Collecting Person-Generated Health Data in the UAS

Arie Kapteyn

The research reported in this presentation has been supported by grants from the Social Security Administration and the National Institute on Aging 5U01AG054580)

UAS "New" Data

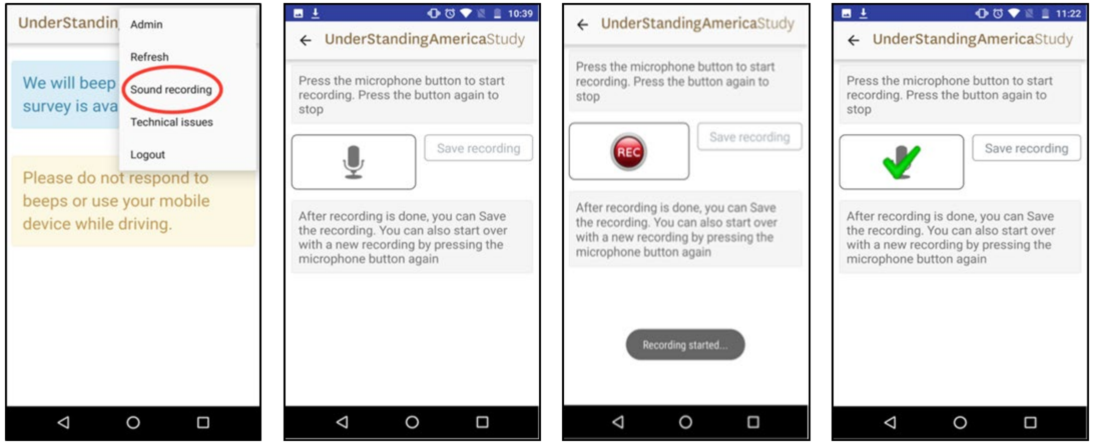
- Linguistic sampling
- Paradata for measuring cognition
- Accelerometers
  - GENEActive/Fitbit in various combinations
  - Sleep and cognition
- Air quality

Linguistic Sampling of Daily Social Experience

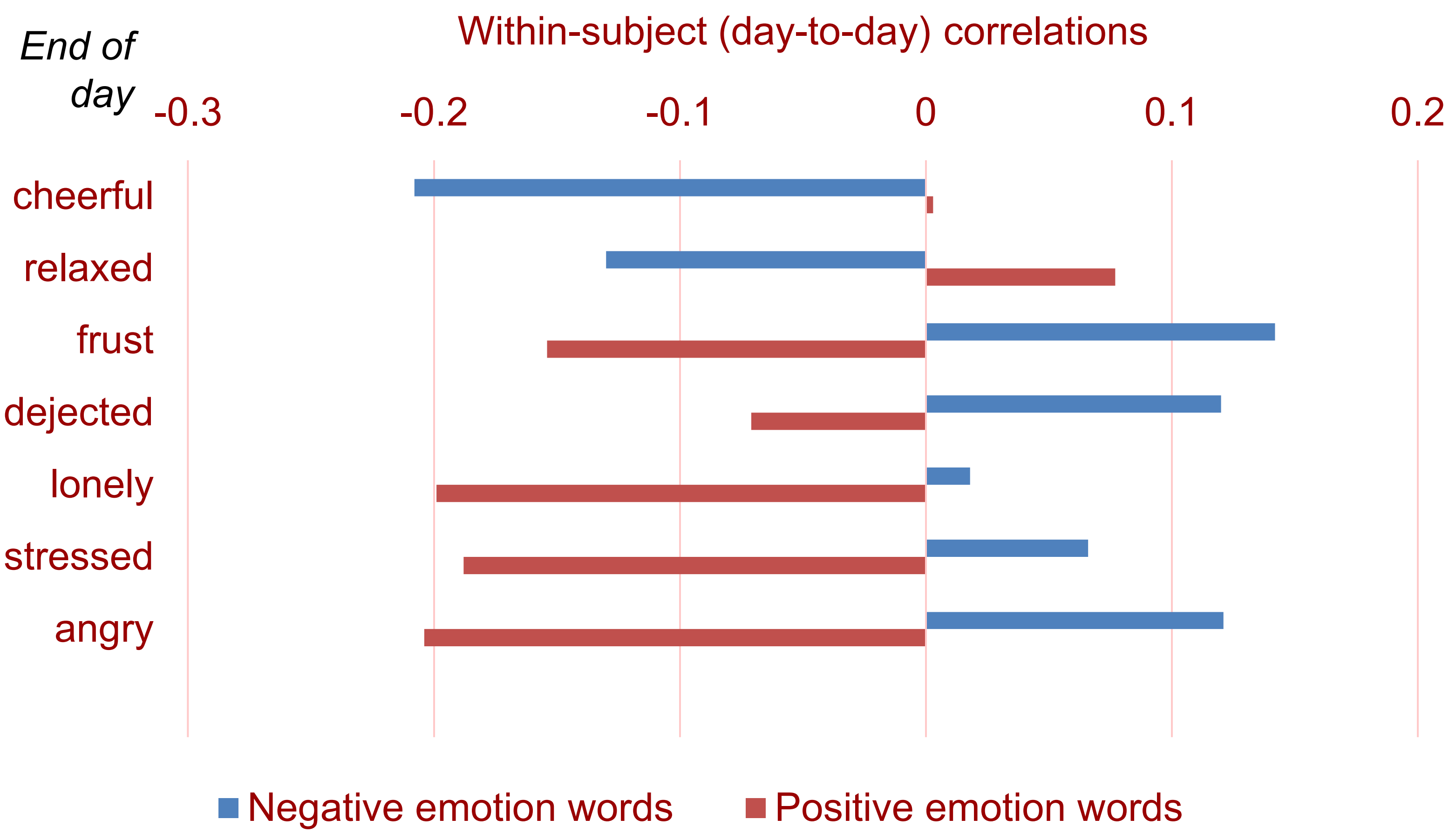
Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). *The development and psychometric properties of LIWC-22*. Austin, TX: University of Texas at Austin. https://www.liwc.app

How we do it:
Linguistic Sampling Instructions

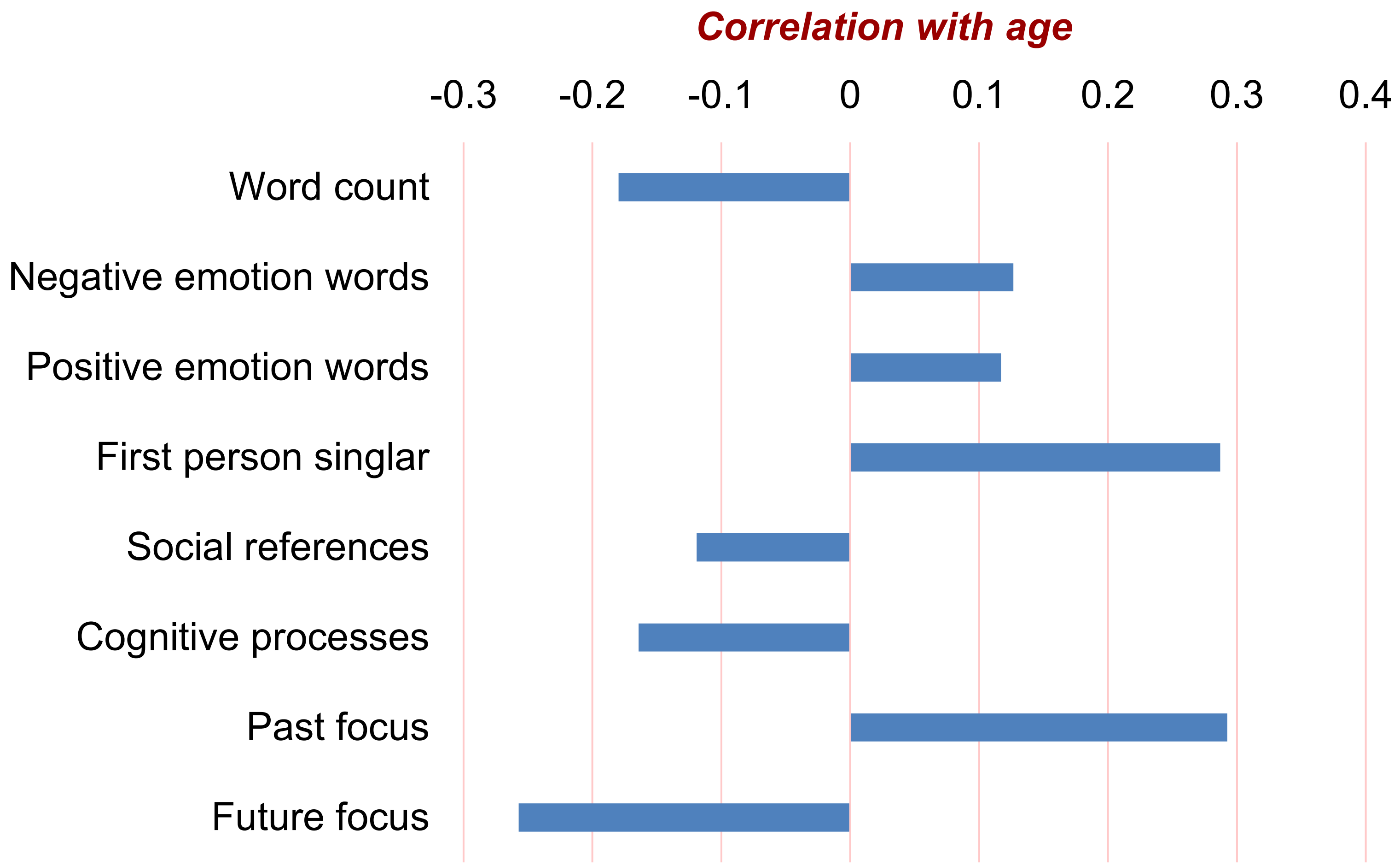
"We would like you to spend 2-5 minutes telling us about your day. We are especially interested in hearing if you interacted with other people today, and if you did, how you felt during those interactions. You do not have to write it up in advance, but can simply talk about what you did today, starting from if you had breakfast alone or with anyone, all the way through to your evening activities with or without other people."



Correlations of Emotion Words with Self-reported End-of-Day Emotion Ratings



Age Differences in Word Use



Analysis of sound files is next

Paradata and Cognition (examples)

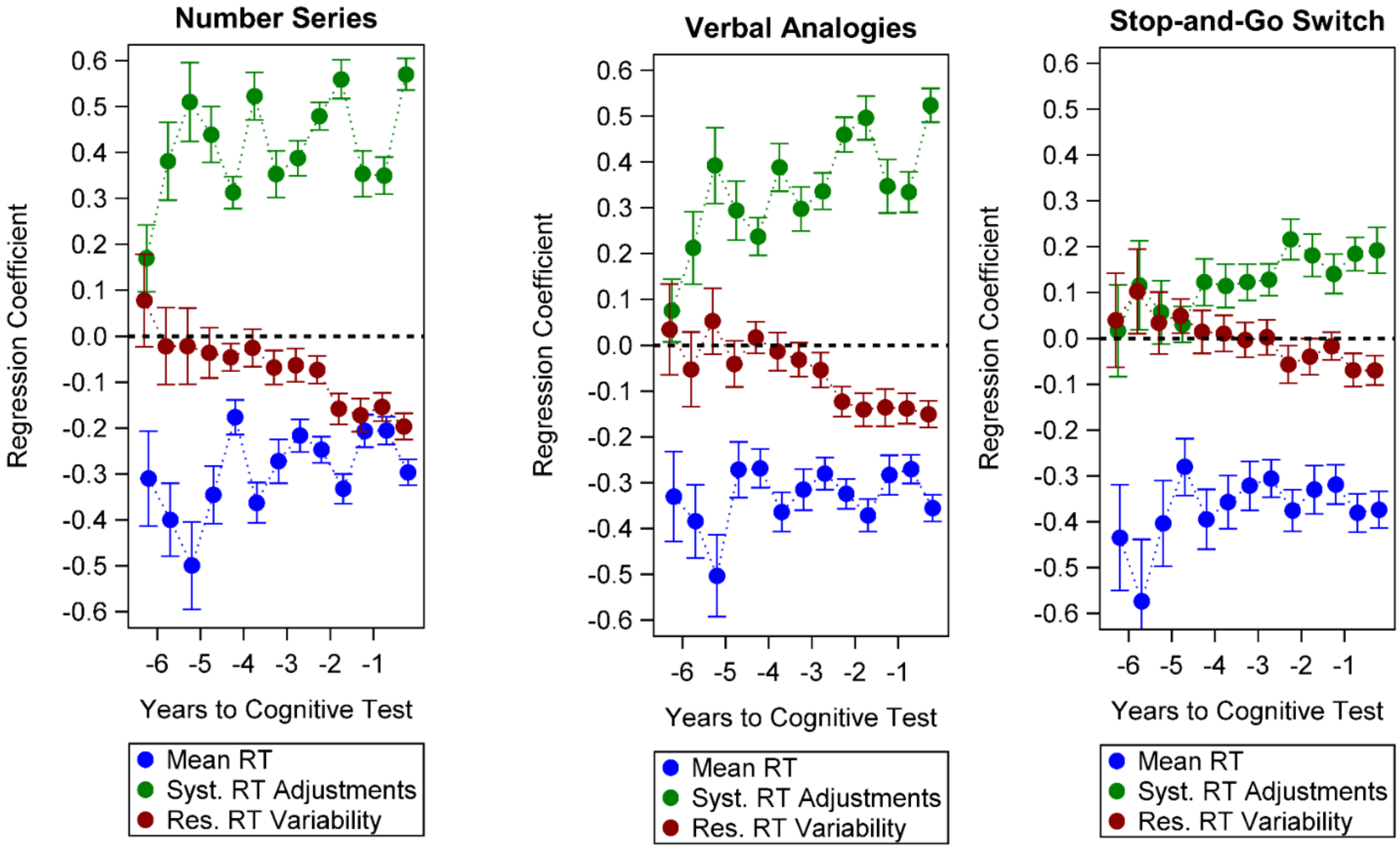
Response times (both means and variances) explain 22-26% of variance in scores on later (up-to-six years) cognitive tests.

Junghaenel, D.U., Schneider, S., Orriens, B., Jin, H., Lee, P.J., Kapteyn, A., Meijer, E., Zelinski, E.M., Hernandez, R., Stone, A.A. (2023). "Inferring cognitive abilities from response times to web-administered survey items in a population-representative sample". *Journal of Intelligence*, 11, 3

Furthermore, mistakes people make in answering questions predict cognitive decline.

Schneider, S., Junghaenel, D.U., Zelinski, E.M., Meijer, E., Stone, A.A. Langa, K.M., Kapteyn, A., "Subtle Mistakes in Self-report Surveys Predict Future Transition to Dementia", *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 2021. **13**(1)

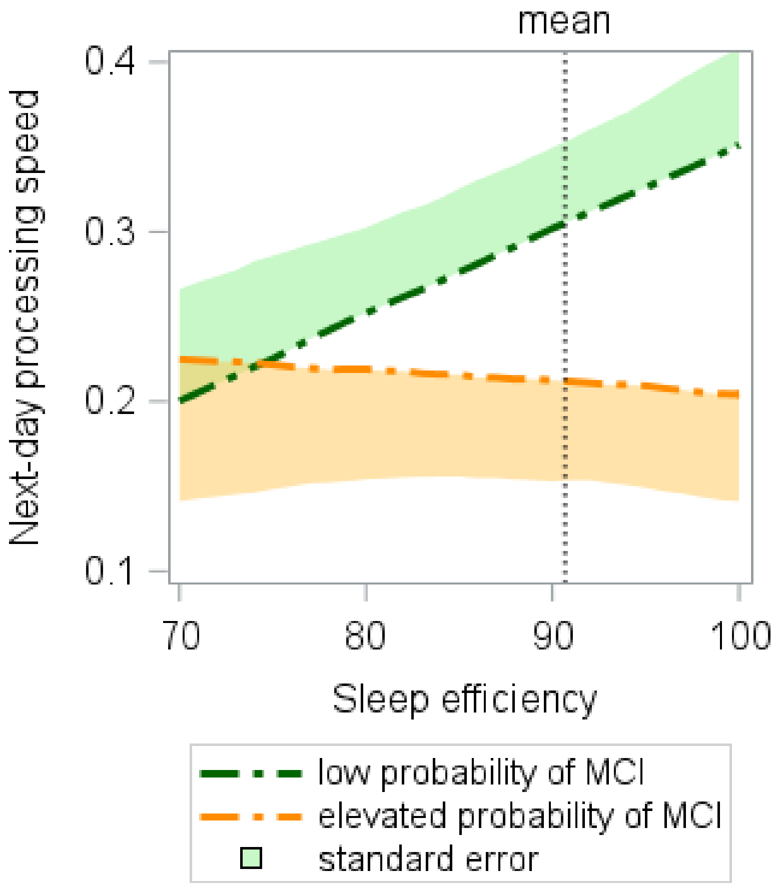
Past response times predict cognitive scores



This motivates the use of paradata as cognitive indicators

Sleep Efficiency and Processing Speed

- A higher sleep efficiency, **measured by Fitbit**, on the previous night predicted better processing speed on the next day for cognitively normal respondents ($p<.05$) but not for those with a higher probability of MCI ($p=.80$)



Measuring Physical Activity

Comparing Physical Activity

- Kapteyn, A., Banks, J., Hamer, M., Smith, J.P., Steptoe, A., Van Soest, A.H.O., Koster, A., Saw H.W., “What They Say and What They Do: Comparing Physical Activity Across U.S., England, and the Netherlands”, *Journal of Epidemiology & Community Health* 72, 2018, 471–476

Our study compared three countries

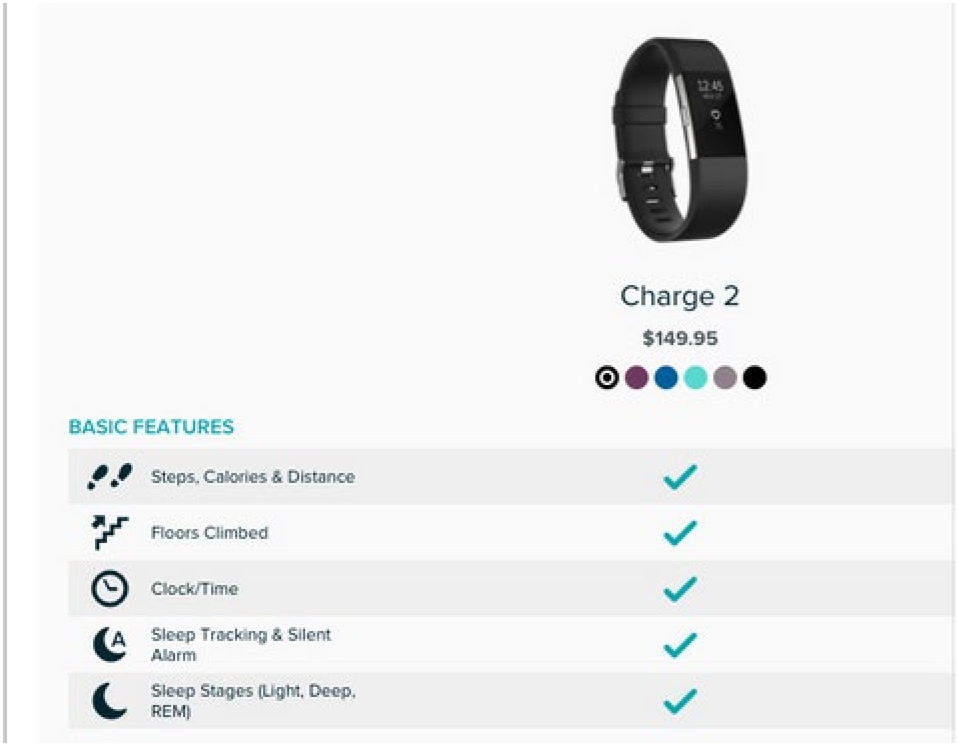
- Netherlands (NL)
- United States (US)
- England
- We use two probability-based Internet panels:
  - Longitudinal Internet Studies for the Social Sciences (LISS, NL)
  - Understanding America Study (UAS, US)
- And one panel of individuals 50+, who are interviewed face-to-face
  - English Longitudinal Study of Aging (ELSA)

Results

- The accelerometers (GENEActiv) show significant and substantial differences across countries and across age groups
- Self-reports don’t show these differences

Pros and Cons of Fitbit

- Measures several features, but no access to raw data.
- Commercially available:
  - May save money
- Provides feedback



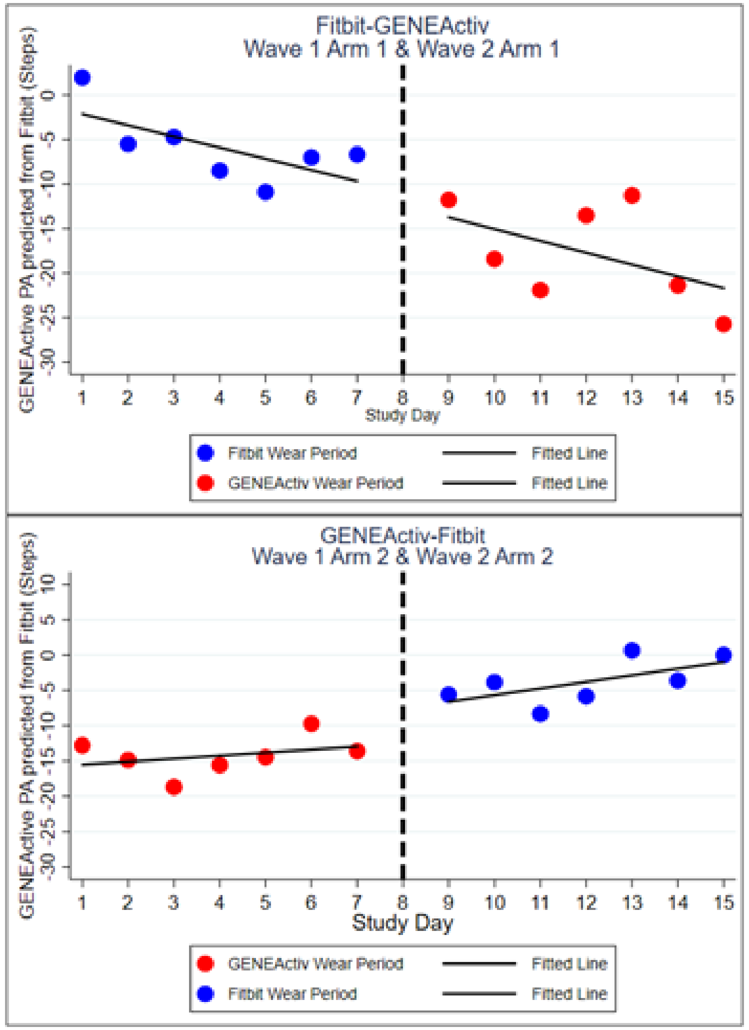
Does Feedback from Activity Trackers influence Physical Activity? Evidence from a Randomized Controlled Trial

Jill Darling
Arie Kapteyn
Htay-Wah Saw

This research is supported by the National Institute on Aging and the Social Security Administration under NIA grant 5U01AG054580

USC Dornsife Center for Economic and Social Research
University of Southern California

Graphical Analysis of Feedback

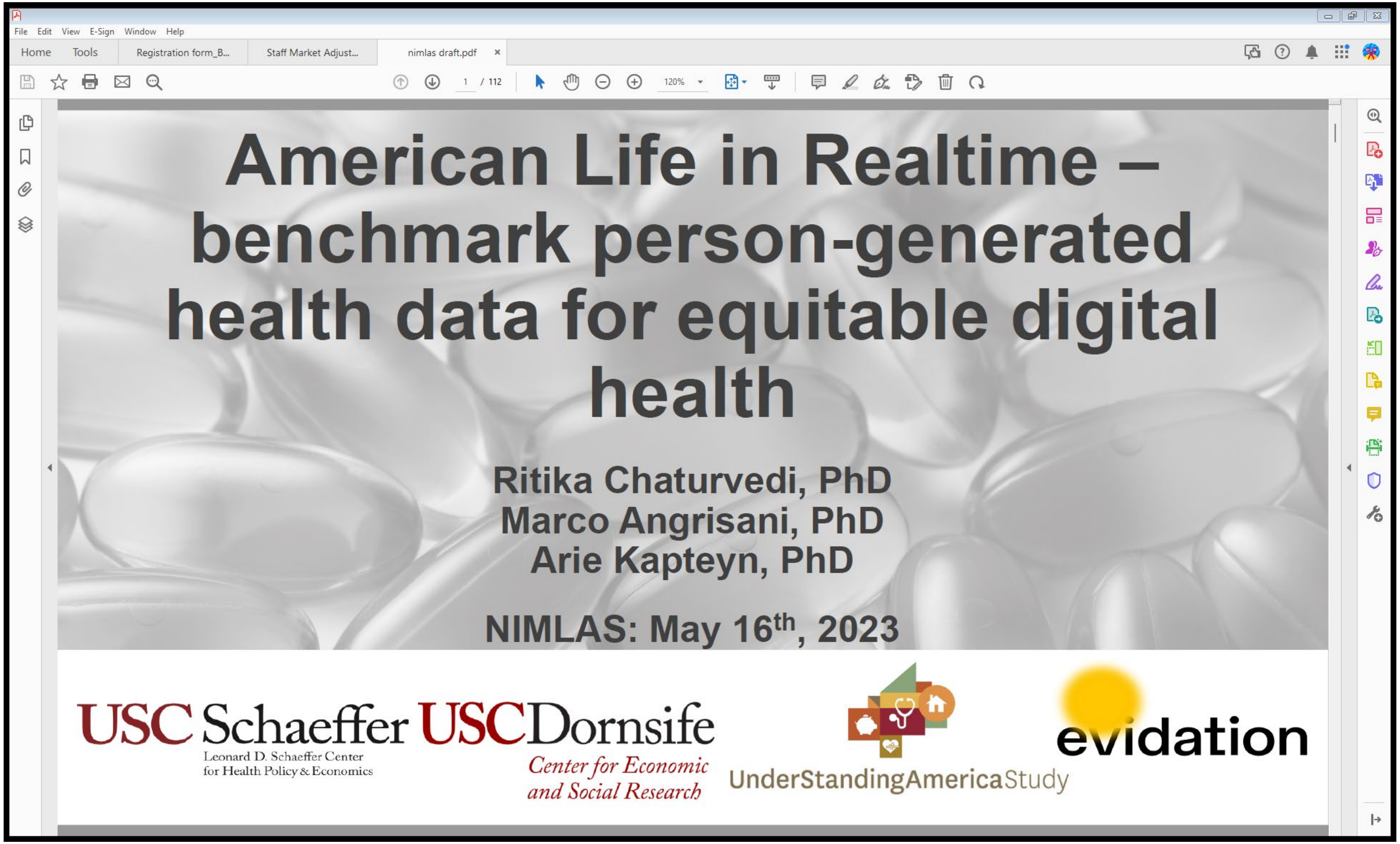


USCDornsife
Center for Economic and Social Research

University of Southern California

The next project relies entirely on Fitbits

- American Life in Realtime (ALiR)
- About 1,000 UAS respondents have received a Fitbit and were asked to wear it for at least a year
- The aim is to use Fitbit as a "digital biomarker" and measure relations in a population representative panel



File Edit View E-Sign Window Help

Home Tools Registration form_B... Staff Market Adjust... nimlas draft.pdf

1 / 112 120%

American Life in Realtime – benchmark person-generated health data for equitable digital health

Ritika Chaturvedi, PhD
Marco Angrisani, PhD
Arie Kapteyn, PhD

NIMLAS: May 16th, 2023

USC Schaeffer
Leonard D. Schaeffer Center for Health Policy & Economics

USC Dornsife
Center for Economic and Social Research

UnderStandingAmericaStudy

evidation

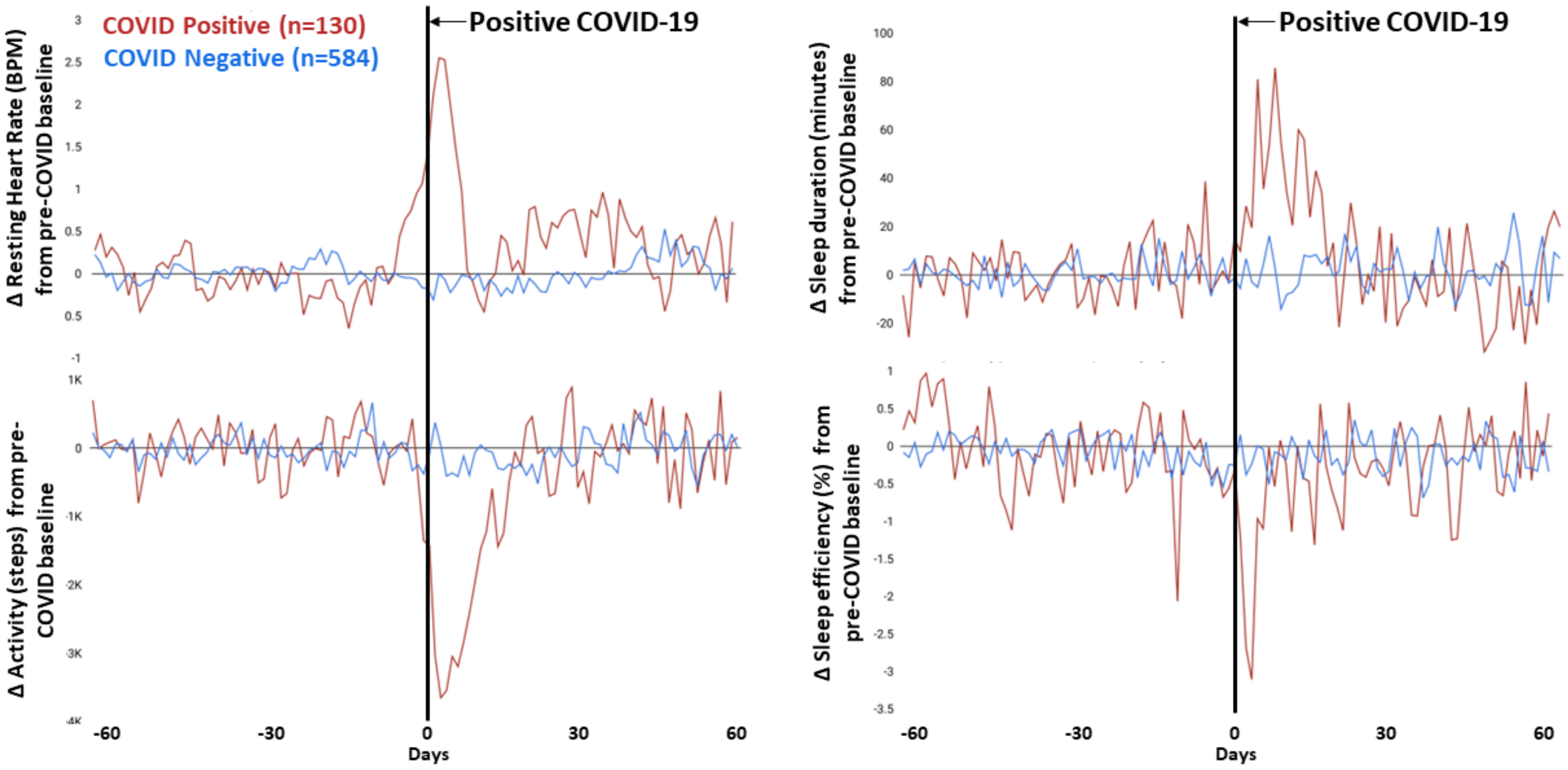
The Motivation of ALiR is to avoid the pitfalls of using large and selective datasets

- There are several large and ambitious projects using volunteer samples, for instance:
  - UK Biobank
  - All of US in the U.S.
- There is doubt how generalizable conclusions from these studies are and how they may reinforce disparities.
- We observe large differences between owners and non-owners of Fitbits

Carol Brayne & Terrie E. Moffitt (2022), “The limitations of large-scale volunteer databases to address inequalities and global challenges in health and aging” *Nature Aging* volume 2, 775–783 (2022)

ALiR picks up COVID infections

Changes in biometrics from individual-specific baselines during COVID infection



We are planning a substantial expansion of the sample size.

Analyses are only just starting.

Measuring Air Quality with Wearable Devices

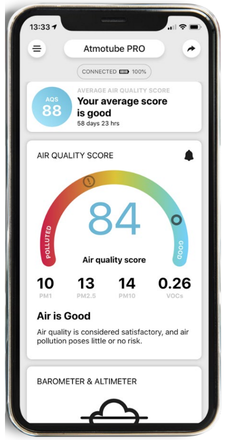
Arie Kapteyn
Center for Economic and Social Research, University of Southern California

Htay-Wah Saw
Center for Economic and Social Research (CESR), University of Southern California &
Michigan Program in Survey and Data Science, University of Michigan-Ann Arbor

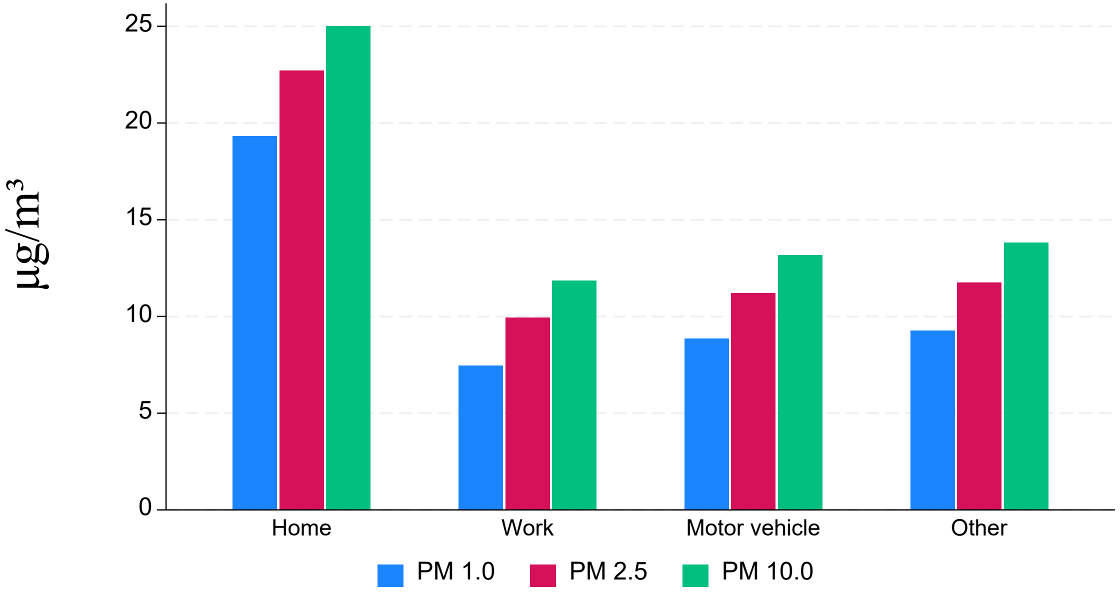
Bas Weerman
Center for Economic and Social Research, University of Southern California

Study Background

- Pilot started in June 2021
- Respondents wear air quality monitor (Atmotube) continuously
- About the device:
  - Bluetooth enabled
  - Communicates with a smartphone app
  - Collect pollution data (PM 1.0, PM 2.5, PM 10.0, VOC) and weather data (temperature, pressure, humidity) at 1-minute intervals
  - App uploads pollution data to server

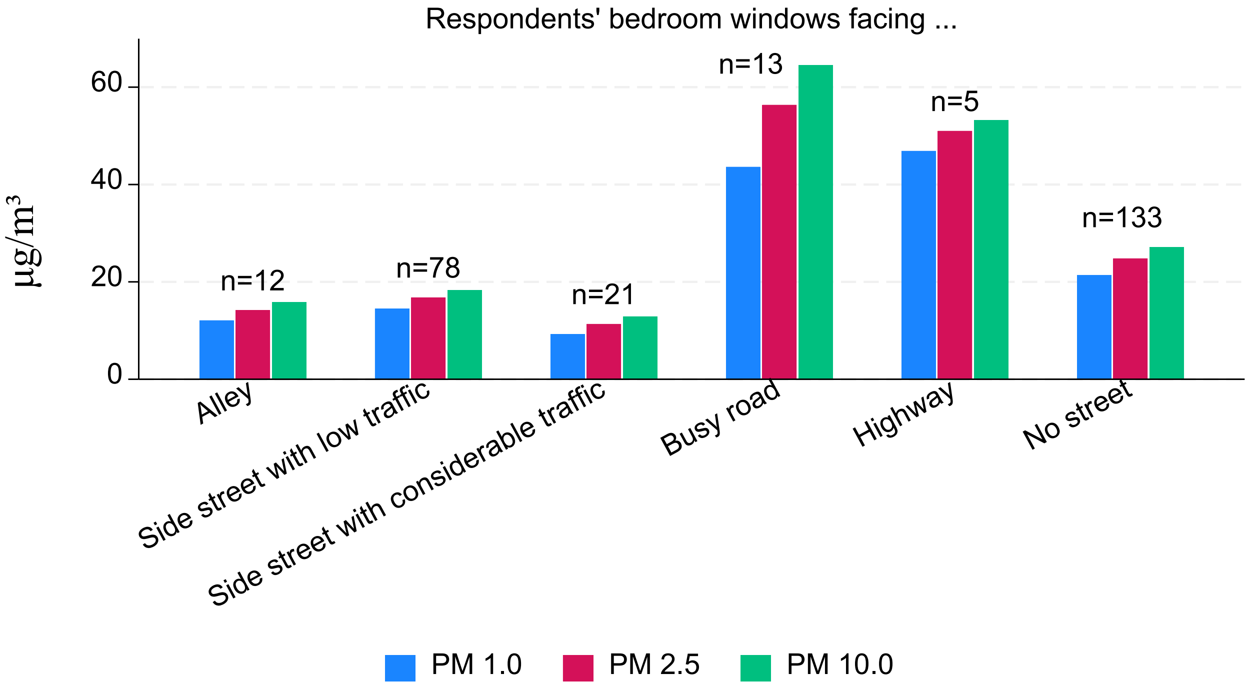


Average Pollution by Respondents' Location Yesterday

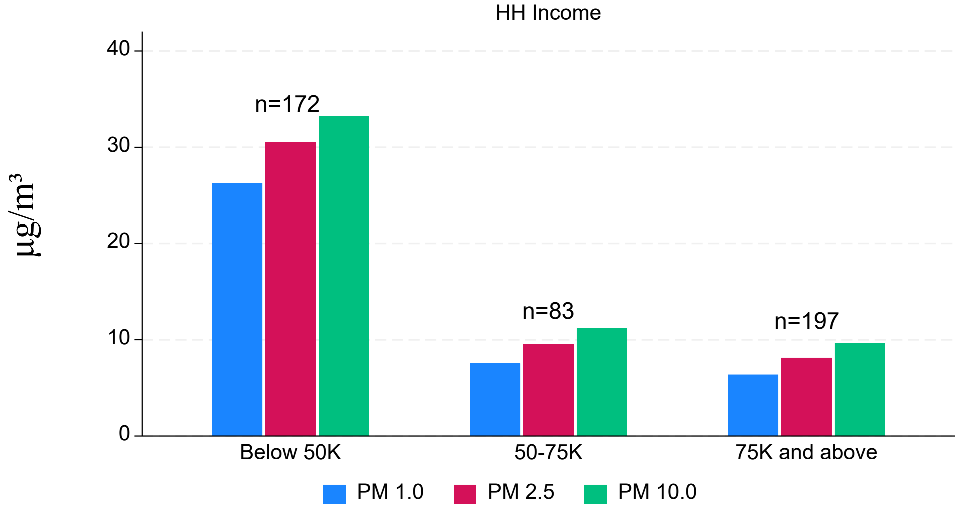


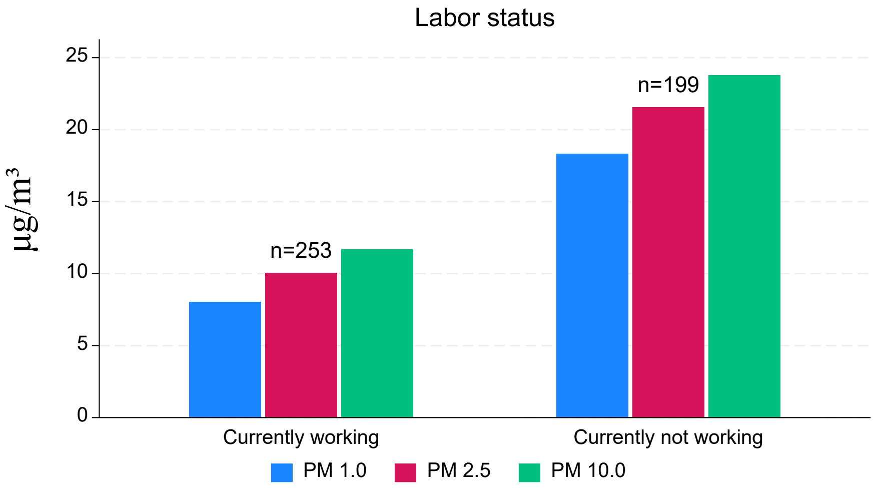
Average pollution by respondents' housing characteristics

- Living next to a busy road may be least healthy
- A graph for living room windows shows similar results

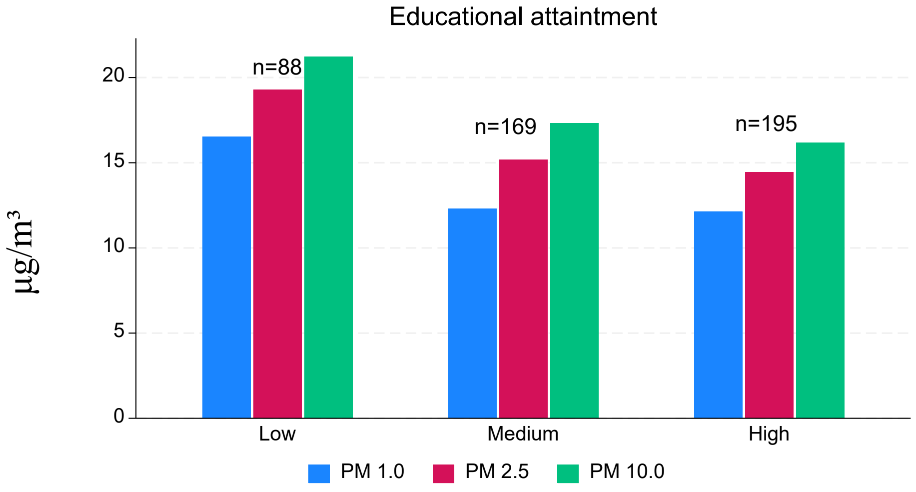


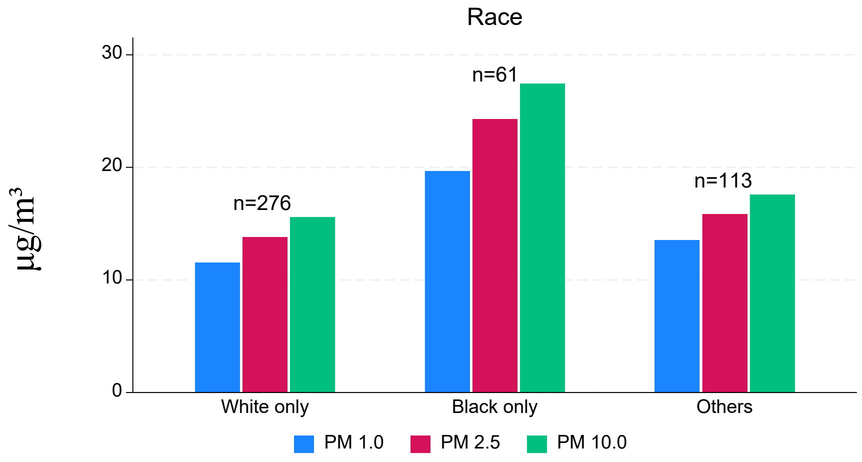
Average pollution by household income and employment status



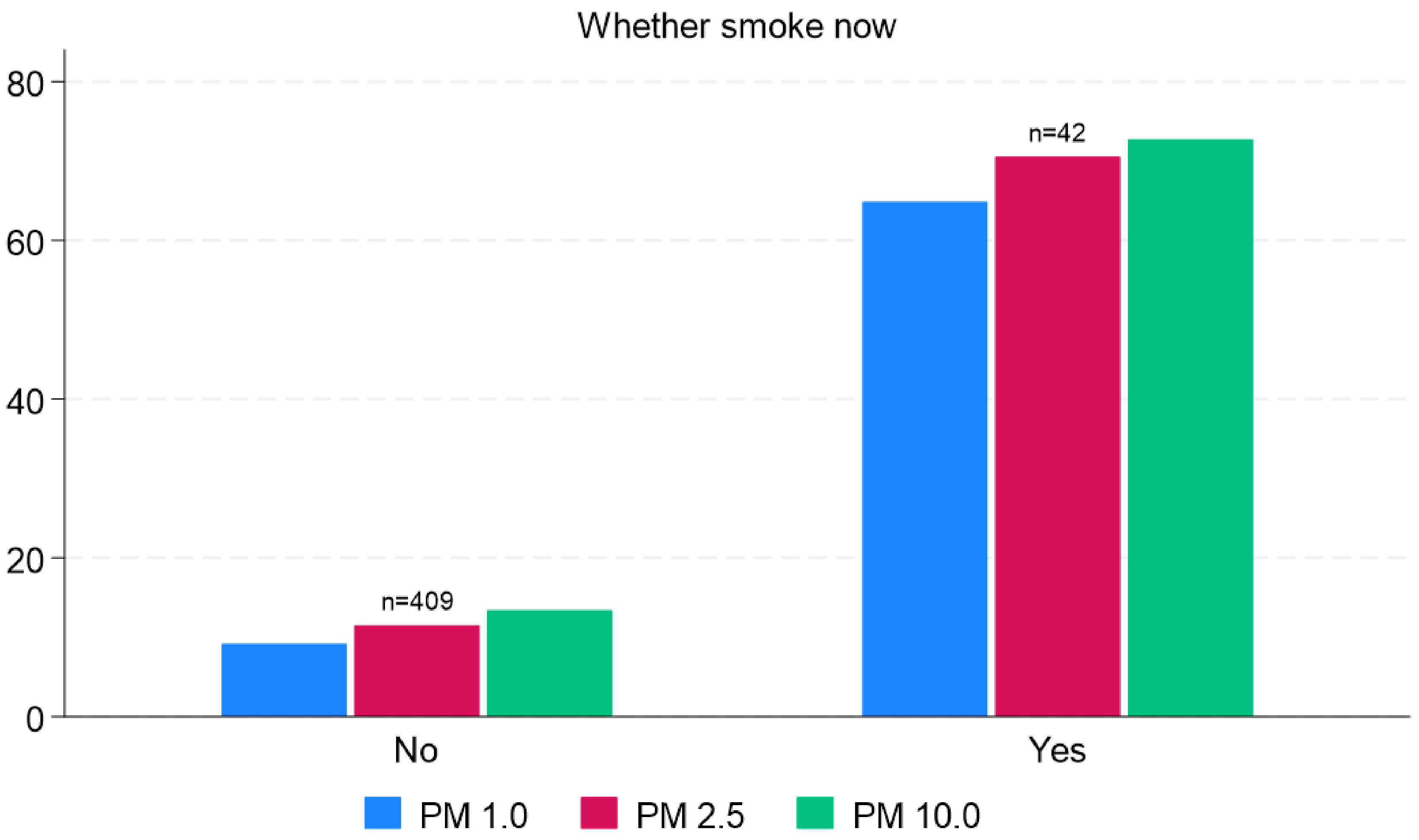


Average pollution by education and race/ethnicity

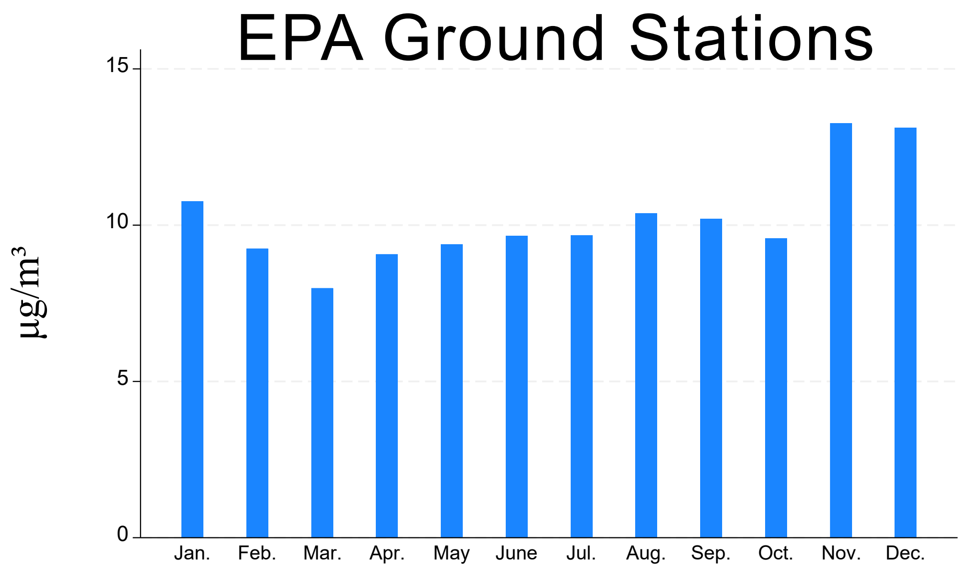


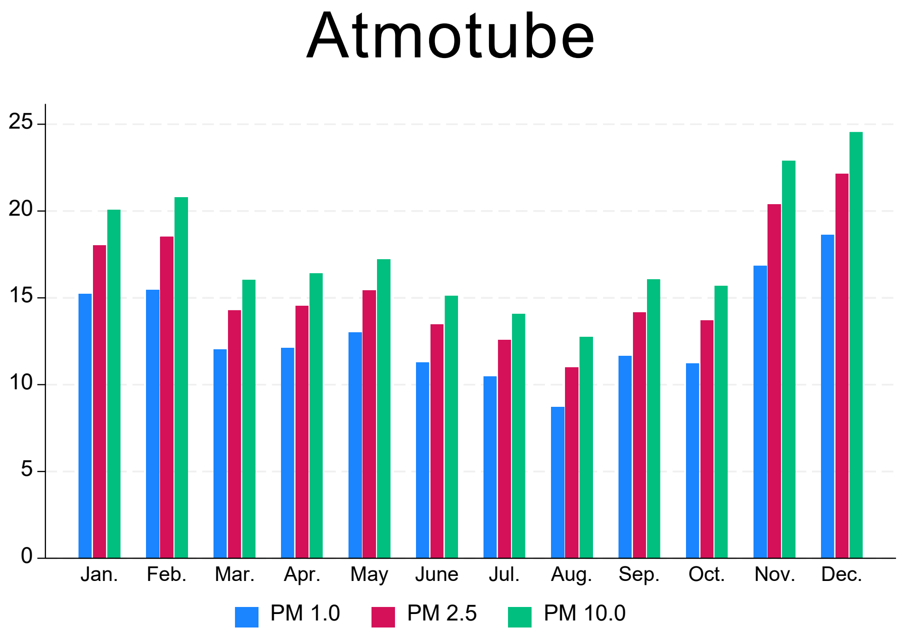


Smoking

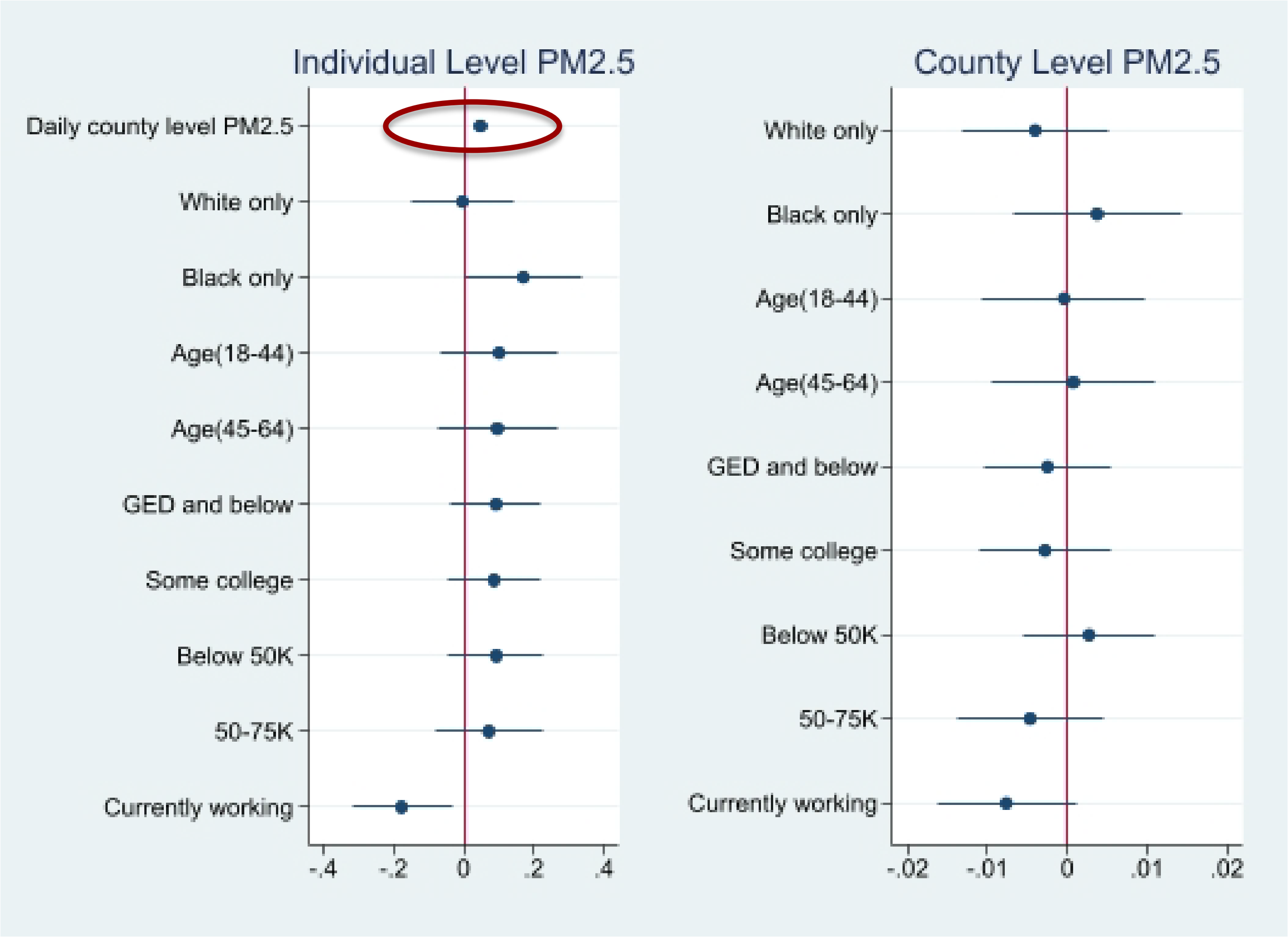


Average pollution by month of the year





Explaining PM2.5 (RE regressions)



First impression

- Not much correlation between individually worn monitors and the EPA monitors
- The EPA monitors don't seem to be very informative about invidual exposure to bad air
  - But note, these are county averages
- We still have to model the air quality by Census Tract

Next Steps

- Recruit up to 1000 respondents, and collect their pollution data up to one year
- Model daily air quality measures by Census Tract based on EPA ground station measures.
- Substantive analysis:
  - relate exposure to air pollution to health and cognitive outcomes, racial and socioeconomic differences in exposure to air pollution

Thank you!

Consent and Compliance

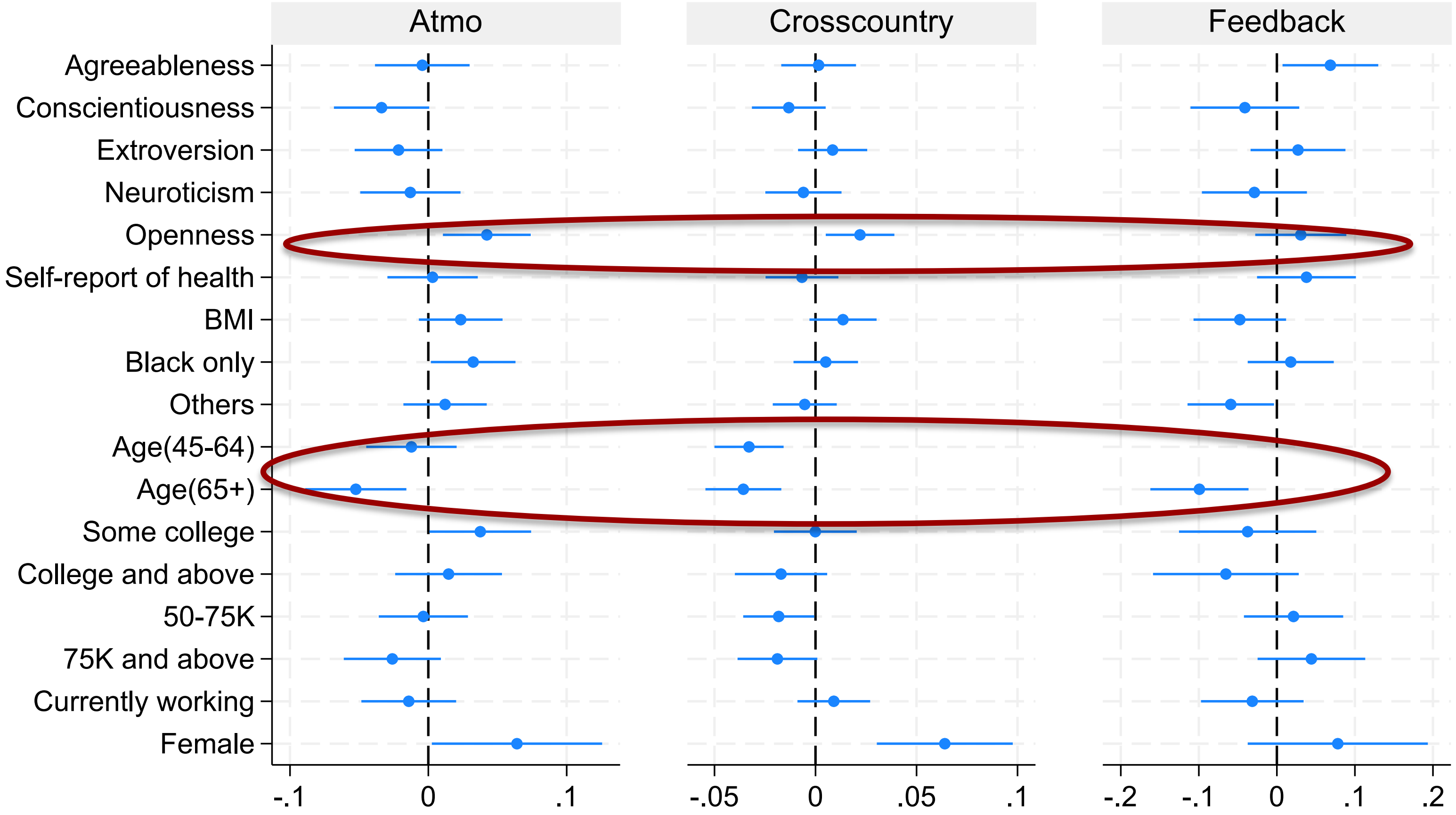
Arie Kapteyn
Center for Economic and Social Research, University of Southern California

Htay-Wah Saw
Center for Economic and Social Research (CESR), University of Southern California &
Michigan Program in Survey and Data Science (MPSDS), University of Michigan-Ann Arbor

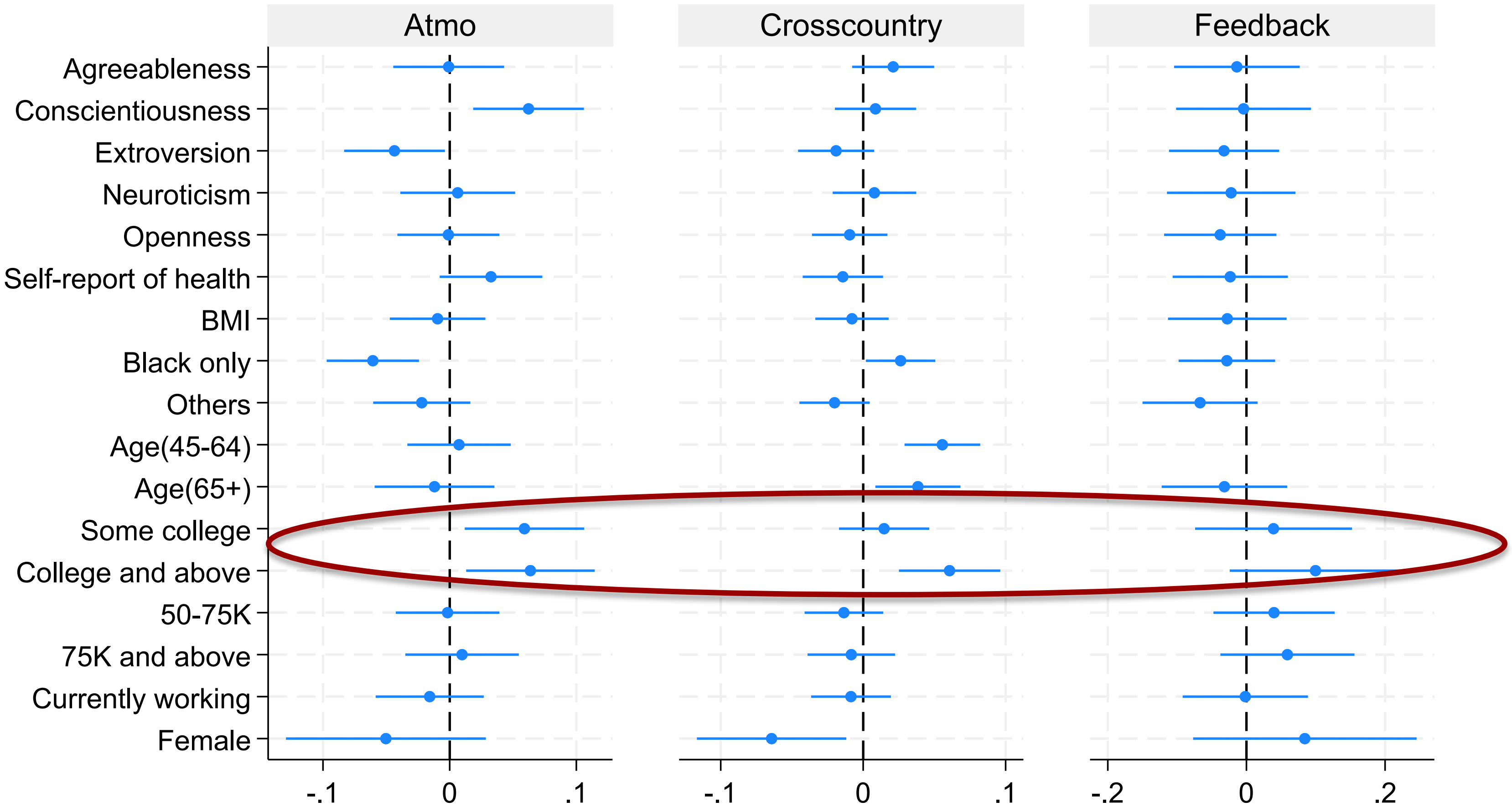
Overview: Three Experiments Involving Three Different Wearables

Experiment	Name	Device	Year	Country
1	Cross-country accelerometry study	GENEActiv	2015-2017	US, UK, Netherlands
2	Feedback experiment study	GENEActiv + Fitbit	2019-2020	US
3	Measuring pollution using wearables	Atmotube Pollution Device	2020-present	US

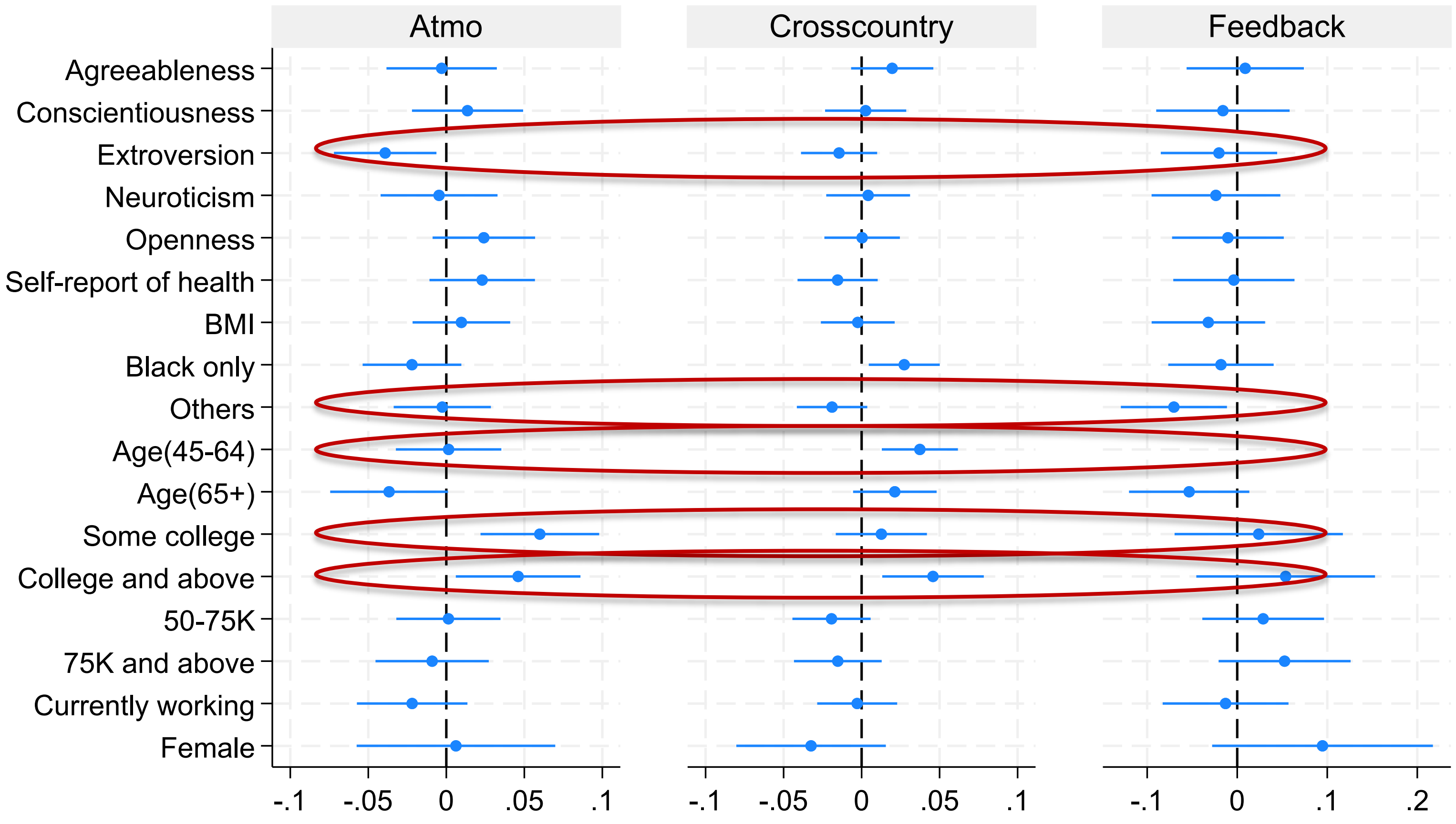
Consent Regression



Data Provision



Total Effects

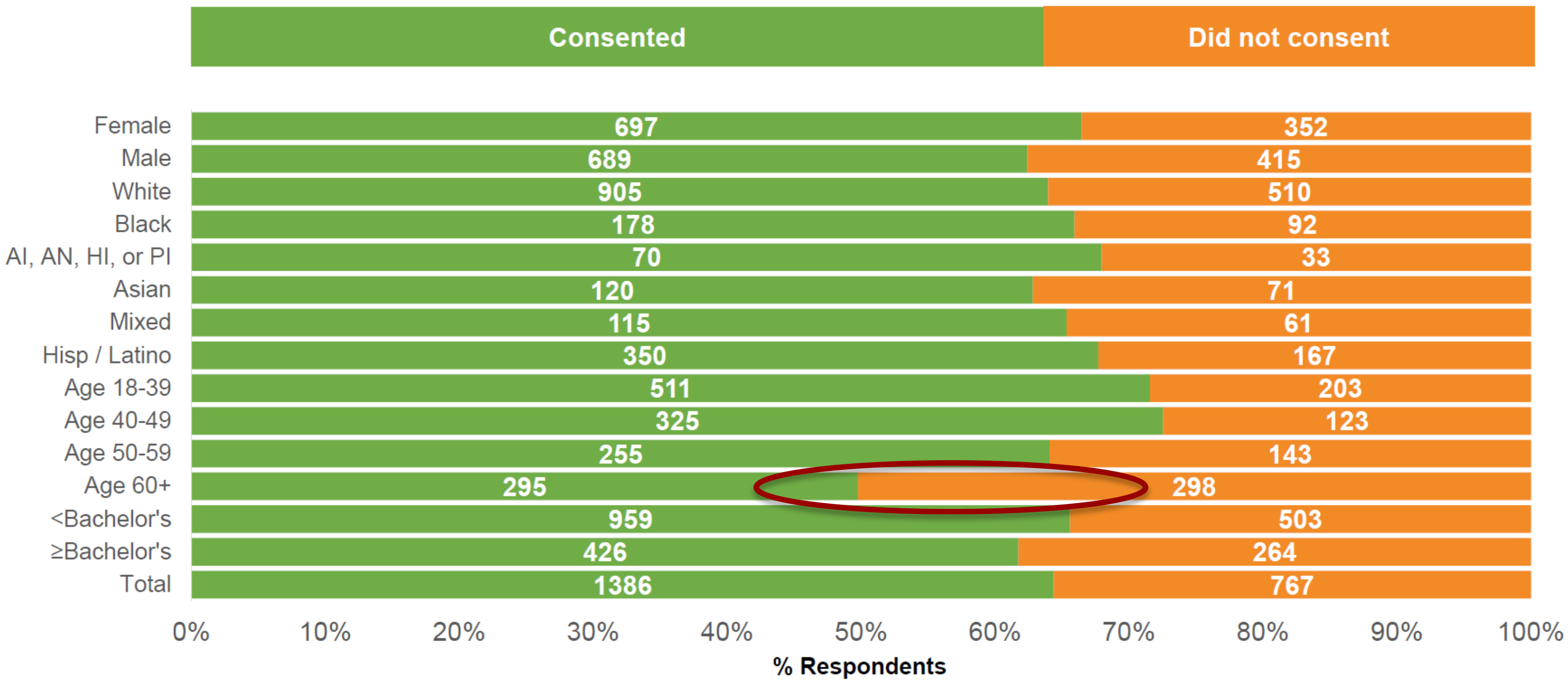


Impressions

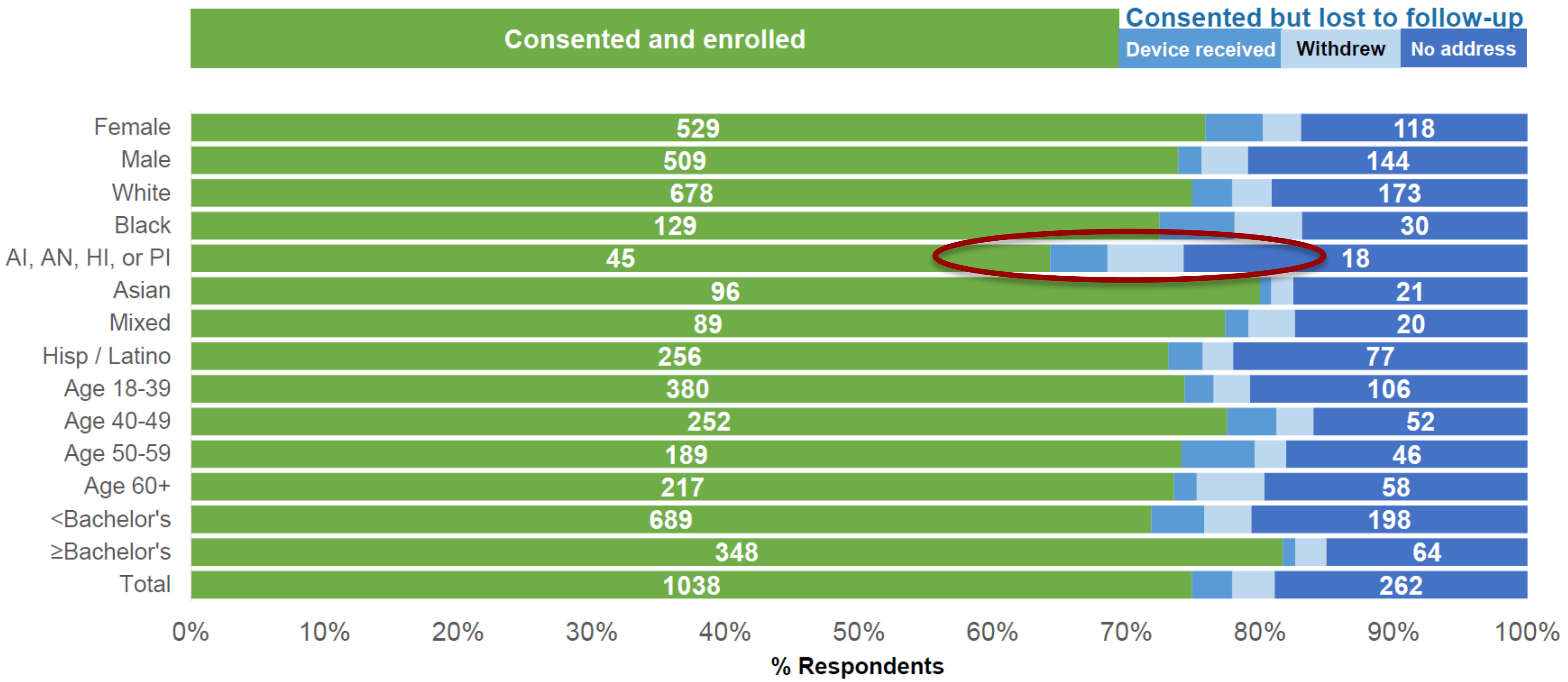
- Older respondents less likely to consent.
- Conscientious respondents less likely to consent and “open” respondents more likely.
  - But total effect of personality on participation unclear
- Males less likely to consent, but total effect seems minor
- Not much of a healthy volunteer effect
- Education affects both consent and data provision
- Ethnicity also affects participation

Results appear consistent with ALiR experiences

Of consent survey respondents, 64% consented, with few demographic differences



Roughly 75% of consenting individuals (50% of consent respondents) enrolled, with few differences



Implications

- To achieve population representation, extra effort to recruit older respondents is warranted.
- Strategies to include minorities and respondents with lower education are needed.

The field is moving fast

- It is important to remember “old fashioned” statistical concepts and not be seduced by large samples and fancy technology, if that technology implies serious coverage error.

- Combing self-reports and non-survey data combines the strength of both.

Thank You!

Email: kapteyn@usc.edu
UAS Web Site: https://uasdata.usc.edu