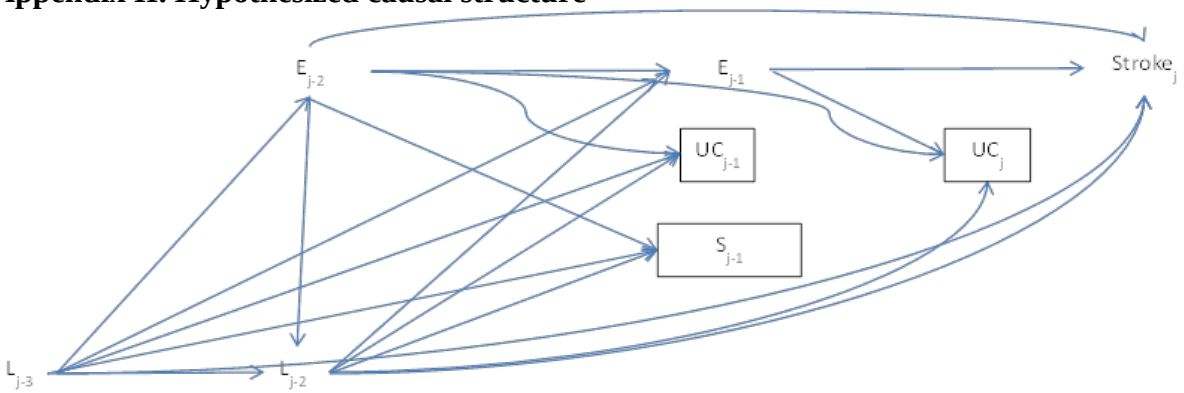
SUPPLEMENTAL MATERIAL

Appendix I. Baseline characteristics of HRS participants included vs excluded from our sample due to missing values at baseline, HRS 1998 (n= 18,766 participants)

Characteristics	Missing N (%) Mean (Std)	Included N (%) Mean (Std)d	Significant difference
Male	1,388 (53.6)	6,712 (41.5)	*
Race/Ethnicity	1,760 (68.1)	12,655 (78.2)	
Non-Hispanic white	474 (18.4)	2,075 (12.9)	
Non-Hispanic black	280 (10.8)	1,151 (7.1)	
Hispanic	69 (2.7)	293 (1.8)	
"Other" race			
Age (years)	68.1 (11.2)	65.9 (9.7)	*
Married	1,687 (65.7)	10,701 (66.1)	
Income/household members			*
>42,988	452 (1.76)	4061 (25.2)	
$42,987 - $23,918	502 (19.6)	4050 (25.2)	
$23,917 - $12,936	636 (24.8)	4048 (25.2)	
<$12,935	973 (38.0)	3932 (24.4)	
Wealth/household members			*
> $253,145	423 (21.7)	4050 (25.9)	
$253,144– $105,509	474 (24.3)	4055 (25.9)	
$105,508– $35,356	532 (27.3)	4059 (26.0)	
<$35,355	522 (26.8)	3474 (22.2)	
Years of education	10.7 (3.9)	12.2 (3.2)	*
CES-D score (continuous)	2.2 (2.1)	1.5 (1.9)	*
CES-D score >=3	382 (14.8)	3669 (22.7)	*
Obese	471 (18.2)	3790 (23.4)	*
Smoker	456 (17.6)	2636 (16.3)	
Elevated blood pressure	1,289 (49.8)	7294 (45.1)	*
Diabetes	444 (17.2)	2117 (13.1)	*
Heart disease	494 (22.76)	3076 (19.0)	*

*Significant difference at $p<0.05$

Appendix II. Hypothesized causal structure



$Stroke_j$: Stroke reported at outcome wave *j*
E_{j-1}: Second exposure wave depressive symptom category associated with outcome at wave *j*
E_{j-2}: First exposure wave depressive symptom associated with outcome at wave *j*
UC_j: Uncensored at outcome wave *j*
S_j: Survival at outcome wave *j*-1
L_{j-2}: Measured time-varying confounders/mediators at outcome wave *j*-2
L_{j-3}: Measured time-varying confounders/mediators at outcome wave *j*-3

Appendix III. Details regarding inverse probability weight construction

The stabilized inverse probability weight $SIPW_{ij}$ for individual *i* at outcome wave *j* (outcomes were assessed at waves 3 to 8) was the product of the inverse probability of survival weight ($IPSW_{ij}$), the inverse probability of exposure to depressive symptoms weight ($IPEW_{ij}$), and the inverse probability of remaining uncensored weight ($IPUCW_{ij}$). At each outcome wave strokes were recorded regardless of whether the participant was alive or dead as long as a proxy participated in the interview wave. Assessment of depressive symptoms, however, occurred only for living and participating respondents, so observations for outcome wave j required the participant to be alive at wave *j*-1 and to be in the study at wave *j*-1 to estimate the inverse probability of exposure. For a participant to be in our sample at outcome wave j, and therefore uncensored at outcome wave j, required the participant to be alive and in the study at wave *j*-1, provided depressive symptoms scores at waves *j*-1 and *j*-2, and had a stroke outcome assessment at wave j.

The $IPSW_{i\bar{j}}$ was estimated as the probability that individual *i* survived through wave *j*-1 given that individual *i* participated in wave *j*-1 and given individual *i*'s depressive symptoms and covariate values at time *j*-2 (Equation 1). The $IPEW_{i\bar{j}}$ was estimated as the probability that individual *i* had elevated depressive symptoms at wave *j*-1 given that individual *i* participated and survived up to wave *j*-1, and given individual *i*'s depressive symptoms and covariate values at time *j*-2 (Equation 2). The $IPUCW_{i\bar{j}}$ for each outcome wave *j* was estimated as the probability that individual *i* remained in the study (via self or proxy interviews) through wave *j*, given that individual *i* participated and survived up to wave j-1, provided depressive symptoms scores at waves *j*-1 and *j*-2, had a stroke outcome assessment at wave j, and given individual *i*'s depressive symptoms and covariate values at time *j*-1 and *j*-2 (Equation 3). Each weight was accumulated across all prior waves *j*.

Therefore $SIPW_{i\bar{j}}$= $IPSW_{i\bar{j}}$X $IPEW_{i\bar{j}}$ X $IPUCW_{i\bar{j}}$ with each component defined as below. Where *S* is survival, *E* is exposure status (i.e., elevated depressive symptoms or not), *UC* is being uncensored, *k* indexes the interview wave, V is a vector of time-constant baseline covariates, M is a vector of time-varying missing values status on depressive symptoms score or stroke, *L* is a vector of time-varying covariates from the first (*j*-2) or second (*j*-1) exposure waves.

(1) $$IPSW_{i\bar{j}} = \Pi_{k=3}^{j} \frac{Pr[S_{ik-1}|E_{ik-2},V_{i0},UC_{ik-1}=1,S_{ik-2}=1\ ]}{Pr[S_{ik-1}|E_{ik-2},L_{ik-2},UC_{ik-1}=1,S_{ik-2}=1]}$$

(2) $$IPEW_{i\bar{j}} = \Pi_{k=3}^{j} \frac{Pr[E_{ik-1}|E_{ik-2},V_{i0},UC_{ik-1}=1,S_{ik-1}=1]}{Pr[E_{ik-1}|E_{ik-2},L_{ik-2},UC_{ik-1}=1,S_{ik-1}=1]}$$

(3) $$IPUCW_{i\bar{j}} = \Pi_{k=3}^{j} \frac{Pr[UC_{ik}|E_{ik-2},E_{ik-1},V_{i0},UC_{ik-1}=1,S_{ik-1}=1,M_{ik-1}=0,M_{ik-2}=0,M_{ik}=0]}{Pr[UC_{ik}|E_{ik-2},E_{ik-1},L_{ik-2},L_{ik-1},UC_{ik-1}=1,S_{ik-1}=1,M_{ik-1}=0,M_{ik-2}=0,M_{ik}=0]}$$

We estimated the numerator and denominator of the weights with pooled logistic regressions. We considered incident strokes starting in 2000 and set the corresponding exposure wave 1 (i.e., 1996) values of depressive symptoms and covariates to zero. Since we required participants to be alive and participating up to that year, the corresponding weight participation and survival weights were set to 1; the estimation of these weights for future waves did not included data from 2000. Both sets of models adjusted for the following baseline covariates: sex, age at enrollment, race, ethnicity, and education. Additionally, models estimating the denominator included time-varying covariates. To avoid collinearity, a subset of covariates was selected from a large pool of plausible confounders by an automated forward stepwise selection process

including all possible time-varying confounders using an entry and staying criteria of p=0.2 (Results shown in Appendix Table IV). Models were required to contain the previously described baseline covariates and depressive symptoms level from past exposure waves. Covariate values were obtained from exposure wave 1 (*j*-2) except when calculating the IPUCW, which also included values from exposure wave 2 (*j*-1). The analytic model estimates the hazard of stroke at outcome wave *j* using only individuals who survived to time *j* -1 and participated until time *j* and accounts for their history of confounders. Individuals who were included in the estimation of the person time specific IPSW but who passed away at that time point had missing values for their IPEW at that wave and their IPUCW the following wave. This resulted in missing values for final weights (SIPW*sample weights) and the exclusion of this observation in the MSM model sample.

SAS code:

```
*time stable baseline covariates;
%let demo_98=male b_ageyr b_ageyr_sq RAEDYRS HS PostHSyrs College nhblack Hispanic
other; run;

/************************************************************
Estimating numerator probabilities and sorting
************************************************************/

*Survival at exposure wave 2 (i-1): Pr (SurvivedEW2|cesdDew1, demo_98, partEW2=1);
proc genmod descending data=hrsipw;
where partEW2=1 and STKwave ne 3;
class STKwave HHIDPN ;
model SurvivedEW2=cesdDew1 &demo_98/dist=binomial link=logit;
repeated subject=hhidpn/ type=un;
output out=SurvivedEW2_num (keep= hhidpn stkwave p_SurvivedEW2_num)
p=p_SurvivedEW2_num;
run;

*Treatment at exposure wave 2 (i-1): Pr(treatedEW2|cesdDew1, demo_98, partEW2=1,
SurvivedEW2=1);
proc genmod descending data=hrsipw;
where partEW2=1 and survivedEW2=1;
class STKwave HHIDPN;
model treatedEW2=cesdDew1 &demo_98 /dist=binomial link=logit;
repeated subject=hhidpn/ type=un;
output out=treatedEW2_num (keep= hhidpn stkwave p_treatedEW2_num)
p=p_treatedEW2_num;
run;

*Participation at outcome wave (i): Pr(partOW|cesdDew1, treatedEW2, demo_98, partEW2=1,
SurvivedEW2=1, stkmiss=0, cesdCFew1=0, cesdCFew2=0);
proc genmod descending data=hrsipw;
```

```
where partEW2=1 and survivedEW2=1 and stkmiss=0 and cesdCFew1=0 and cesdCFew2=0
and STKwave ne 3;
class STKwave HHIDPN;
model partOW=cesdDew1 treatedEW2 &demo_98/dist=binomial link=logit;
repeated subject=hhidpn/ type=un;
output out=partout_num (keep= hhidpn stkwave p_partout_num) p=p_partout_num;
run;

proc sort data=SurvivedEW2_num;        by hhidpn STKwave; run;
proc sort data=treatedEW2_num;         by hhidpn STKwave; run;
proc sort data=partout_num;            by hhidpn STKwave; run;


%let demo_98=male b_ageyr b_ageyr_sq RAEDYRS HS PostHSyrs College nhblack Hispanic
other; run;

%let timevarallEW1=r_ageyrEW1 r_ageyrEW1_sq r_marriedEW1 incomecapEW1_qt
wlthcapEW1_qt r_antidepdEW1 r_drinkdEW1 r_smknowEW1r_obeseEW1 r_heartdEW1
r_hibpdEW1 r_diabdEW1; run;

%let timevarallEW12=&timevarallEW1 r_ageyrEW2 r_ageyrEW2_sq r_marriedEW2
incomecapEW2_qt wlthcapEW2_qt r_antidepdEW2 r_drinkdEW2 r_smknowEW2r_obeseEW2
r_heartdEW2 r_hibpdEW2 r_diabdEW2; run;


/*****************************************************************
Estimating denominator probabilities and sorting
*****************************************************************/
proc logistic data=hrsipw;
where partEW2=1 and STKwave ne 3;
class HHIDPN incomecapEW1_qt wlthcapEW1_qt;
model survivedEW2= cesdDew1 &demo_98 &timevarallEW1
/selection=stepwise slentry=.2 slstay = .2 include=11;
run;

*Survival at exposure wave 2 (i-1): Pr (SurvivedEW2=1|cesdDew1, demo_98, timecovariates,
partEW2=1);
proc genmod descending data=hrsipw;
where partEW2=1 and STKwave ne 3;
class STKwave hhidpn incomecapEW1_qt wlthcapEW1_qt;
model SurvivedEW2= cesdDew1 &demo_98 r_ageyrEW1_sq r_marriedEW1
incomecapEW1_qt wlthcapEW1_qt r_antidepdEW1 r_drinkdEW1 r_smknowEW1 r_obeseEW1
r_heartdEW1 r_hibpdEW1 r_diabdEW1/dist=binomial link=logit;
repeated subject=hhidpn/ type=un;
output out=survivedEW2_denom  (keep= hhidpn stkwave p_SurvivedEW2_denom)
p=p_SurvivedEW2_denom;
```

```
run;

proc logistic descending data=hrsipw;
where partEW2=1 and survivedEW2=1;
class HHIDPN incomecapEW1_qt wlthcapEW1_qt;
model treatedEW2= cesdDew1 &demo_98 &timevarallEW1
/selection=stepwise slentry=.2 slstay = .2 include=11;
run;

*Treatment at exposure wave 2 (i-1): Pr(treatedEW2|cesdDew1, demo_98, timecovariates,
partEW2=1, SurvivedEW2=1);
proc genmod descending data=hrsipw;
where partEW2=1 and survivedEW2=1;
class STKwave hhidpn incomecapEW1_qt wlthcapEW1_qt;
model treatedEW2= cesdDew1 &demo_98 r_ageyrEW1 r_ageyrEW1_sq r_marriedEW1
incomecapEW1_qt wlthcapEW1_qt r_antidepdEW1 r_drinkdEW1 r_smknowEW1 r_obeseEW1
r_heartdEW1 r_hibpdEW1 r_diabdEW1/dist=binomial link=logit;
repeated subject=hhidpn/ type=EXCH;
output out=treatedEW2_denom (keep= hhidpn stkwave p_treatedEW2_denom)
p=p_treatedEW2_denom;
run;

proc logistic descending data=hrsipw;
where partEW2=1 and survivedEW2=1 and stkmiss=0 and cesdCFew1=0 and cesdCFew2=0 and
STKwave ne 3;
class HHIDPN incomecapEW1_qt wlthcapEW1_qt incomecapEW2_qt wlthcapEW2_qt;
model partOW= cesdDew1 treatedEW2 &demo_98 &timevarallEW12
/selection=stepwise slentry=.2 slstay = .2 include=11;
run;

*Participation at outcome wave (i): Pr(partOW|cesdDew1, demo_98, timecovariates,
partEW2=1, SurvivedEW2=1, strokemissing=0, CESDCFew1=0, CESDCRew2=0,
treatedEW2);
proc genmod descending data=hrsipw;
where partEW2=1 and survivedEW2=1 and stkmiss=0 and cesdCFew1=0 and cesdCFew2=0 and
STKwave ne 3;
class STKwave hhidpn incomecapEW1_qt wlthcapEW1_qt incomecapEW2_qt wlthcapEW2_qt;
model partOW=cesdDew1 &demo_98 treatedEW2 r_marriedEW1 incomecapEW1_qt
wlthcapEW1_qt r_smknowEW1 r_diabdEW1 r_ageyrEW2 r_ageyrEW2_sq incomecapEW2_qt
wlthcapEW2_qt r_obeseEW2 /dist=binomial link=logit;
repeated subject=hhidpn/ type=un;
output out=partout_denom (keep= hhidpn stkwave p_partout_denom) p=p_partout_denom;
run;

proc sort data=survivedEW2_denom;        by hhidpn STKwave; run;
proc sort data=treatedEW2_denom;         by hhidpn STKwave; run;
```

```
proc sort data=partout_denom;                by hhidpn STKwave; run;


/*****************************************************************
Merging probabilities and creating weights
*****************************************************************/
proc sort data=hrsipw; by hhidpn STKwave; run;

data hrs_ipw_wtspart1;
merge hrsipw
SurvivedEW2_num          survivedEW2_denom
treatedEW2_num           treatedEW2_denom
partout_num              partout_denom;
by hhidpn STKwave;

if first.hhidpn=1 then firstobs=1;

if firstobs=1 then do;  p_SurvEW2_num=1;
        p_SurvEW2_denom=1;
        survEW2prb_s=1;
        survEW2prb_us=1;
        depEW2prb_s=1;
        depEW2prb_us=1;
        p_partout_num=1;
        p_partout_denom=1;
        partoutprb_s=1;
        partoutprb_us=1;
end;

*Estimate stabilized (s) and unstablized (us) weights for current wave (T) and multiple with prior
waves;
*treatment/depression at wave 2 (i-1);
if depsxEW2=1       then depEW2prb_sT=(p_depsxEW2_num/p_depsxEW2_denom);
if depsxEW2=0       then depEW2prb_sT=((1-p_depsxEW2_num)/(1-p_depsxEW2_denom));
if depsxEW2=1       then depEW2prb_usT=(1/(p_depsxEW2_denom));
if depsxEW2=0       then depEW2prb_usT=(1/(1-p_depsxEW2_denom));
depEW2prb_s=        depEW2prb_s*        depEW2prb_sT;
depEW2prb_us=       depEW2prb_us*       depEW2prb_usT;

if firstobs ne 1 then do;
*survival at exposure wave 2 (i-1);
if survivedEW2=1 then
survEW2prb_sT=(p_survivedEW2_num/p_survivedEW2_denom);*stabilized;
if survivedEW2=0 then survEW2prb_sT=((1-p_survivedEW2_num)/(1-
p_survivedEW2_denom));
if survivedEW2=1 then survEW2prb_usT=(1/p_survivedEW2_denom);*unstabilized;
```

```
if survivedEW2=0 then survEW2prb_usT=(1/(1-p_survivedEW2_denom));
survEW2prb_s=       survEW2prb_s*       survEW2prb_sT;
survEW2prb_us=      survEW2prb_us*      survEW2prb_usT;
retain survEW2prb_s  survEW2prb_us;
end;
*participation in outcome wave (i);
if partOW=1         then partoutprb_sT=(p_partout_num/p_partout_denom);
if partOW=0         then partoutprb_sT=((1-p_partout_num)/(1-p_partout_denom));
if partOW=1         then partoutprb_usT=(1/p_partout_denom);
if partOW=0         then partoutprb_usT=(1/(1-p_partout_denom));
partoutprb_s=       partoutprb_s*       partoutprb_sT;
partoutprb_us=      partoutprb_us*      partoutprb_usT;

retain  partoutprb_s partoutprb_us;
end;
run;

data hrs_ipw_wts;
set  hrs_ipw_wtspart1;
wt_combine_s=  survEW2prb_s      *depEW2prb_s       *partoutprb_s;
wt_combine_us= survEW2prb_us     *depEW2prb_us      *partoutprb_us;
run;
```

Appendix IV. Results from pooled logistic regression models for estimating the denominators of the inverse probability of survival (IPSW), the inverse probability of exposure weights (IPEW), and the inverse probability of participation weights (IPUCW)*

Variable	Model predicting survival at the second exposure wave (for IPSW estimate)				Model predicting exposure by elevated depressive symptoms at the second exposure wave (for IPEW estimates)				Model predicting remaining uncensored at the outcome wave (for IPUCW estimates)			
	OR	95% CI		P-value	OR	95% CI		P-value	OR	95% CI		P-value
Elevated depression score at exposure wave 1	0.58	0.53	0.63	<.0001	1.84	1.74	1.95	<.0001	1.02	0.90	1.15	0.80
Elevated depression score at exposure wave 2	--	--	--	--	--	--	--	--	0.93	0.82	1.05	0.23
Time-constant covariates:												
Baseline									0.99	0.90	1.09	0.86
Male	0.49	0.44	0.53	<.0001	0.69	0.65	0.73	<.0001	0.82	0.71	0.94	<0.01
Baseline age	1.00	0.95	1.05	0.96	0.90	0.87	0.93	<.0001	1.00	1.00	1.00	0.03
Baseline age squared	1.00	1.00	1.00	0.79	1.00	1.00	1.00	<.0001	0.95	0.92	0.99	0.01
Years of education	0.96	0.94	0.99	0.01	0.94	0.93	0.96	<.0001	1.22	1.04	1.44	0.02
High school degree	1.15	1.01	1.32	0.04	0.81	0.75	0.88	<.0001	1.10	1.03	1.18	<.01
Years of higher education	1.00	0.94	1.06	0.98	0.97	0.93	1.00	0.06	1.00	0.80	1.25	0.99
College degree	1.16	0.94	1.42	0.16	0.92	0.80	1.05	0.23	0.78	0.67	0.90	<0.01
Non-Hispanic white	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
Non-Hispanic black	1.23	1.07	1.41	0.01	1.30	1.17	1.45	<.0001	0.78	0.67	0.90	<0.01
Hispanic	1.35	1.11	1.65	0.01	1.34	1.10	1.62	<.0001	0.81	0.66	0.99	0.04
Self-identified "Other" race	1.45	1.03	2.05	0.03	1.84	1.74	1.95	<0.01	0.58	0.43	0.78	<0.01
Exposure wave 1												
Age (linear)	--	--	--	--	0.99	0.99	1.00	<0.01	--	--	--	--
Age (squared)	1.00	1.00	1.00	<.0001	1.00	1.00	1.00	<0.01	--	--	--	--
Marital Status	1.21	1.10	1.33	<0.01	1.04	0.98	1.10	0.23	1.14	1.03	1.28	0.02
Income per capita												
1st quartile	0.67	0.54	0.83	<0.01	1.16	1.04	1.28	0.01	0.81	0.63	1.05	0.11

2nd quartile	0.75	0.65	0.87	<0.01	1.21	1.13	1.29	<.0001	1.01	0.85	1.20	0.91
3rd quartile	0.85	0.74	0.97	0.01	1.08	1.02	1.14	0.01	0.90	0.78	1.03	0.13
4th quartile	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
Wealth per capita												
1st quartile	0.53	0.44	0.65	<.0001	1.16	1.04	1.28	<.0001	1.07	0.80	1.44	0.64
2nd quartile	0.69	0.60	0.78	<.0001	1.21	1.13	1.29	<.0001	1.28	1.05	1.57	0.01
3rd quartile	0.84	0.74	0.95	<0.01	1.08	1.02	1.14	0.41	1.16	1.00	1.35	0.06
4th quartile	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
Psychiatric medication	0.83	0.71	0.97	0.02	1.98	1.83	2.14	<.0001	--	--	--	--
Drinking	1.05	1.03	1.08	<.0001	0.99	0.98	1.00	0.14	--	--	--	--
Smoking	0.54	0.47	0.60	<.0001	1.29	1.21	1.39	<.0001	0.95	0.81	1.11	0.16
Obesity	1.29	1.15	1.44	<.0001	1.08	1.02	1.14	0.01	--	--	--	--
Heart disease	0.63	0.57	0.68	<.0001	1.33	1.26	1.41	<.0001	--	--	--	--
High blood pressure	0.94	0.86	1.02	0.13	1.08	1.03	1.14	<0.01	--	--	--	--
Diabetes	0.64	0.58	0.71	<.0001	1.12	1.05	1.19	<0.01	--	--	--	--
Exposure wave 2												
Age (linear)	--	--	--	--	--	--	--	--	1.33	1.16	1.52	<.0001
Age (squared)	--	--	--	--	--	--	--	--	1.00	1.00	1.00	<.0001
Income per capita												
1st quartile	--	--	--	--	--	--	--	--	1.14	0.94	1.39	0.19
2nd quartile	--	--	--	--	--	--	--	--	1.14	0.97	1.34	0.12
3rd quartile	--	--	--	--	--	--	--	--	1.18	1.02	1.36	0.03
4th quartile	--	--	--	--	--	--	--	--	ref	ref	ref	ref
Wealth per capita												
1st quartile	--	--	--	--	--	--	--	--	0.90	0.71	1.13	0.36
2nd quartile	--	--	--	--	--	--	--	--	0.81	0.67	0.98	0.03
3rd quartile	--	--	--	--	--	--	--	--	0.95	0.81	1.11	0.52
4th quartile	--	--	--	--	--	--	--	--	ref	ref	ref	ref
Obesity	--	--	--	--	--	--	--	--	1.23	1.10	1.38	<0.01

* Pool of possible confounders included the following time-updated variables for exposure wave 1 for all models and exposure wave 2 for IPUCW models: age at interview (yrs), income per capita quartile, wealth per capita quartile, number of days/week respondent

drinks alcohol (0-7 days), current smoking status (yes/no), current psychiatric medication use (yes/no), obesity (BMI>30), and self-report of ever being diagnosed with heart disease, high blood pressure, or diabetes (yes/no for each condition).

Appendix V. Descriptive statistics of the stabilized combined inverse probability weight trimmed at the 99% percentile stratified by wave and overall

Outcome wave	N	Mean	Std Dev	Minimum	Maximum
3	16,178	1.00	0.05	0.78	1.08
4	14,147	0.97	0.16	0.29	1.74
5	12,261	0.96	0.20	0.12	1.74
6	10,914	0.95	0.25	0.06	1.74
7	9,746	0.94	0.28	0.02	1.74
8	8,663	0.93	0.31	0.01	1.74
Overall	71,909	0.96	0.21	0.01	1.74

Appendix VI. Histogram of the stabilized combined inverse probability weight trimmed at the 99% percentile