

Survey Data, Big Data, and Economic Research with the Understanding America Study

PhD in Economics — University of Rome Tor Vergata

Marco Angrisani

Center for Economic and Social Research & Economics Department

University of Southern California

June 2026

Lectures June 16-18 (11-13)

Office Hours June 18 (14-16); June 19, 22, 24, 25 (11-13)

Overview

These lectures introduce students to the Understanding America Study (UAS) as a model for modern economic data collection, and use it as a vehicle to discuss three interrelated themes: how to design and maintain a population-representative survey panel, how research design and question wording shape the quality of what we measure, and how passive and administrative data can be integrated with surveys to unlock new research opportunities.

Lecture 1: The UAS: A Modern Data Collection Infrastructure

Overview. This lecture introduces the UAS as a probability-based, nationally representative internet panel, describing its design principles, recruitment strategy, and the breadth of research it supports. The second part of the lecture uses bank account ownership as a concrete case study in how question design shapes measurement quality and downstream inference.

Topics

1. *What is the UAS?* Sampling design, recruitment, and representativeness. How the UAS addresses the limitations of convenience and opt-in panels.
2. *Panel operations.* Survey administration, response rates, attrition, and respondent compensation. Balancing respondent burden with data richness.
3. *The UAS as a versatile platform.* Core biennial surveys versus topical modules; support for embedded experiments; linkages to administrative records and passive data sources.
4. *The measurement problem in survey economics.* What can go wrong: recall error, reference period effects, question ambiguity, social desirability bias. Classic findings on income and expenditure mismeasurement.
5. *Research application: Bank account ownership.* How does the phrasing of a question about bank account ownership affect measured financial inclusion? What do discrepancies across question formats reveal about respondent cognition and reporting behavior? Policy implications for measuring the unbanked.
6. *Broader lessons on question design.* Designing survey modules for a specific research question; pre-testing and validation; the value of embedding methodological experiments in large panels.

Materials

- UAS Website

- Kapteyn A, Angrisani M, Darling J, and Gutsche T. (2024). “The Understanding America Study (UAS).” *BMJ Open* 14:e088183. (doi)
- Angrisani M and Couper M. (2021). “A Simple Question Goes a Long Way: A Wording Experiment on Bank Account Ownership.” *Journal of Survey Statistics and Methodology*, Vol. 10(5): 1172-1182. (doi)

Lecture 2: Stock Market Return Expectations and Political Polarization

Overview. This lecture presents an ongoing research project that uses the UAS to study how affective polarization — the degree to which individuals hold favorable views of their own party and unfavorable views of the opposition — shapes perceived equity return distributions. The paper illustrates how a rich longitudinal panel can be used to tackle questions at the intersection of household finance, beliefs, and political economy.

Topics

1. *Motivation.* Heterogeneous beliefs in asset pricing: what explains why individuals hold such different views about equity returns? Standard demographic variables explain very little of the variation; what else might matter?
2. *Affective polarization as a driver of belief heterogeneity.* The distinction between ideological and affective polarization; why the intensity of partisan attachment — rather than party membership per se — is the theoretically relevant object.
3. *Data and measurement.* UAS surveys on return expectations (perceived likelihood markets rise, rise $\geq 20\%$, fall $\geq 20\%$) linked to thermometer-based polarization measures, spanning four presidential administrations (Obama, Trump I, Biden, Trump II), 2015–2025.
4. *Main findings.*
 - Large partisan gaps in return expectations that reverse with the party of the president and have grown substantially over time.
 - Affective polarization subsumes the role of party affiliation, political philosophy, media source, and media trust.
 - The relation is stronger among more numerate and financially sophisticated respondents.
 - Partisan belief gaps persist around midterm losses and do not narrow ahead of anticipated election outcomes.
5. *Motivated beliefs versus rational heterogeneous beliefs.* How to distinguish empirically between identity-driven belief distortion and differences in models, priors, or information. What the pattern of evidence supports.
6. *Real consequences.* Polarization predicts actual trading decisions; implications for portfolio behavior and asset pricing.

Materials

- Angrisani, M, Sias, R.W., Sobotka, A, and Turtle, H.J. (2026). “Political Polarization and Stock Market Expectations.” SSRN Working Paper. (doi)

Lecture 3: Integrating Survey and Big Data — Opportunities and Challenges

Overview. Surveys have inherent limits. This lecture examines two UAS-based efforts to augment survey data with passively collected records — financial transaction data (the UASFin

project) and wearable device data (the ALiR project). For each, we ask what new research becomes possible, who participates and who does not, and what the selectivity challenges imply for inference. The lecture concludes with a discussion of the broader frontier of linked survey and administrative data in economics and health.

Topics

1. *Why go beyond surveys?* Recall bias, respondent burden, and the limits of self-report. The promise of passively collected data and the challenges it introduces.
2. *The UASFin project: Financial transaction data.*
 - Design and implementation: partnering with a financial aggregator, the consent and account linkage process.
 - The participation funnel: response, consent, sign-up, and linkage — where sample loss occurs and why.
 - Who participates? Determinants of consent and account linkage; how selection biases compound across participation stages.
 - Research application: comparing electronically recorded and self-reported expenditures. What the discrepancies reveal about recall and reporting behavior; heterogeneity by education and financial literacy.
3. *The ALiR project: Wearable (Fitbit) data.*
 - Design: probability sampling, study-provided hardware, and longitudinal surveys as labels.
 - How provisioning hardware and strategic oversampling address digital access barriers and improve representativeness.
 - Remaining selectivity: older age as a persistent source of differential enrollment despite device provisioning.
 - Generalizability: why models trained on representative data outperform those trained on convenience samples, illustrated with a COVID-19 classification example.
4. *Synthesis.*
 - The complementarity of surveys and passive data: what each does well and where they need each other.
 - Selectivity as a first-order concern, not an afterthought: lessons from UASFin and ALiR for study design.
 - Open questions: informed consent, privacy, and the future of linked data collection in economic and health research.

Materials

- Angrisani M., Kapteyn A., and Samek S. (2023). “Real Time Measurement of Household Electronic Financial Transactions in a Population Representative Panel, *Field Methods*, Vol. 35(4): 303-317. (doi)
- Chaturvedi R., Angrisani M, et al. (2023) “American Life in Realtime: a Benchmark Registry of Health Data for Equitable Precision Health.” *Nature Medicine*, Vol. 29: 283-286. (doi)

- Chaturvedi R., Angrisani M, et al. (2025) “American Life in Realtime: Benchmark, Publicly-available Person-generated Health Data for Equity in Precision Health.” *PNAS Nexus*, Vol. 4(10), pgaf295. (doi)