# Gender Differences in Financial Literacy

Rob Alessie and Tabea Bucher-Koenen

**Boot Camp 2025**
**June 24,2025**

# Fearless Woman: Financial Literacy and Stock Market Participation

Tabea Bucher-Koenen (ZEW, U of Mannheim)
Rob Alessie (U of Groningen)
Annamaria Lusardi (The George Washington University and GFLEC)
Maarten van Rooij (DNB)

# Introduction

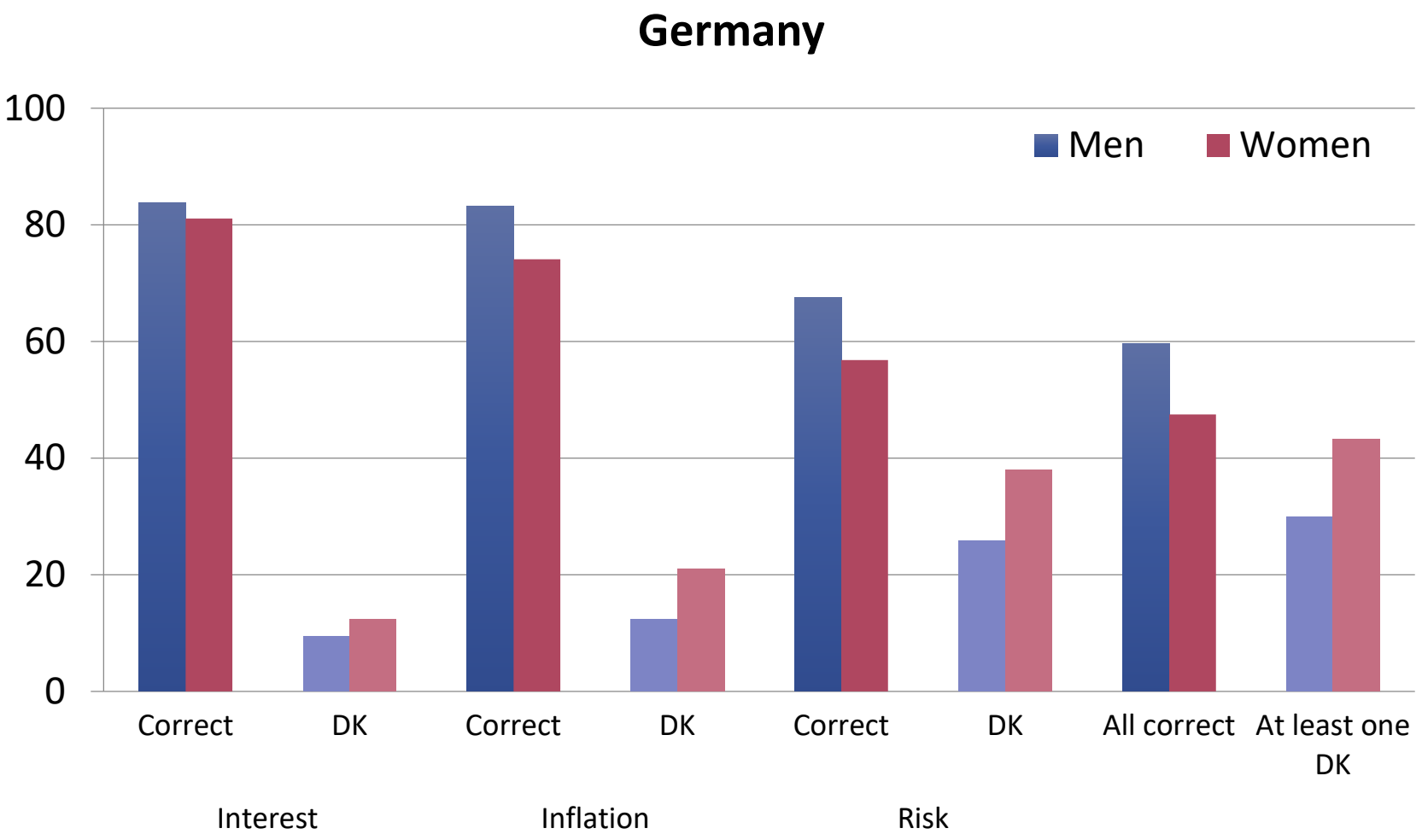
This paper is part of a long term research agenda

- Measuring financial literacy
  - The Big Three
- Assessing the gender gap in financial literacy
  - A consistent finding around the world
- Does the gender gap matter? Examining stock market participation
  - Important for saving and growing wealth
  - Investing is what people identify with “finance”
  - Stock market participation is an important outcome variable in research on financial literacy

# The „Big 3“ financial literacy questions

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*

# Gender Differences in Financial Literacy

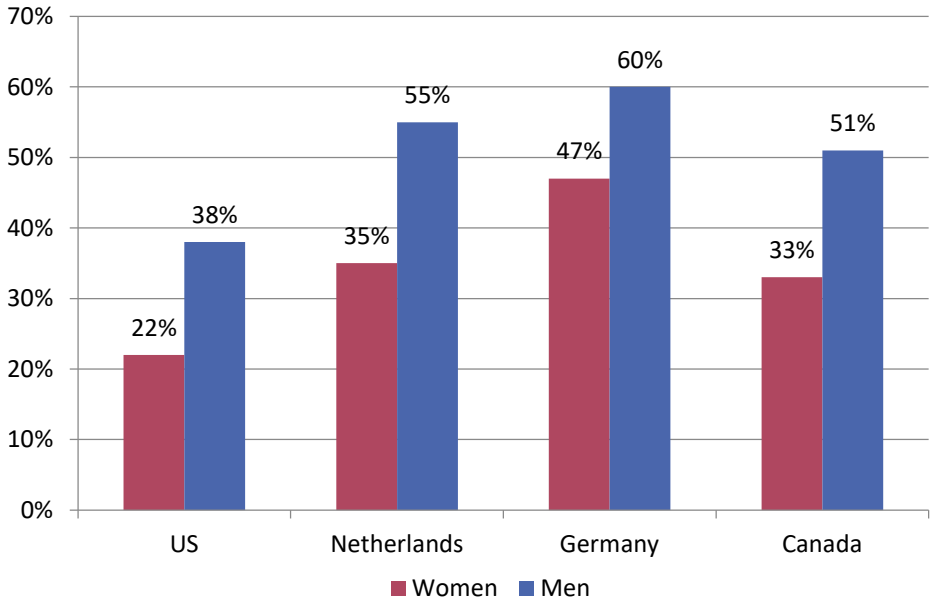



Bucher-Koenen, Lusardi, Alessie, van Rooij (2017) "How financially literate are women? An overview and new insights", *Journal of Consumer Affairs*

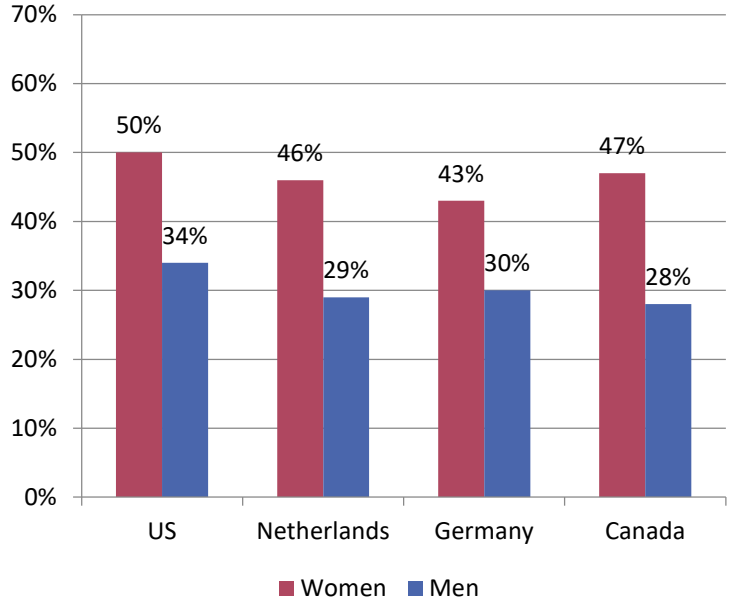
# Gender Differences in Financial Literacy

## Similar findings across countries

**Financial knowledge by gender (% answering 3 Qs correctly)**




**At least one "don't know" answer, by gender**

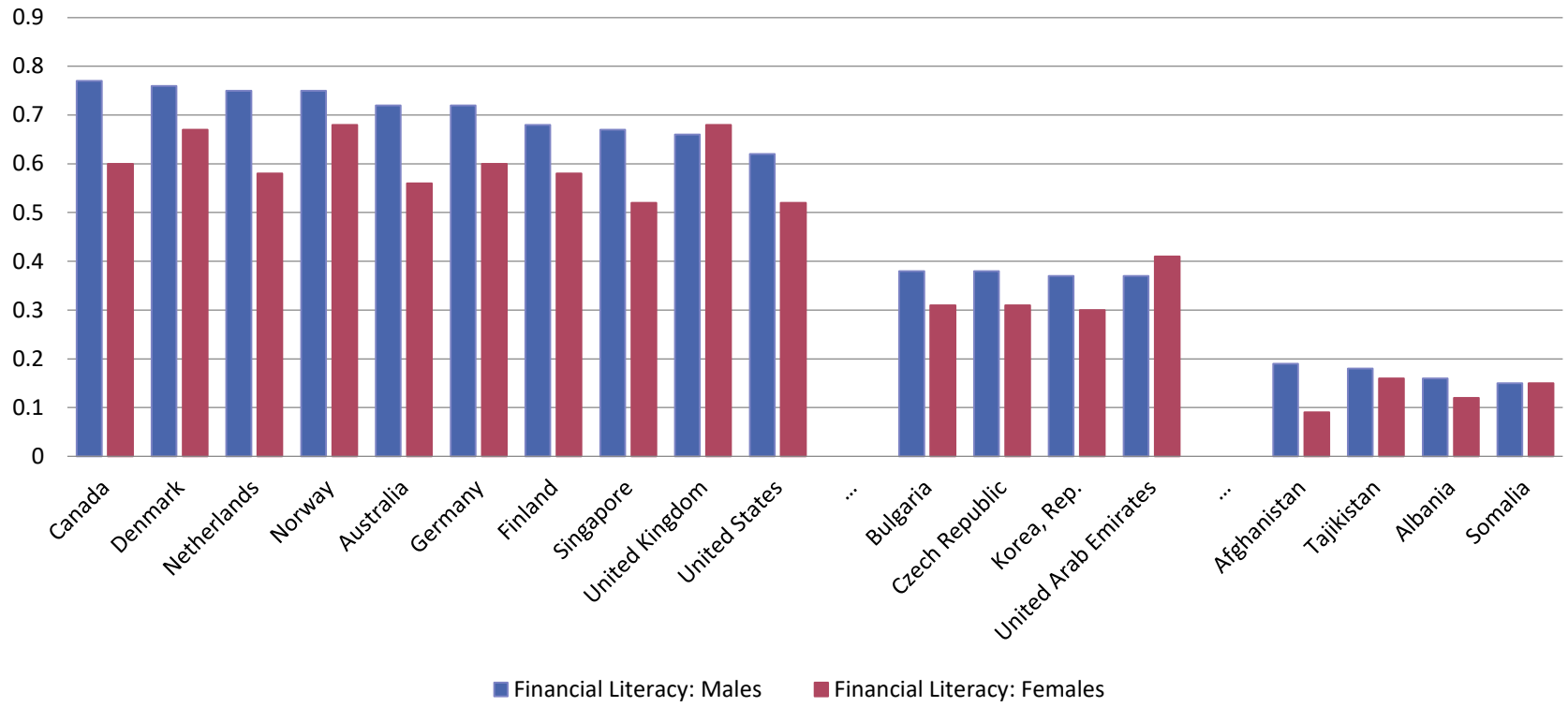



- Very robust findings of large gender differences in financial knowledge
- Women are much more likely to say "I do not know"

Bucher-Koenen, Lusardi, Alessie, van Rooij (2017) "How financially literate are women? An overview and new insights", *Journal of Consumer Affairs*

# Gender Differences in Financial Literacy

## Similar findings across countries – S&P survey




- Similar results for many countries
- Gender gap is persistent over different levels of economic development

# Gender Differences in Financial Literacy

## Striking patterns

- Results persist for **broader sets** of financial literacy questions (Van Rooij et al. 2011a, Lusardi and Mitchell 2009, Lusardi et al. 2009, Bucher-Koenen 2011 )
- Persistent for **different subgroups** of the population (young and old)
- Persistent for **different domains** (pension literacy, economic literacy, debt literacy)

# Research Questions and Contribution

- 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?

**Does how we measure financial literacy affect our understanding and predictions with regard to financial decisions and economic outcomes?**

# Evidence from a Survey Experiment

# The Survey Experiment

Sample and structure of the experiment

- DNB Household Panel (DHS)
- Representative online survey of Dutch households
- We include household heads and their partners, age 18+.

**Module 1: May 2012**

Included the "Big 3" Questions:

Interest | Inflation | Risk

One of the answer options was:

*Do not know*

**Module 2: June/July 2012**

Included the "Big 3" Questions:

Interest | Inflation | Risk

But now, we removed the DK-option:

~~*Do not know*~~

Instead, after each of the 3 questions we asked for confidence levels:

*On a scale from 1 to 7, How confident are you in this answer?*

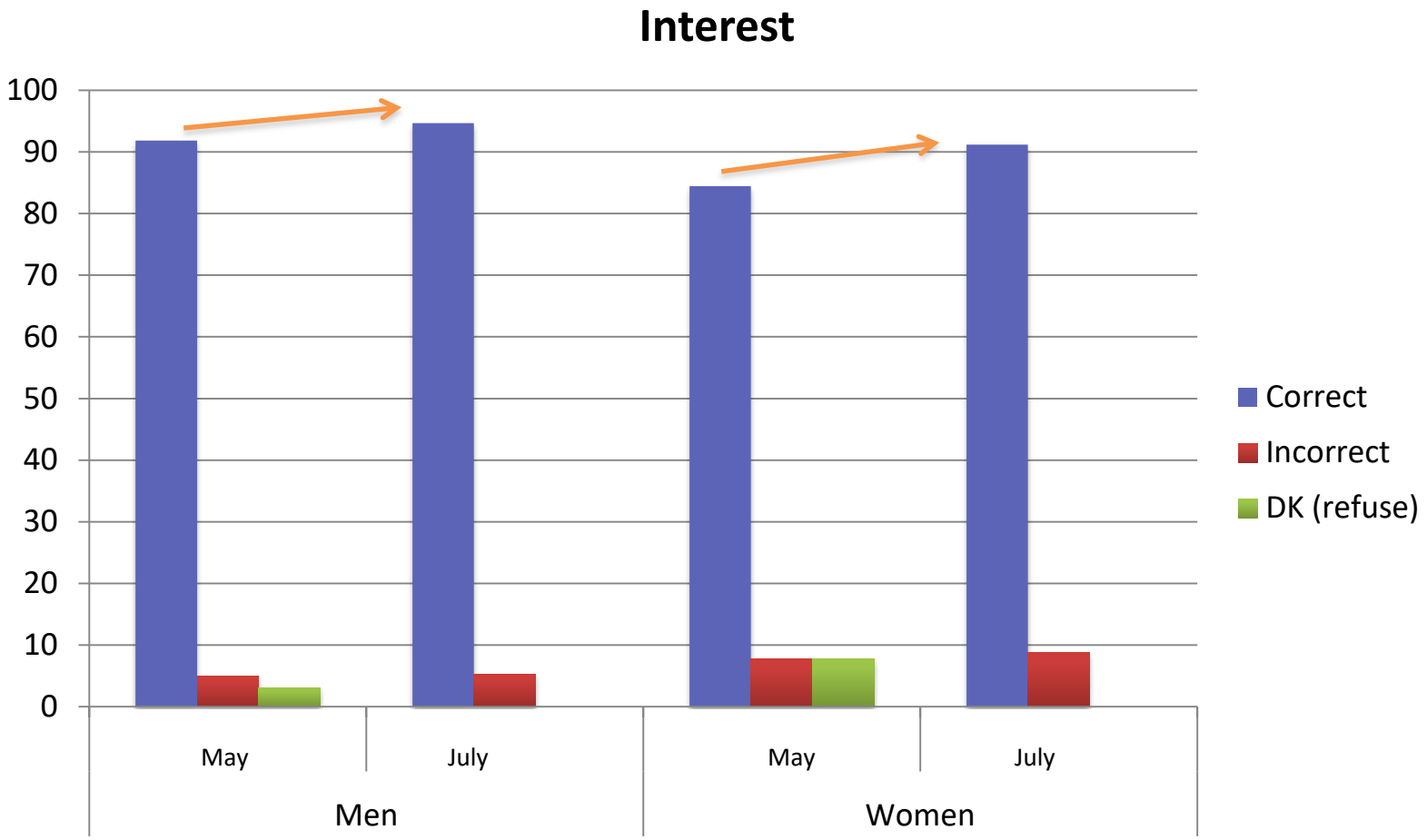
# The Survey Experiment

Additional details on the sample

- **Sample**:
  - Completed both questionnaire modules, N=1532,
  - 861 (56.2%) are men and 671 (43.8%) are women.
- **Attrition**: No significant effects of gender or financial literacy on dropping out after the first module.
- **Learning**: Answers to financial literacy questions in 2nd module for refreshers (N=445) do not differ significantly from participants in both modules.

# Descriptive Statistics

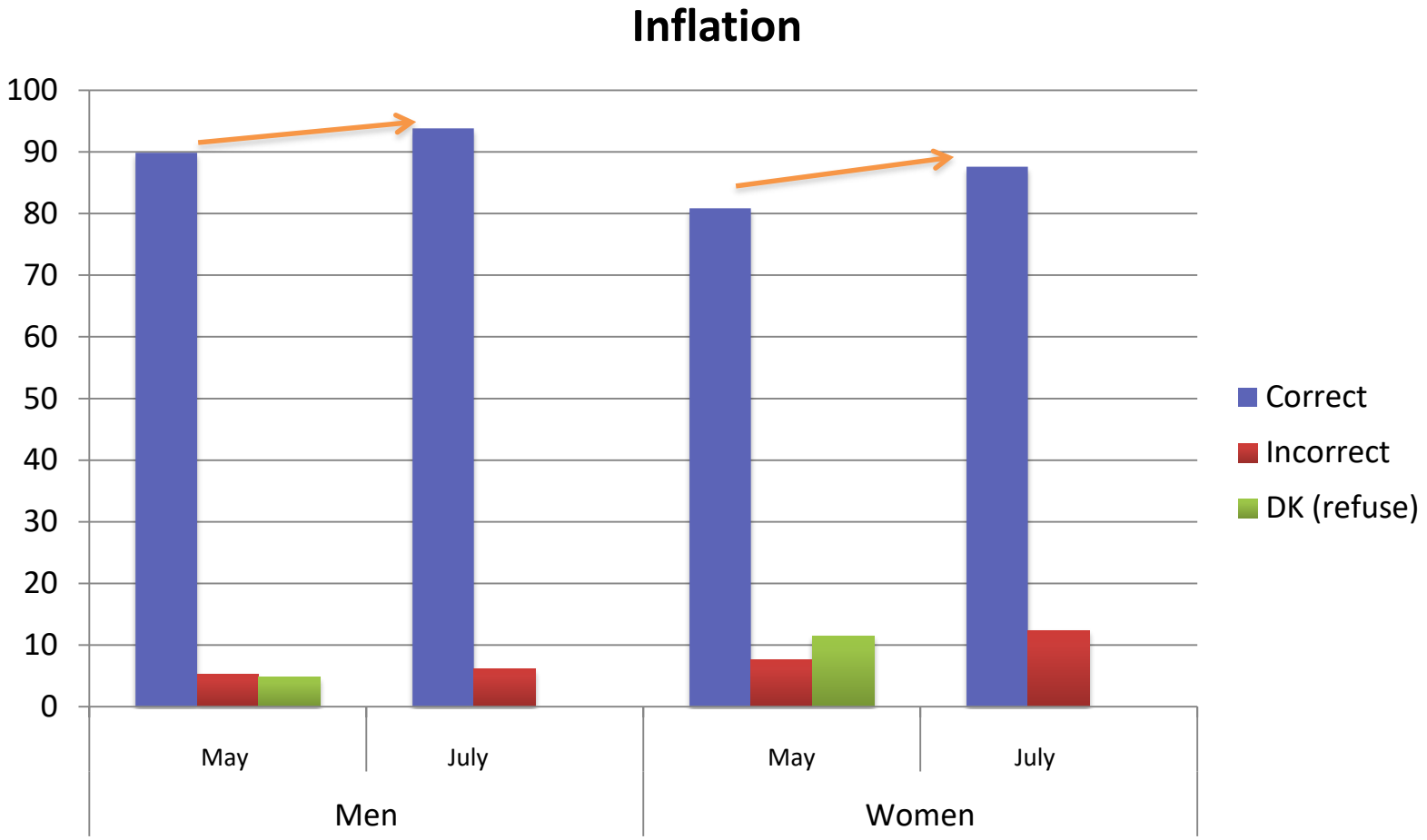
## Comparison of answers in 1st module (May) and 2nd module (July)




Significant improvement in the probability to give a correct answer for men and women (test against random answering). Gender gap decreases from 7.5 to 3.5 pp.

# Descriptive Statistics

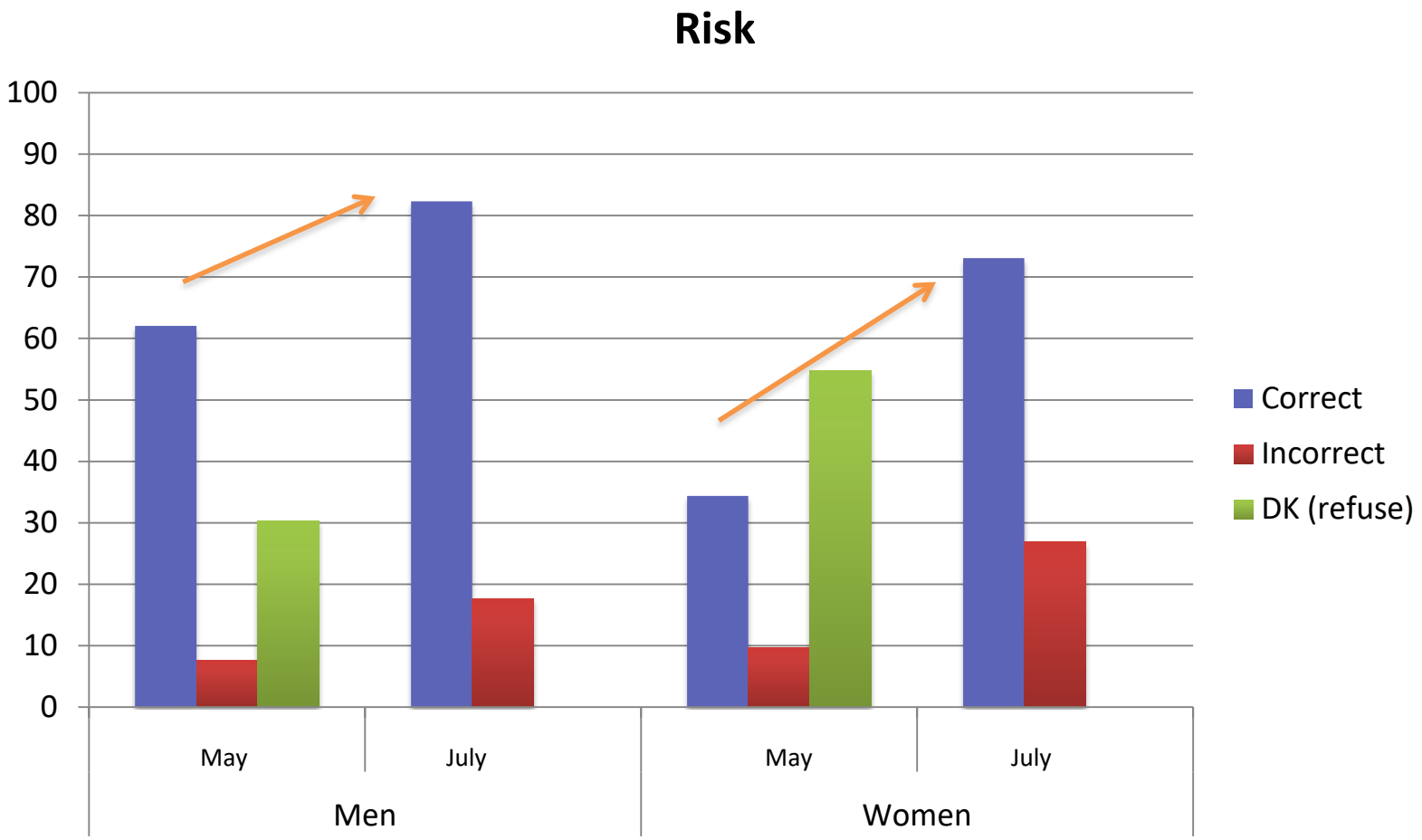
Comparison of answers in 1st module (May) and 2nd module (July)




Significant improvement in the probability to give a correct answer for men and women (test against random answering). Gender gap decreases from 9.2 to 6.2 pp.

# Descriptive Statistics

## Comparison of answers in 1st module (May) and 2nd module (July)




Significant improvement in the probability to give a correct answer for men and women (test against random answering). Gender gap decreases from 27.5 to 9.4 pp.

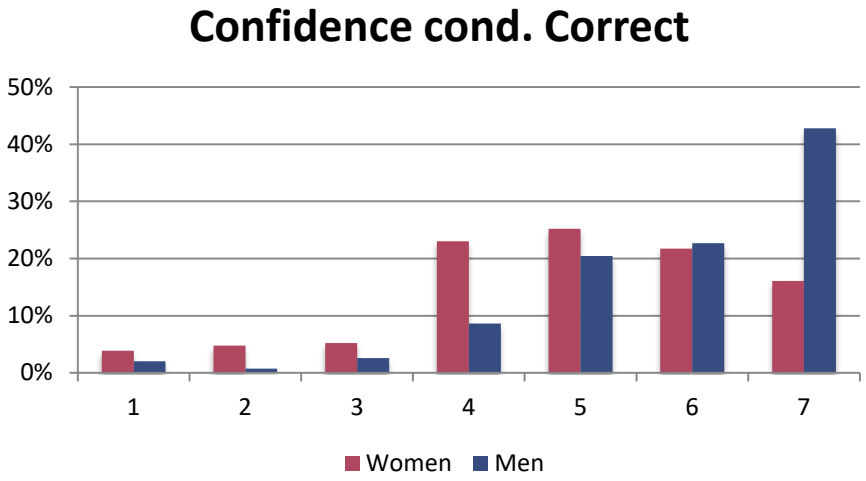
# Descriptive Statistics

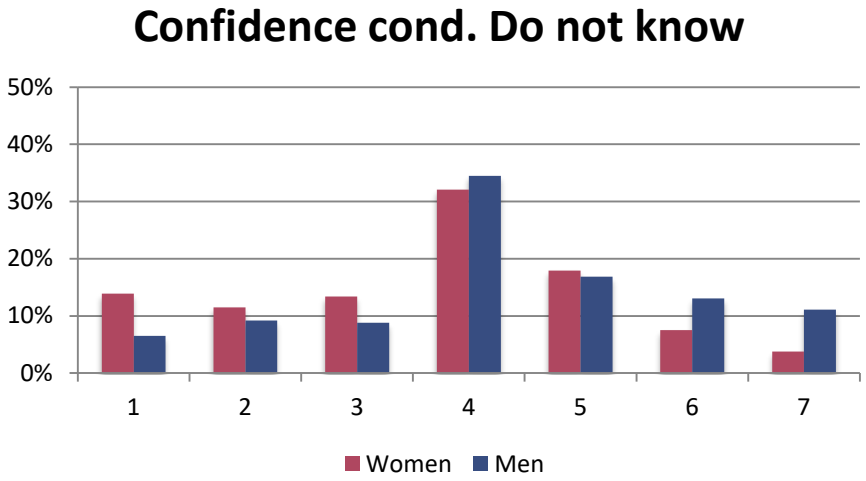
## Consistent and inconsistent answering behavior across modules

| | | Men | | | Women | |
|---|---|---|---|---|---|---|
| | *May* | incorrect | correct | do not know | incorrect | correct | do not know |
| *July* | | | | | | | |
| **A. Interest:** | | | | | | | |
| 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. Inflation:** | | | | | | | |
| 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 |
| **C. Risk Diversification:** | | | | | | | |
| 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 |

# Descriptive Statistics

## Confidence measure conditional on answers in May




**Women report substantially lower confidence levels in module 2 – both when knowing the right answer and when choosing the DK-option in module 1.**

# Issues with directly observed measures

Rationale for developing an econometric latent class model

1. The **May measure** (module 1) corresponds to **Big 3 approach**
   - includes “do not know”-option.
   - reflects both knowledge and *confidence*.
2. On the other hand, the **July measure** (module 2)
   - forces individuals to answer, and therefore is not confounded by confidence.
   - contains measurement error (due to guessing) and is upward biased as a result.
3. On average, women display lower confidence in their answers compared to men irrespective of their chosen answers.

**Econometric model takes these observations into account, deriving an empirical measure of ‘true financial knowledge’**

# Measuring and decomposing financial literacy: A latent class model

# Econometric Model - Definitions

The central latent variable and observable information

We define the following **latent variable for 'true knowledge'** (not observed) for each financial literacy question:

$\tilde{y}_{ik} = 1$ if respondent *i* truly 'knows' the correct answer to literacy question $k$ (*k*=1,2,3),
$\tilde{y}_{ik} = 0$ otherwise.

**Observed proxies** for this variable:

$y_{ik}^{m}$ answer to literacy question *k* in May; 0 (incorrect), 1 (correct), 2 (do not know);

$y_{ik}^{j}$ answer to question *k* in July; 0 (incorrect) and 1 (correct);

$conf_{ik}^{j}$ answer to the confidence question on a scale from 1 to 7.

# Econometric Model - Intuition

Predicted probability of '*true financial literacy*'

Our **goal**: **Predict** the probability that a respondent **truly knows** the answer to literacy question *k* based on background characteristics $x_i$ and on the variables $y_{ik}^m$, $y_{ik}^j$ and $conf_{ik}^j$:

$$P(\tilde{y}_{ik} = 1 | x_i, y_{ik}^m = l_k, y_{ik}^j = m_k, conf_{ik}^j = z_k), k = 1,2,3$$

Summary measure of financial literacy:

$$finlit_i = \sum_{k=1}^{3} P(\tilde{y}_{ik} = 1 | x_i, y_{ik}^m = l_k, y_{ik}^j = m_k, conf_{ik}^j = z_k)$$

# Econometric Model – Approach

The latent class model

Let $g_{ik} = 3 \cdot y_{ik}^{j} + y_{ik}^{m}$, so that it can take on values 0,…,5.

The log-likelihood of our latent class model is based on the conditional multinomial density of $g_{ik}$:

$$P\left(g_{ik} = g \middle| x_i, conf_{ik}^{j} = z_{ik}\right)$$

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

$$\begin{aligned}
&P\left(g_{ik} = g \middle| x_i, conf_{ik}^{j} = z_{ik}\right) \\
&= P\left(g_{ik} = g \middle| \tilde{y}_{ik} = 1, x_i, conf_{ik}^{j} = z_{ik}\right) P\left(\tilde{y}_{ik} = 1 \middle| x_i, conf_{ik}^{j} = z_{ik}\right) \\
&+ P\left(g_{ik} = g \middle| \tilde{y}_{ik} = 0, x_i, conf_{ik}^{j} = z_{ik}\right) P\left(\tilde{y}_{ik} = 0 \middle| x_i, conf_{ik}^{j} = z_{ik}\right) \\
&= \alpha_g^1(x, z_k) P\left(\tilde{y}_i = 1 \middle| x_i, conf_{ik}^{j} = z_{ik}\right) \\
&+ \alpha_g^0(x, z_k) P\left(\tilde{y}_i = 0 \middle| x_i, conf_{ik}^{j} = z_{ik}\right)
\end{aligned}$$

# Latent class model (IV)



- We assume that

  *1.* $P\left(\tilde{y}_{ik} = 1 \middle| x_i, conf_{ik}^{j} = z_k\right) = P(\tilde{y}_{ik} = 1|x_i) = \Phi(x_i'\beta_k)$ (Probit)

  *2.* $P\left(g_{ik} = g \middle| \tilde{y}_{ik} = 1, x_i, conf_{ik}^{j} = z_k\right) = \alpha_g\left(z_k; \gamma_k^1\right)$: Mult. Logit, g=4 ref. group

  $(y_{ik}^{m} = y_{ik}^{j} = 1\ correct\ answers\ in\ May\ and\ July)$

  *3.* $P\left(g_{ik} = g \middle| \tilde{y}_{ik} = 0, x_i, conf_{ik}^{j} = z_k\right) = \alpha_g\left(z_k; ; \gamma_k^0\right)$ (Mult Logit, g=0 ref. group)

- Then we can write

$$P\left(g_{ik} = g \middle| x_i, conf_{ik}^{j} = z_k\right) = \alpha_g\left(z_k; \gamma_k^1\right)\Phi(x_i'\beta_k) + \alpha_g\left(z_k; \gamma_k^0\right)\Phi(-x_i'\beta_k)$$

- Identification problem

the parameter vector $(\gamma_k^{1\prime}, \gamma_k^{0\prime}, \beta_k')'$ is observationally equivalent with $(\gamma_k^{0\prime}, \gamma_k^{1\prime}, -\beta_k{}')'$ in the sense that they both result in the same probability distribution of observable data.

# Latent class model (V): Identifying assumptions

1. $\alpha_0^1(z_k) = P\left(g_{ik} = 0 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = P\left(y_i^m = 0, y_i^j = 0 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = 0, z_k = 1,..,7$
   (if a resp truly knows the answer to FL question, he/she will not pick a wrong answer twice.)

2. $\alpha_1^1(z_k) = P\left(g_{ik} = 1 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = P\left(y_i^m = 1, y_i^j = 0 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = 0, z_k = 1,..,7$
3. $\alpha_3^1(z_k) = P\left(g_{ik} = 3 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = P\left(y_i^m = 0, y_i^j = 1 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = 0, z_k = 1,..,7$
   (conditional on true knowledge, resp will not answer correctly in May and incorrectly in July or vice versa )

4. $\alpha_2^1(z) = P\left(g_{ik} = 2 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = P\left(y_i^m = 2, y_i^j = 0 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = 0, z_k = 1,..,7$
   (Resp. with true knowledge who pick a "dk" response in May, would never answer incorrectly in July.)

5. $\alpha_4^0(z) = P\left(g_{ik} = 4 \middle| \tilde{y}_{ik} = 0, conf_{ik}^j = z_k\right) = P\left(y_i^m = 1, y_i^j = 1 \middle| \tilde{y}_{ik} = 0, conf_{ik}^j = z_k\right) = 0, z_k = 6,7$
   (Given that resp. doesn't have true knowledge ($\tilde{y}_{ik} = 0$) and given high confidence ($conf_{ik}^j = 6,7$), the probability of giving the correct answer twice is 0.)

# Econometric Model – Final Outcome

Empirical estimate of 'true' financial literacy

Once we estimate the parameters, for each financial literacy question, we can calculate:

$$P\left(\tilde{y}_{ik}=1\middle|g_{ik}=g,conf_{ik}^{j}=z_{ik},x_i\right)=\frac{\alpha_g^1(z_{ik};\gamma^1)\Phi(x_i'\beta)}{\alpha_g^1(z_{ik};\gamma^1)\Phi\left(x_i'\beta\right)+\alpha_g^0(z_{ik};\gamma^0)\Phi(-x_i'\beta)}$$

This can be interpreted as *the posterior probability of having true knowledge* (our latent variable) which results after updating using the information from the two surveys (Bayes' rule).

And we can compute our measure of financial literacy:

$$finlit_i=\sum_{k=1}^{3}P(\tilde{y}_{ik}=1|g_{ik}=g,conf_{ik}^{j}=z_k,x_i)$$

# Latent class model (VII)

- Notice that the posterior distribution of $\tilde{y}_{ik}$ is degenerate if the following conditions are met:
  - $P\left(\tilde{y}_{ik} = 1 \middle| g_{ik} = g, x_i, conf_{ik}^{j}\right) = 1$ if $\alpha_g^0\left(conf_{ik}^{j}; \gamma_k^0\right) = 0$
  - $P\left(\tilde{y}_{ik} = 1 \middle| g_{ik} = g, x_i, conf_{ik}^{j}\right) = 0$ if $\alpha_g^1\left(conf_{ik}^{j}; \gamma_k^1\right) = 0$

- So, $\tilde{y}_{ik} = 0$ with certainty if
  - respondents answer inconsistently over time (once correctly, once incorrectly),
  - answer incorrectly two times, or
  - pick the "do not know" answer in the May module and an incorrect answer in the July module.

- $\tilde{y}_{ik} = 1$ with certainty if he/she answers the financial literacy questions correctly two times (with a high conf level in July confidence level)

# Latent class model VIII

- For respondents who provide a ”DK” answer in May and a correct one in July, the LCM is used to predict the probability of true knowledge,

$$0 < P\left(\tilde{y}_{ik} = 1 \middle| g_{ik} = 5, x_i, conf_{ik}^{j}\right) < 1$$

# Results

# Overview of Results

## Financial literacy and gender gap using different measures

| | Total | Gender Difference (Men-Women) |
|---|---|---|
| **Panel A: May measure** | | |
| Interest | 88.6 | 7.5 |
| Inflation | 85.8 | 9.2 |
| Risk | 49.9 | 27.5 |
| Financial literacy measure | 2.24 | 0.45 |
| **Panel B: July measure** | | |
| Interest | 93.2 | 3.5 |
| Inflation | 91 | 6.2 |
| Risk | 78.3 | 9.4 |
| Financial literacy measure | 2.62 | 0.19 |
| **Panel C: true financial literacy** | | |
| Interest | 87.6 | 5.7 |
| Inflation | 86.3 | 8.8 |
| Risk | 63.8 | 13.8 |
| Financial literacy measure | 2.38 | 0.28 |

# Multivariate Regression Results

The gender gap in financial literacy (OLS regression)

| | May | July | True literacy |
|---|---|---|---|
| **Panel A. Only gender** | | | |
| Female | -0.442*** | -0.190*** | -0.284*** |
| | (0.0386) | (0.0291) | (0.0352) |
| Adjusted $R^2$ | 0.067 | 0.024 | 0.035 |
| **Panel B. With controls for age, income, education, marital status** | | | |
| Female | -0.361*** | -0.147*** | -0.225*** |
| | (0.0394) | (0.0301) | (0.0362) |
| Adjusted $R^2$ | 0.156 | 0.094 | 0.143 |

# Economic Consequences (OLS)

## Effects of different fl-measures on stock market participation

| | No controls | May | July | True literacy |
|---|---|---|---|---|
| Financial Literacy | | 0.090*** | 0.055*** | 0.067*** |
| | | (0.0105) | (0.0097) | (0.0101) |
| Gender | -0.136*** | -0.046*** | -0.072*** | -0.065*** |
| | (0.0207) | (0.0212) | (0.0213) | (0.0213) |
| Controls+ | no | yes | yes | yes |
| N | 1532 | 1532 | 1532 | 1532 |
| Adjusted $R^2$ | 0.022 | 0.137 | 0.117 | 0.122 |

Controls+: Age, income, education, marital status

# Economic Consequences (IV)

Taking potential reverse causality/omitted variables into account

- **Instrument**: Economics in high school
- **3 groups**: None, some, DK

| | May | July | True literacy |
|---|---|---|---|
| Financial Literacy | 0.192*** | 0.222*** | 0.204*** |
| | (0.0671) | (0.0842) | (0.0751) |
| Gender | -0.003 | -0.031 | -0.024 |
| | (0.0369) | (0.0308) | (0.0325) |
| First stage F-stats | 14.19 | 9.19 | 11.26 |

Further controls: Age, income, education, marital status

# Financial Literacy and Underconfindence

Quantifying underconfidence and its economic effects

- **Underconfidence** can be defined directly from our model
- Specifically, we calculate the **prob of true knowledge conditional on a DK-answer** in the first wave

$$\text{und\_conf} = \sum_{k=1}^{3} \mathrm{P}(\tilde{y}_{ik} = 1 | y_{ik}^{m} = 2, \text{conf}_{ik} = z, x_i) \cdot I(y_{ik}^{m} = 2)$$

| | OLS I | OLS II | GMM I | GMM II |
|---|---|---|---|---|
| Financial Literacy | 0.067*** | 0.070*** | 0.183** | 0.180** |
| true literacy | (0.0101) | (0.0100) | (0.082) | (0.0705) |
| Underconfidence | | -0.062*** | -0.056 | -0.066*** |
| | | (0.0094) | (0.113) | (0.0099) |
| Gender | -0.065*** | -0.047** | -0.015 | -0.013 |
| | (0.0213) | (0.0211) | (0.0368) | (0.0318) |
| $R^2$ | 0.132 | 0.150 | 0.094 | 0.098 |

# Using DKs as Proxy

## Effects of different fl-measures on stock market participation

| | True Finlit | True+ Underconf | May Finlit | May Finlit + # of DKs |
|---|---|---|---|---|
| Financial Literacy | 0.0672*** | 0.0707*** | 0.0901*** | 0.0666*** |
| | (0.0101) | (0.0100) | (0.0105) | (0.0187) |
| Gender | -0.0646*** | -0.044** | -0.0461** | -0.0443** |
| | (0.0213) | (0.0212) | (0.0212) | (0.0213) |
| Controls+ | yes | yes | yes | yes |
| N | 1532 | 1532 | 1532 | 1532 |
| Adjusted $R^2$ | 0.122 | 0.140 | 0.137 | 0.138 |

Controls+: Age, income, education, marital status

# Conclusion

Main insights

***Financial knowledge and confidence***

- We differentiate two channels for the observed gender gap in financial literacy: a gap in *knowledge (2/3)* and a gap in *confidence (1/3)*

- We are able to estimate whether a respondent *truly knows* the correct answer and therefore get a better measure that matters for behavior

**Financial literacy and confidence matter**

- They both explain stock market participation

# Conclusion

Policy implications

- Financial literacy matters
- Need to improve the levels of financial literacy, in particular among women
- **More research (!)** necessary to understand how to also instill confidence, in particular among women.
- *Fearless Girl* symbolizes this suggestion

*Financially, women on average know less than men –*
*but they know more than they think they know.*