P. Kuśmierczyk, R. Kurach, M. Kośny

On methodological problems with measuring (financial) knowledge

Are women truly less financially literate than men?

This work was supported by:
the National Science Centre, Poland under Grant DEC-2017/25/B/HS4/00186 and by the Ministry of Education and Science in Poland under the programme "Regional Initiative of Excellence" 2019 - 2023 project number 015/RID/2018/19 total funding amount 10 721 040,00 PLN".

Starting point

There are multiple ways in which financial literacy could be measured:

- BIG3
- questionnairies with more questions
- repeated questionnairies (panels)
- experimental methods.

The studies dedicated to measuring financial literacy probably do it more accurately. But that comes at the cost.

As economists we advocate using short and simple measures, which could be easily introduced in any economic studies. Simple, unidimensional measures have their drawbacks, yet economists use them a lot, due to their utility (e.g. risk aversion parameter to measure risk attitude, discount rate to measure time preferences).

So, we advocate using BIG3, potentially, though, with some modifications.

BIG 3 measure (Lusardi and Mitchell, 2004)

Q1. [COMPOUNDED 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?

a) More than $102
b) Exactly $102
c) Less than $102
d) Do not know
e) Refuse to answer

Q2. [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?

a) More than today
b) Exactly the same
c) Less than today
d) Do not know
e) Refuse to answer

Q3. [DIVERSIFICATION] 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."

a) True
b) False
c) Do not know
d) Refuse to answer

$$BIG3 = \sum_{j=1}^{M} s_{1j}$$

where: s_{1j} takes value 1 if a respondent marks the correct answer in question j and 0 otherwise

Identified problems

1. Knowledge is commonly associated with the proportion of correct answers. Therefore incorrect, „don't know" answers or refusals are all treated equally (as a demonstration of the lack of knowledge). But this is a very simplified approach that might provide a biased glimpse on financial knowledge.
2. Using **different modeling of knowledge**, we might get an insight into its structure, and the role that self-confidence, or misinformation play in manifestation of knowledge. This enables us to **design a modified BIG3 measure**. The ultimate test of a measure's usefulness should be external (its applicability to predict financial choices)
3. Is the commonly believed and reported **gender gap** (lower financial knowledge of women compared to men) a fact or an artifact, resulting from applying an overly simplistic model? In particular – does the „don't know" answer necessarily mean complete lack of knowledge?
4. There are **huge differences in reported results** of BIG3, when looking at various countries' data or even the same country's data from various years. What is the reason for that? Would different modeling shed some light on potential reasons of these differences? (a problem not addressed in this presentation)

General model of knowledge

Assumptions:

- Let w_{ij} denote a tuple that shows the i-th respondent's conviction concerning the "truthfulness" of each potential response in j-th question. The tuple's elements show the probability of choosing each answer, if the respondent decides to pick one. Let w_{ij}^* denote the maximal value of the tuple's elements.
- Let r_i denote a reservation level of a respondent; (s)he only marks the answer when $w_{ij}^* \geq r_i$, i.e. if her conviction concerning the truthfulness of at least one answer is higher than the reservation level. If she marks the answer, she picks one according to probability vector w_{ij}
- We will interpret $c_i = 1 - r_i$ as the i-th respondent confidence level
- To simplify notation, we will assume, that the 1st element in the tuple describes the respondent's conviction concerning the correct answer

General model of knowledge

In short: respondents answer result from their knowledge and self-confidence.

Examples:

- Let $w_{ij} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. This shows a respondent that is completely uninformed. If $r_i = 0.5$ she marks „don't know" (she has to be at least 50% certain about some answer to mark it. If $r_i = 0$ she always marks the asnwer (100% confidence), and in this case she would just shoot it.
- Let $w_{ij} = (1,0,0)$. This shows a respondent that has full knowledge. Independent of her confidence she marks the correct answer for sure.
- Let $w_{ij} = (0,1,0)$. This shows a respondent that has misinformed (sure of he wrong answer). Independent of her confidence she marks the false answer for sure.

The model is not estimable in the general case. In order to make estimations we need to introduce some additional assumptions that will simplify the model.

Partial information model - description

Let us assume that w_{ij} can only take 3 values:

$$w_{i,j} = \begin{cases} \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) & uninformed \\ \left(\frac{1}{2}, \frac{1}{2}, 0\right) \ or \ \left(\frac{1}{2}, 0, \frac{1}{2}\right) & partially\ informed \\ (1,0,0) & fully\ informed \end{cases}$$

and let us assume that r_i also takes only 2 values:

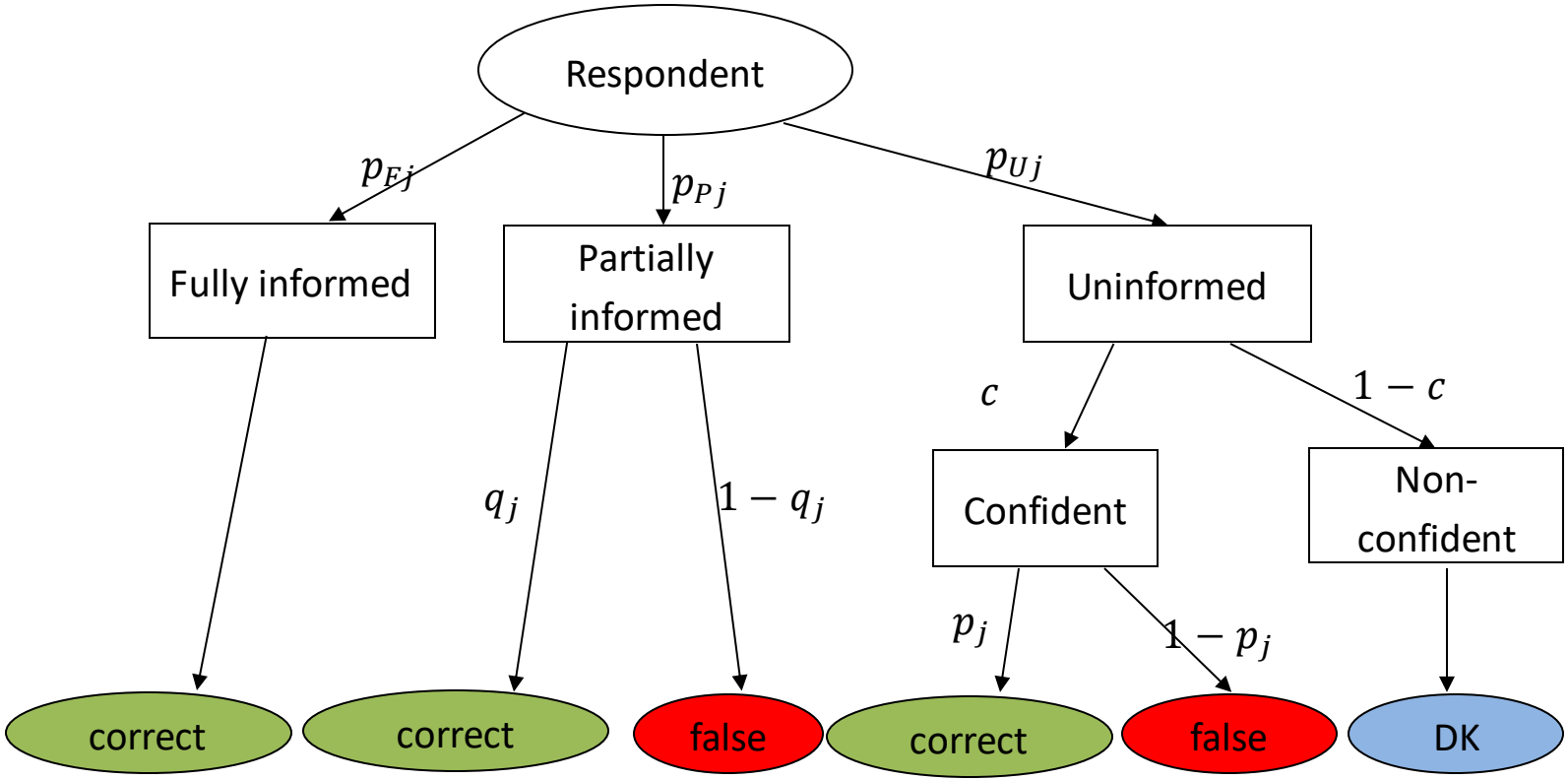
$$r_i = \begin{cases} 0 & confident \\ 0{,}5 & non-confident \end{cases}$$

So, the respondent either knows the correct answer, and then always marks the correct answer.

Or (s)he is partially informed, and then she always shoots, marking the correct answer with p. 1/2.

Or (s)he is uninformed, and then, when confident, (s)he shoots the answer, and when not confident marks DK.

Partial information model - graph



Misinformation model - description

Let us assume that w_{ij} can only take 3 values:

$$w_{i,j} = \begin{cases} \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) & uninformed \\ (0,1,0)\ or\ (0,0,1) & misinformed \\ (1,0,0) & well-informed \end{cases}$$

and let us assume that r_i also takes only 2 values:

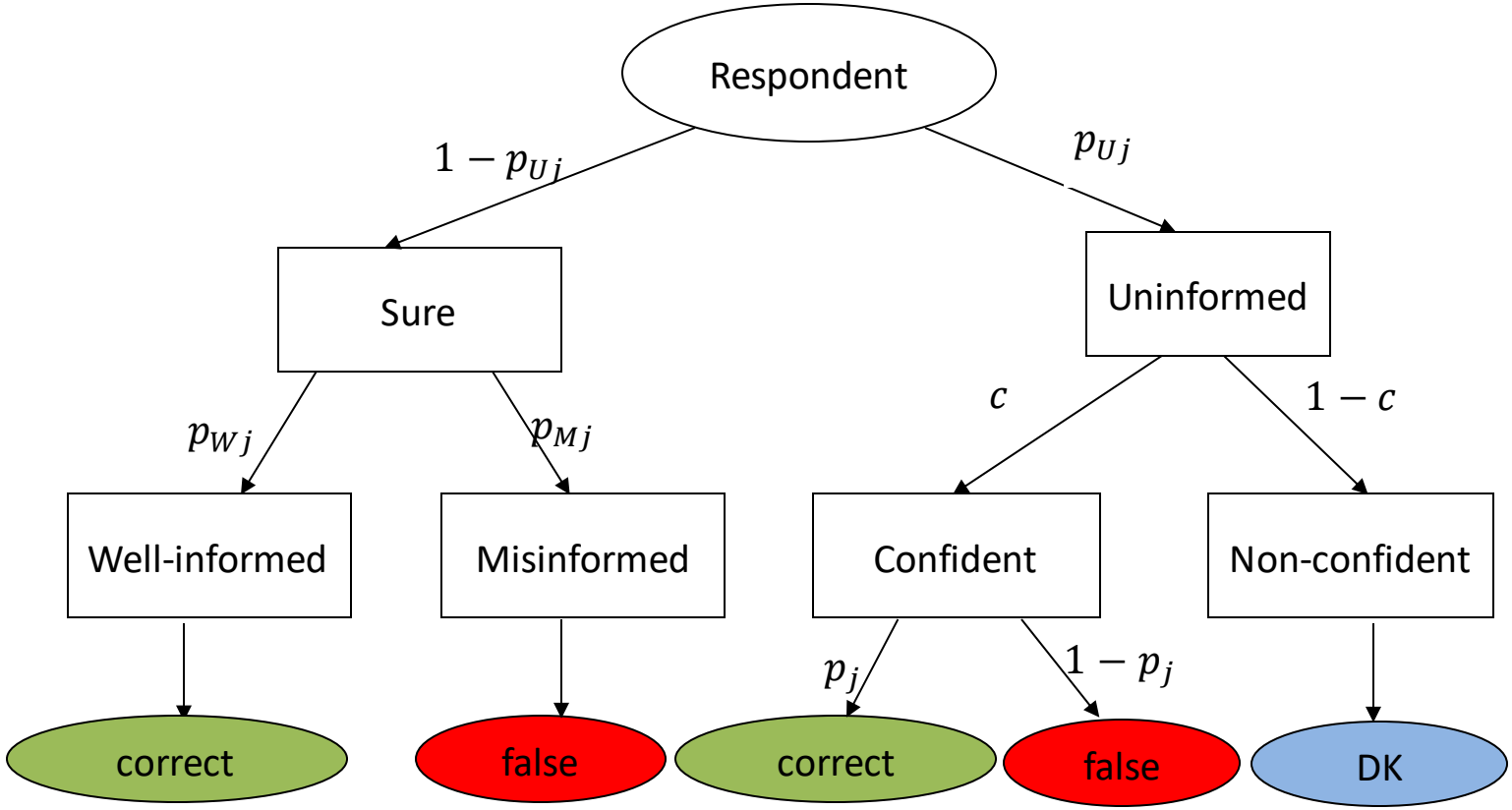
$$r_i = \begin{cases} 0 & confident \\ 0{,}5 & non-confident \end{cases}$$

So, the respondent either knows the correct answer, and then always marks the correct answer.

Or (s)he is misinformed and then (s)he always marks the wrong answer.

Or (s)he is uninformed, and then, when confident, (s)he shoots the answer, and when not confident (s)he marks DK.

Misinformation model - graph



Estimation of the parameters

In case of both models using aggregated data requires estimation of 3 unknowns on the basis of 2 equations.

The third unknown can be estimated once we assume it is question-independent and use the detailed answers of respondents. Consider the misinformation model. We assumed that the variable *c* is question-independent. Hence, we assume that self-confidence is a person's characteristics, that will affect their responses in case of any question.

Let us denote by $\boldsymbol{C_i}$ the binomial variable that takes 1 if a respondent is self-confident and 0 otherwise.

Let us denote by $\boldsymbol{\pi_i}$ the *i*-th respondent profile of answers, assuming 3 possible answers: correct (1), false (0), and „don't know" (2).

In total there is as many as $\boldsymbol{3^M}$ profiles (with *M* – the numer of questions).

For each possible profile of answers let us calculate $\boldsymbol{E(C_i|\pi_i)}$.

Using the total probability formula we calculate $\boldsymbol{c = E_i[E(C_i|\pi_i)]}$

Define a modified measure of financial literacy, e.g.: $\boldsymbol{BIG3m_i = \sum_{j=1}^{M} \mathrm{E}\left(X_{Wij}|\boldsymbol{\pi_i}\right)}$

Interpretation

The models calculate probabilities that a respondent is knowledgeable conditional on her profile of answers, i.e. conditional on all her answers.

As an example consider two profiles of answers:

- $\pi_A = (1,2,2)$
- $\pi_B = (1,0,0)$

Using standard approach, both respondents have the same value of BIG3 (1 correct answer), but once we take into account the conditional probabilities, the respondent A is much more likely to know the answer to the first question than the respondent B. The first one for sure is not confident, and so we know she didn't shoot the first answer, whereas in case of the second respondent it is much more likely.

So, the modified BIG3 measure would provide a higher value for respondent A.

Applying the models to Polish data

We applied the models to financial literacy studies conducted in Poland in years 2011-2016 (data obtained from Bank of Austria), and our own study from 2019.

We will only focus on the data concerning the alleged gender gap.

We present results concerning:

- original gap (no modeling)
- partial information model
- misinformation model

The numbers in the tables show the differences between values for males and females.

Original gap (no modeling)

Data	GAP (original)		
	Q1	**Q2**	**Q3**
6_2011	0,0209	-0,013	0,016
6_2012	0,031	0,0237	0,075
6_2013	0,0339	0,0405	-0
6_2014	0,0137	0,0357	0,024
6_2015	0,0705	0,0216	0,053
6_2016	0,0197	0,024	0,073
2019	0,0221	0,0609	0,027

Comments:

- men have significantly higher financial knowledge in 6 (out of 21) cases
- women never reported to have significantly higher financial knowledge

Partial information model

Data	confidence	GAP (full knowledge)		
		Q1	Q2	Q3
6_2011	0,047981	0,035433	-0,02667	-0,0023
6_2012	0,019209	-0,00572	0,000827	0,022508
6_2013	0,000627	0,033562	0,020592	-0,01231
6_2014	0,084868	0,002568	0,019015	-0,02498
6_2015	0,038474	0,086879	0,005939	-0,00956
6_2016	0,046949	0,023318	0,042777	0,060696
2019	0,060901	0,02434	0,03499	-0,07831

Comments:

- men have significantly higher confidence level (all differences positive, and half of them significant)
- men have significantly higher full knowledge in 3 (out of 21) cases
- women have significantly higher full knowledge in 1 case

Misinformation model

Data	confidence	GAP conditional (well-informed)			GAP total		
		Q1	Q2	Q3	Q1	Q2	Q3
6_2011	0,048	0,035	-0,019	-2E-06	0,012406	-0,02432	-0,00229
6_2012	0,019	-0,09	-0,035	-0,002	0,039358	0,028125	0,095342
6_2013	0,009	-0,01	0,004	-0,157	0,041139	0,054963	-0,04995
6_2014	0,084	0,007	0,017	-0,266	-0,00263	0,021771	-0,04709
6_2015	0,053	0,057	0,017	-0,238	0,073597	0,01219	0,025665
6_2016	0,056	0,032	0,103	-0,125	0,008942	0,004115	0,028372
2019	0,061	0,012	-0,01	-0,06	0,019833	0,074972	0,026854

Comments:

- men have significantly higher confidence level (all differences positive, and most of them significant)
- solid grounds to reject the hypothesis that women are less financially literate when considering the benefits from diversification
- when women are sure of the answer (conditional probabilities) they are usually right – **isn't it a symptom of making better decisions?**

Sumary and further work plan

1. We propose modeling that can be applied to any existing data.
2. The models allow for reaching different interpetation of the past results, including such potential factors as self-confidence, partial knowledge, misinformation etc.
3. At this moment it is impossible to judge which modeling of knowledge is the most insightful. We believe that each of them can bring something in. Ultimately, we should search for the model that leads to the the most efficient measure of financial knowledge (using external test → financial decisions).
4. We challange the common belief that women have a lower financial knowledge than men do. Women do mark the correct answers less often, but once we allow for a different interpretation (different model of knowledge) this does not necessarily translate into „less knowledge".
5. We have collected a huge set of data from the CEE countries and we intend to „remodel" them. These sets of data include additional economic variables and we will search how they correlate with our modified measures.

Thank you for the attention!

pawel.kusmierczyk@ue.wroc.pl