

# Gender, confidence and financial literacy

**Tabea Bucher-Koenen (MEA, Netspar)**

**Rob Alessie (Groningen, Netspar)**

**Annamaria Lusardi (GWB, NBER, Netspar)**

**Maarten van Rooij (DNB, Netspar)**

Preliminary

Cherry Blossom conference, GFLEC, 2017

# Motivation

We find a substantial **gender gap** in financial literacy.

Women are significantly more likely to **answer** with „**do not know**“.

This **gap persists** across different **countries**, **subgroups** (young and old), and knowledge **domains** (pension literacy, economic literacy, debt literacy).

see Bucher-Koenen, Lusardi, Alessie, van Rooij (2017) „How financially literate are women? An overview and new insights“, NBER Working Paper, w20793, forthcoming in: *Journal of Consumer Affairs*.

# Research Question

---

What lies behind the **gender gap** in financial literacy?

Why do women answer with „**do not know**“ more frequently?

Is it due to a lack of **knowledge** or lack of **confidence** (see Beyer, 1990; Odean, 2001)?

---

# Survey Experiment

# Survey Experiment

## SAMPLE:

- DNB Household Panel (DHS)
- Online survey representative of Dutch-speaking households
- We include panel members who are household heads and their partners, age 18 and older

## DESIGN:

- 3 financial literacy questions asked twice
- First survey (May 2012): Financial literacy questions including a “Do not know/refuse” option
- Second survey (June /July 2012): Same questions without a “Do not know option”, after each question ask for confidence in the answer

# Survey Experiment

Set Up Week 1 (May 2012):

- 1) Interest:** *Suppose you had \$100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow? **More than \$102** / Exactly \$102 / Less than \$102 / Do not know/ Refuse to answer*
- 2) Inflation:** *Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account? More than today / Exactly the same / **Less than today** / Do not know / Refuse to answer*
- 3) Risk:** *Please tell me whether this statement is true or false. “Buying a single company’s stock usually provides a safer return than a stock mutual fund.” True / **False** / Do not know / Refuse to answer*

# Survey Experiment

Set Up Week 2 (June/July 2012):

- Questions without the “Do not know” and “refuse to answer” options (DK)
- After each question – **Confidence:**  
*On a scale from 1 to 7, How confident are you in this answer?*  
*1-not confident at all ... 7- completely confident*

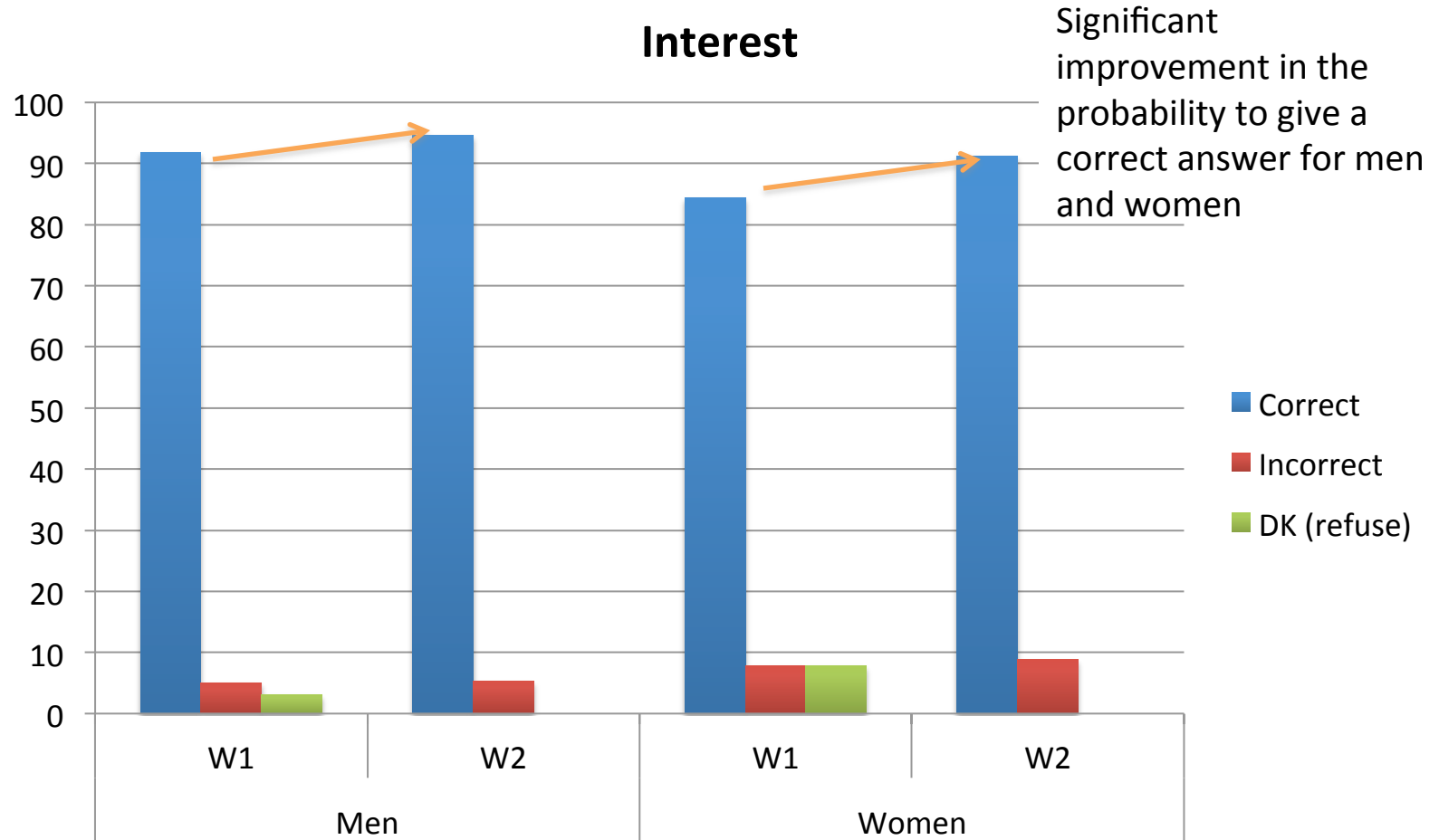
# Survey Experiment

- Week 1: N= 1,748
- Week 2: N= 1,973 (incl. refresher)
- Sample for the analysis: complete questionnaire in both weeks, N=1,532; 861 (56.2%) are men and 671 (43.8%) are women.
- **Attrition:** no sign. effects of gender or financial literacy on dropping out after week 1
- **Learning:** answers to financial literacy questions in week 2 for refreshers (N=445) do not differ significantly from participants in both weeks.



# Descriptives

## Answers Week 1 and Week 2



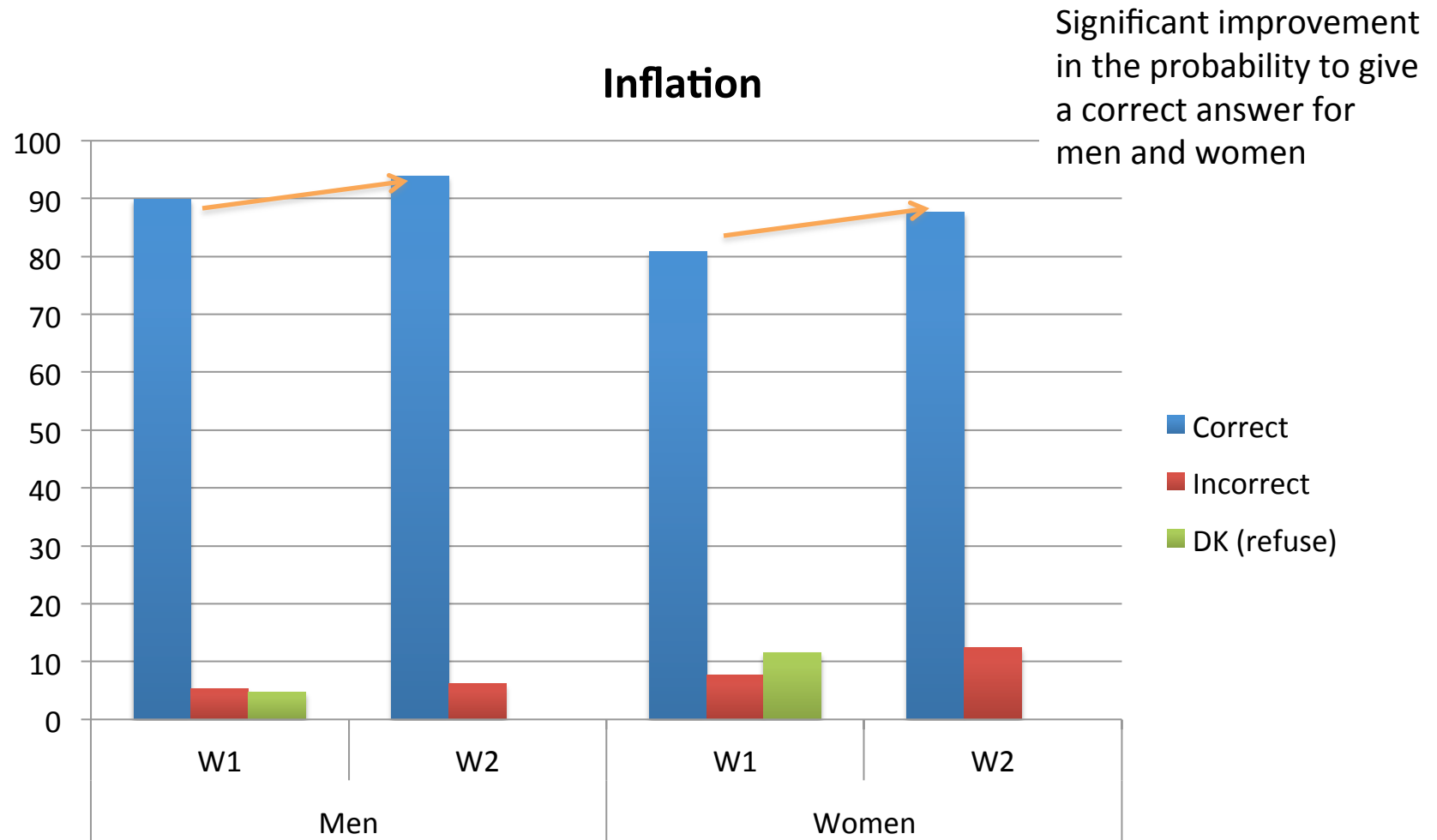
# Descriptives

**Table 1: Descriptive Results (Observations: Men – 861, Women – 671, Total - 1532)**

	Survey 1: May 2012			Survey 2: July 2012		
<b>A. Interest:</b>	Men	Women	Total	Men	Women	Total
More than 102 euro	91.9	84.4	88.6	94.7	91.2	93.2
Exactly 102 euro	3.0	4.0	3.5	3.7	6.0	4.7
Less than 102 euro	2.0	3.9	2.8	1.6	2.8	2.2
Do not know	2.8	6.7	4.5	-	-	-
Refuse	0.4	1.0	0.7	-	-	-
Total	100	100	100	100	100	100

# Descriptives

## Answers Week 1 and Week 2



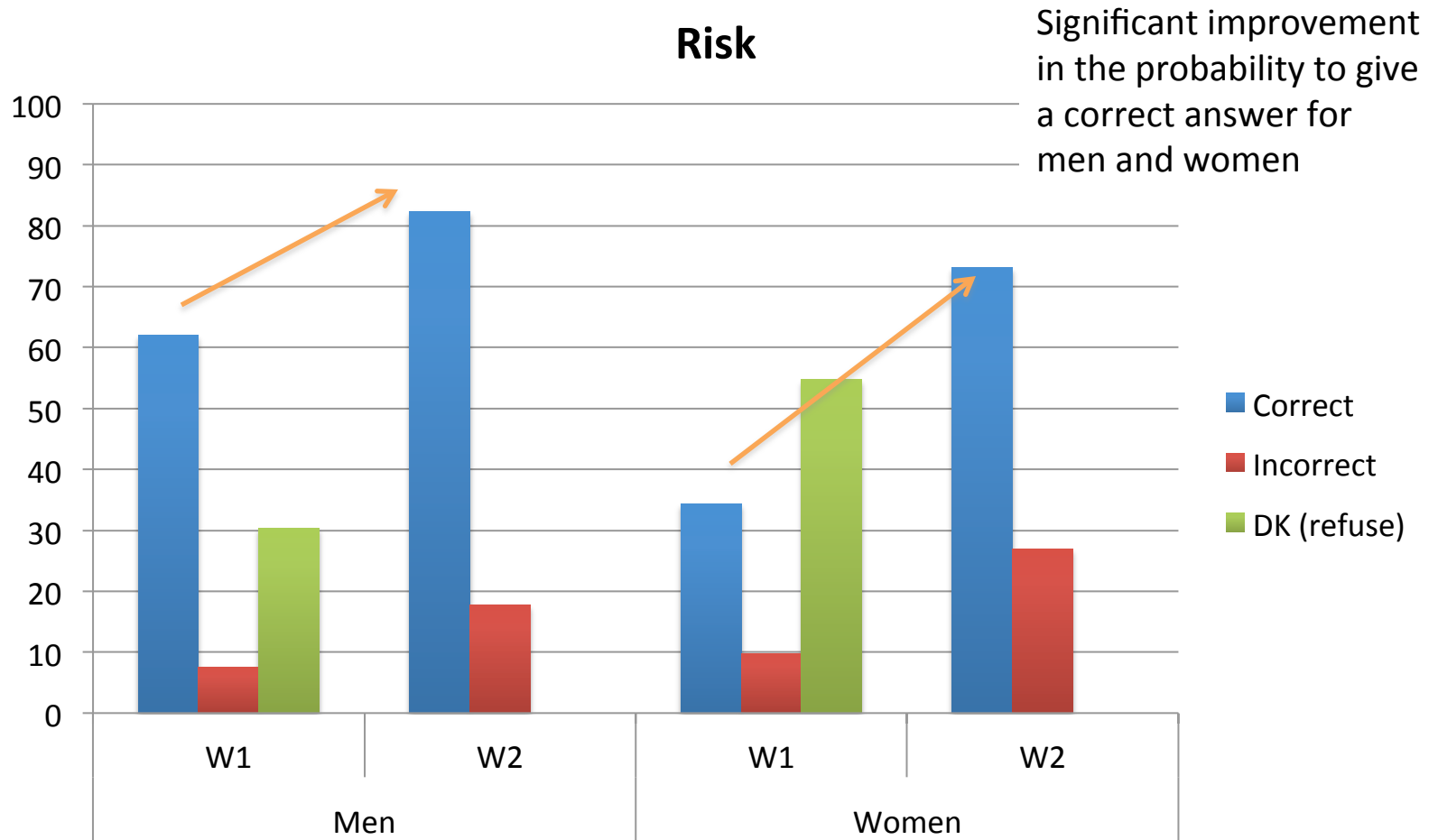
# Descriptives

**Table 1: Descriptive Results (Observations: Men – 861, Women – 671, Total - 1532)**

	Survey 1: May 2012			Survey 2: July 2012		
<b>B. Inflation:</b>	Men	Women	Total	Men	Women	Total
More	2.1	2.4	2.2	2.2	2.7	2.4
Exactly the same	3.3	5.4	4.2	4.1	9.8	6.6
Less	89.8	80.6	85.8	93.7	87.5	91.0
Do not know	4.7	10.7	7.3	-	-	-
Refuse	0.2	0.9	0.5	-	-	-
Total	100	100	100	100	100	100

# Descriptives

## Answers Week 1 and Week 2



# Descriptives

**Table 1: Descriptive Results (Observations: Men – 861, Women – 671, Total - 1532)**

	Survey 1: May 2012			Survey 2: July 2012		
	Men	Women	Total	Men	Women	Total
<b>C. Risk Diversification:</b>						
Incorrect 'right'	7.6	9.7	8.5	17.7	27.0	21.7
Correct 'false'	61.9	34.4	49.9	82.4	73.0	78.3
Do not know	30.1	54.7	40.9	-	-	-
Refuse	0.5	1.2	0.8	-	-	-
Total	100	100	100	100	100	100
<b>D. Overall No. of correct answers:</b>						
	Men	Women	Total	Men	Women	Total
0	3.6	6.6	4.9	0.5	0.8	0.6
1	7.3	16.8	11.5	3.3	6.9	4.8
2	31.0	47.2	38.1	21.4	32.3	26.2
3	58.1	29.4	45.5	74.9	60.1	68.4

# Descriptives

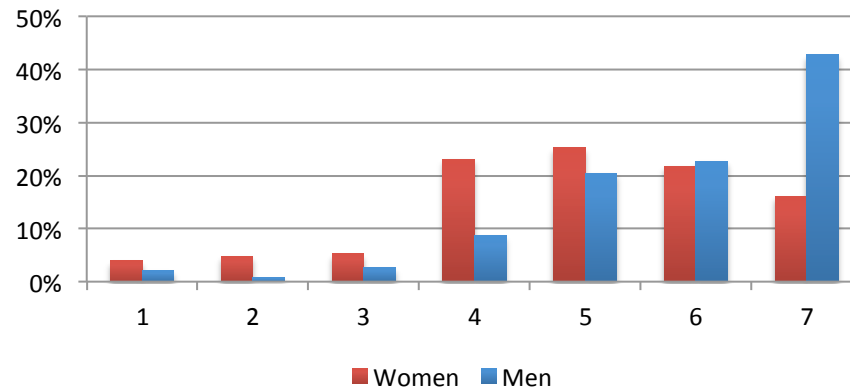
## Consistent and inconsistent answering across waves

<b>A. Interest:</b>		<b>Men</b>			<b>Women</b>		
<i>Survey May</i>	<b>incorrect</b>	<b>correct</b>	<b>don't know</b>	<b>incorrect</b>	<b>correct</b>	<b>don't know</b>	
<i>Survey July</i>							
incorrect	23.26	3.54	29.63	28.3	4.95	30.77	
correct	76.74	96.46	70.37	71.7	95.05	69.23	
Total	100	100	100	100	100	100	
<b>B. Inflation:</b>							
incorrect	41.3	2.72	33.33	30.77	7.02	38.46	
correct	58.7	97.28	66.67	69.23	92.98	61.54	
Total	100	100	100	100	100	100	
<b>C. Risk Diversification:</b>							
incorrect	38.46	10.32	27.38	47.69	12.55	32.27	
correct	61.54	89.68	72.62	52.31	87.45	67.73	
Total	100	100	100	100	100	100	

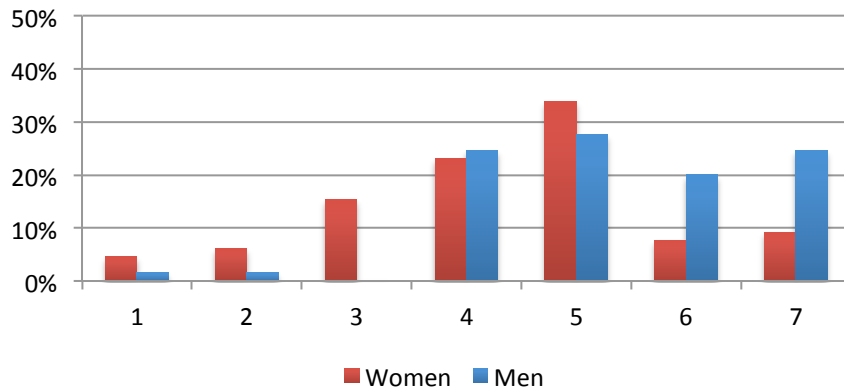
# Descriptives

What about confidence conditional on being correct, incorrect, do not know? e.g. Risk Q

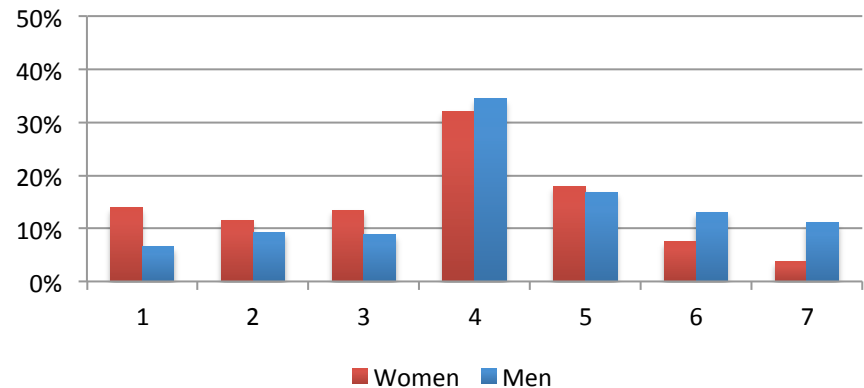
### Confidence cond. Correct



### Confidence cond. Incorrect



### Confidence cond. Do not know





# Summary

- Financial literacy scores in May reflect both knowledge and confidence in answering.
- In July, respondents are forced to answer, providing a knowledge measure that is not confounded by confidence.
- However, at the same time people who do not know the answer are forced to guess an answer, thus the July measure contains measurement error and is upward biased

---

# **Measuring financial literacy: a latent class model**

# Model

## Stylized facts that are the basis of our model for measuring financial literacy

1. Answering behavior changes when deleting the DK option.
2. Respondents who gave a DK answer in May are correct with a high likelihood in July.
3. Fraction of correct answers increases also due to gambling.
4. DK responses in May are correlated with confidence July.

## A model of true financial literacy (I)

- Latent variable
  - $y_{\downarrow ik} = 1$  if respondent  $i$  'knows' the correct answer to question  $k$  ("true knowledge");
  - $y_{\downarrow ik} = 0$  otherwise
- We observe
  1.  $y_{\downarrow ik \uparrow m} = 0$  if incorrect answer given in May; =1 if correct answer; =2 if DK
  2.  $y_{\downarrow ik \uparrow j} = 0$  if incorrect answer given in July; =1 if correct answer;
  3.  $conf_{\downarrow ik \uparrow j}$  is the answer to the confidence question in likert scale from 1 to 7

## A model of true financial literacy (II)

- Our **goal: predict** the probability that a respondent **truly knows** the answer to literacy question  $k$  based on background characteristics  $x \downarrow i$  and on the variables  $y \downarrow ik \uparrow m$ ,  $y \downarrow ik \uparrow j$  and  $conf \downarrow ik \uparrow j$ :

$$P(y \downarrow ik = 1 | x \downarrow i, y \downarrow ik \uparrow m = l \downarrow k, y \downarrow ik \uparrow j = m \downarrow k, conf \downarrow ik \uparrow j = z \downarrow k), k=1, 2, 3 \quad (1)$$

- Our measure of financial literacy:

$$finlit \downarrow i = \sum_{k=1}^3 P(y \downarrow ik = 1 | x \downarrow i, y \downarrow ik \uparrow m =$$

## Latent class model (III)

- Let  $g_{lik} = g = 3 y_{lik}^j + y_{lik}^m$ ,  $y_{lik}^j = 0, 1$ ;  $y_{lik}^m = 0, 1, 2$ ;  $g = 0, \dots, 5$
- The log-likelihood of our latent class model is based on the conditional multinomial density of  $g_{lik}$ :

$$P_{g_{lik}} = g_{x_{li}}, \text{conf}_{lik}^j = z_{lk}$$

- This conditional probability can be written as a weighted average of two multinomial probabilities:

$$P_{g_{lik}} = g_{x_{li}}, \text{conf}_{lik}^j = z_{lk} = P_{g_{lik}} = g_{y_{lik}} = 1, x_{li}, \text{conf}_{lik}^j = z_{lk} P_{y_{lik}} = 1, x_{li}, \text{conf}_{lik}^j$$

## Latent class model (IV)

- We assume that

$$1. P_{y \downarrow ik} = 1, x \downarrow i, \text{conf} \downarrow ik \uparrow j = z \downarrow k = P_{y \downarrow ik} = 1, x \downarrow i \\ = \Phi(x \downarrow i \uparrow \beta) \text{ (Probit)}$$

$$2. P_{g \downarrow ik} = g, y \downarrow ik = 1, x \downarrow i, \text{conf} \downarrow ik \uparrow j = z \downarrow k = \alpha \downarrow g \uparrow 1 \\ (x, z \downarrow k) = \alpha \downarrow g \uparrow 1 (z \downarrow k) \text{ (Multinomial Logit)}$$

$$3. P_{g \downarrow ik} = g, y \downarrow ik = 0, x \downarrow i, \text{conf} \downarrow ik \uparrow j = z \downarrow k = \alpha \downarrow g \uparrow 0 \\ (x, z \downarrow k) = \alpha \downarrow g \uparrow 0 (z \downarrow k) \text{ (Multinomial Logit)}$$

- Then we can write

$$P_{g \downarrow ik} = g, x \downarrow i, \text{conf} \downarrow ik \uparrow j = z \downarrow k = \alpha \downarrow g \uparrow 1 (z \downarrow k) \Phi(x \downarrow i \uparrow \beta) + \alpha \downarrow g \uparrow 0 (z \downarrow k) \Phi(-x \downarrow i \uparrow \beta)$$

## Latent class model (V): Identifying assumptions

- We have made the following additional assumptions:

- $$1. \alpha \downarrow 0 \uparrow 1 (z \downarrow k) = P g \downarrow ik = 0 y \downarrow ik = 1, \text{conf} \downarrow ik \uparrow j = z \downarrow k = P y \downarrow i \uparrow m = 0, y \downarrow i \uparrow j = 0 y \downarrow i = 1, \text{conf} \downarrow ik \uparrow j = z \downarrow k = 0, z \downarrow k = 1, \dots, 7$$
- $$2. \alpha \downarrow 1 \uparrow 1 (z \downarrow k) = P g \downarrow ik = 1 y \downarrow ik = 1, \text{conf} \downarrow ik \uparrow j = z \downarrow k = P y \downarrow i \uparrow m = 1, y \downarrow i \uparrow j = 0 y \downarrow i = 1, \text{conf} \downarrow ik \uparrow j = z \downarrow k = 0, z \downarrow k = 1, \dots, 7$$
- $$3. \alpha \downarrow 2 \uparrow 1 (z) = P g \downarrow ik = 2 y \downarrow ik = 1, \text{conf} \downarrow ik \uparrow j = z \downarrow k = P y \downarrow i \uparrow m = 2, y \downarrow i \uparrow j = 0 y \downarrow i = 1, \text{conf} \downarrow ik \uparrow j = z \downarrow k = 0, z \downarrow k = 1, \dots, 7$$



## Latent class model (VI)

- Once we have estimated the parameters we can compute  $P(y_{lik} = 1 | g_{lik} = g, conf_{lik}^j = z_{lk}, x_{li})$  by applying Bayes' rule:

$$P(y_{lik} = 1 | g_{lik} = g, conf_{lik}^j = z_{lk}, x_{li}) = \frac{\alpha_{g=1}(z_{lk}; \gamma^1) \Phi(x_{li} | \beta)}{\alpha_{g=1}(z_{lk}; \gamma^1) \Phi(x_{li} | \beta) + \alpha_{g=0}(z_{lk}; \gamma^0) \Phi(-x_{li} | \beta)}$$

- ..and we can compute our measure of financial literacy:

$$finlit_{li} = \sum_{k=1}^3 P(y_{lik} = 1 | g_{lik} = g, conf_{lik}^j = z_{lk}, x_{li})$$

# Results

## Financial literacy using different measures

	Prob (May)	Prob (July)	Prob (True literacy)
Interest question	88.58%	93.15%	87.60%
Inflation question	85.77%	90.99%	86.26%
Risk question correct	49.87%	78.26%	64.01%
Average number of correct answers	2.24	2.62	2.38

# Results

## The gender gap in financial literacy (OLS regression)

	May	July	True literacy
<b>Panel A. Only gender</b>			
Female	-0.441***	-0.189***	-0.282***
	(0.0386)	(0.0291)	(0.0351)
<i>R</i> <sup>2</sup>	0.068	0.024	0.035
<b>Panel B. With controls for age, income, education, marital status</b>			
Female	-0.359***	-0.146***	-0.221***
	(0.0393)	(0.0301)	(0.0360)
<i>R</i> <sup>2</sup>	0.167	0.103	0.154

# Summary

- The financial literacy scores in May reflect both knowledge and confidence in answering.
- The July measure is likely to be a noisy proxy for true knowledge as respondents who do not know the answer are required to guess an answer.
- The estimated “true knowledge” measure minimizes both the measurement error and the bias due to confidence which particularly makes a difference for the women.

# OLS Results

Does it matter for stock market participation?

	No controls	May	July	True knowledge
Financial Literacy		0.0901***	0.0549***	0.0671***
		(0.0105)	(0.00970)	(0.0101)
Gender	-0.136***	-0.0461***	-0.0715***	-0.0646***
	(0.0207)	(0.0212)	(0.0213)	(0.0213)
Controls+	no	yes	yes	yes
N	1532	1532	1532	1532
R2	0.023	0.147	0.126	0.132

+controls for age, income, education, marital status

# IV Results

Instrument: Economics in high school  
(3 groups: no education, some education, DK)

	May	July	True knowledge
Financial Literacy	0.192***	0.222***	0.204***
	(0.0671)	(0.0842)	(0.0751)
Gender	-0.00335	-0.0310	-0.0235
	(0.0369)	(0.0308)	(0.0325)
First stage <i>F</i> -stats	14.19	9.189	11.26

# 5. Conclusion

## Summary

- Gender gap in financial literacy decreases but does not disappear when deleting the “Do not know option”.
- Gender gap is both: a gap in knowledge and a gap in confidence.
- Gender gap in knowledge is exacerbated in descriptive statistics and there is a bias in OLS regressions.
- We develop a strategy to consistently estimate whether a respondent truly knows the correct answer and therefore improve the measurement of financial literacy.

# Conclusion

---

## Policy

Important from a policy perspective for the design of **financial education programs**.

*Women know less than men but they know more than they think they know.*



---

**Thank you!**

---