // marginal effects

svy linearized : logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.proxy i.proxy#i.male i.intnum i.proxy#i.intnum i.intever i.edu

margins, dydx(year10) //marginal effects

\* Regressions with non-parametric year trend

recode iwyear(1995/1999=1 "1995-99")(2000/2004=2 "2000-04")(2005/2009=3 "2005-09")(2010/2015=4 "2010-14"), gen(yearcat)

svy linearized : logistic dem i.yearcat i.race\_eth i.male age10 age10sq i.phone i.proxy i.proxy#i.male

svy linearized : logistic dem i.yearcat i.race\_eth i.male age10 age10sq i.phone i.proxy i.proxy#i.male i.intnum i.proxy#i.intnum

svy linearized : logistic dem i.yearcat i.race\_eth i.male age10 age10sq i.phone i.proxy i.proxy#i.male i.intnum i.proxy#i.intnum i.intever

svy linearized : logistic dem i.yearcat i.race\_eth i.male age10 age10sq i.phone i.proxy i.proxy#i.male i.intnum i.proxy#i.intnum i.intever i.edu

\* Regressions for self-respondents

recode proxy(0=1)(1=0), gen(self)

svy linearized, subpop(self): logistic dem year10 i.race\_eth i.male age10 age10sq i.phone

svy linearized, subpop(self): logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.intnum

svy linearized, subpop(self): logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.intnum i.intever

svy linearized, subpop(self): logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.intnum i.intever i.edu

\* Regressions for proxies

svy linearized, subpop(proxy): logistic dem year10 i.race\_eth i.male age10 age10sq i.phone

svy linearized, subpop(proxy): logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.intnum

svy linearized, subpop(proxy): logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.intnum i.intever

svy linearized, subpop(proxy): logistic dem year10 i.race\_eth i.male age10 age10sq i.phone i.intnum i.intever i.edu