WEB MATERIAL

Uncovering Heterogeneous Associations Between Disaster-Related Trauma and Subsequent Functional Limitations: A Machine-Learning Approach

Koichiro Shiba, PhD*; Adel Daoud, PhD; Hiroyuki Hikichi, PhD; Aki Yazawa, PhD; Jun Aida, PhD; Katsunori Kondo, PhD; Ichiro Kawachi, PhD

* Corresponding author. Email: shiba_k@g.harvard.edu

Table of Contents

Web Appendix 1. R code for the main analyses.

```
library(tidyverse)
library(patchwork)
library(ggsci)
library(broom)
library(gtsummary)
library(grf)
library(ltmle)
library(SuperLearner)
library(sandwich)
```

1. Setup

```
df.imputed.2013 <- read.csv("df.imputed.2013.csv")
df.imputed.2016 <- read.csv("df.imputed.2016.csv")


demographic.factors  <- c("age_10_meibo","sex_10_meibo","mari5st10","hous2lo1
0")
ses.factors          <- c("educ5_10","empl3pt10","s_eqincome_x10")

health.factors       <- c("srh_4_10","bmi_10","adl_3_10","iadl10","iadl10.ins
trument","iadl10.cognitive","iadl10.social","s_gds_x10",
                          "dgns2ca10","dgns2rd10","dgns2st10","dgns2ht10","dg
ns2dm10","dgns2ob10","dgns2hl10","dgns2op10","dgns2an10",
                          "dgns2fx10","dgns2ge10","dgns2lv10","dgns2ps10","dg
ns2vi10","dgns2he10",
                          "dgns2bw10","dgns2sl10","dgns2uk10","dgns2ot10")


psychosocial.factors <- c("fot_5_10","cmn5cntr10","cmn5at10","cmnt6sp10","cmn
t6hb10","meet6fr10","num5fr10","lsnd2no10","lstn2no10","card2no10","wcar2no10
","acqu4_10","soc10","negative.events")
behavioral.factors   <- c("smok4_10","alcl3_10","fq7prt10","fq7veg10","gout6f
q10","hoby2_10")

treatment            <- c("hloss")
outcomes             <- c("iadl13","iadl16","adl3ra13","adl3ra16","svy_physic
al_dis13","svy_physical_dis16")

covariates <- c(demographic.factors,ses.factors,health.factors,psychosocial.f
actors,behavioral.factors)
```

2. TMLE with SuperLearner for ATEs

```
SL.library <- c("SL.glm","SL.xgboost","SL.nnet") # specify candicate estimato
rs for SuperLearner

# function to extract estimates from the output object of ltmle
extract.estimates.ltmle <- function(fit){
```

```r
  output.TMLE <- summary(fit)
  output.TMLE <- output.TMLE$effect.measures
  ATE.TMLE    <- output.TMLE$ATE %>% as.data.frame()
  RR.TMLE     <- output.TMLE$RR %>% as.data.frame()
  treatment   <- output.TMLE$treatment %>% as.data.frame()
  control     <- output.TMLE$control %>% as.data.frame()
  summary.TMLE<- rbind(ATE.TMLE,RR.TMLE,treatment,control)
  summary.TMLE <- summary.TMLE %>% filter(long.name == "Additive Treatment Ef
fect")
  return(summary.TMLE)
}


# Outcomes in 2013
## ADL
ate.ltmle.2013.adl <- df.imputed.2013 %>%
  select(covariates,hloss,adl3ra13) %>%
  ltmle(
    Anodes = "hloss",
    Ynodes = "adl3ra13",
    abar = list(1,0),
    SL.library = SL.library
  )
## IADL
ate.ltmle.2013.iadl <- df.imputed.2013 %>%
  select(covariates,hloss,iadl13) %>%
  ltmle(
    Anodes = "hloss",
    Ynodes = "iadl13",
    abar = list(1,0),
    SL.library = SL.library
  )
## Certified physical disability levels
ate.ltmle.2013.ltc <- df.imputed.2013 %>%
  select(covariates,hloss,svy_physical_dis13) %>%
  ltmle(
    Anodes = "hloss",
    Ynodes = "svy_physical_dis13",
    abar = list(1,0),
    SL.library = SL.library
  )

# Create a dataframe that contains all ATE estimates
ATE.estimates.2013 <- rbind(
  ate.ltmle.2013.adl %>% extract.estimates.ltmle() %>% mutate(outcome = "ADL"
, year = 2013),
  ate.ltmle.2013.iadl %>% extract.estimates.ltmle() %>% mutate(outcome = "IAD
L", year = 2013),
  ate.ltmle.2013.ltc %>% extract.estimates.ltmle() %>% mutate(outcome = "Phys
ical Disability", year = 2013)
```

```
) %>%
  rename(CI.low = CI.2.5., CI.up = CI.97.5.) %>%
  select(outcome,year,estimate,CI.low,CI.up,pvalue) %>%
  mutate_if(is.numeric,round,3)

# Outcomes in 2016

## ADL
ate.ltmle.2016.adl <- df.imputed.2016 %>%
  select(covariates,hloss,adl3ra16) %>%
  ltmle(
    Anodes = "hloss",
    Ynodes = "adl3ra16",
    abar = list(1,0),
    SL.library = SL.library
  )

## IADL
ate.ltmle.2016.iadl <- df.imputed.2016 %>%
  select(covariates,hloss,iadl16) %>%
  ltmle(
    Anodes = "hloss",
    Ynodes = "iadl16",
    abar = list(1,0),
    SL.library = SL.library
  )

## Certified physical disability levels
ate.ltmle.2016.ltc <- df.imputed.2016 %>%
  select(covariates,hloss,svy_physical_dis16) %>%
  ltmle(
    Anodes = "hloss",
    Ynodes = "svy_physical_dis16",
    abar = list(1,0),
    SL.library = SL.library
  )

# Create a dataframe that contains all ATE estimates
ATE.estimates.2016 <- rbind(
  ate.ltmle.2016.adl %>% extract.estimates.ltmle() %>% mutate(outcome = "ADL"
, year = 2016),
  ate.ltmle.2016.iadl %>% extract.estimates.ltmle() %>% mutate(outcome = "IAD
L", year = 2016),
  ate.ltmle.2016.ltc %>% extract.estimates.ltmle() %>% mutate(outcome = "Phys
ical Disability", year = 2016)
) %>%
  rename(CI.low = CI.2.5., CI.up = CI.97.5.) %>%
  select(outcome,year,estimate,CI.low,CI.up,pvalue) %>%
  mutate_if(is.numeric,round,3)
```

2.1. Figure 2

```
# Prepare a dataframe for plotting
df.p =
  bind_rows(
  ATE.estimates.2013 %>% mutate(year = "2013"), # change the variable "year"
to a character string
  ATE.estimates.2016 %>% mutate(year = "2016")
) %>%
  mutate(
    outcome = factor(outcome, levels = c("IADL","ADL","Physical Disability"))
  )

# Create a vector of CI results
ci_str <- c()
for (i in seq_along(df.p$CI.low)){
  ci_str <- c(ci_str, str_interp("(${format(round(df.p$CI.low[i],2),nsmall =2
)}, ${format(round(df.p$CI.up[i],2),nsmall = 2)})"))
}


df.p %>%
  ggplot(aes(x = outcome, y = estimate, ymin = CI.low, ymax = CI.up, shape =
year)) +
  geom_errorbar(position = position_dodge(-0.4),
                width = 0.2) +
  geom_point(position = position_dodge(-0.4),
             size = 2.5) +
  geom_hline(yintercept = 0,
             linetype = "dashed") +
  coord_flip(clip = "off") +
  ylab("Estimated Popualtion Average Effect of Home Loss") +
  xlab("") +
  scale_shape_discrete(name = "Year") +
  theme_classic() +
  theme(legend.position = "bottom",
        axis.line.y = element_line(color = "transparent"),
        axis.text.y = element_text(hjust = 0,size = 12,color="black"),
        text = element_text(color = "black",size = 12, family = "serif"),
        axis.ticks.y = element_blank(),
        axis.text.x = element_text(size = 12, color = "black"),
        legend.text = element_text(size = 12, color = "black"),
        plot.margin = unit(c(2.8,10,0,1),"lines")
        )
```

3. Generalized Random Forest for CATEs

```
# Reference: https://bookdown.org/halflearned/ml-ci-tutorial/hte-i-binary-tre
atment.html#via-grf
## Set up for cross-fitting
```

```
num.rankings <- 5
num.folds <- 20

n.2013 = df.imputed.2013 %>% nrow()
n.2016 = df.imputed.2016 %>% nrow()

folds.2013 <- sort(seq(n.2013) %% num.folds) + 1
folds.2016 <- sort(seq(n.2016) %% num.folds) + 1

## Outcomes in 2013
set.seed(777)

forest.iadl2013 <- causal_forest(df.imputed.2013[covariates], # covariate mat
rix
                       df.imputed.2013[,outcomes[1]], # outcome vector
                       df.imputed.2013[,treatment[1]], # exposure vector
                       num.trees = 2000, # grow 2000 trees
                       tune.parameters = "all",
                       clusters = folds.2013) # tune parameters

preds.iadl2013 <- predict(forest.iadl2013, estimate.variance = F) %>% # predi
ct CATEs
  mutate(outcome = "IADL", year = "2013")


forest.adl2013 <- causal_forest(df.imputed.2013[covariates],
                       df.imputed.2013[,outcomes[3]],
                       df.imputed.2013[,treatment[1]],
                       num.trees = 2000,
                       tune.parameters = "all",
                       clusters = folds.2013)

preds.adl2013 <- predict(forest.adl2013, estimate.variance = F) %>%
  mutate(outcome = "ADL", year = "2013")



forest.disability2013 <- causal_forest(df.imputed.2013[covariates],
                       df.imputed.2013[,outcomes[5]],
                       df.imputed.2013[,treatment[1]],
                       num.trees = 2000,
                       tune.parameters = "all",
                       clusters = folds.2013)

preds.disability2013 <- predict(forest.disability2013, estimate.variance = F)
 %>%
  mutate(outcome = "Physical Disability", year = "2013")

preds.2013 <- rbind(preds.adl2013,preds.iadl2013,preds.disability2013)
```

```
## Outcomes in 2016
forest.iadl2016 <- causal_forest(df.imputed.2016[covariates],
                       df.imputed.2016[,outcomes[2]],
                       df.imputed.2016[,treatment[1]],
                       num.trees = 2000,
                       tune.parameters = "all",
                       clusters = folds.2016)

preds.iadl2016 <- predict(forest.iadl2016, estimate.variance = F) %>%
  mutate(outcome = "IADL", year = "2016")


forest.adl2016 <- causal_forest(df.imputed.2016[covariates],
                       df.imputed.2016[,outcomes[4]],
                       df.imputed.2016[,treatment[1]],
                       num.trees = 2000,
                       tune.parameters = "all",
                       clusters = folds.2016)

preds.adl2016 <- predict(forest.adl2016, estimate.variance = F) %>%
  mutate(outcome = "ADL", year = "2016")

forest.disability2016 <- causal_forest(df.imputed.2016[covariates],
                       df.imputed.2016[,outcomes[6]],
                       df.imputed.2016[,treatment[1]],
                       num.trees = 2000,
                       tune.parameters = "all",
                       clusters = folds.2016)

preds.disability2016 <- predict(forest.disability2016, estimate.variance = F)
 %>%
  mutate(outcome = "Physical Disability", year = "2016")


preds.2016 <- rbind(preds.adl2016,preds.iadl2016,preds.disability2016)
```

3.1 Test for heterogeneity

3.1.1 Comparison of those who were above/below median CATEs

```
test.heterogeneity <- function(preds,data,exposure,outcome){
  tau.hat <- preds$predictions
  high_effect = tau.hat > median(tau.hat)

  ate.high = data %>%
    filter(high_effect) %>%
    select(covariates,exposure,outcome) %>%
    ltmle(
```

```
      Anodes = exposure,
      Ynodes = outcome,
      abar = list(1,0),
      SL.library = SL.library
    ) %>%
    extract.estimates.ltmle() %>%
    mutate(subset = "high")

  ate.low = data %>%
    filter(!high_effect) %>%
    select(covariates,exposure,outcome) %>%
    ltmle(
      Anodes = exposure,
      Ynodes = outcome,
      abar = list(1,0),
      SL.library = SL.library
    ) %>%
    extract.estimates.ltmle() %>%
    mutate(subset = "low")

  ates <- rbind(ate.high, ate.low)
  ate.diff <- ate.high$estimate - ate.low$estimate
  ci.low <- (ate.high$estimate - ate.low$estimate) - qnorm(0.975)*sqrt(ate.hi
gh$std.dev^2 + ate.low$std.dev^2)
  ci.up <- (ate.high$estimate - ate.low$estimate) + qnorm(0.975)*sqrt(ate.hig
h$std.dev^2 + ate.low$std.dev^2)
  p.value <- 2*pnorm(-abs((ate.high$estimate - ate.low$estimate)/sqrt(ate.hig
h$std.dev^2 + ate.low$std.dev^2)))
  test.result <- cbind(ate.diff,ci.low,ci.up,p.value) %>% as.tibble() %>% mut
ate_if(is.numeric,round,3)
  output<- list()
  output[[1]] <- ates
  output[[2]] <- test.result
  return(output)
}


test1 <- test.heterogeneity(preds.iadl2013,df.imputed.2013,"hloss","iadl13")
test2 <- test.heterogeneity(preds.iadl2016,df.imputed.2016,"hloss","iadl16")
test3 <- test.heterogeneity(preds.adl2013,df.imputed.2013,"hloss","adl3ra13")
test4 <- test.heterogeneity(preds.adl2016,df.imputed.2016,"hloss","adl3ra16")
test5 <- test.heterogeneity(preds.disability2013,df.imputed.2013,"hloss","svy
_physical_dis13")
test6 <- test.heterogeneity(preds.disability2016,df.imputed.2016,"hloss","svy
_physical_dis16")

median.ates <- list(test1[[1]],test2[[1]],test3[[1]],test4[[1]],test5[[1]],te
st6[[1]]) %>% bind_rows()
median.ates.pvalues <- list(test1[[2]],test2[[2]],test3[[2]],test4[[2]],test5
[[2]],test6[[2]]) %>% bind_rows()
```

3.1.2. Calibration test

```
calibration_test <- function(forest, year, outcome){
  test_calibration(forest) %>%
    tidy() %>%
    mutate(year = year,
           outcome= outcome) %>%
    select(outcome, year, term, estimate, p.value)
}

rbind(
  calibration_test(forest.iadl2013,"2013","IADL"),
  calibration_test(forest.iadl2016,"2016","IADL"),
  calibration_test(forest.adl2013,"2013","ADL"),
  calibration_test(forest.adl2016,"2016","ADL"),
  calibration_test(forest.disability2013,"2013","Physical Disability"),
  calibration_test(forest.disability2016,"2016","Physical Disability")
) %>%
  mutate_if(is.numeric,round,2) %>%
  filter(term == "differential.forest.prediction")
```

3.1.3. Calibration plot

```
# Reference: https://bookdown.org/halflearned/ml-ci-tutorial/hte-i-binary-tre
atment.html#via-grf

ATE.ranking.2013 = function(forest, outcome){
  tau.hat = predict(forest)$predictions
  e.hat <- forest$W.hat # P[W=1|X]
  m.hat <- forest$Y.hat # E[Y|X]
  W = df.imputed.2013[,"hloss"]
  Y = df.imputed.2013[,outcome]

  ranking <- rep(NA, n.2013)
  for (fold in seq(num.folds)) {
    tau.hat.quantiles <- quantile(tau.hat[folds.2013 == fold], probs = seq(0,
 1, by=1/num.rankings))
    ranking[folds.2013 == fold] <- cut(tau.hat[folds.2013 == fold], tau.hat.q
uantiles, include.lowest=TRUE,labels=seq(num.rankings))
  }

  mu.hat.0 <- m.hat - e.hat * tau.hat        # E[Y|X,W=0] = E[Y|X] - e(X)*tau
(X)
  mu.hat.1 <- m.hat + (1 - e.hat) * tau.hat  # E[Y|X,W=1] = E[Y|X] + (1 - e(X
))*tau(X)

  # AIPW scores
  aipw.scores <- tau.hat + W / e.hat * (Y -  mu.hat.1) - (1 - W) / (1 - e.hat
) * (Y -  mu.hat.0)
  ols <- lm(aipw.scores ~ 0 + factor(ranking))
```

```
  forest.ate <- data.frame("aipw", paste0("Q", seq(num.rankings)), coeftest(o
ls, vcov=vcovHC(ols, "HC2"))[,1:2])
  colnames(forest.ate) <- c("method", "ranking", "estimate", "std.err")
  rownames(forest.ate) <- NULL # just for display
  forest.ate %>%
    mutate(year = "2013")
}


ATE.ranking.2016 = function(forest, outcome){
  tau.hat = predict(forest)$predictions
  e.hat <- forest$W.hat
  m.hat <- forest$Y.hat
  W = df.imputed.2016[,"hloss"]
  Y = df.imputed.2016[,outcome]

  ranking <- rep(NA, n.2016)
  for (fold in seq(num.folds)) {
    tau.hat.quantiles <- quantile(tau.hat[folds.2016 == fold], probs = seq(0,
 1, by=1/num.rankings))
    ranking[folds.2016 == fold] <- cut(tau.hat[folds.2016 == fold], tau.hat.q
uantiles, include.lowest=TRUE,labels=seq(num.rankings))
  }

  mu.hat.0 <- m.hat - e.hat * tau.hat        # E[Y|X,W=0] = E[Y|X] - e(X)*tau
(X)
  mu.hat.1 <- m.hat + (1 - e.hat) * tau.hat  # E[Y|X,W=1] = E[Y|X] + (1 - e(X
))*tau(X)

  # AIPW scores
  aipw.scores <- tau.hat + W / e.hat * (Y -  mu.hat.1) - (1 - W) / (1 - e.hat
) * (Y -  mu.hat.0)
  ols <- lm(aipw.scores ~ 0 + factor(ranking))
  forest.ate <- data.frame("aipw", paste0("Q", seq(num.rankings)), coeftest(o
ls, vcov=vcovHC(ols, "HC2"))[,1:2])
  colnames(forest.ate) <- c("method", "ranking", "estimate", "std.err")
  rownames(forest.ate) <- NULL # just for display
  forest.ate %>%
    mutate(year = "2016")
}

p1 = rbind(
  ATE.ranking.2013(forest.iadl2013,"iadl13"),
  ATE.ranking.2016(forest.iadl2016,"iadl16")
) %>%
  ggplot(aes(x = ranking, y = estimate, color=year)) +
  geom_point(position=position_dodge(0.2)) +
  geom_errorbar(aes(ymin=estimate-1.96*std.err, ymax=estimate+1.96*std.err),
width=.2, position=position_dodge(0.2)) +
  ylab("Average CATE") + xlab("Ranking") +
```

```
  ggtitle("IADL outcome") +
  theme_bw() +
  theme(legend.position="bottom", legend.title = element_blank())

p2 = rbind(
  ATE.ranking.2013(forest.adl2013,"adl3ra13"),
  ATE.ranking.2016(forest.adl2016,"adl3ra16")
) %>%
  ggplot(aes(x = ranking, y = estimate, color=year)) +
  geom_point(position=position_dodge(0.2)) +
  geom_errorbar(aes(ymin=estimate-1.96*std.err, ymax=estimate+1.96*std.err),
width=.2, position=position_dodge(0.2)) +
  ylab("Average CATE") + xlab("Ranking") +
  ggtitle("ADL outcome") +
  theme_bw() +
  theme(legend.position="bottom", legend.title = element_blank())

p3 = rbind(
  ATE.ranking.2013(forest.disability2013,"svy_physical_dis13"),
  ATE.ranking.2016(forest.disability2016,"svy_physical_dis16")
) %>%
  ggplot(aes(x = ranking, y = estimate, color=year)) +
  geom_point(position=position_dodge(0.2)) +
  geom_errorbar(aes(ymin=estimate-1.96*std.err, ymax=estimate+1.96*std.err),
width=.2, position=position_dodge(0.2)) +
  ylab("Average CATE") + xlab("Ranking") +
  ggtitle("Disability level") +
  theme_bw() +
  theme(legend.position="bottom", legend.title = element_blank())

p1+p2+p3 # use the patchwork package to combine two plots
```

3.2. Create Figure 3

```
# Create density plots for estimated CATEs.
p.density.2013.dis <- preds.2013 %>%
  mutate(
    outcome = factor(outcome, levels = c("Physical Disability","ADL","IADL"))
  ) %>%
  ggplot(aes(x = predictions, fill = outcome)) +
  geom_density(alpha = 0.5) +
  geom_vline(xintercept = 0, linetype = "dashed")+
  xlab("Estimated Conditional Average Treatment Effect of Home Loss") +
  ylab("Density") +
  labs(fill = "Outcome Assessed",color = "Outcome Assessed") +
  theme_bw() +
  scale_fill_nejm() +
  ggtitle("(A) Functional Disability 2.5 Years After the Earthquake Onset") +
  theme(text = element_text(family = "serif",size = 12),
        title = element_text(size = 12),
```

```
        axis.text = element_text(size = 12),
        plot.title = element_text(hjust = 0.5)) +
  scale_x_continuous(limits = c(-1.5,0.5))

p.density.2016.dis <- preds.2016 %>%
  mutate(
    outcome = factor(outcome, levels = c("Physical Disability","ADL","IADL"))
  ) %>%
  ggplot(aes(x = predictions, fill = outcome)) +
  geom_density(alpha = 0.5) +
  geom_vline(xintercept = 0, linetype = "dashed")+
  xlab("Estimated Conditional Average Treatment Effect of Home Loss") +
  ylab("Density") +
  labs(fill = "Outcome Assessed",color = "Outcome Assessed") +
  theme_bw() +
  scale_fill_nejm() +
  ggtitle("(B) Functional Disability 5.5 Years After the Earthquake Onset") +
  theme(text = element_text(family = "serif",size = 12),
        title = element_text(size = 12),
        axis.text = element_text(size = 12),
        plot.title = element_text(hjust = 0.5)) +
  scale_x_continuous(limits = c(-1.5,0.5))

p.density.2013.dis/p.density.2016.dis # use the patchwork package to combine
two plots
```

4. Compare extreme cases

```
compare.extremecases.2013 <- function(outcome){
  preds <- paste("preds",outcome,sep=".")

  data <- df.imputed.2013 %>%
    mutate(folds = folds.2013,
           preds = (!!as.name(preds))) %>%
    group_by(folds) %>% # extract extreme groups by folds
    mutate(cate.percentile = ifelse(preds > quantile(preds, probs = 0.9),"Abo
ve 90 percentile (Resilient)",ifelse(preds < quantile(preds, probs = 0.1), "B
elow 10 percentile (Vulnerable)",NA))) %>%
    filter(!is.na(cate.percentile)) %>%
    ungroup()


  # use the gtsummary package to create a table
  table <- data %>%
    select(preds,covariates,cate.percentile) %>%
    tbl_summary(
      by = cate.percentile,
      statistic = list(all_continuous() ~ "{mean} ({sd})"),
      type = list(c(cum.cond,adl_3_10) ~ "continuous"),
      digits = all_continuous() ~ 2,
```

```
    ) %>%
    add_p(
      test = list(all_continuous() ~ "wilcox.test",
                  all_categorical() ~ "fisher.test")
    ) %>%
    bold_labels()
  return(table)
}

compare.extremecases.2016 <- function(outcome){
  preds <- paste("preds",outcome,sep=".")

  data <- df.imputed.2016 %>%
    mutate(folds = folds.2016,
           preds = (!!as.name(preds))) %>%
    group_by(folds) %>%
    mutate(cate.percentile = ifelse(preds > quantile(preds, probs = 0.9),"Abo
ve 90 percentile (Resilient)",ifelse(preds < quantile(preds, probs = 0.1), "B
elow 10 percentile (Vulnerable)",NA))) %>%
    filter(!is.na(cate.percentile)) %>%
    ungroup()

  table <- data %>%
    select(preds,covariates,cate.percentile) %>%
    tbl_summary(
      by = cate.percentile,
      statistic = list(all_continuous() ~ "{mean} ({sd})"),
      type = list(c(cum.cond,adl_3_10) ~ "continuous"),
      digits = all_continuous() ~ 2,
    ) %>%
    add_p(
        test = list(all_continuous() ~ "wilcox.test",
                    all_categorical() ~ "fisher.test")
    ) %>%
    bold_labels()
  return(table)
}


# use the functions above to create a table for both outcome assessment years
compare.acrossyears <- function(outcome){
  tbl.2013 = compare.extremecases.2013(outcome)
  tbl.2016 = compare.extremecases.2016(outcome)
  tbl_merge(
  tbls = list(tbl.2013,tbl.2016),
  tab_spanner = c("2013","2016")
)
}

# Merge CATE estimates from grf
```

```r
df.imputed.2013$preds.iadl <- preds.iadl2013$predictions
df.imputed.2013$preds.adl    <- preds.adl2013$predictions
df.imputed.2013$preds.disability    <- preds.disability2013$predictions

df.imputed.2016$preds.iadl <- preds.iadl2016$predictions
df.imputed.2016$preds.adl    <- preds.adl2016$predictions
df.imputed.2016$preds.disability    <- preds.disability2016$predictions

compare.acrossyears("iadl")
compare.acrossyears("adl")
compare.acrossyears("disability")
```

5. Variable importance

```r
# Show top three variables that were most often used in growing trees.
rbind(
  variable_importance(forest.iadl2013) %>%
    tidy() %>%
    mutate(varname = covariates) %>%
    arrange(desc(x)) %>%
    mutate_if(is.numeric,round,3) %>%
    slice(1:3) %>%
    mutate(outcome = "IADL",
           year    = 2013),

  variable_importance(forest.iadl2016) %>%
    tidy() %>%
    mutate(varname = covariates) %>%
    arrange(desc(x)) %>%
    mutate_if(is.numeric,round,3) %>%
    slice(1:3) %>%
    mutate(outcome = "IADL",
           year    = 2016),

  variable_importance(forest.adl2013) %>%
    tidy() %>%
    mutate(varname = covariates) %>%
    arrange(desc(x)) %>%
    mutate_if(is.numeric,round,3) %>%
    slice(1:3) %>%
    mutate(outcome = "ADL",
           year    = 2013),

  variable_importance(forest.disability2013) %>%
    tidy() %>%
    mutate(varname = covariates) %>%
    arrange(desc(x)) %>%
    mutate_if(is.numeric,round,3) %>%
    slice(1:3) %>%
    mutate(outcome = "Disability",
```

```
          year    = 2013)

)
```

6. Heat maps

```
# IADL in 2013
df.imputed.2013 %>%
  select(age_10_meibo, soc10, s_eqincome_x10,preds.iadl) %>%
  mutate(age.cat = gtools::quantcut(age_10_meibo, 4),
         soc.cat = gtools::quantcut(soc10,4),
         income.med = ifelse(s_eqincome_x10 >= median(s_eqincome_x10),"Income
 ≥ median (202,000 yen)","Income < median (202,000 yen)")) %>%
  group_by(age.cat,soc.cat,income.med) %>%
  summarise(mean = mean(preds.iadl), sd = sd(preds.iadl)) %>%
  mutate(result = paste0(signif(mean,3),"\n","(",signif(sd,3),")")) %>% # cre
ate labels
  ggplot(aes(x = age.cat, y = soc.cat)) +
  geom_tile(aes(fill = mean)) +
  geom_text(aes(label = result), family = "serif") +
  scale_fill_gradient2(name = "Mean CATEs", limits = c(-0.95,0.1), breaks = c
(-0.9,-0.6,-0.3,0)) +
  facet_wrap(vars(income.med)) +
  xlab("Age") +
  ylab("Sense of Coherence Score") +
  theme(legend.position = "bottom",
        legend.key.width = unit(2,"cm"),
        text = element_text(size = 12, family = "serif"),
        strip.text = element_text(size = 12, family = "serif"),
        axis.text = element_text(size = 12, family = "serif")) +
  guides(fill = guide_colorbar(reverse = T)) +
  ggtitle("A) Outcome: Instrumental Activities of Daily Living in 2013")

# IADL in 2016
df.imputed.2016 %>%
  select(age_10_meibo, soc10, meet6fr10,preds.iadl) %>%
  mutate(age.cat = gtools::quantcut(age_10_meibo, 4),
         soc.cat = gtools::quantcut(soc10,4),
         meet.med = ifelse(meet6fr10 > 3,"Meeting friends ≤ 1-2 times a month
","Meeting friends ≥ once a week")) %>%
  group_by(age.cat,soc.cat,meet.med) %>%
  summarise(mean = mean(preds.iadl), sd = sd(preds.iadl)) %>%
  mutate(result = paste0(signif(mean,3),"\n","(",signif(sd,3),")")) %>%
  ggplot(aes(x = age.cat, y = soc.cat)) +
  geom_tile(aes(fill = mean)) +
  geom_text(aes(label = result), family = "serif") +
  scale_fill_gradient2(name = "Mean CATEs", limits = c(-0.95,0.1), breaks = c
(-0.9,-0.6,-0.3,0)) +
  facet_wrap(vars(meet.med)) +
  xlab("Age") +
```

```
  ylab("Sense of Coherence Score")  +
  theme(legend.position = "bottom",
        legend.key.width = unit(2,"cm"),
        text = element_text(size = 12, family = "serif"),
        strip.text = element_text(size = 12, family = "serif"),
        axis.text = element_text(size = 12, family = "serif")) +
  guides(fill = guide_colorbar(reverse = T)) +
  ggtitle("B) Outcome: Instrumental Activities of Daily Living in 2016")

# ADL in 2013
df.imputed.2013 %>%
  select(age_10_meibo, gout6fq10, s_gds_x10,preds.adl) %>%
  mutate(age.cat = gtools::quantcut(age_10_meibo, 4),
         gout.cat = ifelse(gout6fq10 == 1, "Almost everyday",ifelse(gout6fq10
 == 2, "2-3 times a week", "≤ Once a week")),
         dep.cat = ifelse(s_gds_x10 >= 5, "Depressed","Not depressed")
  ) %>%
  group_by(age.cat,gout.cat,dep.cat) %>%
  summarise(mean = mean(preds.adl), sd = sd(preds.adl)) %>%
  mutate(result = paste0(signif(mean,3),"\n","(",signif(sd,3),")")) %>%
  ggplot(aes(x = age.cat, y = gout.cat)) +
  geom_tile(aes(fill = mean)) +
  geom_text(aes(label = result), family = "serif") +
  scale_fill_gradient2(name = "Mean CATEs",
                       low = "navy",
                       limits = c(-0.25,0),
                       breaks = c(-0.25,-0.2,-0.15,-0.10,-0.05,0)) +
  scale_y_discrete(labels = function(x) str_wrap(x, width = 10)) +
  facet_wrap(vars(dep.cat)) +
  xlab("Age") +
  ylab("Frequency of going out")   +
  theme(legend.position = "bottom",
        legend.key.width = unit(2,"cm"),
        text = element_text(size = 12, family = "serif"),
        strip.text = element_text(size = 12, family = "serif"),
        axis.text = element_text(size = 12, family = "serif")) +
  guides(fill = guide_colorbar(reverse = T)) +
  ggtitle("A) Outcome: Activities of Daily Living in 2013")

# Disability levels in 2013
df.imputed.2013 %>%
  select(age_10_meibo, iadl10, s_eqincome_x10, preds.disability) %>%
  mutate(age.cat = gtools::quantcut(age_10_meibo, 4),
         iadl.cat = gtools::quantcut(iadl10,4),
         income.med = ifelse(s_eqincome_x10 >= median(s_eqincome_x10),"Income
 ≥ median (202,000 yen)","Income < median (202,000 yen)")
  ) %>%
  group_by(age.cat,iadl.cat, income.med) %>%
  summarise(mean = mean(preds.disability), sd = sd(preds.disability)) %>%
  mutate(result = paste0(signif(mean,3),"\n","(",signif(sd,3),")")) %>%
```

```
ggplot(aes(x = age.cat, y = iadl.cat)) +
geom_tile(aes(fill = mean)) +
geom_text(aes(label = result), family = "serif") +
scale_fill_gradient2(name = "Mean CATEs",
                     low = "green4",
                     limits = c(-0.6,0),
                     breaks = c(-0.6,-0.4,-0.2,0)) +
facet_wrap(vars(income.med)) +
xlab("Age") +
ylab("IADL")   +
theme(legend.position = "bottom",
      legend.key.width = unit(2,"cm"),
      text = element_text(size = 12, family = "serif"),
      strip.text = element_text(size = 12, family = "serif"),
      axis.text = element_text(size = 12, family = "serif")) +
guides(fill = guide_colorbar(reverse = T)) +
ggtitle("B) Outcome: Levels of Certified Physical Disability in 2013")
```

Web Appendix 2. Details on our imputation approach

To impute missing data, we used the MissForest, which is a non-parametric imputation approach based on the random forest algorithm.(1) Random forest algorithm learns ensembles of regression or classification trees, each tree fitting a different resampled population and covariate set, to estimate and reduce model variance.(2) Each tree learns a set of rules (e.g., age≥75 vs. age<75) which partition the population of units into different leaves of the tree. The predicted outcome for a new unit is the average of outcomes for observed units assigned to the same leaf; the prediction of the forest is the average of the predictions of all trees. MissForest first trains a random forest on observed values and then use the trained random forest to predict (i.e., impute) missing values; it repeats this process iteratively until the difference between the newly imputed data and the previous one increases for the first time. MissForest allows imputation of missing data with mixed variable types (continuous, binary, and categorical). For the MissForest algorithm, we used all the variables (i.e., the 55 pre-disaster characteristics, home loss, and the three functional limitation outcomes in 2013 and 2016).

Web Appendix 3. On the issue of potential selection bias due to selective attrition.

Methods to "correct" for selection bias (e.g., inverse probability weighting) require information of home loss status both among the censored individuals and the uncensored individuals. Unfortunately, home loss status was measured only among people who participated in the 2013 wave. In other words, we cannot correct for the sample attrition due to non-participation in the 2013 follow-up survey or deaths that occurred before the 2013 survey wave because we do not have the exposure (home loss) information among those censored individuals; hence, censoring weights could not be computed.

To assess how likely is selection bias due to selective attrition, we compared pre-disaster characteristics among 1) the whole sample disaster survivors (n=4,299; baseline study participants who did not have disability before the disaster and were alive on 2011/3/11 and not killed by the tsunami), 2) the analytic sample for the outcomes in 2013 (n=3,350), and 3) the analytic sample for the outcomes in 2016 (n=2,664). We also checked the distribution of distance from the coast to home address in 2010 (<1,000m versus ≥1,000m), which is highly correlated with the degree of property damage from the tsunami in 2011 (please see Web Figure 8), as a proxy of home loss. This variable was available among survivors who did not participate in the post-disaster waves; hence, it provides an approximation of how attrition might be associated with the level of the exposure (home loss). The result is shown in Web Table 12. We found that property damage -- proxied by pre-disaster distance of the individual's home address from the coast -- was slightly less prevalent in the samples with greater attrition, suggesting that the exposed individuals (people who suffered housing damage) were more likely to drop out over time. Moreover, we found that attrition was likely associated with pre-disaster characteristics that may also increase the risk of future functional limitations; for instance, compared to the sample of disaster survivors with least attrition (n=4,299), the group with the greatest attrition in 2016 (n=2,664) tended to be younger, less likely to be depressed, and had higher IADL before the disaster. Although the difference in pre-disaster characteristics across study samples was overall small, this kind of selective attrition will likely underestimate the true causal effect of home loss, as demonstrated by Shiba et al (2021).(3)

Web Table 1. Criteria for levels of physical disability in the Japanese long-term care insurance system.

Disability Rank	Outcome value	Criteria
	9	No physical disability/not requesting care services.
J		Individuals in this rank have some functional disability but is almost completely independent in everyday activities and can go outside unassisted.
J1	8	Can go outside using public transportations, etc.
J2	7	Can go outside to the immediate neighborhood.
A		Individuals in this rank are normally independent in daily living activities at home but cannot go out without assistance.
A1	6	Can go outside if assisted and is out of bed for most of the day.
A2	5	Rarely go outside and get in and out of bed frequently during the daytime.
B		Individuals in this rank require some assistance in performing activities indoors and spend times mostly in bed during the daytime; however, they can keep a seated position.
B1	4	Can independently use a wheelchair to get off a bed for meals and toileting.
B2	3	Need assistance for using a wheelchair to move.
C		Individuals in this rank are in bed all day and require assistance for toileting, eating, and changing clothes.
C1	2	Can turn over in a bed without support.
C2	1	Cannot turn over in a bed independently.

Web Table 2. Levels of housing damage and criterion certifying by local governments

Grade	Criterion [a]
No damage	Not affected.
Partial	Under 20% structural damage or inundation below the floor.
Minor	20% to 40% structural damage or inundation above the floor.
Major	40% to 50% structural damage or inundation approximately 1 meter above the floor.
Complete destruction (Home loss)	Over 50% structural damage, inundation up to ceiling in the first floor, or completely washed away. Uninhabitable beyond repair.

[a] Structural damage was observed in roof, walls, and foundation.

Web Table 3. List of baseline covariates.

Variable Type	Measurement	Coding
Demographic	Sex	1= Man; 2 = Women
Demographic	Age	As continuous
Demographic	Marital status	1 = Married; 2 = Widowed; 3 = Divorced; 4 = Single; 5 = Others
Demographic	Living alone	0 = No; 1 = Yes
Socioeconomic	Education attainment	1= Less than 6 years; 2 = 6-9 years; 3 = 10-12 years; 4 = 13-15 years; 5 = Others
Socioeconomic	Employment status	1= Working; 2 = Retired; 3 = Never worked
Socioeconomic	Equalized household income	As continuous
Health	Self-rated health	1= Very good; 2 = Good; 3 = Not good; 4 = Bad
Health	Body Mass Index	As continuous
Health	Depressive Symptoms	As continuous
Health	ADL	1= support needed completely, 2 = support needed partially, and 3 = no help needed
Health	IADL (total scores as well as subscales of instrumental, intellectual, and social IADL)	As continuous
Health	All self-reported diagnosis of 20 major diseases (cancer, heart diseases, stroke, hypertension, diabetes, obesity, hyperlipidemia, osteoporosis, arthritis, fracture, respiratory diseases, gastrointestinal diseases, liver diseases, psychiatric diseases, dysphagia, visual impairment, hearing loss, dysuria, insomnia, and other)	0 = No; 1 = Yes for each disease.
Psychosocial	Trust in local people (Can you trust your local people?)	1 = Very much; 2 = Moderate; 3 = Neutral; 4 = Not very much; 5 = Not at all
Psychosocial	Mutual help in your community	1 = Very much; 2 = Moderate; 3 = Neutral; 4 = Not very much; 5 = Not at all
Psychosocial	How much do you attach your community?	1 = Very much; 2 = Moderate; 3 = Neutral; 4 = Not very much; 5 = Not at all
Psychosocial	Participation in sport clubs	1 = Everyday; 2 = 2-3times a week; 3 = once a week; 4 = 1-2 times a month; 5 = a few times a year; 6 = not at all
Psychosocial	Participation in hobby clubs	1 = Everyday; 2 = 2-3 times a week; 3 = once a week; 4 = 1-2 times a month; 5 = a few times a year; 6 = not at all
Psychosocial	Meeting friends	1 = Every day; 2 =2-3 times a year; 3 = once a week; 4 = 1-2 times a month; 5 = a few times a year, 6 = Not at all
Psychosocial	How many friends did you meet in the past month?	1 = 0; 2 = 1-2 friends; 3 = 3-5 friends; 4 = 6-9 friends; 5 = 10 or more friends
Psychosocial	Received emotional support	0 = Yes; 1 = No
Psychosocial	Providing emotional support	0 = Yes; 1 = No
Psychosocial	Received care support	0 = Yes; 1 = No
Psychosocial	Giving care support	0 = Yes; 1 = No
Psychosocial	Communication with neighbors	1 = Very much; 2 = Moderate; 3 = Minimum; 4 = Not at all

Psychosocial	Count of 6 negative life events	As continuous
Psychosocial	Sense of coherence	As continuous
Behavioral	Current smoking status	1 = Never; 2 = Quit; 3 = Smoking
Behavioral	Current alcohol drinking	1 = Drinking; 2 = Quit; 3 = No drinking
Behavioral	Frequency of fish and meat consumption	1 = Twice a day or more often; 2 = once a day; 3 = 4-6 times a week; 4 = 2-3 times a week; 5 = once a week; 6 = less than once a week; 7 = not at all
Behavioral	Eating vegetable or fruit	1 = Twice a day or more often; 2 = once a day; 3 = 4-6 times a week; 4 = 2-3 times a week; 5 = once a week; 6 = less than once a week; 7 = not at all
Behavioral	Frequency of going out	1 = Everyday; 2 = 2-3 times a week; 3 = once a week; 4 = 1-2 times a month; 5 = a few times a year; 6 = not at all
Behavioral	Having hobby	1 = Yes; 2 = No

All covariates were measured at the 2010 wave, 7 months prior to the disaster onset. These variables were used as covariates to control for confounding and also to estimate effect heterogeneity.

Web Table 4. Distributions of the functional limitation outcomes in 2013 and 2016 among the analytic samples (n = 3,350 for 2013 and n = 2,264 for 2016).

	2013		2016	
	Home loss	No home loss	Home loss	No home loss
Outcomes	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Certified physical disability	8.42 (1.46)	8.73 (0.98)	8.35 (1.41)	8.58 (1.17)
ADL	2.74 (0.59)	2.90 (0.39)	2.64 (0.70)	2.74 (0.58)
Total IADL	10.1 (3.54)	11.3 (2.48)	10.7 (3.33)	11.2 (2.53)

Abbreviations: ADL, Activity of Daily Living; IADL, Instrumental Activity of Daily Living; SD, standard deviation

IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. IADL scores ranged from 0-13 points, where smaller scores indicate lower functional independence. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). Levels of certified physical disability ranged from 1 (“Cannot roll over in a bed independently”) to 9 (“no physical disability/not requesting care services”), where smaller values indicate greater levels of disability. Thus, lower values for these outcomes indicate greater functional limitation.

Web Table 5. Estimates of conditional average treatment effects of home loss on functional limitation in 2013 and 2016.[a]

Year	Outcome	Conditional Average Treatment Effects[b]			
		Mean	SD	Min	Max
2013	Total IADL	-0.68	0.27	-1.95	0.06
	Physical Disability	-0.20	0.21	-1.4	0.08
	ADL	-0.10	0.07	-0.41	0.02
2016	Total IADL	-0.35	0.41	-2.49	0.37
	Physical Disability	-0.18	0.13	-0.79	0.18
	ADL	-0.07	0.08	-0.48	0.10

Abbreviations: ADL, Activity of Daily Living; IADL, Instrumental Activity of Daily Living.

[a] All models were adjusted for the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. IADL scores ranged from 0-13 points, where smaller scores indicate lower functional independence. Levels of certified physical disability ranged from 1 (“Cannot roll over in a bed independently”) to 9 (“no physical disability/not requesting care services”), where smaller values indicate greater levels of disability. Thus, lower values for these outcomes indicate greater functional limitation.

[b] Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave.

Web Table 6. Evaluation of causal forest fit and test of heterogeneity.[a]

Year	Outcome	Comparison of ATEs among those who were above versus below the median CATE[b]	Best linear predictor analysis[c]			
			Mean forest		Differential forest	
		p for heterogeneity	estimate	p-value	estimate	P for heterogeneity
2013	Total IADL	<0.001	1.04	0.03	1.40	0.02
	Physical Disability	<0.001	1.11	<0.01	1.76	0.01
	ADL	<0.001	1.10	<0.01	0.91	0.01
2016	Total IADL	<0.001	1.20	0.17	2.01	0.02
	Physical Disability	<0.001	1.24	0.08	1.81	<.001
	ADL	<0.001	1.28	0.20	1.93	0.10

Abbreviations: ADL, Activity of Daily Living; IADL, Instrumental Activity of Daily Living

[a] All models were adjusted for the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. IADL scores ranged from 0-13 points, where smaller scores indicate lower functional independence. Levels of certified physical disability ranged from 1 ("Cannot roll over in a bed independently") to 9 ("no physical disability/not requesting care services"), where smaller values indicate greater levels of disability. Thus, lower values for these outcomes indicate greater functional limitation. Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave.

[b] We estimated average treatment effects among those who were above/below the median CATE. We then computed p-value for the difference between the two estimated average effects as a test for heterogeneity.

[c] Best linear predictor analysis was implemented by using the test_calibration() function in the grf package. In the BLP analysis, if the coefficient for mean forest is closer to 1, then it indicates that the average forest prediction is correct. If the coefficient for the differential forest prediction is closer to 1, then it indicates the forest prediction adequately captures the underlying heterogeneity. The single-sided test of a coefficient for differential forest >0 is used as an omnibus test for heterogeneity.

Web Table 7. Pre-disaster characteristics of people at bottom 10% vs. top 10% of the estimated conditional average treatment effect of home loss on certified physical disability level in 2013 and 2016.[a]

Baseline Characteristics	Year: 2013 (n = 3,350)			Year: 2016 (n = 2,664)		
	Resilient (n = 340)	Vulnerable (n = 340)	p-value[b]	Resilient (n = 280)	Vulnerable (n = 280)	p-value[b]
CATE Estimates, mean (SD)[e]	0.00 (0.02)	-0.69 (0.19)	<0.001	-0.01 (0.04)	-0.43 (0.08)	<0.001
Age, mean (SD)	68.64 (2.69)	83.08 (5.04)	<0.001	68.60 (3.14)	79.98 (3.41)	<0.001
Gender, n (%)			<0.001			<0.001
Men	121 (36%)	229 (67%)		109 (39%)	179 (64%)	
Women	219 (64%)	111 (33%)		171 (61%)	101 (36%)	
Marital status, n (%)			<0.001			<0.001
Married	334 (98%)	148 (44%)		244 (87%)	178 (64%)	
Widowed	6 (1.8%)	180 (53%)		27 (9.6%)	93 (33%)	
Divorced	0 (0%)	3 (0.9%)		7 (2.5%)	5 (1.8%)	
Single	0 (0%)	7 (2.1%)		1 (0.4%)	3 (1.1%)	
Others	0 (0%)	2 (0.6%)		1 (0.4%)	1 (0.4%)	
Living alone, n (%)			<0.001			0.2
No	337 (99%)	306 (90%)		266 (95%)	257 (92%)	
Yes	3 (0.9%)	34 (10%)		14 (5.0%)	23 (8.2%)	
Education, n (%)			<0.001			<0.001
Less than 6 years	1 (0.3%)	17 (5.0%)		1 (0.4%)	3 (1.1%)	
6-9 years	74 (22%)	175 (51%)		60 (21%)	113 (40%)	
10-12 years	174 (51%)	114 (34%)		148 (53%)	105 (38%)	
13 years or more	86 (25%)	29 (8.5%)		71 (25%)	55 (20%)	
Others	5 (1.5%)	5 (1.5%)		0 (0%)	4 (1.4%)	
Job, n (%)			<0.001			<0.001
Working	100 (29%)	17 (5.0%)		90 (32%)	15 (5.4%)	
Retired	216 (64%)	174 (51%)		173 (62%)	177 (63%)	
Never worked	24 (7.1%)	149 (44%)		17 (6.1%)	88 (31%)	
Household income, mean (SD)	234.45 (141.51)	231.60 (125.21)	0.4	284.21 (129.83)	242.74 (107.89)	<0.001
Depressive symptoms, n (%)			<0.001			<0.001
Mild/severe depressive symptoms	61 (18%)	128 (38%)		22 (7.9%)	94 (34%)	
No depressive symptoms	279 (82%)	212 (62%)		258 (92%)	186 (66%)	
Self-rated health, n (%)			<0.001			<0.001
Very good	63 (19%)	23 (6.8%)		55 (20%)	28 (10%)	
Good	260 (76%)	212 (62%)		215 (77%)	196 (70%)	
Not good	17 (5.0%)	85 (25%)		8 (2.9%)	49 (18%)	
Bad	0 (0%)	20 (5.9%)		2 (0.7%)	7 (2.5%)	
Body mass index, mean (SD)	25.58 (2.51)	22.79 (3.15)	<0.001	24.40 (2.84)	22.69 (2.43)	<0.001
Total IADL, mean (SD)	12.48 (0.78)	9.60 (3.32)	<0.001	12.18 (1.40)	12.30 (1.06)	0.8
Instrumental IADL, mean (SD)[e]	4.96 (0.20)	3.82 (1.50)	<0.001	4.66 (0.78)	4.94 (0.27)	<0.001
Intellectual IADL, mean (SD)[e]	3.80 (0.45)	3.02 (1.18)	<0.001	3.87 (0.46)	3.79 (0.47)	0.001
Social IADL, mean (SD)[e]	3.73 (0.53)	2.76 (1.36)	<0.001	3.65 (0.63)	3.58 (0.82)	0.7
ADL, mean (SD)	3.00 (0.05)	2.93 (0.33)	<0.001	2.99 (0.08)	3.00 (0.06)	0.6
# of Treatment for major diseases, mean (SD)	1.02 (1.02)	2.07 (1.60)	<0.001	1.16 (1.04)	1.70 (1.37)	<0.001
Trust in local people, n (%)			<0.001			<0.001
Very much	25 (7.4%)	71 (21%)		33 (12%)	47 (17%)	
Moderate	245 (72%)	171 (50%)		204 (73%)	156 (56%)	
Neutral	66 (19%)	82 (24%)		41 (15%)	71 (25%)	

Not very much	4 (1.2%)	14 (4.1%)		2 (0.7%)	2 (0.7%)	
Not at all	0 (0%)	2 (0.6%)		0 (0%)	4 (1.4%)	
Mutual help in community, n (%)			<0.001			<0.001
Very much	16 (4.7%)	48 (14%)		14 (5.0%)	36 (13%)	
Moderate	201 (59%)	167 (49%)		187 (67%)	140 (50%)	
Neutral	108 (32%)	90 (26%)		70 (25%)	82 (29%)	
Not very much	12 (3.5%)	25 (7.4%)		7 (2.5%)	20 (7.1%)	
Not at all	3 (0.9%)	10 (2.9%)		2 (0.7%)	2 (0.7%)	
Attachment to community, n (%)			<0.001			0.021
Very much	79 (23%)	122 (36%)		73 (26%)	103 (37%)	
Moderate	208 (61%)	150 (44%)		172 (61%)	136 (49%)	
Neutral	38 (11%)	44 (13%)		26 (9.3%)	26 (9.3%)	
Not very much	14 (4.1%)	18 (5.3%)		8 (2.9%)	13 (4.6%)	
Not at all	1 (0.3%)	6 (1.8%)		1 (0.4%)	2 (0.7%)	
Participation in sport clubs, n (%)			<0.001			0.008
Everyday	2 (0.6%)	1 (0.3%)		4 (1.4%)	2 (0.7%)	
A few times a week	34 (10%)	16 (4.7%)		19 (6.8%)	44 (16%)	
Once a week	40 (12%)	13 (3.8%)		23 (8.2%)	26 (9.3%)	
1-2 times a month	25 (7.4%)	9 (2.6%)		19 (6.8%)	9 (3.2%)	
A few times a year	25 (7.4%)	6 (1.8%)		12 (4.3%)	13 (4.6%)	
Not at all	214 (63%)	295 (87%)		203 (72%)	186 (66%)	
Participation in hobby clubs, n (%)			<0.001			<0.001
Everyday	2 (0.6%)	0 (0%)		4 (1.4%)	1 (0.4%)	
A few times a week	31 (9.1%)	24 (7.1%)		11 (3.9%)	45 (16%)	
Once a week	46 (14%)	23 (6.8%)		24 (8.6%)	43 (15%)	
1-2 times a month	56 (16%)	30 (8.8%)		43 (15%)	65 (23%)	
A few times a year	60 (18%)	12 (3.5%)		39 (14%)	22 (7.9%)	
Not at all	145 (43%)	251 (74%)		159 (57%)	104 (37%)	
Frequency of meeting friends, n (%)			<0.001			<0.001
Every day	25 (7.4%)	46 (14%)		13 (4.6%)	45 (16%)	
2-3 times a week	83 (24%)	64 (19%)		41 (15%)	76 (27%)	
Once a week	64 (19%)	55 (16%)		51 (18%)	65 (23%)	
1-2 times a month	96 (28%)	62 (18%)		91 (32%)	58 (21%)	
A few times a year	70 (21%)	49 (14%)		66 (24%)	32 (11%)	
Not at all	2 (0.6%)	64 (19%)		18 (6.4%)	4 (1.4%)	
# of friends interacted last month, n (%)			<0.001			<0.001
0	8 (2.4%)	50 (15%)		15 (5.4%)	8 (2.9%)	
1-2 friends	35 (10%)	92 (27%)		57 (20%)	30 (11%)	
3-5 friends	84 (25%)	85 (25%)		96 (34%)	75 (27%)	
6-9 friends	57 (17%)	32 (9.4%)		36 (13%)	50 (18%)	
10 or more friends	156 (46%)	81 (24%)		76 (27%)	117 (42%)	
Received emotional social support, n (%)	10 (2.9%)	36 (11%)	<0.001	7 (2.5%)	7 (2.5%)	>0.9
Provision of emotional support, n (%)	3 (0.9%)	62 (18%)	<0.001	6 (2.1%)	15 (5.4%)	0.073
Received care support, n (%)	6 (1.8%)	12 (3.5%)	0.2	7 (2.5%)	2 (0.7%)	0.2
Provision of care support, n (%)	5 (1.5%)	106 (31%)	<0.001	12 (4.3%)	33 (12%)	0.002
Communication with neighbors, n (%)			0.2			<0.001
Very much	84 (25%)	78 (23%)		32 (11%)	87 (31%)	
Moderate	193 (57%)	176 (52%)		197 (70%)	153 (55%)	
Minimum	60 (18%)	79 (23%)		51 (18%)	39 (14%)	
Not at all	3 (0.9%)	7 (2.1%)		0 (0%)	1 (0.4%)	
# of negative events in the past year, mean (SD)	0.54 (0.73)	0.64 (0.72)	0.042	0.46 (0.62)	0.68 (0.75)	<0.001

Sense of coherence, mean (SD)	23.54 (3.60)	21.84 (4.21)	<0.001	24.00 (2.99)	21.82 (4.26)	<0.001
Current smoking status, n (%)			<0.001			<0.001
Never	157 (46%)	244 (72%)		143 (51%)	198 (71%)	
Quit	121 (36%)	76 (22%)		96 (34%)	64 (23%)	
Smoking	62 (18%)	20 (5.9%)		41 (15%)	18 (6.4%)	
Current alcohol drinking, n (%)			<0.001			<0.001
Drinking	225 (66%)	74 (22%)		187 (67%)	56 (20%)	
Quit	8 (2.4%)	13 (3.8%)		6 (2.1%)	9 (3.2%)	
Never	107 (31%)	253 (74%)		87 (31%)	215 (77%)	
Frequency of fish and meat consumption, n (%)			0.5			0.2
Twice a day or more often	45 (13%)	45 (13%)		29 (10%)	39 (14%)	
Once a day	130 (38%)	138 (41%)		145 (52%)	124 (44%)	
4-6 times a week	82 (24%)	64 (19%)		62 (22%)	56 (20%)	
2-3 times a week	69 (20%)	76 (22%)		41 (15%)	56 (20%)	
Once a week	10 (2.9%)	10 (2.9%)		3 (1.1%)	3 (1.1%)	
Less than once a week	4 (1.2%)	4 (1.2%)		0 (0%)	2 (0.7%)	
Not at all	0 (0%)	3 (0.9%)				
Frequency of fruits and vegetable consumption, n (%)			0.049			0.4
Twice a day or more often	185 (54%)	187 (55%)		166 (59%)	175 (62%)	
Once a day	100 (29%)	101 (30%)		84 (30%)	80 (29%)	
4-6 times a week	42 (12%)	25 (7.4%)		25 (8.9%)	16 (5.7%)	
2-3 times a week	10 (2.9%)	22 (6.5%)		5 (1.8%)	8 (2.9%)	
Once a week	2 (0.6%)	1 (0.3%)				
Less than once a week	0 (0%)	2 (0.6%)		0 (0%)	1 (0.4%)	
Not at all	1 (0.3%)	2 (0.6%)				
Frequency of going out, n (%)			<0.001			<0.001
Every day	230 (68%)	90 (26%)		176 (63%)	127 (45%)	
2-3 times a week	96 (28%)	104 (31%)		88 (31%)	98 (35%)	
Once a week	14 (4.1%)	52 (15%)		8 (2.9%)	41 (15%)	
1-2 times a month	0 (0%)	50 (15%)		7 (2.5%)	10 (3.6%)	
A few times a year	0 (0%)	16 (4.7%)		1 (0.4%)	3 (1.1%)	
Not at all	0 (0%)	28 (8.2%)		0 (0%)	1 (0.4%)	
Having hobby, n (%)			<0.001			<0.001
Yes	232 (68%)	146 (43%)		165 (59%)	222 (79%)	
No	108 (32%)	194 (57%)		115 (41%)	58 (21%)	

Abbreviations: ADL, Activity of Daily Living; CATE, Conditional average treatment effect; IADL, Instrumental Activity of Daily Living

[a] Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated via the generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. Levels of certified physical disability ranged from 1 ("Cannot roll over in a bed independently") to 9 ("no physical disability/not requesting care services"), where smaller values indicate greater levels of disability. Thus, decrease in this outcome indicate increased functional limitation. Top 10 % of the distributions of CATEs were labeled as a "Resilient" group because they showed weaker associations between home loss and decreased outcome. Bottom 10 % of the distributions of CATEs were labeled as a "Vulnerable" group because they showed greater associations between home loss and decreased outcome.

[b] P-values for between-group differences. We used Wilcoxon rank sum test for continuous variables and Fisher's exact test for categorical variables.

Web Table 8. Pre-disaster characteristics of people at bottom 10% vs. top 10% of the estimated conditional average treatment effect of home loss on total instrumental activities of daily living scores in 2013 and 2016.[a]

Baseline Characteristics	Year: 2013 (n = 3,350)			Year: 2016 (n = 2,664)		
	Resilient (n = 340)	Vulnerable (n = 340)	p-value[b]	Resilient (n = 280)	Vulnerable (n = 280)	p-value[b]
CATE Estimates, mean (SD)[e]	-0.32 (0.08)	-1.20 (0.19)	<0.001	0.17 (0.07)	-1.16 (0.27)	<0.001
Age, mean (SD)	68.86 (3.17)	81.39 (4.87)	<0.001	69.34 (2.97)	79.50 (4.68)	<0.001
Gender, n (%)			<0.001			<0.001
Men	131 (39%)	206 (61%)		118 (42%)	190 (68%)	
Women	209 (61%)	134 (39%)		162 (58%)	90 (32%)	
Marital status, n (%)			<0.001			<0.001
Married	330 (97%)	154 (45%)		260 (93%)	146 (52%)	
Widowed	6 (1.8%)	173 (51%)		15 (5.4%)	122 (44%)	
Divorced	4 (1.2%)	6 (1.8%)		4 (1.4%)	5 (1.8%)	
Single	0 (0%)	6 (1.8%)		0 (0%)	5 (1.8%)	
Others	0 (0%)	1 (0.3%)		1 (0.4%)	2 (0.7%)	
Living alone, n (%)			<0.001			<0.001
No	332 (98%)	299 (88%)		271 (97%)	238 (85%)	
Yes	8 (2.4%)	41 (12%)		9 (3.2%)	42 (15%)	
Education, n (%)			<0.001			0.001
Less than 6 years	1 (0.3%)	9 (2.6%)		1 (0.4%)	5 (1.8%)	
6-9 years	73 (21%)	140 (41%)		64 (23%)	102 (36%)	
10-12 years	179 (53%)	131 (39%)		152 (54%)	125 (45%)	
13 years or more	84 (25%)	54 (16%)		61 (22%)	46 (16%)	
Others	3 (0.9%)	6 (1.8%)		2 (0.7%)	2 (0.7%)	
Job, n (%)			<0.001			<0.001
Working	94 (28%)	25 (7.4%)		82 (29%)	16 (5.7%)	
Retired	217 (64%)	192 (56%)		177 (63%)	184 (66%)	
Never worked	29 (8.5%)	123 (36%)		21 (7.5%)	80 (29%)	
Household income, mean (SD)	233.85 (139.35)	254.94 (118.20)	<0.001	233.32 (147.77)	254.00 (117.63)	0.002
Depressive symptoms, n (%)			<0.001			<0.001
Mild/severe depressive symptoms	48 (14%)	106 (31%)		24 (8.6%)	104 (37%)	
No depressive symptoms	292 (86%)	234 (69%)		256 (91%)	176 (63%)	
Self-rated health, n (%)			<0.001			<0.001
Very good	71 (21%)	30 (8.8%)		14 (5.0%)	34 (12%)	
Good	249 (73%)	241 (71%)		234 (84%)	197 (70%)	
Not good	18 (5.3%)	60 (18%)		30 (11%)	39 (14%)	
Bad	2 (0.6%)	9 (2.6%)		2 (0.7%)	10 (3.6%)	
Body mass index, mean (SD)	25.55 (2.75)	22.16 (2.60)	<0.001	25.29 (2.79)	22.76 (2.31)	<0.001
Total IADL, mean (SD)	11.57 (1.62)	11.95 (1.61)	<0.001	12.07 (1.23)	12.01 (1.69)	0.2
Instrumental IADL, mean (SD)[e]	4.78 (0.57)	4.72 (0.69)	0.056	4.89 (0.39)	4.75 (0.67)	0.003
Intellectual IADL, mean (SD)[e]	3.43 (0.76)	3.75 (0.60)	<0.001	3.60 (0.67)	3.76 (0.59)	<0.001
Social IADL, mean (SD)[e]	3.34 (0.91)	3.50 (0.82)	0.053	3.58 (0.70)	3.52 (0.93)	>0.9
ADL, mean (SD)	3.00 (0.05)	2.99 (0.13)	0.3	3.00 (0.06)	3.00 (0.00)	0.3
# of Treatment for major diseases, mean (SD)	0.99 (1.01)	1.74 (1.36)	<0.001	1.30 (1.21)	1.76 (1.45)	<0.001
Trust in local people, n (%)			<0.001			<0.001
Very much	42 (12%)	55 (16%)		26 (9.3%)	48 (17%)	
Moderate	223 (66%)	198 (58%)		224 (80%)	140 (50%)	
Neutral	53 (16%)	82 (24%)		29 (10%)	76 (27%)	

Not very much	22 (6.5%)	5 (1.5%)		0 (0%)	12 (4.3%)	
Not at all				1 (0.4%)	4 (1.4%)	
Mutual help in community, n (%)			0.018			<0.001
Very much	17 (5.0%)	34 (10%)		13 (4.6%)	36 (13%)	
Moderate	198 (58%)	203 (60%)		219 (78%)	134 (48%)	
Neutral	97 (29%)	88 (26%)		42 (15%)	83 (30%)	
Not very much	25 (7.4%)	15 (4.4%)		6 (2.1%)	20 (7.1%)	
Not at all	3 (0.9%)	0 (0%)		0 (0%)	7 (2.5%)	
Attachment to community, n (%)			0.2			<0.001
Very much	96 (28%)	108 (32%)		65 (23%)	106 (38%)	
Moderate	187 (55%)	187 (55%)		190 (68%)	128 (46%)	
Neutral	49 (14%)	33 (9.7%)		21 (7.5%)	29 (10%)	
Not very much	8 (2.4%)	11 (3.2%)		4 (1.4%)	14 (5.0%)	
Not at all	0 (0%)	1 (0.3%)		0 (0%)	3 (1.1%)	
Participation in sport clubs, n (%)			0.03			0.11
Everyday	2 (0.6%)	1 (0.3%)		3 (1.1%)	2 (0.7%)	
A few times a week	31 (9.1%)	30 (8.8%)		18 (6.4%)	36 (13%)	
Once a week	35 (10%)	20 (5.9%)		25 (8.9%)	25 (8.9%)	
1-2 times a month	27 (7.9%)	15 (4.4%)		15 (5.4%)	9 (3.2%)	
A few times a year	23 (6.8%)	16 (4.7%)		16 (5.7%)	11 (3.9%)	
Not at all	222 (65%)	258 (76%)		203 (72%)	197 (70%)	
Participation in hobby clubs, n (%)			0.051			<0.001
Everyday	5 (1.5%)	2 (0.6%)		1 (0.4%)	2 (0.7%)	
A few times a week	30 (8.8%)	30 (8.8%)		17 (6.1%)	45 (16%)	
Once a week	44 (13%)	32 (9.4%)		34 (12%)	35 (12%)	
1-2 times a month	61 (18%)	52 (15%)		45 (16%)	58 (21%)	
A few times a year	39 (11%)	25 (7.4%)		38 (14%)	16 (5.7%)	
Not at all	161 (47%)	199 (59%)		145 (52%)	124 (44%)	
Frequency of meeting friends, n (%)			<0.001			<0.001
Every day	29 (8.5%)	52 (15%)		1 (0.4%)	54 (19%)	
2-3 times a week	66 (19%)	73 (21%)		57 (20%)	75 (27%)	
Once a week	50 (15%)	69 (20%)		58 (21%)	55 (20%)	
1-2 times a month	81 (24%)	83 (24%)		73 (26%)	53 (19%)	
A few times a year	80 (24%)	52 (15%)		76 (27%)	28 (10%)	
Not at all	34 (10%)	11 (3.2%)		15 (5.4%)	15 (5.4%)	
# of friends interacted last month, n (%)			<0.001			<0.001
0	30 (8.8%)	11 (3.2%)		11 (3.9%)	11 (3.9%)	
1-2 friends	59 (17%)	61 (18%)		64 (23%)	28 (10%)	
3-5 friends	78 (23%)	112 (33%)		81 (29%)	84 (30%)	
6-9 friends	38 (11%)	51 (15%)		37 (13%)	41 (15%)	
10 or more friends	135 (40%)	105 (31%)		87 (31%)	116 (41%)	
Received emotional social support, n (%)	13 (3.8%)	15 (4.4%)	0.8	7 (2.5%)	9 (3.2%)	0.8
Provision of emotional support, n (%)	11 (3.2%)	27 (7.9%)	0.011	5 (1.8%)	20 (7.1%)	0.003
Received care support, n (%)	7 (2.1%)	5 (1.5%)	0.8	7 (2.5%)	5 (1.8%)	0.8
Provision of care support, n (%)	13 (3.8%)	65 (19%)	<0.001	9 (3.2%)	45 (16%)	<0.001
Communication with neighbors, n (%)			0.008			<0.001
Very much	59 (17%)	89 (26%)		44 (16%)	87 (31%)	
Moderate	209 (61%)	198 (58%)		191 (68%)	154 (55%)	
Minimum	71 (21%)	50 (15%)		44 (16%)	38 (14%)	
Not at all	1 (0.3%)	3 (0.9%)		1 (0.4%)	1 (0.4%)	
# of negative events in the past year, mean (SD)	0.46 (0.64)	0.70 (0.76)	<0.001	0.43 (0.55)	0.67 (0.75)	<0.001

Sense of coherence, mean (SD)	24.34 (2.72)	21.26 (4.55)	<0.001	23.72 (2.50)	20.39 (4.89)	<0.001
Current smoking status, n (%)			<0.001			<0.001
Never	162 (48%)	230 (68%)		139 (50%)	211 (75%)	
Quit	124 (36%)	91 (27%)		105 (38%)	55 (20%)	
Smoking	54 (16%)	19 (5.6%)		36 (13%)	14 (5.0%)	
Current alcohol drinking, n (%)			<0.001			<0.001
Drinking	199 (59%)	86 (25%)		157 (56%)	61 (22%)	
Quit	6 (1.8%)	14 (4.1%)		5 (1.8%)	9 (3.2%)	
Never	135 (40%)	240 (71%)		118 (42%)	210 (75%)	
Frequency of fish and meat consumption, n (%)			0.007			0.037
Twice a day or more often	35 (10%)	44 (13%)		21 (7.5%)	38 (14%)	
Once a day	121 (36%)	150 (44%)		114 (41%)	117 (42%)	
4-6 times a week	71 (21%)	73 (21%)		53 (19%)	62 (22%)	
2-3 times a week	94 (28%)	67 (20%)		78 (28%)	55 (20%)	
Once a week	16 (4.7%)	5 (1.5%)		12 (4.3%)	7 (2.5%)	
Less than once a week	3 (0.9%)	1 (0.3%)		2 (0.7%)	1 (0.4%)	
Not at all	0 (0%)	0 (0%)		0 (0%)	0 (0%)	
Frequency of fruits and vegetable consumption, n (%)			0.001			<0.001
Twice a day or more often	170 (50%)	196 (58%)		139 (50%)	177 (63%)	
Once a day	92 (27%)	105 (31%)		91 (32%)	83 (30%)	
4-6 times a week	52 (15%)	25 (7.4%)		38 (14%)	11 (3.9%)	
2-3 times a week	23 (6.8%)	14 (4.1%)		12 (4.3%)	9 (3.2%)	
Once a week	3 (0.9%)	0 (0%)		0 (0%)	0 (0%)	
Less than once a week	0 (0%)	0 (0%)		0 (0%)	0 (0%)	
Not at all	0 (0%)	0 (0%)		0 (0%)	0 (0%)	
Frequency of going out, n (%)			<0.001			<0.001
Every day	217 (64%)	129 (38%)		167 (60%)	111 (40%)	
2-3 times a week	97 (29%)	119 (35%)		91 (32%)	113 (40%)	
Once a week	12 (3.5%)	59 (17%)		14 (5.0%)	33 (12%)	
1-2 times a month	8 (2.4%)	24 (7.1%)		6 (2.1%)	15 (5.4%)	
A few times a year	5 (1.5%)	4 (1.2%)		2 (0.7%)	4 (1.4%)	
Not at all	1 (0.3%)	5 (1.5%)		0 (0%)	4 (1.4%)	
Having hobby, n (%)			0.038			0.009
Yes	230 (68%)	203 (60%)		173 (62%)	203 (72%)	
No	110 (32%)	137 (40%)		107 (38%)	77 (28%)	

Abbreviations: ADL, Activity of Daily Living; CATE, Conditional average treatment effect; IADL, Instrumental Activity of Daily Living

[a] Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated via the generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. Scores ranged from 0-13 points for total IADL, 0-5 points for instrumental IADL, 0-4 points for intellectual IADL, and 0-4 points for social IADL, where smaller scores indicate lower functional independence. Thus, decrease in this outcome indicate increased functional limitation. Top 10 % of the distributions of CATEs were labeled as a “Resilient” group because they showed weaker associations between home loss and decreased outcome. Bottom 10 % of the distributions of CATEs were labeled as a “Vulnerable” group because they showed greater associations between home loss and decreased outcome.

[b] P-values for between-group differences. We used Wilcoxon rank sum test for continuous variables and Fisher’s exact test for categorical variables.

Web Table 9. Pre-disaster characteristics of people at bottom 10% vs top 10% of the estimated conditional average treatment effect of home loss on activities of daily living scores in 2013 and 2016.[a]

Baseline Characteristics	Year: 2013 (n = 3,350)			Year: 2016 (n = 2,664)		
	Resilient (n = 340)	Vulnerable (n = 340)	p-value[b]	Resilient (n = 280)	Vulnerable (n = 280)	p-value[b]
CATE Estimates, mean (SD)[e]	-0.02 (0.01)	-0.24 (0.04)	<0.001	0.03 (0.02)	-0.23 (0.06)	<0.001
Age, mean (SD)	68.18 (2.62)	81.57 (4.65)	<0.001	68.53 (2.92)	80.78 (3.42)	<0.001
Gender, n (%)			<0.001			<0.001
Men	101 (30%)	247 (73%)		114 (41%)	178 (64%)	
Women	239 (70%)	93 (27%)		166 (59%)	102 (36%)	
Marital status, n (%)			<0.001			<0.001
Married	334 (98%)	150 (44%)		255 (91%)	160 (57%)	
Widowed	3 (0.9%)	178 (52%)		14 (5.0%)	109 (39%)	
Divorced	3 (0.9%)	4 (1.2%)		6 (2.1%)	7 (2.5%)	
Single	0 (0%)	7 (2.1%)		3 (1.1%)	3 (1.1%)	
Others	0 (0%)	1 (0.3%)		2 (0.7%)	1 (0.4%)	
Living alone, n (%)			<0.001			0.001
No	337 (99%)	291 (86%)		269 (96%)	248 (89%)	
Yes	3 (0.9%)	49 (14%)		11 (3.9%)	32 (11%)	
Education, n (%)			<0.001			<0.001
Less than 6 years	1 (0.3%)	12 (3.5%)		0 (0%)	6 (2.1%)	
6-9 years	50 (15%)	172 (51%)		44 (16%)	114 (41%)	
10-12 years	178 (52%)	116 (34%)		148 (53%)	107 (38%)	
13 years or more	109 (32%)	37 (11%)		86 (31%)	49 (18%)	
Others	2 (0.6%)	3 (0.9%)		2 (0.7%)	4 (1.4%)	
Job, n (%)			<0.001			<0.001
Working	109 (32%)	17 (5.0%)		83 (30%)	16 (5.7%)	
Retired	222 (65%)	179 (53%)		182 (65%)	182 (65%)	
Never worked	9 (2.6%)	144 (42%)		15 (5.4%)	82 (29%)	
Household income, mean (SD)	270.69 (151.07)	235.72 (113.53)	0.003	260.29 (116.96)	257.05 (126.67)	0.5
Depressive symptoms, n (%)			<0.001			<0.001
Mild/severe depressive symptoms	33 (9.7%)	134 (39%)		24 (8.6%)	84 (30%)	
No depressive symptoms	307 (90%)	206 (61%)		256 (91%)	196 (70%)	
Self-rated health, n (%)			<0.001			0.03
Very good	77 (23%)	26 (7.6%)		39 (14%)	30 (11%)	
Good	253 (74%)	213 (63%)		221 (79%)	210 (75%)	
Not good	10 (2.9%)	81 (24%)		18 (6.4%)	32 (11%)	
Bad	0 (0%)	20 (5.9%)		2 (0.7%)	8 (2.9%)	
Body mass index, mean (SD)	24.87 (2.87)	23.03 (3.05)	<0.001	24.02 (2.25)	23.10 (2.63)	<0.001
Total IADL, mean (SD)	12.39 (0.90)	10.42 (3.08)	<0.001	12.03 (1.46)	12.21 (1.40)	0.047
Instrumental IADL, mean (SD)[e]	4.82 (0.43)	4.18 (1.39)	<0.001	4.71 (0.69)	4.87 (0.50)	<0.001
Intellectual IADL, mean (SD)[e]	3.82 (0.46)	3.29 (1.02)	<0.001	3.84 (0.51)	3.77 (0.56)	0.035
Social IADL, mean (SD)[e]	3.74 (0.49)	2.96 (1.27)	<0.001	3.48 (0.79)	3.59 (0.83)	0.014
ADL, mean (SD)	3.00 (0.00)	2.94 (0.29)	<0.001	3.00 (0.06)	3.00 (0.06)	>0.9
# of Treatment for major diseases, mean (SD)	0.94 (1.01)	2.14 (1.65)	<0.001	1.16 (1.12)	1.78 (1.38)	<0.001
Trust in local people, n (%)			<0.001			<0.001

Very much	12 (3.5%)	66 (19%)		13 (4.6%)	57 (20%)	
Moderate	275 (81%)	152 (45%)		186 (66%)	166 (59%)	
Neutral	46 (14%)	110 (32%)		70 (25%)	50 (18%)	
Not very much	7 (2.1%)	8 (2.4%)		11 (3.9%)	5 (1.8%)	
Not at all	0 (0%)	4 (1.2%)		0 (0%)	2 (0.7%)	
Mutual help in community, n (%)			<0.001			<0.001
Very much	4 (1.2%)	47 (14%)		2 (0.7%)	45 (16%)	
Moderate	236 (69%)	154 (45%)		146 (52%)	161 (57%)	
Neutral	80 (24%)	106 (31%)		102 (36%)	60 (21%)	
Not very much	20 (5.9%)	26 (7.6%)		26 (9.3%)	12 (4.3%)	
Not at all	0 (0%)	7 (2.1%)		4 (1.4%)	2 (0.7%)	
Attachment to community, n (%)			<0.001			<0.001
Very much	75 (22%)	115 (34%)		33 (12%)	118 (42%)	
Moderate	239 (70%)	154 (45%)		183 (65%)	140 (50%)	
Neutral	14 (4.1%)	51 (15%)		47 (17%)	16 (5.7%)	
Not very much	11 (3.2%)	14 (4.1%)		16 (5.7%)	5 (1.8%)	
Not at all	1 (0.3%)	6 (1.8%)		1 (0.4%)	1 (0.4%)	
Participation in sport clubs, n (%)			<0.001			0.3
Everyday	0 (0%)	2 (0.6%)		1 (0.4%)	1 (0.4%)	
A few times a week	33 (9.7%)	20 (5.9%)		23 (8.2%)	38 (14%)	
Once a week	42 (12%)	26 (7.6%)		28 (10%)	33 (12%)	
1-2 times a month	31 (9.1%)	8 (2.4%)		15 (5.4%)	10 (3.6%)	
A few times a year	30 (8.8%)	8 (2.4%)		17 (6.1%)	15 (5.4%)	
Not at all	204 (60%)	276 (81%)		196 (70%)	183 (65%)	
Participation in hobby clubs, n (%)			<0.001			0.001
Everyday	5 (1.5%)	0 (0%)		2 (0.7%)	2 (0.7%)	
A few times a week	25 (7.4%)	32 (9.4%)		19 (6.8%)	41 (15%)	
Once a week	34 (10%)	34 (10%)		36 (13%)	42 (15%)	
1-2 times a month	65 (19%)	36 (11%)		45 (16%)	64 (23%)	
A few times a year	53 (16%)	18 (5.3%)		32 (11%)	21 (7.5%)	
Not at all	158 (46%)	220 (65%)		146 (52%)	110 (39%)	
Frequency of meeting friends, n (%)			<0.001			<0.001
Every day	38 (11%)	41 (12%)		1 (0.4%)	51 (18%)	
2-3 times a week	76 (22%)	91 (27%)		61 (22%)	77 (28%)	
Once a week	64 (19%)	52 (15%)		46 (16%)	61 (22%)	
1-2 times a month	86 (25%)	59 (17%)		77 (28%)	57 (20%)	
A few times a year	74 (22%)	43 (13%)		76 (27%)	27 (9.6%)	
Not at all	2 (0.6%)	54 (16%)		19 (6.8%)	7 (2.5%)	
# of friends interacted last month, n (%)			<0.001			0.044
0	5 (1.5%)	47 (14%)		16 (5.7%)	8 (2.9%)	
1-2 friends	29 (8.5%)	86 (25%)		55 (20%)	34 (12%)	
3-5 friends	88 (26%)	83 (24%)		70 (25%)	85 (30%)	
6-9 friends	59 (17%)	41 (12%)		41 (15%)	47 (17%)	
10 or more friends	159 (47%)	83 (24%)		98 (35%)	106 (38%)	
Received emotional social support, n (%)	9 (2.6%)	29 (8.5%)	0.001	13 (4.6%)	11 (3.9%)	0.8
Provision of emotional support, n (%)	6 (1.8%)	50 (15%)	<0.001	10 (3.6%)	20 (7.1%)	0.09
Received care support, n (%)	10 (2.9%)	9 (2.6%)	>0.9	8 (2.9%)	3 (1.1%)	0.2

Provision of care support, n (%)	3 (0.9%)	98 (29%)	<0.001	11 (3.9%)	37 (13%)	<0.001
Communication with neighbors, n (%)			0.016			<0.001
Very much	64 (19%)	77 (23%)		36 (13%)	93 (33%)	
Moderate	222 (65%)	184 (54%)		175 (62%)	155 (55%)	
Minimum	52 (15%)	73 (21%)		67 (24%)	31 (11%)	
Not at all	2 (0.6%)	6 (1.8%)		2 (0.7%)	1 (0.4%)	
# of negative events in the past year, mean (SD)	0.64 (0.78)	0.65 (0.71)	0.7	0.59 (0.70)	0.67 (0.73)	0.2
Sense of coherence, mean (SD)	23.68 (3.60)	21.72 (4.13)	<0.001	24.13 (2.68)	21.94 (4.10)	<0.001
Current smoking status, n (%)			<0.001			<0.001
Never	139 (41%)	264 (78%)		138 (49%)	199 (71%)	
Quit	127 (37%)	62 (18%)		87 (31%)	63 (22%)	
Smoking	74 (22%)	14 (4.1%)		55 (20%)	18 (6.4%)	
Current alcohol drinking, n (%)			<0.001			<0.001
Drinking	222 (65%)	62 (18%)		153 (55%)	72 (26%)	
Quit	11 (3.2%)	10 (2.9%)		5 (1.8%)	7 (2.5%)	
Never	107 (31%)	268 (79%)		122 (44%)	201 (72%)	
Frequency of fish and meat consumption, n (%)			0.074			0.3
Twice a day or more often	54 (16%)	40 (12%)		37 (13%)	39 (14%)	
Once a day	141 (41%)	131 (39%)		133 (48%)	114 (41%)	
4-6 times a week	69 (20%)	66 (19%)		58 (21%)	52 (19%)	
2-3 times a week	67 (20%)	80 (24%)		47 (17%)	67 (24%)	
Once a week	7 (2.1%)	16 (4.7%)		3 (1.1%)	6 (2.1%)	
Less than once a week	2 (0.6%)	3 (0.9%)		2 (0.7%)	2 (0.7%)	
Not at all	0 (0%)	4 (1.2%)				
Frequency of fruits and vegetable consumption, n (%)			0.2			0.3
Twice a day or more often	202 (59%)	188 (55%)		159 (57%)	166 (59%)	
Once a day	92 (27%)	103 (30%)		90 (32%)	85 (30%)	
4-6 times a week	36 (11%)	27 (7.9%)		27 (9.6%)	18 (6.4%)	
2-3 times a week	8 (2.4%)	17 (5.0%)		4 (1.4%)	9 (3.2%)	
Once a week	2 (0.6%)	2 (0.6%)		0 (0%)	1 (0.4%)	
Less than once a week	0 (0%)	1 (0.3%)		0 (0%)	1 (0.4%)	
Not at all	0 (0%)	2 (0.6%)				
Frequency of going out, n (%)			<0.001			<0.001
Every day	245 (72%)	94 (28%)		186 (66%)	116 (41%)	
2-3 times a week	77 (23%)	111 (33%)		77 (28%)	110 (39%)	
Once a week	17 (5.0%)	27 (7.9%)		11 (3.9%)	33 (12%)	
1-2 times a month	1 (0.3%)	56 (16%)		6 (2.1%)	14 (5.0%)	
A few times a year	0 (0%)	23 (6.8%)		0 (0%)	5 (1.8%)	
Not at all	0 (0%)	29 (8.5%)		0 (0%)	2 (0.7%)	
Having hobby, n (%)			<0.001			<0.001
Yes	236 (69%)	180 (53%)		185 (66%)	221 (79%)	
No	104 (31%)	160 (47%)		95 (34%)	59 (21%)	

Abbreviations: ADL, Activity of Daily Living; CATE, Conditional average treatment effect; IADL, Instrumental Activity of Daily Living
[a] Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated via the generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). Thus, decrease in this outcome indicate increased functional limitation. Top 10 % of the distributions of CATEs were

labeled as a “Resilient” group because they showed weaker associations between home loss and decreased outcome. Bottom 10 % of the distributions of CATEs were labeled as a “Vulnerable” group because they showed greater associations between home loss and decreased outcome.

[b] P-values for between-group differences. We used Wilcoxon rank sum test for continuous variables and Fisher’s exact test for categorical variables.

Web Table 10. Sensitivity analysis for associations between home loss and functional limitation in 2013[a]

Outcome	Year	Average Treatment Effects[b]		Conditional Average Treatment Effects[c]			
		Estimate	95% CI	Mean	SD	Min	Max
Total IADL	Main analysis	-0.50	(-0.56, -0.44)	-0.68	0.27	-1.95	0.06
	Sensitivity analysis[d]	-0.55	(-0.60, -0.50)	-0.56	0.16	-1.25	-0.08
Physical Disability	Main analysis	-0.09	(-0.13, -0.06)	-0.20	0.21	-1.40	0.08
	Sensitivity analysis[d]	-0.13	(-0.16, -0.10)	-0.13	0.06	-0.41	-0.03
ADL	Main analysis	-0.05	(-0.07, -0.04)	-0.10	0.07	-0.41	0.02
	Sensitivity analysis[d]	-0.04	(-0.05, -0.03)	-0.06	0.03	-0.17	0.03

Abbreviations: ADL, Activity of Daily Living; IADL, Instrumental Activity of Daily Living.

[a] All models were adjusted for the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. IADL scores ranged from 0-13 points, where smaller scores indicate lower functional independence. Levels of certified physical disability ranged from 1 (“Cannot roll over in a bed independently”) to 9 (“no physical disability/not requesting care services”), where smaller values indicate greater levels of disability. Thus, lower values for these outcomes indicate greater functional limitation.

[b] Average treatment effects were estimated via the doubly-robust targeted maximum likelihood estimation. Models were estimated via the SuperLearner using generalized linear models, gradient boosting machine, and neural net as candidate estimators.

[c] Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave.

[d] We restricted the analytic sample to those who participated in both 2013 and 2016 waves.

Web Table 11. Variable importance ranking

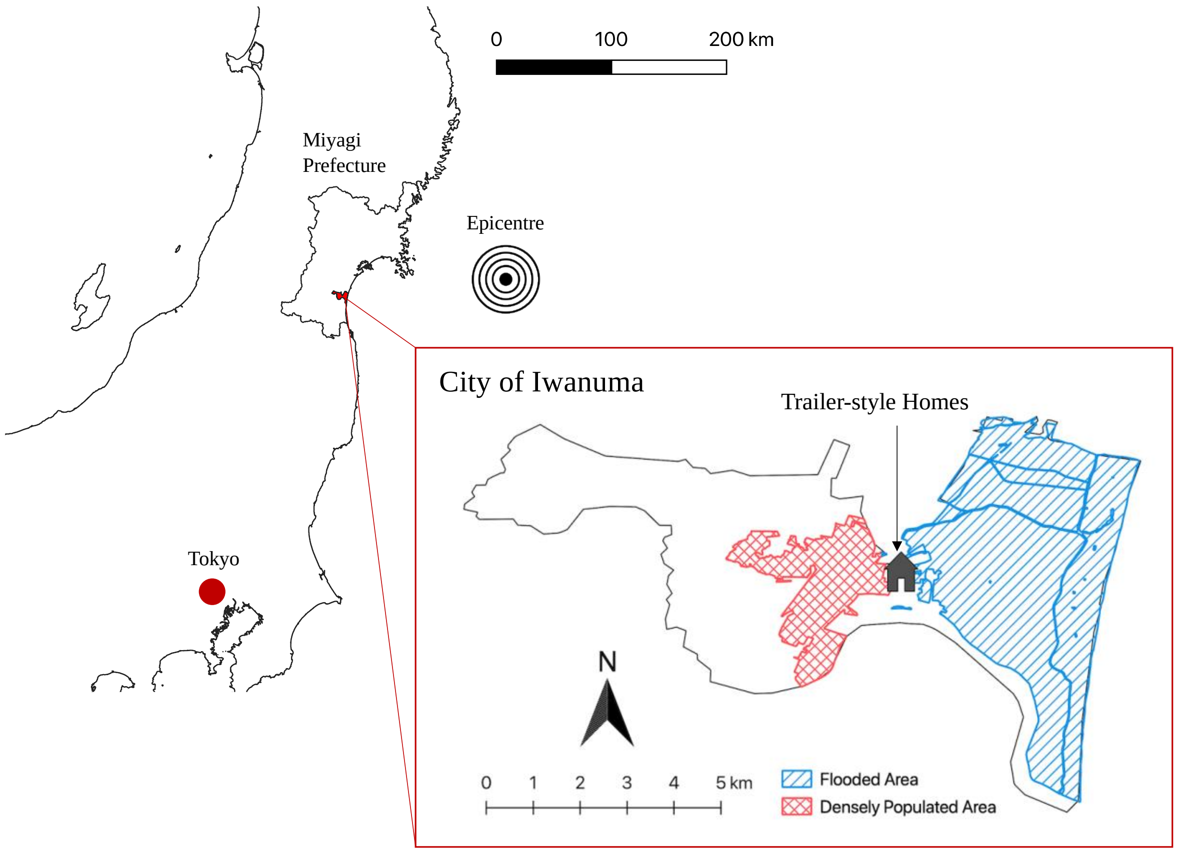
Outcome	Year	Rank[a]	Pre-disaster characteristics	Proportion[b]	Spearman's Correlation[c]
2013	Physical Disability	1	Age	0.15	-0.71
		2	IADL	0.11	0.33
		3	Income[d]	0.07	-0.01
	Total IADL	1	Age	0.10	-0.57
		2	Sense of coherence scores	0.10	0.32
		3	BMI	0.09	0.34
	ADL	1	Frequency of going out	0.15	-0.31
		2	IADL	0.11	0.25
		3	Age	0.09	-0.68
2016	Physical Disability	1	Age	0.13	-0.62
		2	Income[d]	0.11	0.16
		3	Sense of coherence scores	0.07	0.27
	Total IADL	1	Sense of coherence scores	0.15	0.33
		2	Age	0.13	-0.54
		3	Self-rated Health	0.07	-0.06
	ADL	1	Age	0.14	-0.65
		2	Income[d]	0.09	0.06
		3	Frequency of meeting friends	0.08	0.18

Abbreviations: ADL, Activity of Daily Living; BMI, Body Mass Index; IADL, Instrumental Activity of Daily Living.
[a]Variable importance ranking identifies variables that were most often used in splitting the trees.
[b]For each variable, the proportion of trees that used the variable in training is shown.
[c]Spearman's correlation coefficients were calculated for the relationship between each pre-disaster characteristic and predicted CATE.
[d]Income was log-transformed.

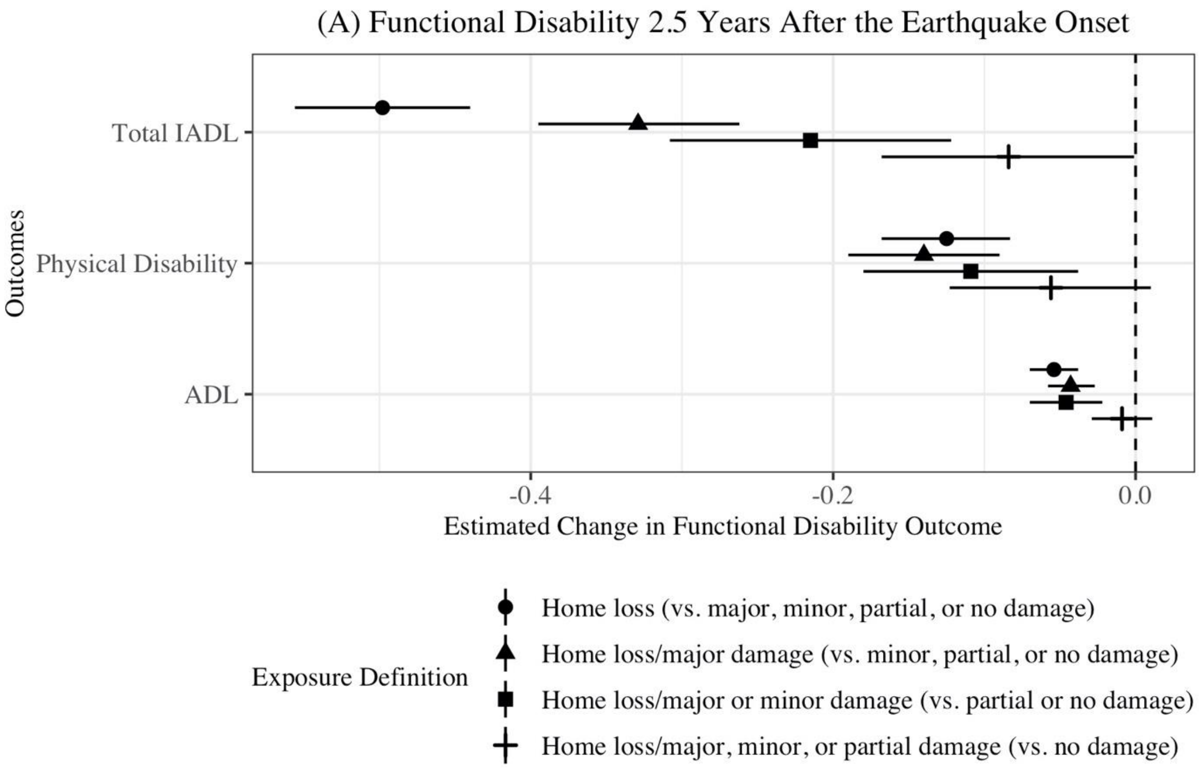
Web Table 12. Comparison of pre-disaster characteristics across study samples.

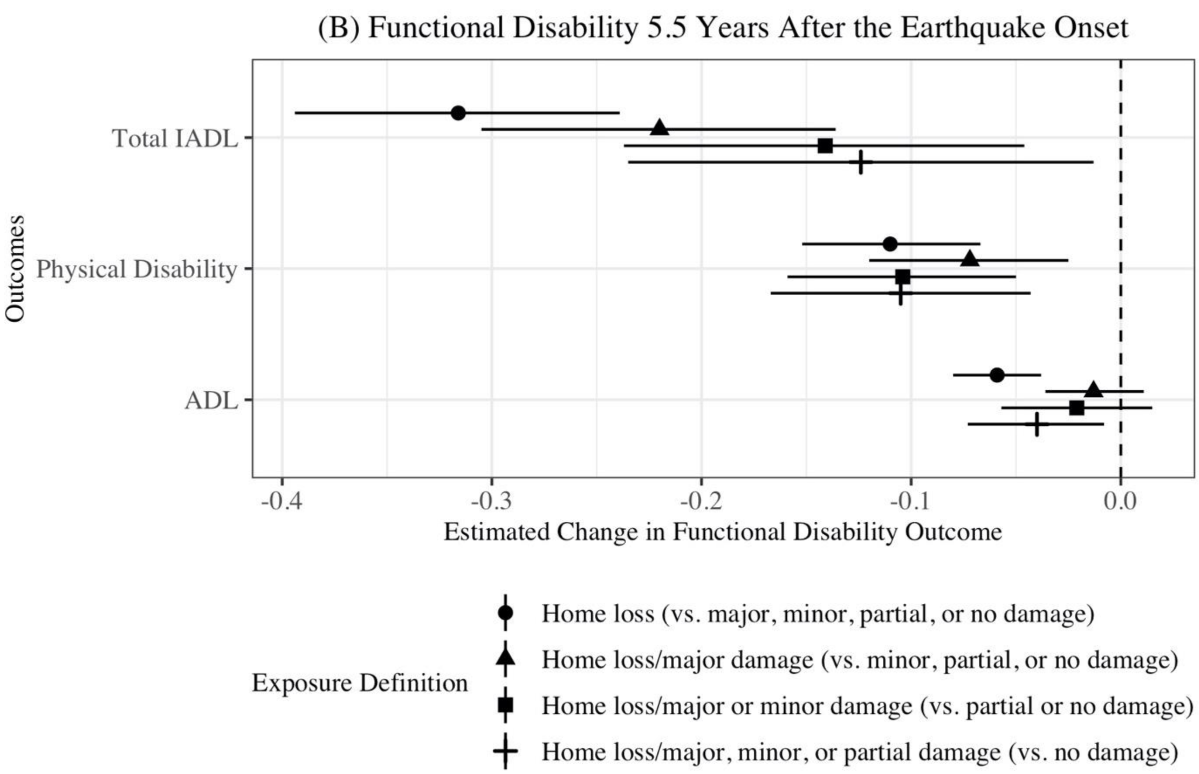
Baseline Characteristics	Disaster survivors (n=4,299)	Sample for the 2013 outcomes (n=3,350)	Sample for the 2016 outcomes (n=2,664)
Distance from the coast			
<1,000m	201 (4.7%)	131 (3.9%)	91 (3.4%)
≥1,000m	4,098 (95%)	3,219 (96%)	2,573 (97%)
Home loss	NA	149 (4.4%)	107 (4.0%)
Age (years), mean (SD)	74 (6)	73.2 (6.0)	72.5 (5.5)
Gender, n (%)			
Men	2,348 (55%)	1,857 (55%)	1,486 (56%)
Women	1,951 (45%)	1,493 (45%)	1,178 (44%)
Marital status, n (%)			
Married	2,962 (72%)	2,364 (73%)	1,946 (75%)
Widowed	958 (23%)	733 (23%)	533 (21%)
Divorced	124 (3.0%)	83 (2.6%)	66 (2.6%)
Single	56 (1.4%)	39 (1.2%)	28 (1.1%)
Others	30 (0.7%)	18 (0.6%)	15 (0.6%)
Living alone, n (%)			
No	3,778 (91%)	2,979 (91%)	2,391 (92%)
Yes	381 (9.2%)	281 (8.6%)	213 (8.2%)
Education, n (%)			
Less than 6 years	61 (1.5%)	33 (1.0%)	21 (0.8%)
6-9 years	1,426 (34%)	1,103 (34%)	830 (32%)
10-12 years	1,777 (43%)	1,417 (44%)	1,166 (45%)
13 years or more	841 (20%)	676 (21%)	572 (22%)
Others	43 (1.0%)	26 (0.8%)	17 (0.7%)
Job, n (%)			
Working	675 (18%)	550 (19%)	462 (19%)
Retired	2,434 (65%)	1,892 (64%)	1,531 (65%)
Never worked	651 (17%)	520 (18%)	378 (16%)
Household income [10,000 yen], mean (SD)[a]	230 (146)	231 (141)	232 (138)
Depressive symptoms, n (%)[b]			
Mild/severe depressive symptoms	1,141 (31%)	857 (30%)	663 (28%)
No depressive symptoms	2,539 (69%)	2,039 (70%)	1,673 (72%)
Self-rated health, n (%)			
Bad	518 (12%)	417 (13%)	343 (13%)
Not good	2,922 (69%)	2,336 (71%)	1,891 (72%)
Good	651 (15%)	458 (14%)	334 (13%)
Very good	125 (3.0%)	75 (2.3%)	48 (1.8%)
Body mass index (m/kg^2), mean (SD)	23.5 (3.2)	23.6 (3.1)	23.58 (3.02)
Total IADL, mean (SD)[c]	11.64 (2.25)	11.88 (1.80)	12.02 (1.59)
Instrumental IADL, mean (SD)[c]	4.64 (0.96)	4.73 (0.75)	4.78 (0.65)
Intellectual IADL, mean (SD)[c]	3.60 (0.81)	3.66 (0.69)	3.71 (0.62)
Social IADL, mean (SD)[c]	3.38 (1.01)	3.46 (0.91)	3.51 (0.87)
ADL, mean (SD)[d]	2.96 (0.25)	2.98 (0.15)	2.99 (0.10)
# of Treatment for major diseases, mean (SD)[e]	2.15 (1.41)	2.08 (1.33)	2.02 (1.27)

Note. Abbreviations: ADL, Activity of Daily Living; IADL, Instrumental Activity of Daily Living
[a] Annual household income was divided by the square root of the number of household members to account for household size.
[b] We used the Geriatric Depression Scale (range: 0-15 points; higher scores indicate more depressive symptoms) to assess depressive symptoms. We used a cut-off of ≥5 points to define moderate/severe depressive symptoms.
[c] IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. Scores ranged from 0-13 points for total IADL, 0-5 points for instrumental IADL, 0-4 points for intellectual IADL, and 0-4 points for social IADL, where smaller scores indicate lower functional independence.
[d] ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed).
[e] We calculated counts of current treatment for major diseases, including cancer, heart diseases, stroke, hypertension, diabetes, obesity, hyperlipidemia, osteoporosis, arthritis, fracture, respiratory diseases, gastrointestinal diseases, liver diseases, psychiatric diseases, dysphagia, visual impairment, hearing loss, dysuria, and insomnia.



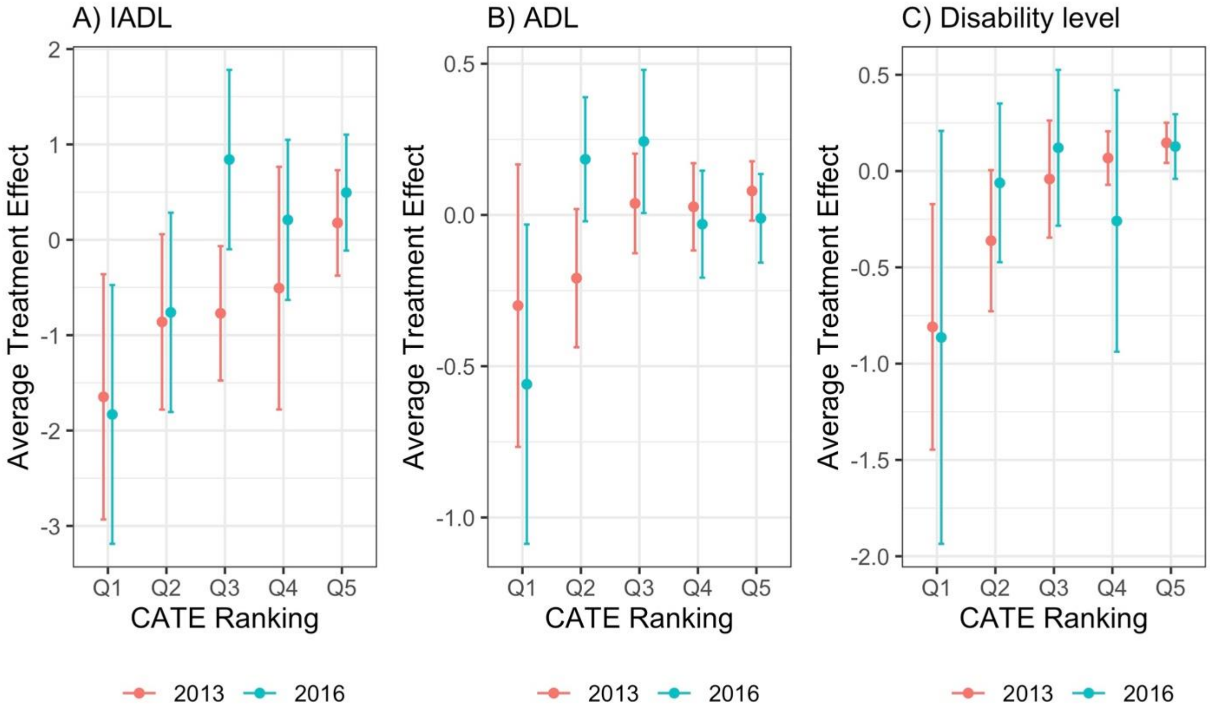
Web Figure 1. Map of Iwanuma city.





Web Figure 2. Sensitivity analysis for population average treatment effects of housing damage on functional limitation in 2013 and 2016.

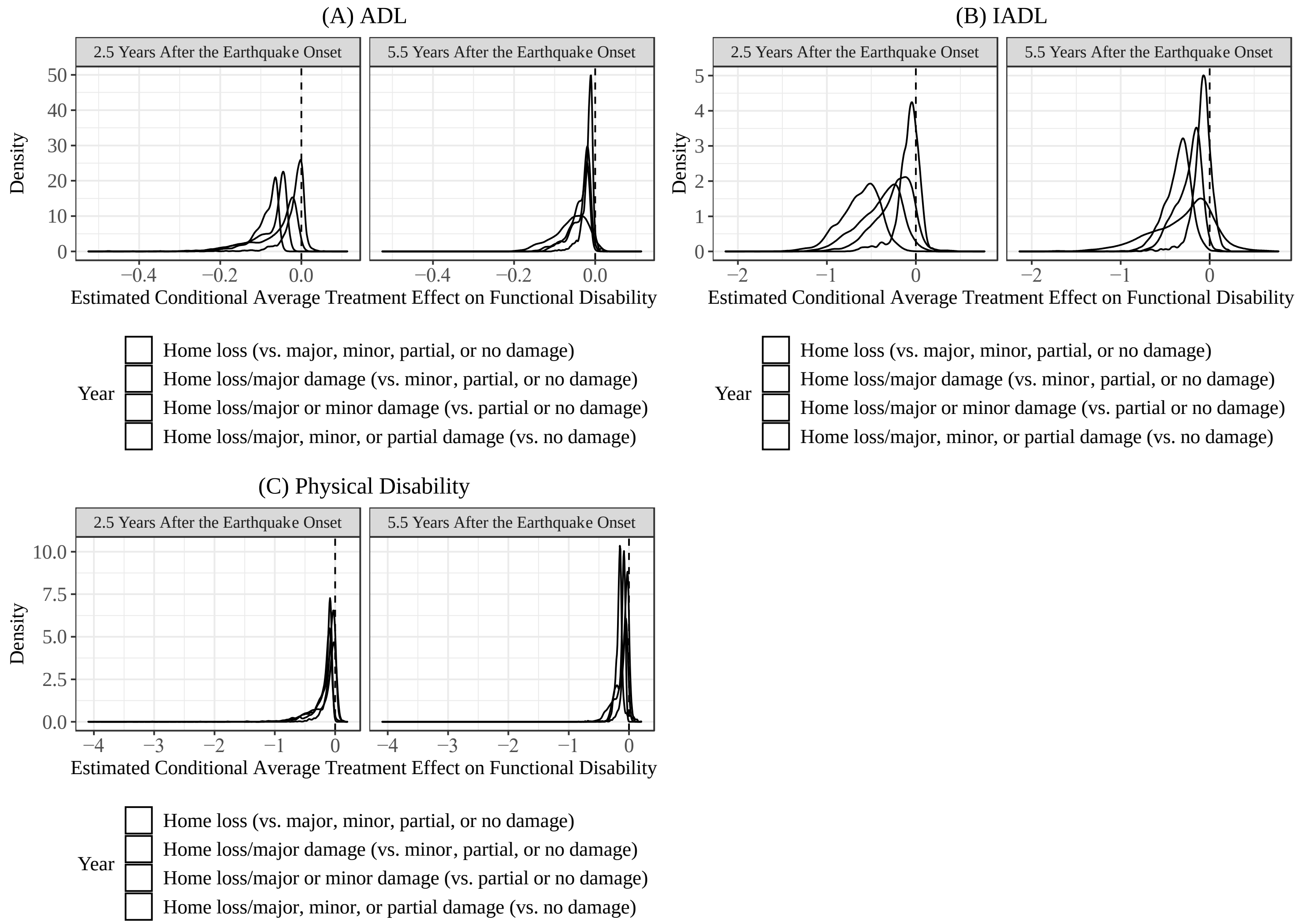
Abbreviations: ADL, Activity of Daily Living; IADL, Instrumental Activity of Daily Living.
Population average effects (i.e., average treatment effects) of the exposures were estimated via the doubly-robust targeted maximum likelihood estimation. Models were estimated data-adaptively via the SuperLearner using generalized linear models, gradient boosting machine, and neural net as candidate estimators. All models were adjusted for the 55 pre-disaster demographic and socioeco-nomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. IADL scores ranged from 0-13 points, where smaller scores indicate lower functional independence. Levels of certified physical disability ranged from 1 ("Cannot roll over in a bed independently") to 9 ("no physical disability/not requesting care services"), where smaller values indicate greater levels of disability. Thus, decrease in these outcomes indicate increased functional limitation.



Web Figure 3. Causal forest calibration plots for outcomes A) IADL, B) ADL, and C) physical disability level.

Note. Abbreviations: ADL, Activity of Daily Living; IADL, Instrumental Activity of Daily Living

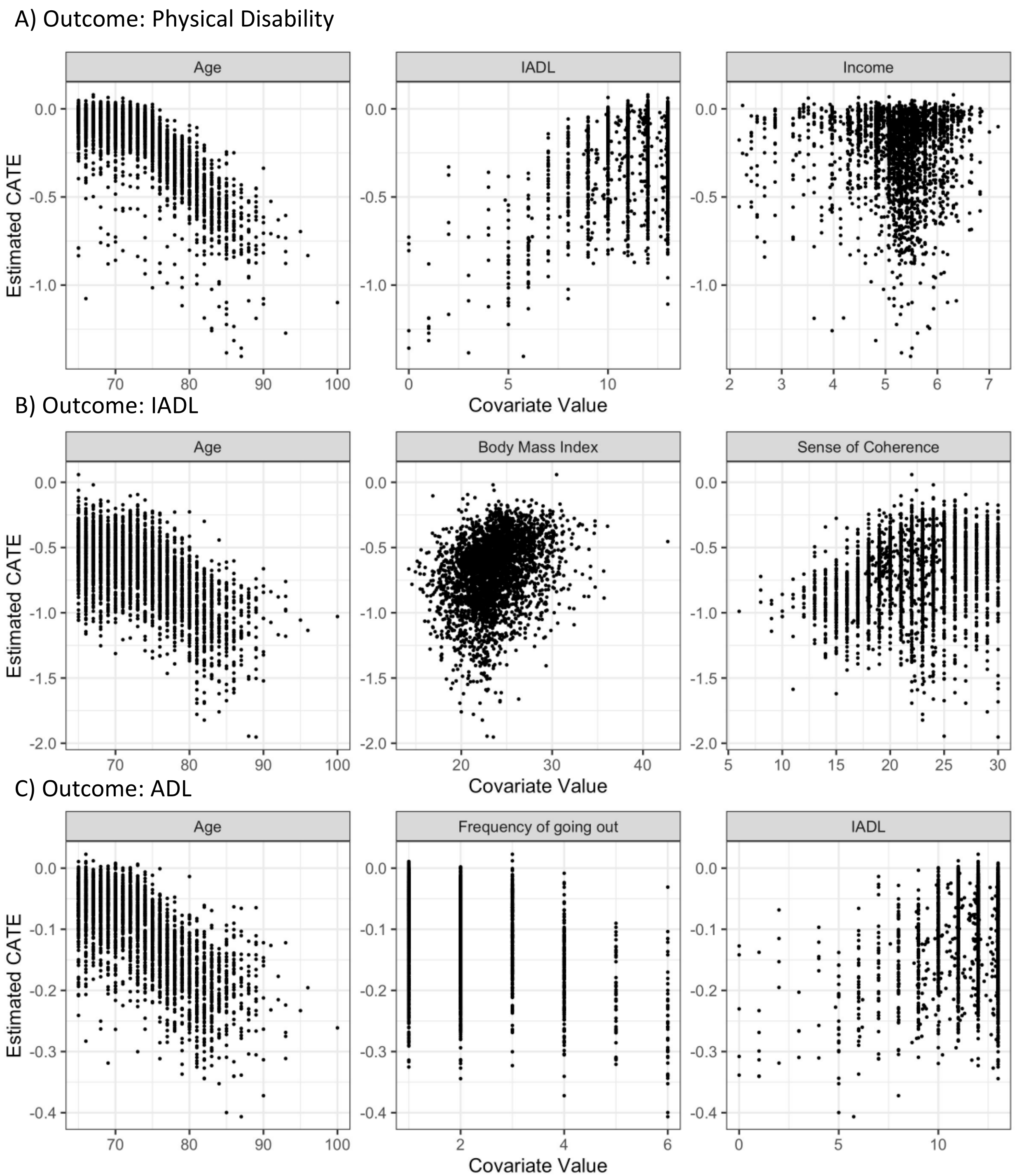
We followed the instructions provided in the GRF online tutorial material (https://bookdown.org/halflearned/ml-ci-tutorial/hte-i-binary-treatment.html#via-grf). CATE quintiles were determined within each fold used for cross-fitting and aggregated across folds. Monotonically increasing trend indicates a good fit. Average treatment effects of the exposure were estimated via the doubly-robust targeted maximum likelihood estimation. Models were estimated data-adaptively via the SuperLearner using generalized linear models, gradient boosting machine, and neural net as candidate estimators. All models were adjusted for the 55 pre-disaster demographic and socioeco-nomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. IADL scores ranged from 0-13 points, where smaller scores indicate lower functional independence. Levels of certified physical disability ranged from 1 (“Cannot roll over in a bed independently”) to 9 (“no physical disability/not requesting care services”), where smaller values indicate greater levels of disability. Thus, decrease in these outcomes indicate increased functional limitation.



Web Figure 4. Sensitivity analysis for conditional average treatment effects of housing damage on functional limitation in 2013 and 2016.

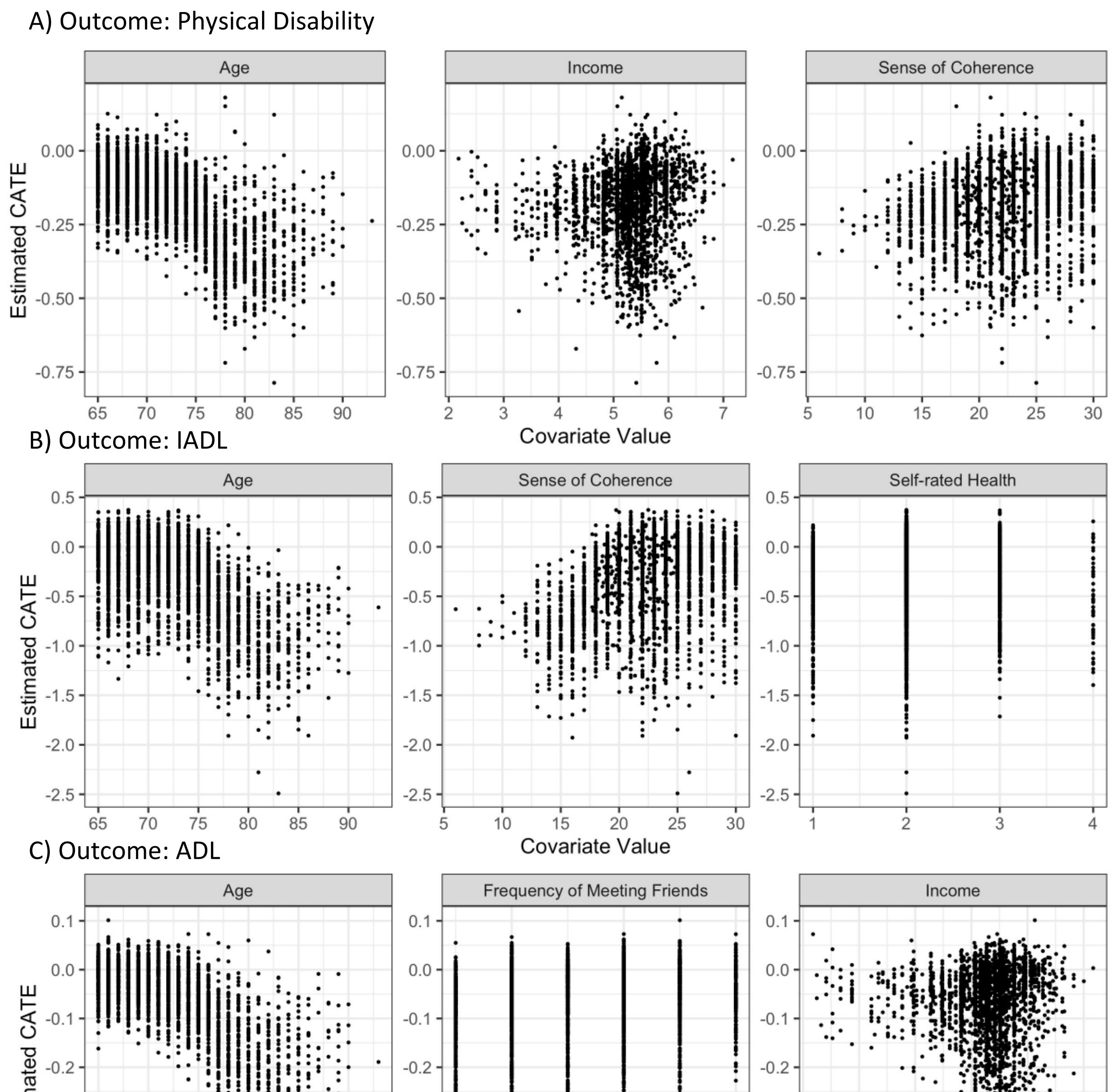
Abbreviations: ADL, Activity of Daily Living; IADL, Instrumental Activity of Daily Living.

Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. Scores ranged from 0-13 points for total IADL, where smaller scores indicate lower functional independence. Levels of certified physical disability ranged from 1 (“Cannot roll over in a bed independently”) to 9 (“no physical disability/not requesting care services”), where smaller values indicate greater levels of disability. Thus, decrease in these outcomes indicate increased functional limitation.



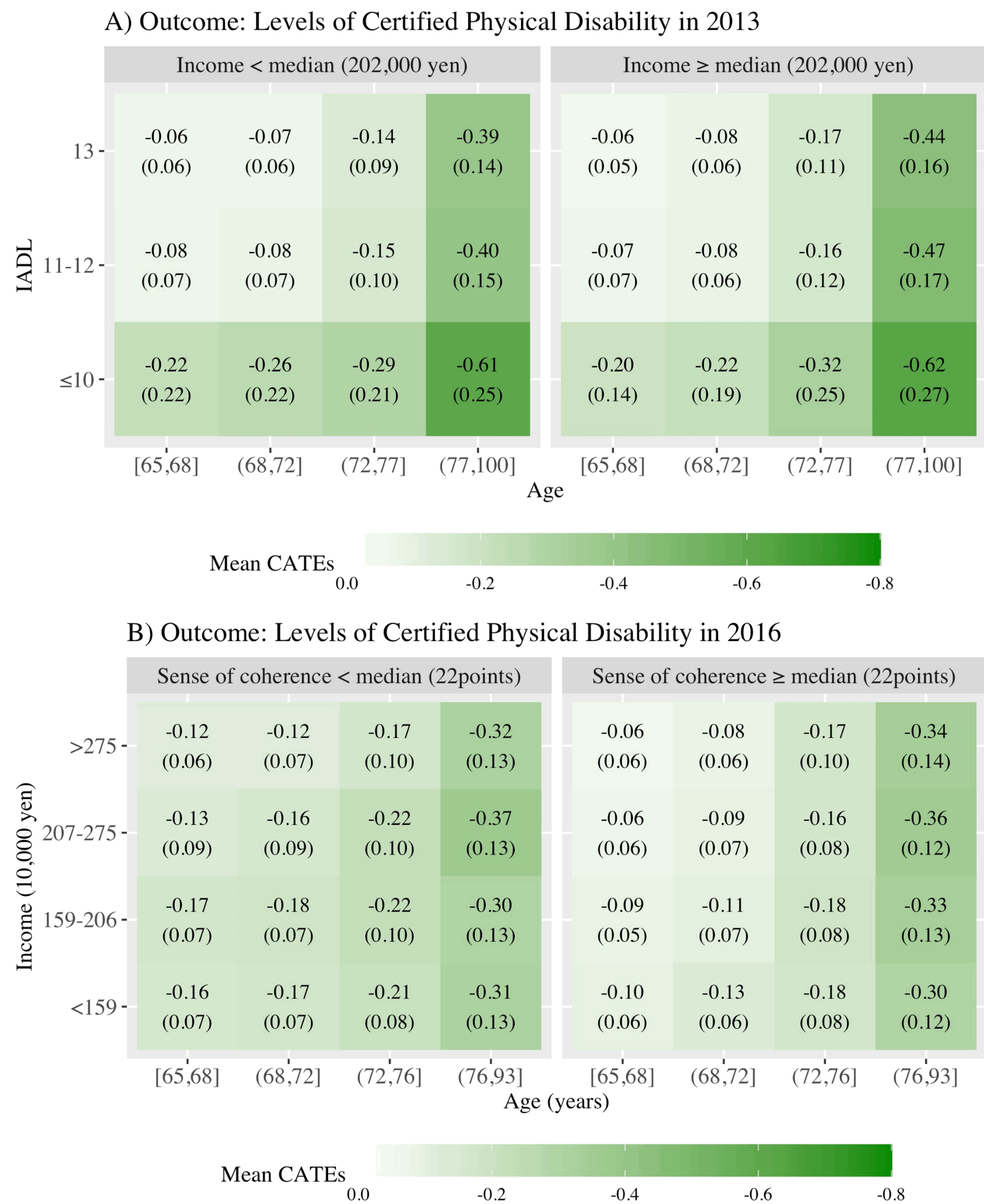
Web Figure 5. Distribution of estimated conditional average treatment effects of home loss on functional limitation outcomes in 2013 by covariate values.

The three covariates were chosen based on the variable importance ranking shown in Web Table 11. Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. Decrease in the outcomes indicates increased functional limitation; hence, smaller CATEs indicates more adverse impacts of home loss.



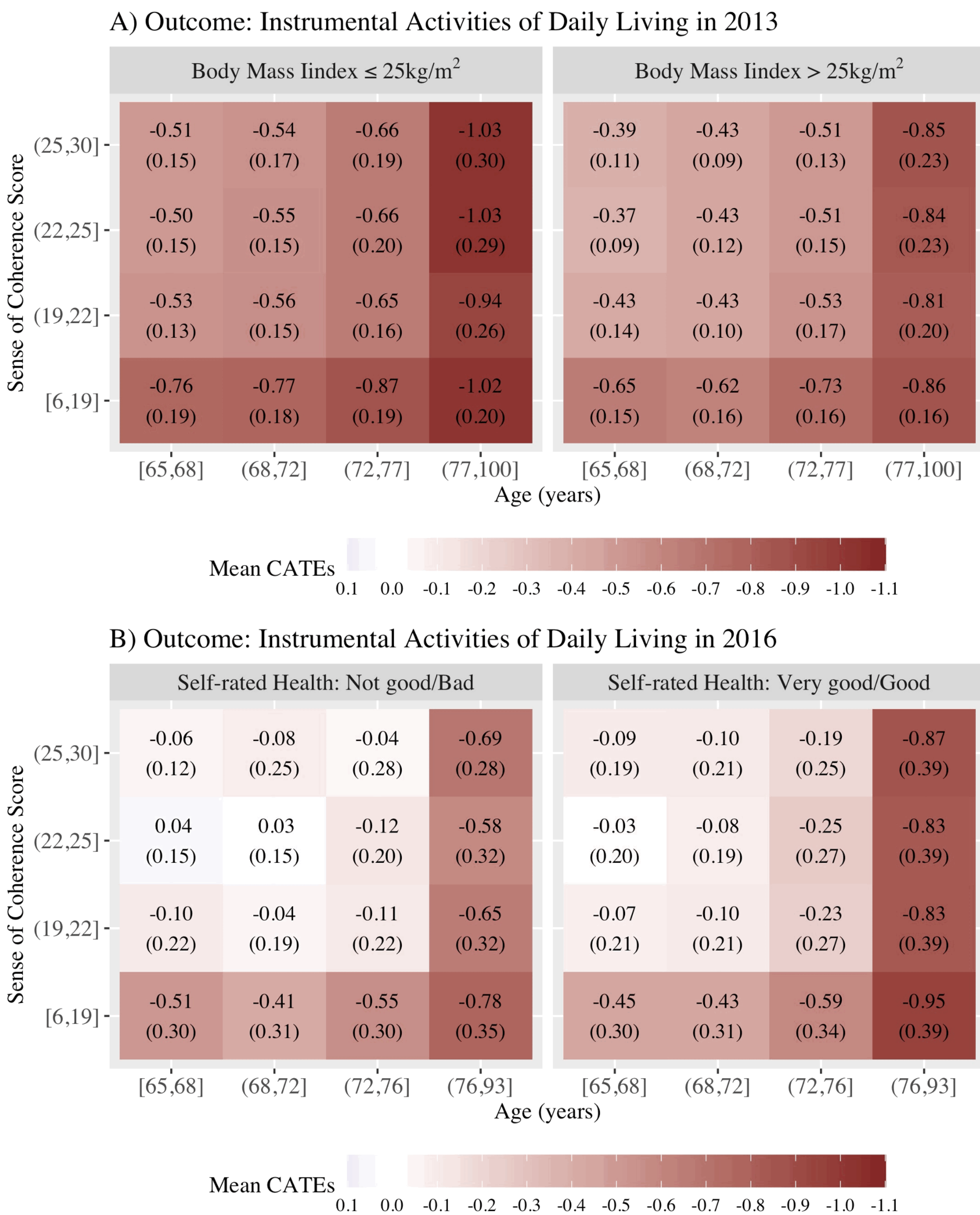
Web Figure 6. Distribution of estimated conditional average treatment effects of home loss on functional limitation outcomes in 2016 by covariate values.

The three covariates were chosen based on the variable importance ranking shown in Web Table 11. Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. Decrease in the outcomes indicates increased functional limitation; hence, smaller CATEs indicates more adverse impacts of home loss.



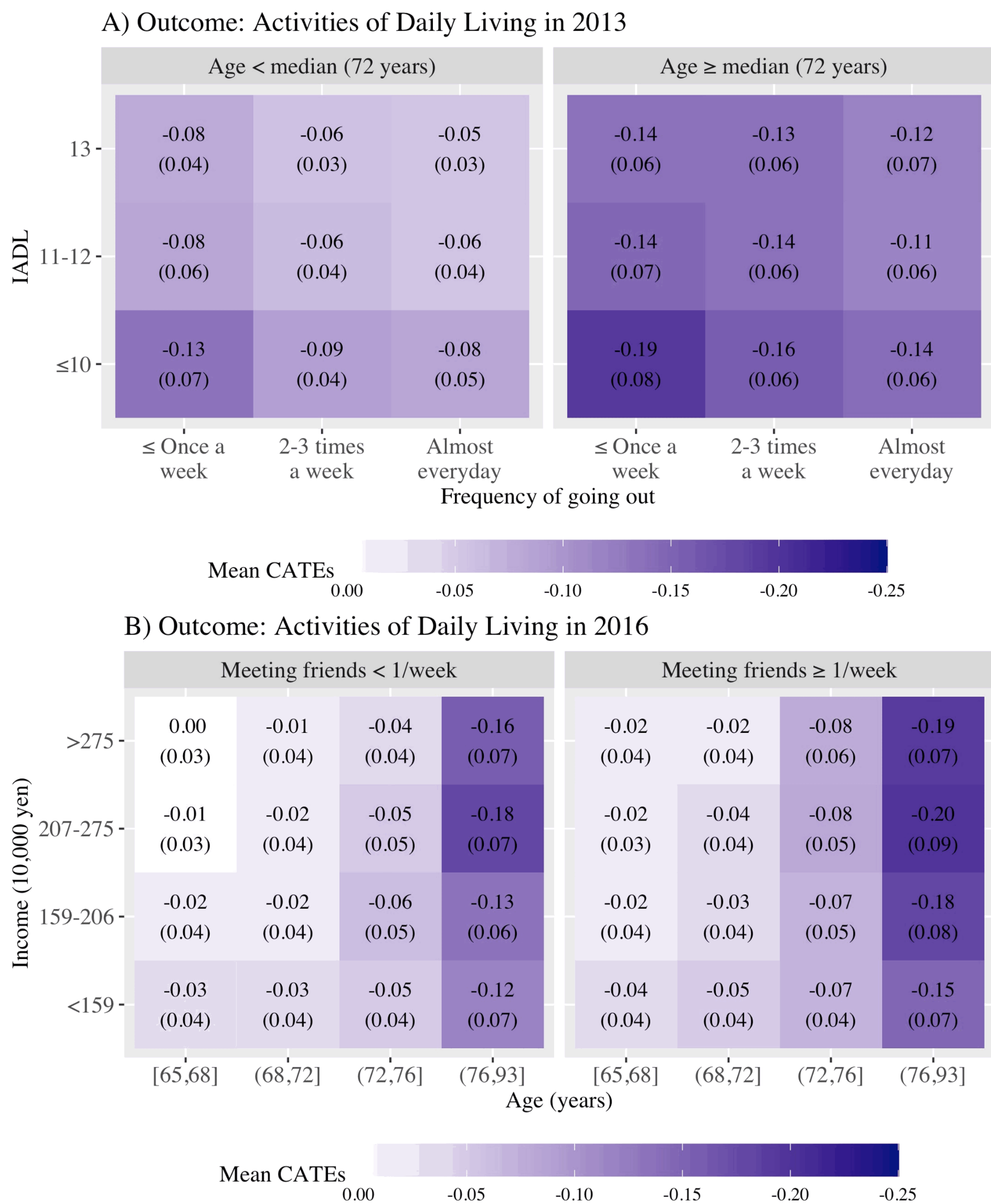
Web Figure 7. Heatmap showing the distribution of estimated conditional average treatment effects of home loss on physical disability levels in 2013 and 2016.

The three covariates were chosen based on the variable importance ranking shown in Web Table 11. Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. Levels of certified physical disability ranged from 1 ("Cannot roll over in a bed independently") to 9 ("no physical disability/not requesting care services"), where smaller values indicate greater levels of disability. Thus, decrease in the outcome indicates increased functional limitation.



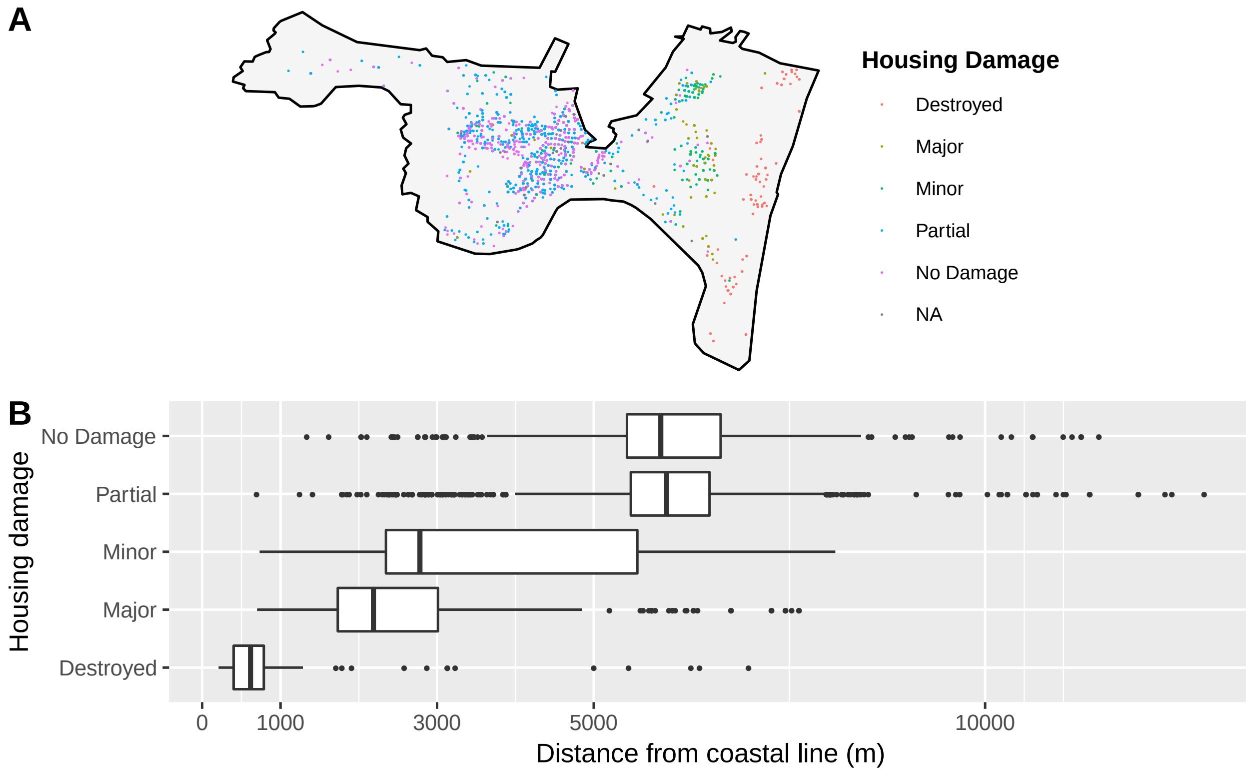
Web Figure 8. Heatmap showing the distribution of estimated conditional average treatment effects of home loss on instrumental activities of daily living in 2013 and 2016.

The three covariates were chosen based on the variable importance ranking shown in Web Table 11. Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. Scores ranged from 0-13 points for total IADL, where smaller scores indicate lower functional independence. Thus, decrease in the outcome indicates increased functional limitation.



Web Figure 9. Heatmap showing the distribution of estimated conditional average treatment effects of home loss on activities of daily living in 2013 and 2016.

The three covariates were chosen based on the variable importance ranking shown in Web Table 11. Heterogeneous effects (i.e., conditional average treatment effects: CATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. ADL had three levels (1= support needed completely, 2 = support needed partially, and 3 = no help needed). Thus, decrease in the outcome indicates increased functional limitation.



Web Figure 10. Relationship between degree of property damage and distance from the coastal line. A) Geographic distribution of housing damage in 2013; B) Association between pre-earthquake distance from the coast and housing damage.

References

1. Stekhoven DJ, Stekhoven MDJ. Package 'missForest.' *cran.r-project.org › web › packages › missForestcran.r-project.org › web › packages › missForest*. 2012. https://cran.r-project.org/web/packages/missForest/missForest.pdf.

2. Breiman L. Random Forests. *Mach Learn*. 2001;45(1):5-32. doi:10.1023/A:1010933404324.

3. Shiba K, Kawahara T, Aida J, et al. Causal Inference in Studying the Long-term Health Effects of Disasters: Challenges and Potential Solutions. *Am J Epidemiol*. doi:10.1093/aje/kwab064.