

Capturing adults' familiarity with financial fraud through natural language processing: Relationships with targeting, engagement, and victimization

Duygu Basaran Sahin, Olivia Valdes, Gary R. Mottola, Andrew M. Parker, Kelly Hyde, Sierra Smucker, Emily Hoch

February 28, 2025

CIPHER 2025

Outline

- Overview of the project
- Background on financial fraud & motivation
- Data & sample
- Measures
- Methods & results
- Next steps & limitations

Acknowledgements

- This project was conducted with financial support from the FINRA Investor Education Foundation. The results, interpretations, conclusions and opinions provided herein are those of the authors and do not necessarily reflect the views of FINRA or the FINRA Investor Education Foundation.
- Duygu Basaran Sahin acknowledges support from the RAND Center for the Study of Aging (5T32AG000244-27), funded by the National Institute on Aging. The findings and conclusions expressed are solely those of the authors, and do not represent the views of any agency of the federal government, National Institute of Aging or RAND Corporation.
- RAND fielded the survey to its American Life Panel on [October 18-27, 2023]. To learn more information about the panels, visit www.rand.org/panels

Overview of the project

- Collaborative project between RAND and the FINRA Investor Education Foundation
- Goal: understand what 40+ population knows about financial fraud
- Data: RAND American Life Panel
- Novelty: Open-ended questions on financial fraud knowledge in a nationally representative online panel & natural language processing application to textual data



Background & motivation

BACKGROUND

- Financial fraud is a persistent and growing problem in the U.S.
- Financial fraud has non-financial costs: individual & societal
- **Knowledge of fraud** can be an important protective mechanism against victimization

MOTIVATION

- Limited scholarship on what types of fraud schemes and tactics people know / are familiar with
- If we know what people know we can better
 - Assess their exposure to risk
 - Create targeted educational and informative trainings
 - Inform policy-makers

Research questions

RQ1: What financial fraud schemes are known and salient to U.S. adults?

RQ2: What demographic, socioeconomic, and psychosocial factors are associated with knowledge of financial fraud schemes?

RQ3: How does knowledge of financial fraud schemes relate to personal fraud experiences, including being targeted by, engaging with, and losing money to fraudsters?

Data & analytical sample

DATA

- ALP is an online, probability-based panel with approximately 11,000 respondents aged 18 or older
- RAND ALP fielded our survey in October 2023
- **1,509** out of 2,568 invited respondents completed the survey (58.7% completion rate)
- **28 closed- and open-ended questions** to adults 21 and older, with a focus on population 40+

SAMPLE DESCRIPTIVES (N= 1,506)

- Mean age 60
- 57% female
- 74% NH White, 9% NH Black, 12% Hispanic/Latino/a, 5% NH AI/AN/Asian/PI/Other
- 61% married
- 14% high school or less, 20% some college, 66% college or more

72%
targeted
by a
fraudster

54 %
engaged
with a
fraudster

34 %
lost
money

Measures: How we defined fraud

“as a situation in which an individual or what appeared to be an organization used **deception to intentionally take your money**. We refer to the people behind these schemes as ‘fraudsters.’ A fraudster could be someone you know (professionally or personally) or someone you don’t know.”

“Being a victim of fraud is **different** from spending money on something legitimate, and then feeling disappointed in the result, **such as poor or delayed service, a product that arrived late or did not meet your expectations**, or even that an individual or organization could have been more honest in their presentation of a product or service.”

Measures: How we defined fraud cont'

Based on this definition, can you list any schemes or tactics of fraudsters?

These can be things you've heard about or that you've experienced.

1)

2)

3)

4)

5)

I don't know

Source: RAND's American Life Panel

Measures: How we defined fraud cont'

Based on this definition, can you list any schemes or tactics of fraudsters? These can be things you've heard about or that you've experienced.

1) You have a package that needs to be shipped

2) We are from the IRS

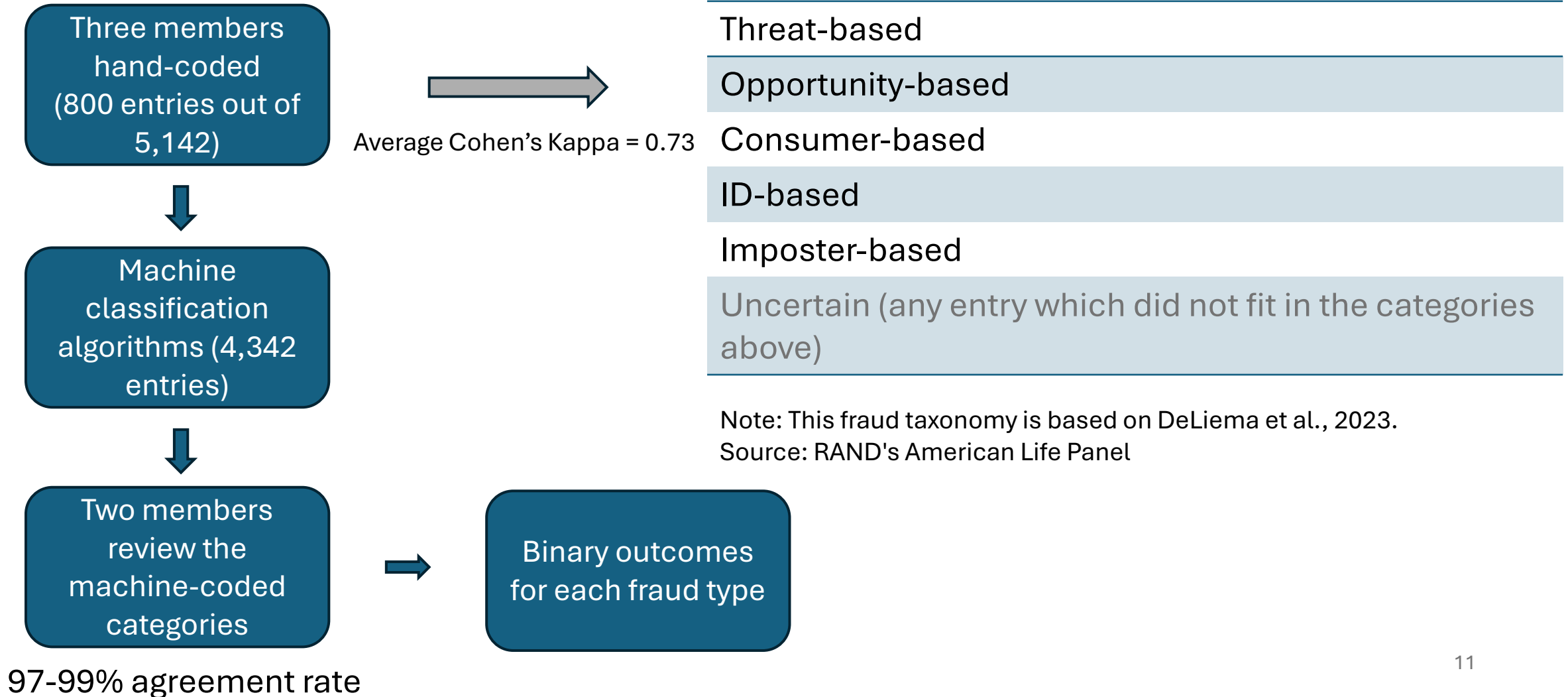
3) This is about your medical condition

4) This call is from social security

5) You won a free trip

I don't know

Methods Step 1: Natural Language Processing (RQ1)



Results RQ1: Financial fraud knowledge prevalence

Fraud Type	Respondent Prevalence	Brief Definition	Verbatim Examples
ID-based	50.0%	A clear reference to using personal identifiable information or theft of the victim's identity	Calls asking for details of financial instruments; someone has gotten into your account; spoofing
Threat-based	19.8%	A clear reference to fear, panic, or urgency to avoid something bad	There is a virus on your computer; panicking someone to force a quick decision
Opportunity-based	17.4%	An opportunity, often capitalizing on a feeling of excitement	Get rich quick; ponzi schemes; you've won a sweepstakes- send us money to get your prize
Consumer-based	16.0%	Purporting to sell a service or product, or collecting money for a product you have or would buy	Call to extend car warranty but just takes your money; telemarketing; Sales fraud
Imposter-based	13.9%	A clear reference to impersonation of a company, entity, or person (other than the victim)	Pretending to be a debt collector; law enforcement impersonators;
N	N=1,506		

Methods Step 2: Logistic regression analysis (RQ2)

Binary outcomes

Mentions threat-based

Mentions opportunity-based

Mentions consumer-based

Mentions ID-based

Mentions imposter-based

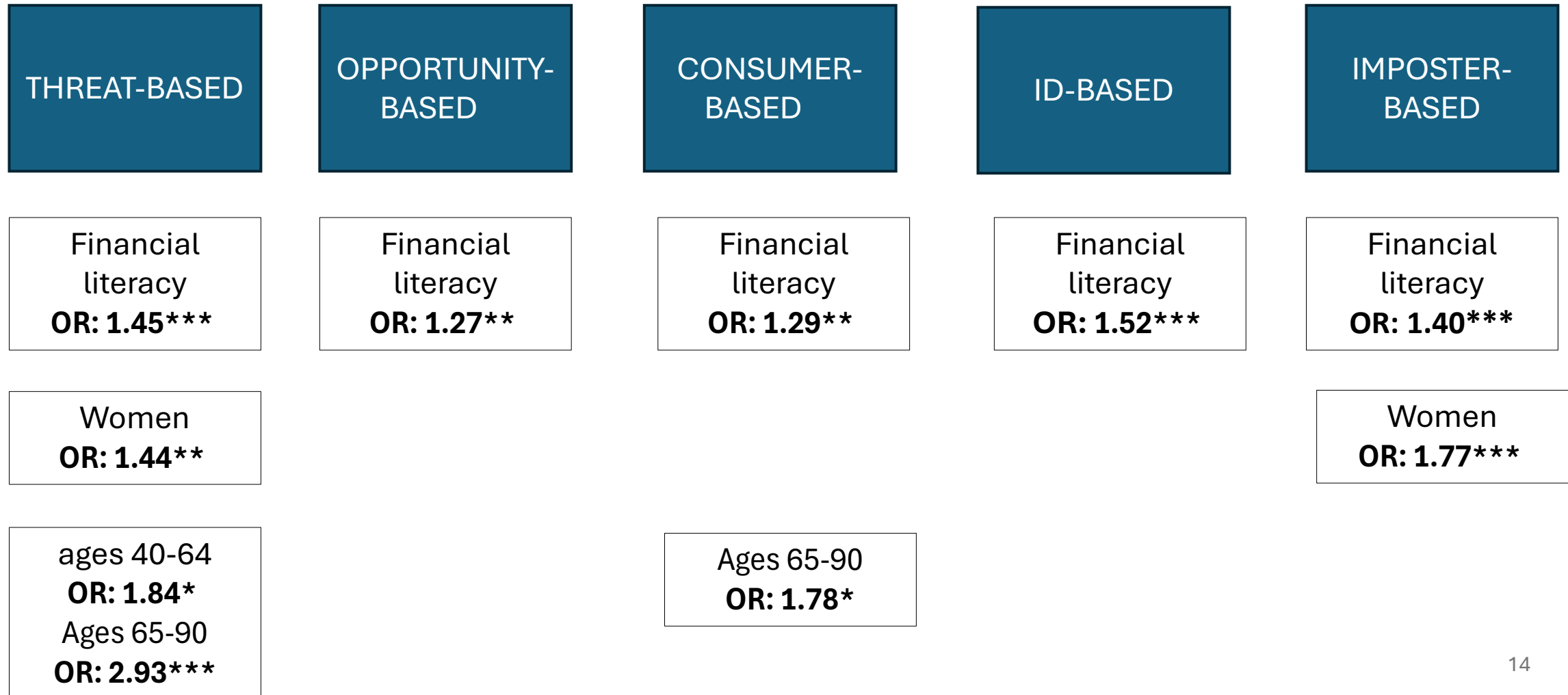
N= 1,506

Predictors

- Demographic: age, race/ethnicity, gender, marital status
- Socioeconomic: education, household income, objective financial literacy
- Psychosocial: depressive symptoms, loneliness

Results RQ2: Correlates of financial fraud knowledge

Source: RAND's American Life Panel, N=1,506 (weighted), * p<0.05, ** p<0.01, *** p<0.001



Next steps & limitations

Next steps

- Personal protection behavior
- Personal narratives about respondents' own financial fraud experiences

Limitations

- Cross-sectional & associations – reverse causality
- Self-reports not claims

Thank you!

For further questions & feedback
basaransahin@rand.org