



Does Overconfidence Affect Financial Behaviors?

Minghao Yang

MA Program of Computational Social Science (MACSS), University of Chicago

Email: mhyang@uchicago.edu, GitHub: WMhYang

May 31, 2020

**MASTERS IN
COMPUTATIONAL
SOCIAL SCIENCE**
THE UNIVERSITY OF CHICAGO

Research Question

Does overconfidence in financial literacy affect households' financial behaviors?

Evidence from:

- Retirement readiness
- Precautionary savings
- Financial market participation

Data

2012, 2015, and 2018 National Finance Capability Study (NFCSS), which covers 80,164 households with sample weights to mimic the national population

• Financial behaviors & Demographic characteristics

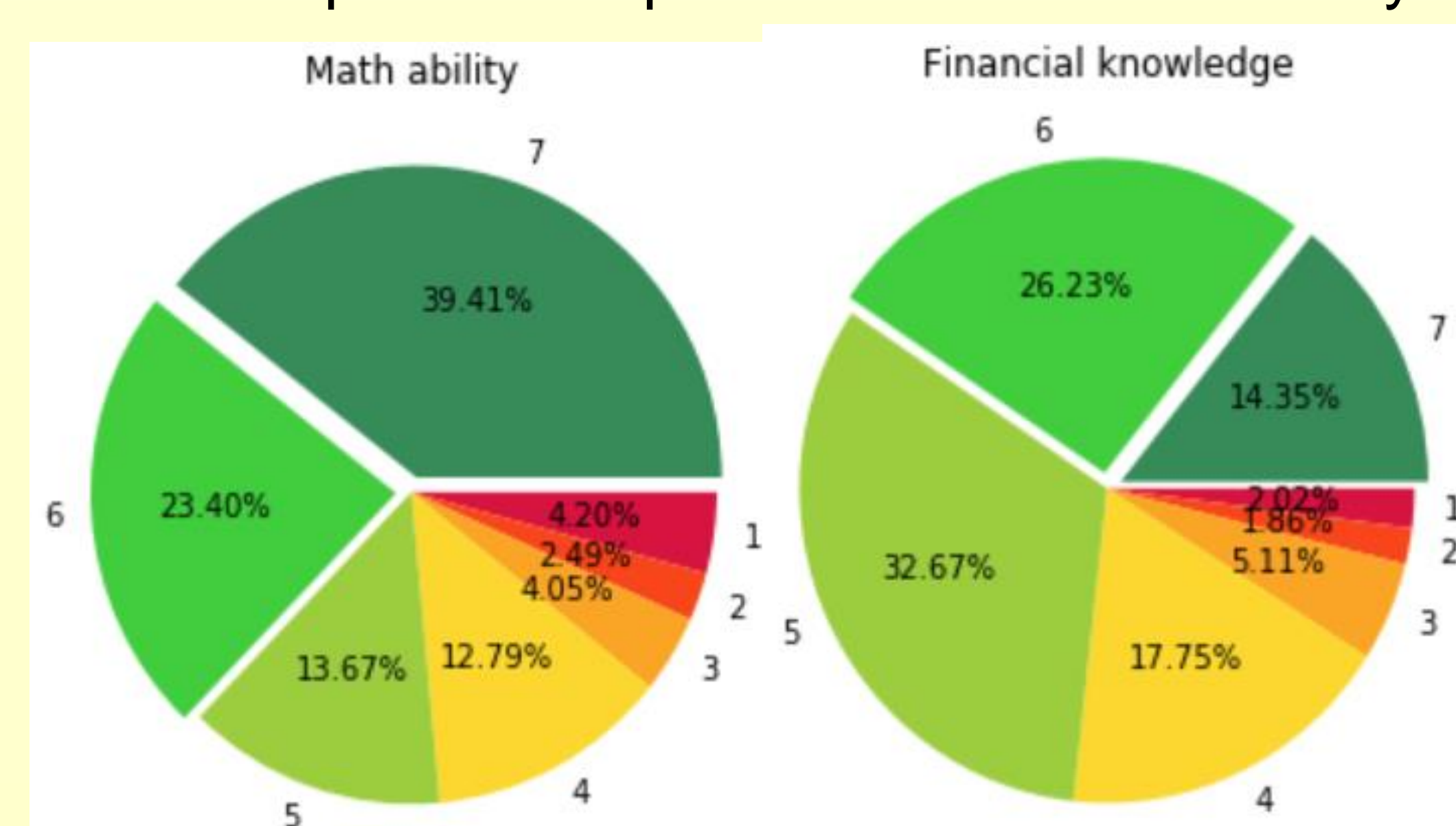
T1: Sum. stat. – Financial behaviors & demo. chars.

Variables	10 th pct.	Median	90 th pct.	Mean	S.D.
Readiness	0	0	1	0.309	0.462
Precaution	0	0	1	0.449	0.497
Participation	0	0	1	0.314	0.464
Female	0	1	1	0.514	0.500
Age	20	50	70	46.34	16.52
Nonwhite	0	0	1	0.350	0.477
Married	0	1	1	0.523	0.499
Income	7500	42500	125000	62054	49232
High School	1	1	1	0.954	0.210
College	0	0	1	0.355	0.479

• Perceived financial literacy

Larger numbers for higher literacy
Households tend to perceive high levels of financial literacy

F1: Composition of perceived financial literacy

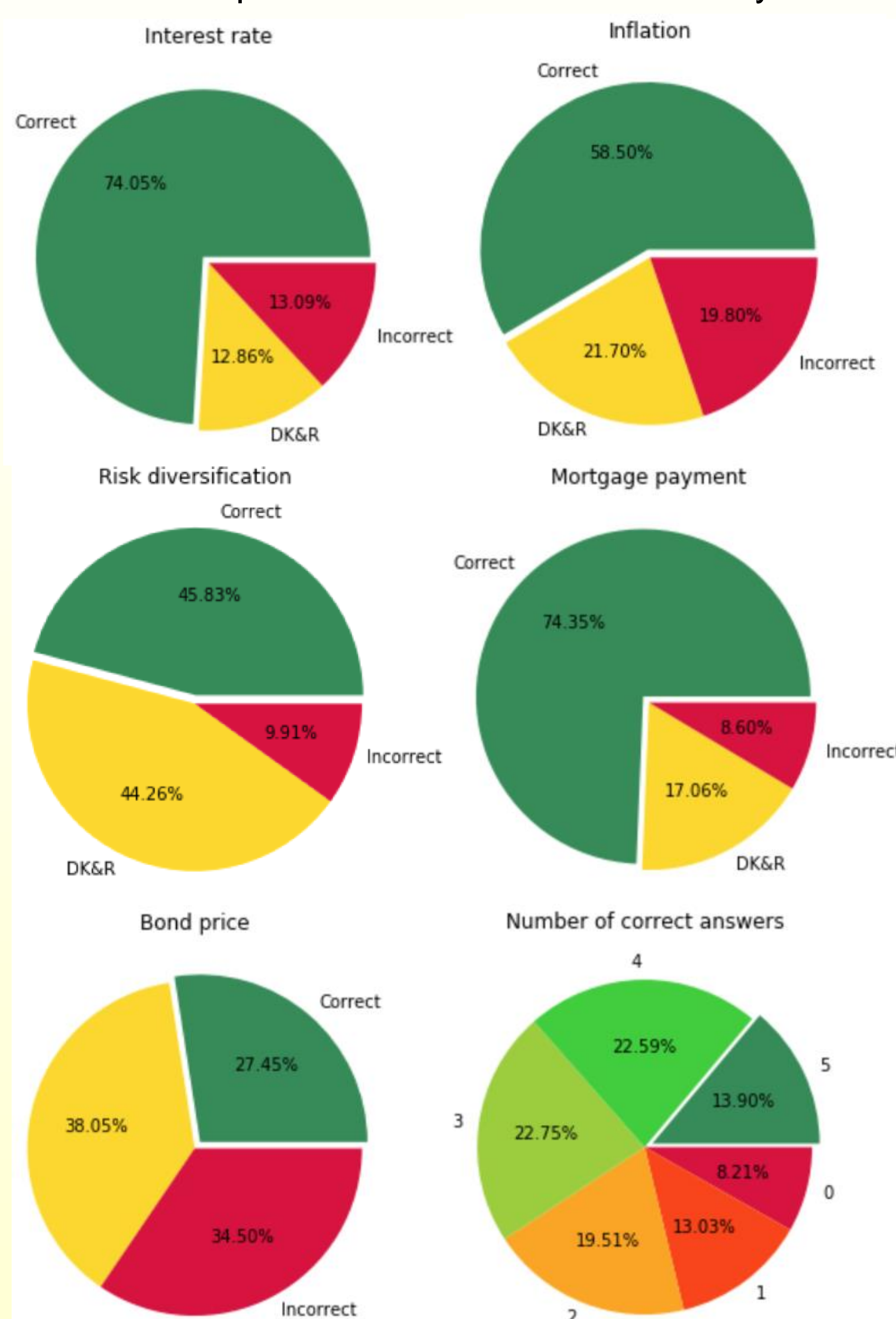


Data (Con't)

• True financial literacy

No more than 15% households answer all the "Big Five" questions correctly

F2: Composition of true financial literacy



Methods & Results

• Overconfidence measures (ML based)

Learning set: Unambiguous overconfident (858) or not overconfident (7,506) households

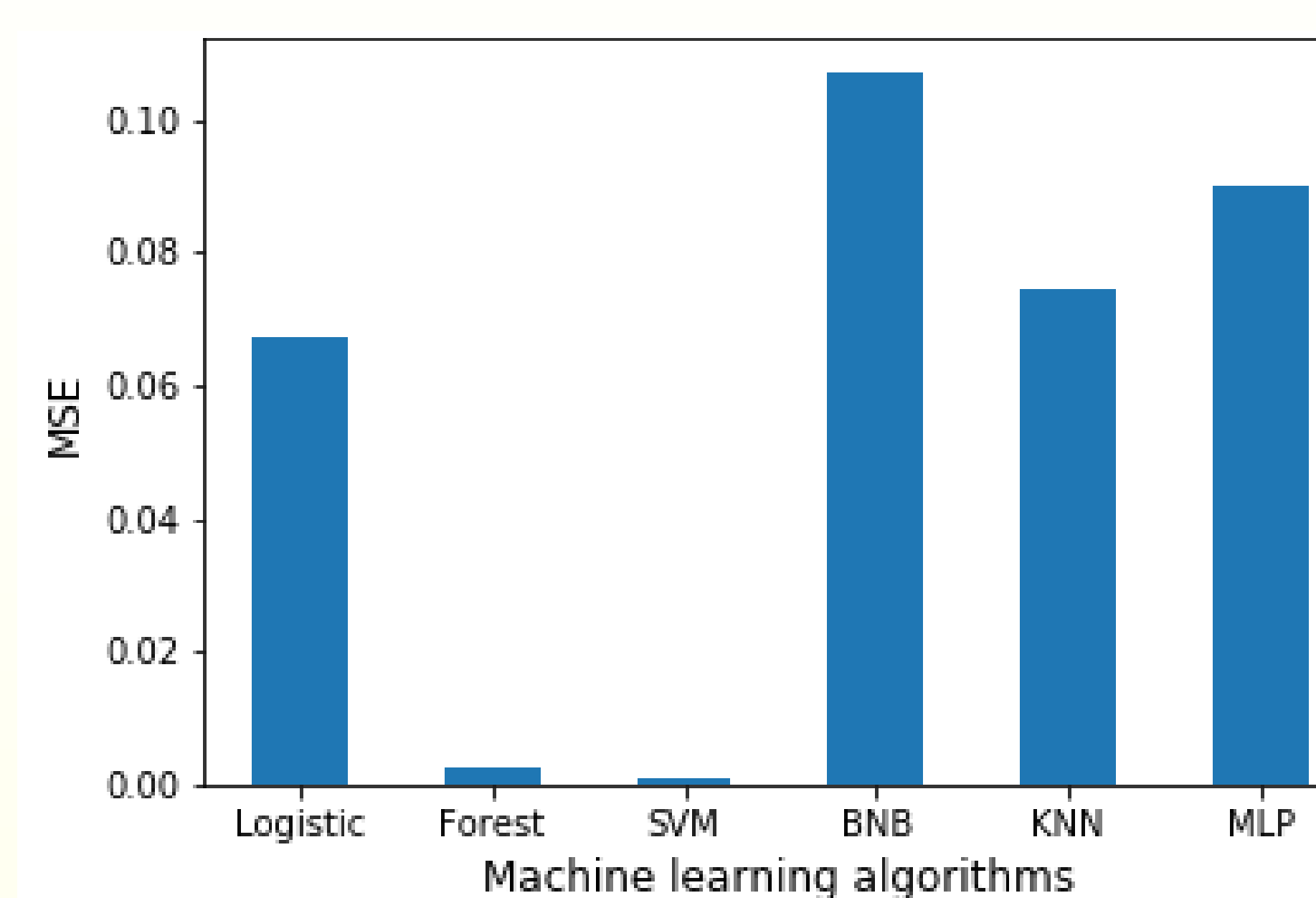
Features: Demographic characteristics, perceived & true financial literacy

Classifiers: Logistic, Random Forest, SVM, Bernoulli NB, KNN, MLP

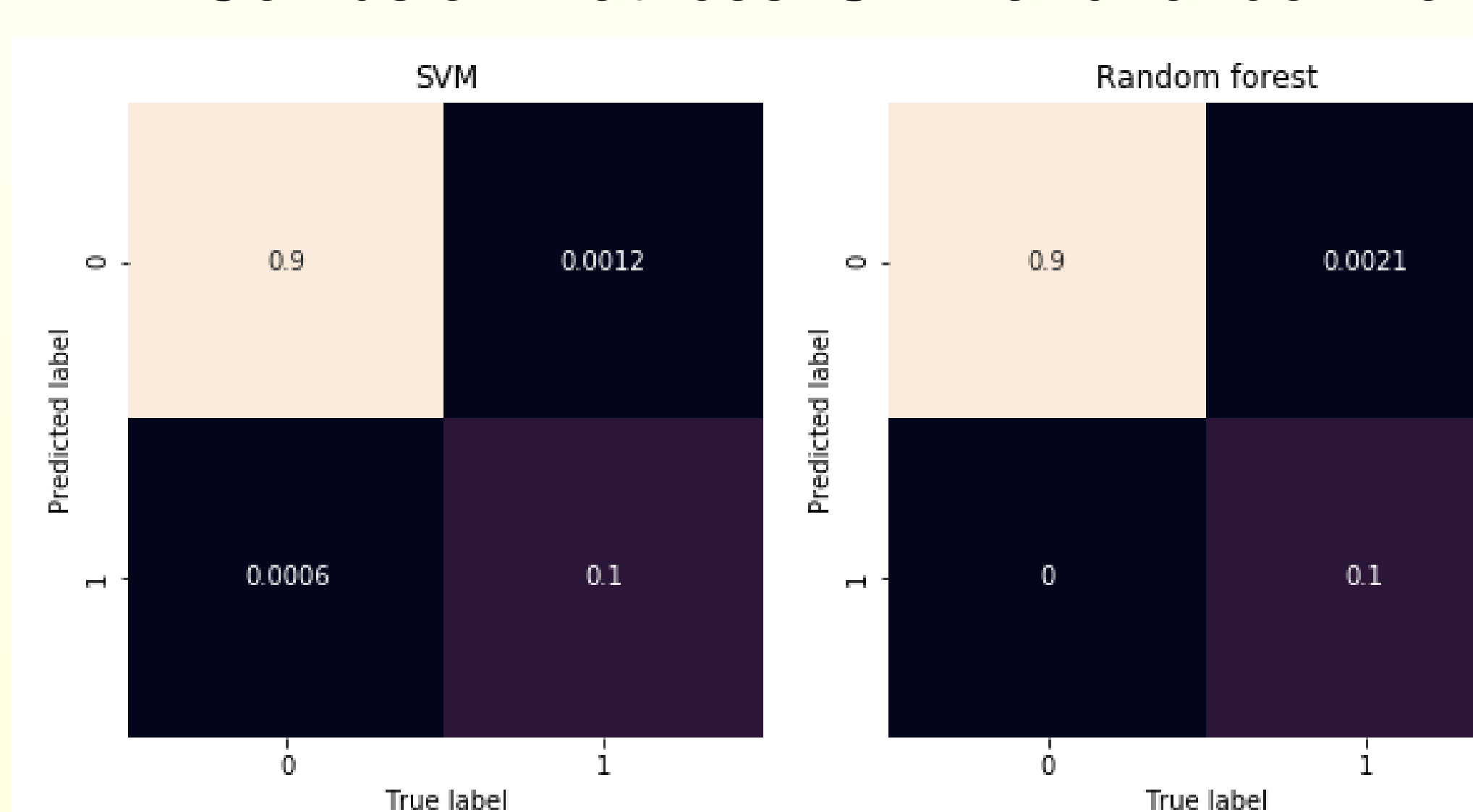
Out of sample prediction: Probabilities

Methods & Results (Con't)

F3: MSEs of different classifiers



F4: Confusion matrices: SVM and random forest



• True financial literacy measure

Factor analysis on the "Big Five" questions using principal component method

Calculate normalized factor scores

T2: Sum. stat. – Overconfidence & True financial literacy

Variables	10 th pct.	Median	90 th pct.	Mean	S.D.
Overconfidence					
SVM	≈ 0	0.133	≈ 1	0.392	0.426
Forest	0.029	0.203	0.455	0.234	0.170
True literacy	0.214	0.630	1	0.580	0.299

• The effect of overconfidence

Dependent variables: Financial behaviors
Independent variables: Overconfidence, True literacy, Demographic characteristics, Year dummies, State dummies
Logit regression

$$\Pr(y_{it} = 1 | \mathbf{X}_{it}, \beta_0, \beta_1, \varepsilon_{it}) = F(\beta_0 + \mathbf{X}_{it}\beta_1 + \varepsilon_{it})$$

where $F(x) = e^x / (1 + e^x)$

Methods & Results (Con't)

T3: Overconfidence and retirement readiness

Dept. Var.:	(1)	(2)
Readiness	SVM	Forest
Overconfidence	0.142*** (0.00534)	0.477*** (0.0210)
True literacy	0.344*** (0.00834)	0.441*** (0.0126)

T4: Overconfidence and precautionary savings

Dept. Var.:	(1)	(2)
Precaution	SVM	Forest
Overconfidence	0.152*** (0.00542)	0.459*** (0.0219)
True literacy	0.313*** (0.00835)	0.389*** (0.0129)

T5: Overconfidence and financial mkt. participation

Dept. Var.:	(1)	(2)
Participation	SVM	Forest
Overconfidence	0.141*** (0.00547)	0.485*** (0.0213)
True literacy	0.377*** (0.00854)	0.475*** (0.0128)

Conclusion

Overconfidence in financial literacy has a decent effect on financial behaviors of households with similar true literacy

Overconfidence ↑ by 1 std. div. :

- Pr(readiness) ↑ by **6.1% - 8.1%**
- Pr(precaution) ↑ by **6.5% - 7.8%**
- Pr(participation) ↑ by **6.0% - 8.2%**

Limitations

- Unbalanced overconfidence classification
- Potential endogeneity

Acknowledgements

I sincerely thank Dr. Richard Evans for the wonderful courses on these ML methods. I also thank all classmates for their feedbacks.