A Reconsideration of Gender Differences in Risk Attitudes

Paolo Crosetto and Antonio Filippin

INRA - GAEL, Grenoble
and
Università degli Studi di Milano

ASFEE Conference, Besançon
May 23rd, 2014

Gender differences: a stylized fact

Are evolution and biology dictating that women are more risk averse than men? Or is the gender gap in risk aversion an outcome of child-rearing practices?
(Bertrand 2011)

Our subject pool is atypical in the sense that the female subjects were generally less risk averse than the male subjects.
(Anderson and Mellor, 2009)

My main concern is that your paper gives the impression as if there were no gender differences in risk attitudes. I think you must make very clear in the paper that this is not true.
(An anonymous referee for another paper, 2014)

Where does this evidence come from?

- surveys by Eckel and Grossman (2008) and Croson and Gneezy (2009). But coverage is limited (16 and 10 papers).
- **Eckel Grossman:** Sizeable differences emerge both in the experiment presenting the task (Eckel and Grossman, 2002, 2008a) as well as in almost all the later replications.
- **Investment Game:** A survey by Charness and Gneezy (2012) reports that males invest sizeably, systematically, and significantly more than females.
- self-reported SOEP scale (Dohmen et al. 2011)

But...

- psychology survey: only somewhat more than half of >150 studies reported by Birnes et al. (1999)
- **BRET:** In the BRET task the absence of gender differences is a robust result (Crosetto and Filippin, 2013b). New task but $N > 1000$.
- **Holt and Laury:** A survey for the most popular elicitation task is missing.

We thougth this result is very interesting and worth digging deeper...

Summary

Take home message

- Gender differences are less ubiquitous than what depicted to be.
- Observing a gender gap is task-dependent.
- Some features of the task correlate with the likelihood of observing gender differences.

How do we do it?

1. Survey of the literature
2. In-depth analysis of data from the replications of Holt and Laury (today)
3. Experimental evidence (another paper - if we have time)

Comprehensive analysis of the HL task

Article	N	significant gender difference	lab?	subjects	country	type of evidence	p-value
?	212	mixed	lab	students	U.S.	text	-
?	845	yes	lab	non-student	U.S.	t-test	< 0.05
?	188	yes	lab	-	Spain	Mann-Whitney	0.0027
?	801	yes	lab	labor force	Canada	coefficient	0.001
?	80	mixed	lab	students	U.S.	various	-
?	6496	mixed	field	high school	Italy	various	-
?	90	no	lab	students	Denmark	coefficient	0.38
?	140	no	lab	students	U.S.	coefficient	0.54
?	120	no	lab	students	U.S.	coefficient	0.891
?	213	no	field	rural	China	Wilcoxon	0.14
?	74	no	lab	students	India	coefficient	0.644
?	57	no	lab	general	Greece	coefficient	> 0.05
?	232	no	lab	students	U.S.	coefficient	0.586
?	345	no	lab	students	CN, F, Niger, U.S.	text	-
?	178	no	lab	students	U.S.	text	-
?	108	no	lab	students	Colombia	coefficient	0.78
?	204	no	lab	students	Germany	text	-
?	144	no	lab	students+	France	coefficient	0.19
?	127	no	lab	students	Germany	text	-
?	48	no	lab	students	Spain	correlation	-
?	144	no	lab	students	U.S.	text	-

Table : Gender results as reported in the HL literature

Why a comprehensive analysis of HL task

It is difficult to draw a conclusion about the HL task relying upon the evidence available in the literature.

1. Low number of contributions reporting about gender differences
2. Results not easy to compare, even when reported (non parametric or parametric tests, multivariate regressions...)
3. Tasks not always homogeneous
4. Treatment of inconsistencies not homogeneous.

We decided to go beyond a meta-analysis and collect the data of the replications

Replications of HL

Published HL replications as of Jan 31st, 2013:	**118**	
of which:		
Not recording gender or single gender	16	
Duplicate dataset	8	
Universe of reference	**94**	100%
of which:		
No response or not sharing the data	40	42.5%
Final dataset	**54**	57.5%
of which:		
Microdata (shared or available online)	48	
Summary statistics	6	

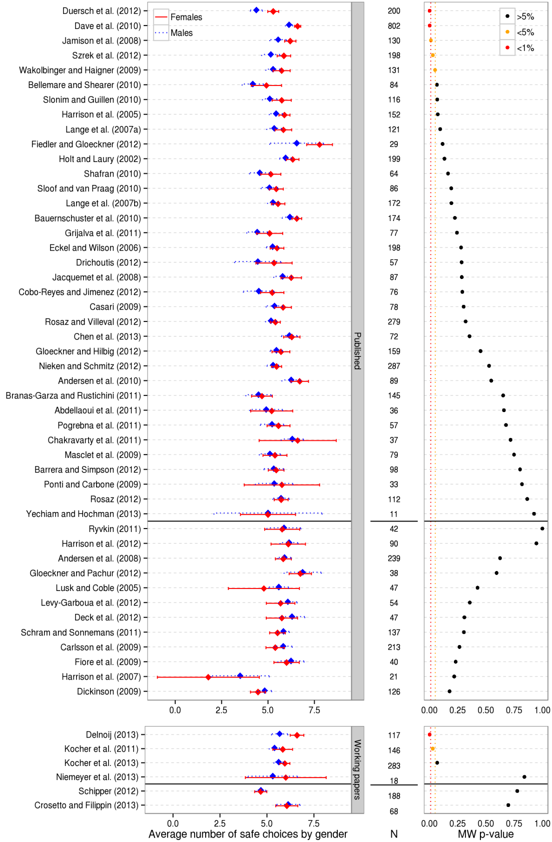
The original Holt and Laury (2002) is included in the 118 replications.

Table : Dataset of HL replications

Subjects in the sample

	Detail	Consistent subjects			Inconsistent subjects		
		Males	Females	Total	Males	Females	Total
Microdata	*full*	2119	2205	4324	411	502	913
# safe + consistency	*partial*	504	408	912	64	98	162
# safe only		375	324	699	3	1	4
Summary statistics	*summary*	413	359	772	-	-	
Total		3411	3296	**6707**	478	601	**1079**

Table : Subjects in the sample by consistency and type of data. Published papers only.


Females
Males
Duersch et al. (2012)
Dave et al. (2010)
Jamison et al. (2008)
Szrek et al. (2012)
Wakolbinger and Haigner (2009)
Bellemare and Shearer (2010)
Slonim and Guillen (2010)
Harrison et al. (2005)
Lange et al. (2007a)
Fiedler and Gloeckner (2012)
Holt and Laury (2002)
Shafran (2010)
Sloof and van Praag (2010)
Lange et al. (2007b)
Bauernschuster et al. (2010)
Grijalva et al. (2011)
Eckel and Wilson (2006)
Drichoutis (2012)
Jacquemet et al. (2008)
Cobo-Reyes and Jimenez (2012)
Casari (2009)
Rosaz and Villeval (2012)
Chen et al. (2013)
Gloeckner and Hilbig (2012)
Nieken and Schmitz (2012)
Andersen et al. (2010)
Branas-Garza and Rustichini (2011)
Abdellaoui et al. (2011)
Pogrebna et al. (2011)
Chakravarty et al. (2011)
Masclet et al. (2009)
Barrera and Simpson (2012)
Ponti and Carbone (2009)
Rosaz (2012)
Yechiam and Hochman (2013)
Ryvkin (2011)
Harrison et al. (2012)
Andersen et al. (2008)
Gloeckner and Pachur (2012)
Lusk and Coble (2005)
Levy-Garboua et al. (2012)
Deck et al. (2012)
Schram and Sonnemans (2011)
Carlsson et al. (2009)
Fiore et al. (2009)
Harrison et al. (2007)
Dickinson (2009)
Published
0.0
2.5
5.0
7.5
Delnoij (2013)
Kocher et al. (2011)
Kocher et al. (2013)
Niemeyer et al. (2013)
Schipper (2012)
Crosetto and Filippin (2013)
Working papers
Average number of safe choices by gender
200
802
130
198
131
84
116
152
121
29
199
64
86
172
174
77
198
57
87
76
78
279
72
159
287
89
145
36
57
37
79
98
33
112
11
42
90
239
38
47
54
47
137
213
40
21
126
117
146
283
18
188
68
N
>5%
<5%
<1%
0.00
0.25
0.50
0.75
1.00
MW p-value

Analysis of the papers separately

Very good coverage (58%) of the universe of papers replicating the HL task, larger sample than contributions reporting about gender differences (54 vs. 21).

Analysis includes only subjects making consistent choices

Of the 54 datasets replicating the classic HL task only 5 display gender differences reaching at least the 5% of significance (8 out of 64 considering also the non-Scopus papers).

Merging the dataset: inconsistencies

	Inconsistent choices		*% of inconsistent subjects*		
	Number	out of	Males	Females	Total
Switching from B to A	973	6962	12.1	15.8	14.0
Always Option A	100	6334	1.8	1.3	1.6
Always Option B	6	383	1.4	1.7	1.5
Total	**1079**	-			

Note. For each type if inconsistency the maximum number of observations (out of) has been computed separately, including only the studies in which each event can possibly happen.

Table : Summary statistics of inconsistent subjects by type and gender

- About 14% of the subjects make inconsistent choices
- Females are more likely to make inconsistent choices (numeracy? No, see below)
- Inconsistent subjects make less safe choices on average (5.15 Vs. 5.63) and significantly so ($p < 0.001$)
- why? random choice most likely (Anderssen et al 2013).

Merging the dataset: gender differences

	Mean	St.Dev	N
Whole sample	**5.63**	**1.91**	**5935**
Males	5.47	1.89	2998
Females	5.78	1.91	2937
Microdata	**5.73**	**1.96**	**4324**
Males	5.59	1.94	2119
Females	5.87	1.97	2205

Merging the dataset: gender differences

	Dep. var.: number of safe choices			
	(1)	(2)	(3)	(4)
female	.311***	.315***	.280***	.288***
realmoney		.013***	.020***	
realmoney^2/100		-.0.04***	-.0.07***	
exchange/100		.010	-.002	
randomorder		.361***	.311***	
fixed effects	no	no	no	yes
R^2	.007	.019	.024	.095
N	5935	5935	4324	5935

Table : Determinants of the number of safe choices

Maximum likelihood model

We assume that subjects are EU maximizers with CRRA preferences $U(x) = x^r$, and that they can make an evaluation error μ when comparing the utility of the two lotteries A (safe) and B (risky). The probability of choosing the safe lottery is

$$Prob(S) = \frac{EU_A^{\frac{1}{\mu}}}{EU_B^{\frac{1}{\mu}} + EU_A^{\frac{1}{\mu}}}, \quad \text{and } EU_i = \sum_j p_j (x_j)^r,$$

The probability converges to $\frac{1}{2}$ as $\mu \to \infty$, and, as $\mu \to 0$, to 1 if $EU_A > EU_B$ and to 0 if $EU_A < EU_B$.
We estimate over the whole dataset a structural model of choice using maximum likelihood and clustering standard errors by subject.

Maximum likelihood results

		Coeff.		*St.Err.*
r	constant	0.640	***	(0.0179)
	female	-0.0633	**	(0.0203)
	realmoney/100	-0.457	***	(0.1028)
	realmoney^2/100	0.00158	***	(0.0003)
	randomorder	-0.0950	*	(0.0392)
	exchange/1000	0.00348		(0.0313)
μ	constant	0.229	***	(0.0073)
	female	-0.0135		(0.0085)
	realmoney/100	-0.19	***	(0.0247)
	realmoney^2/100	0.000658	***	(0.0000)
	randomorder	0.012		(0.0160)
	exchange/1000	0.00861		(0.1160)
N decisions				52735
n subjects				5237
Log-likelihood				-23494.025
Wald χ^2 p-value				0.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. t statistics in parentheses.

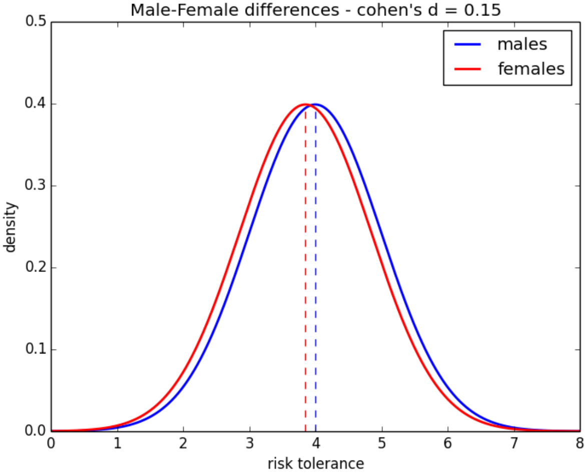
Merging the dataset 2

Focus on gender differences:

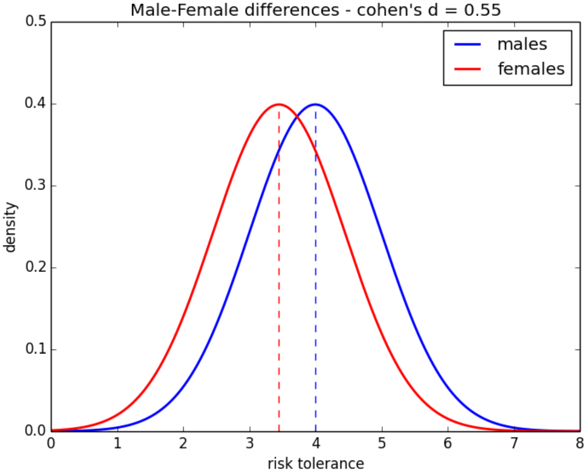
- Pooling all the microdata boosts the statistical power and gender differences become significant ($p < 0.001$)
- Results are confirmed in a Maximum Likelihood specification allowing for decisions with errors.
- The magnitude is very small, however: Cohen's $d = 0.15$ (about 15% of a standard deviation), below the threshold commonly used to define a small effect (0.2)
- The same statistic computed for the Investment Game and for the EG gives a Cohen's $d = 0.55$, above the threshold commonly used to define a medium effect (0.5)

Punch line: the likelihood of observing gender differences as well as their magnitude show clearly different patterns across elicitation methods. This seems to reflect some measurement error induced by the (features of the) tasks rather than genuine differences in risk aversion.
These findings are confirmed by our own experimental replications of the aforementioned tasks.

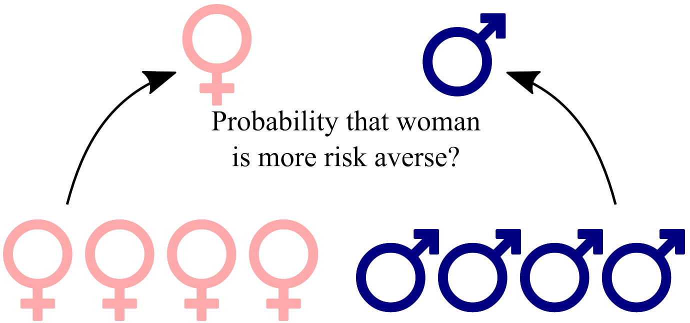
Cohen's D - I


Male-Female differences - cohen's d = 0.15
males
females
density
risk tolerance
0.0
0.1
0.2
0.3
0.4
0.5
0
1
2
3
4
5
6
7
8

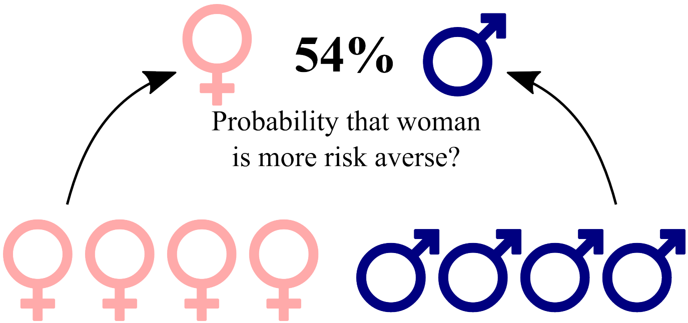
Cohen's D - II


Male-Female differences - cohen's d = 0.55
males
females
density
risk tolerance
0.0
0.1
0.2
0.3
0.4
0.5
0
1
2
3
4
5
6
7
8

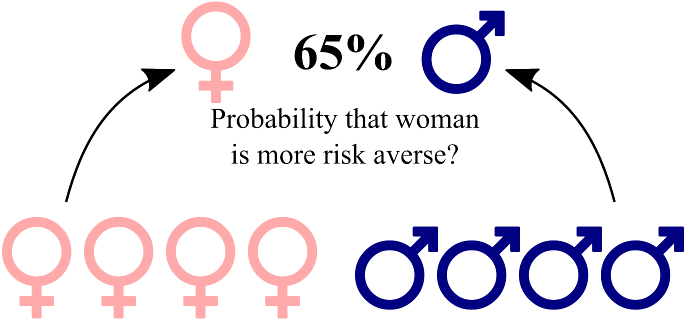
Implication



Implication - cohen's D = 0.15



Implication - cohen's D = 0.55



Interpretation

Risk attitudes and measurement error

Risk attitudes are a latent construct.

Each elicitation method provides an assessment of risk attitudes with some measurement error that depends on the specific characteristics of each task.

Instability of findings across elicitation methods is a robust result (not only about gender).

Next research question: What drives such an instability as far as gender differences are concerned?

Risk Elicitation Tasks

Risk Elicitation Tasks, I: Eckel Grossman

A single choice among five (six) 50-50 lotteries – chosen lottery is played.

Lottery 1	50%	A	4 Euro
	50%	B	4 Euro
Lottery 2	50%	A	6 Euro
	50%	B	3 Euro
Lottery 3	50%	A	8 Euro
	50%	B	2 Euro
Lottery 4	50%	A	10 Euro
	50%	B	1 Euro
Lottery 5	50%	A	12 Euro
	50%	B	0 Euro

Subject select a lottery.
Risk-neutral should select lottery 5, risk lover as well.

Risk Elicitation Tasks, II: The Investment Game

Endowment (4€)

How much would you like to invest (X)?

Safe account
4 - X

Risky investment
50%: 2.5*X ; 50%: 0

Subject self-select into a lottery.
Risk-neutral should invest all, as $E(\text{risky}) = 1.25 > 1$.

Risk Elicitation Tasks, III: Holt and Laury

Ten binary lottery choices – risk attitude as switching point

1/10 prob. of 4.0 Euro	9/10 prob. of 3.2 Euro	**A**	○	○	**B**	1/10 prob. of 7.7 Euro	9/10 prob. of 0.2 Euro
2/10 prob. of 4.0 Euro	8/10 prob. of 3.2 Euro	**A**	○	○	**B**	2/10 prob. of 7.7 Euro	8/10 prob. of 0.2 Euro
3/10 prob. of 4.0 Euro	7/10 prob. of 3.2 Euro	**A**	○	○	**B**	3/10 prob. of 7.7 Euro	7/10 prob. of 0.2 Euro
4/10 prob. of 4.0 Euro	6/10 prob. of 3.2 Euro	**A**	○	○	**B**	4/10 prob. of 7.7 Euro	6/10 prob. of 0.2 Euro
5/10 prob. of 4.0 Euro	5/10 prob. of 3.2 Euro	**A**	○	○	**B**	5/10 prob. of 7.7 Euro	5/10 prob. of 0.2 Euro
6/10 prob. of 4.0 Euro	4/10 prob. of 3.2 Euro	**A**	○	○	**B**	6/10 prob. of 7.7 Euro	4/10 prob. of 0.2 Euro
7/10 prob. of 4.0 Euro	3/10 prob. of 3.2 Euro	**A**	○	○	**B**	7/10 prob. of 7.7 Euro	3/10 prob. of 0.2 Euro
8/10 prob. of 4.0 Euro	2/10 prob. of 3.2 Euro	**A**	○	○	**B**	8/10 prob. of 7.7 Euro	2/10 prob. of 0.2 Euro
9/10 prob. of 4.0 Euro	1/10 prob. of 3.2 Euro	**A**	○	○	**B**	9/10 prob. of 7.7 Euro	1/10 prob. of 0.2 Euro
10/10 prob. of 4.0 Euro	0/10 prob. of 3.2 Euro	**A**	○	○	**B**	10/10 prob. of 7.7 Euro	0/10 prob. of 0.2 Euro

One row is selected and played.
Risk aversion measured by the nr of safe choices (risk-neutral switches in row 5)
Inconsistent choices are likely.

Risk Elicitation Tasks, IV: The Bomb Risk Elicitation Task (BRET)

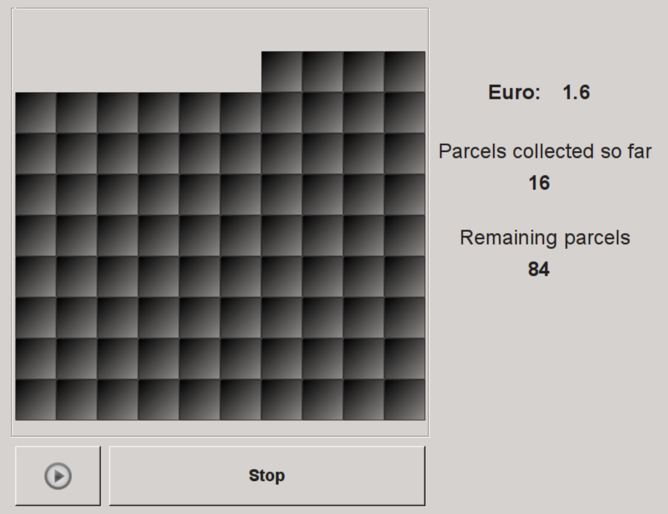


Figure : The BRET interface after 16 seconds

Determinants of gender differences

Interesting correlation between the likelihood of observing gender differences and

- the availability of a safe alternative in the choice set
- changing the expected value of the alternatives keeping fixed the probabilities $(50\% - 50\%)$

.

	Highest sure amount	Focality	Probabilities	Gender gap found
HL	Low outcome safe lottery	Low	Moving	Pooling the results
EG	Outcome degenerate lottery	High	$50\% - 50\%$	Always
CGP	Endowment	High	$50\% - 50\%$	Always
BRET	None	None	Moving	Never

Determinants of gender differences

- Unfortunately, the results available in the literature do not allow to disentangle the two explanatory factors proposed because they are correlated.

.

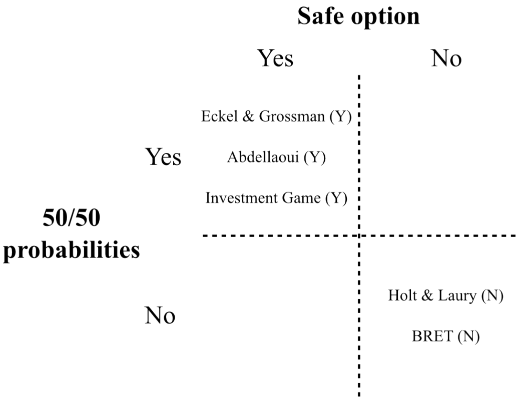


Figure : Results by gender of different tasks

Possible explanations

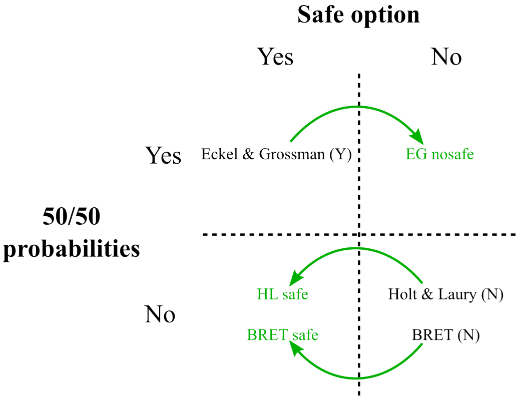
Several determinants may drive the observed behaviour:

- **Certainty effect**. Certainty effect has already been shown to cause failures of EUT predictions (Andreoni and Sprenger, 2012)
- **Loss aversion**. Safe amount can act as an Endeogenous Reference Point (Koszegi and Rabin, 2007) around which subjects could evaluate the other outcomes.
- **Probability weighting** (Fehr-Duda et al. 2006)
- **Salience theory** (Bordalo et al. 2012)
- **Regret** (Loomes and Sudgen 1982)
- **... any idea?**

Experimental manipulation to try to disentangle the characteristics

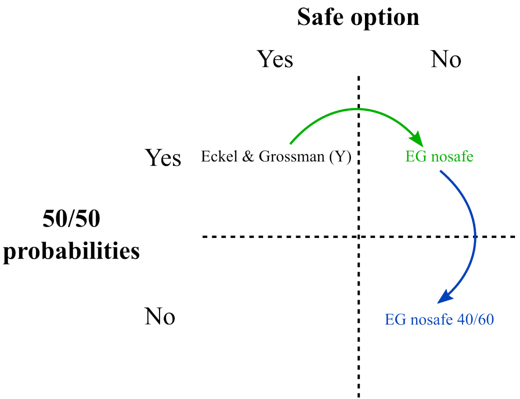
Experimental manipulation: safe alternative

Some of the tasks are manipulated adding/removing the safe option



Experimental manipulation: 50-50 probabilities

One task can also be manipulated removing the 50-50 probabilities. It is not possible instead to fix the probabilities to 50-50 in the two task where they are moving.



Manipulations

We aim at disentangling which characteristic of the task generates a gender difference

How do we do it? Safe option

1. HL with and without a safe option
2. BRET with and without a safe option
3. EG with and without a safe option

How do we do it? Changing Probabilities

1. EG with 40-60 probabilities
2. EG with moving probabilities (not incentivized)

Availability of a safe alternative

HL with safe option

Treatment 1: Both Risky

1/10 prob. of 4.0 Euro	9/10 prob. of 3.2 Euro	A	B	1/10 prob. of 7.7 Euro	9/10 prob. of 0.2 Euro
2/10 prob. of 4.0 Euro	8/10 prob. of 3.2 Euro	A	B	2/10 prob. of 7.7 Euro	8/10 prob. of 0.2 Euro
3/10 prob. of 4.0 Euro	7/10 prob. of 3.2 Euro	A	B	3/10 prob. of 7.7 Euro	7/10 prob. of 0.2 Euro
4/10 prob. of 4.0 Euro	6/10 prob. of 3.2 Euro	A	B	4/10 prob. of 7.7 Euro	6/10 prob. of 0.2 Euro
5/10 prob. of 4.0 Euro	5/10 prob. of 3.2 Euro	A	B	5/10 prob. of 7.7 Euro	5/10 prob. of 0.2 Euro
6/10 prob. of 4.0 Euro	4/10 prob. of 3.2 Euro	A	B	6/10 prob. of 7.7 Euro	4/10 prob. of 0.2 Euro
7/10 prob. of 4.0 Euro	3/10 prob. of 3.2 Euro	A	B	7/10 prob. of 7.7 Euro	3/10 prob. of 0.2 Euro
8/10 prob. of 4.0 Euro	2/10 prob. of 3.2 Euro	A	B	8/10 prob. of 7.7 Euro	2/10 prob. of 0.2 Euro
9/10 prob. of 4.0 Euro	1/10 prob. of 3.2 Euro	A	B	9/10 prob. of 7.7 Euro	1/10 prob. of 0.2 Euro
10/10 prob. of 4.0 Euro	0/10 prob. of 3.2 Euro	A	B	10/10 prob. of 7.7 Euro	0/10 prob. of 0.2 Euro

Figure : The classic HL with two risky lotteries

Treatment 2: One Safe

3.3 Euro in any case	A			B	7.7 Euro with prob. 10%	0.2 Euro with prob. 90%
3.4 Euro in any case	A			B	7.7 Euro with prob. 20%	0.2 Euro with prob. 80%
3.5 Euro in any case	A			B	7.7 Euro with prob. 30%	0.2 Euro with prob. 70%
3.5 Euro in any case	A			B	7.7 Euro with prob. 40%	0.2 Euro with prob. 60%
3.6 Euro in any case	A			B	7.7 Euro with prob. 50%	0.2 Euro with prob. 50%
3.7 Euro in any case	A			B	7.7 Euro with prob. 60%	0.2 Euro with prob. 40%
3.7 Euro in any case	A			B	7.7 Euro with prob. 70%	0.2 Euro with prob. 30%
3.8 Euro in any case	A			B	7.7 Euro with prob. 80%	0.2 Euro with prob. 20%
3.9 Euro in any case	A			B	7.7 Euro with prob. 90%	0.2 Euro with prob. 10%
4 Euro in any case	A			B	7.7 Euro with prob. 100%	0.2 Euro with prob. 0%

Figure : The HL with a safe amount against a lottery

The Treatments

Every lottery proposed as Option A in the **Both Risky** has been replaced with the corresponding safe amount, i.e. the certainty equivalent for an agent characterized by CRRA utility function of the form x^r and by a degree of risk aversion that would determine the switching point towards Option B in that row (virtually identical to the expected value).

Treatment **One Safe** increases the visibility of a safe alternative as compared to Treatment **Both Risky**, but such a manipulation is rather weak. In fact, the multiple price format is likely to maintain a comparison of risky alternatives always present across rows. Moreover, direct comparability across treatment imposes to substitute each Option A with a different safe amount. Both factors are likely to dilute the impact of the introduction of a safe choice in every row.

Procedure

Both treatments have been incentivized using a pay-one-at-random scheme and took place in the laboratory of the Max Planck Institute for Economics in Jena, Germany, between May 2012 and July 2013.

Pure between-subject design:
Both Risky: 147 participants (63 males, 84 females)
One Safe: 165 participants (79 males, 86 females)

26 inconsistent choices in the Both Risky, 6 inconsistent choices in the One Safe (eliminated from the analysis).

Results

	Treatment 1: Both Risky			Treatment 2: One safe		
	N	Safe choices	Mann-Whitney	N	Safe choices	Mann-Whitney
Males	74	5.70	p=0.18	76	5.66	p=0.004
Females	77	6.02		83	6.24	
Cohen's d		0.18			0.42	

Table : Results by gender of the Both Risky HL and the One Safe HL treatments

Results support our conjecture.

The availability of a safe option causes an increase of the size and the significance of the gender differences.

BRET with a safe option

BRET: rules

We developed the 'Bomb' Risk Elicitation Task (BRET)

- Subjects are shown a field with 100 boxes.
- Are told that under one of the boxes lies a *time bomb*
- Their task is to decide how many boxes to collect.

- Once the task is over, the position of the bomb is determined (hence the *time bomb*).
- If bomb collected → earnings equal zero.
- If bomb not collected → earnings equal to number of boxes collected.

BRET: interface

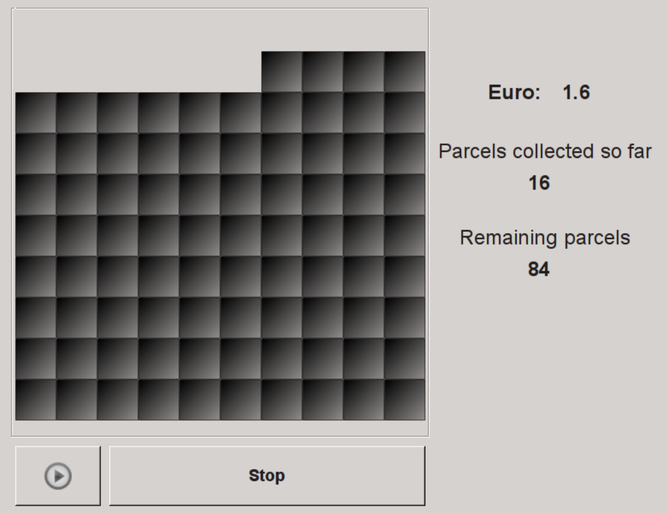


Figure : The BRET interface after 16 seconds

BRET: under the hood

- Theoretically, the task amounts to choosing the preferred among 101 lotteries.
- Each lottery is characterized by:

$$L^k = \begin{cases} 0 & \frac{k}{100} \\ k & \frac{100-k}{100} \end{cases}$$

- The 101 lotteries are all summarized by the parameter k...
- ...that is also governing probabilities.
- Example: at $k = 20$, $L = \{20\% : 0 \,;\, 80\% : 20\}$

There is no safe amount

Safe-BRET: interface

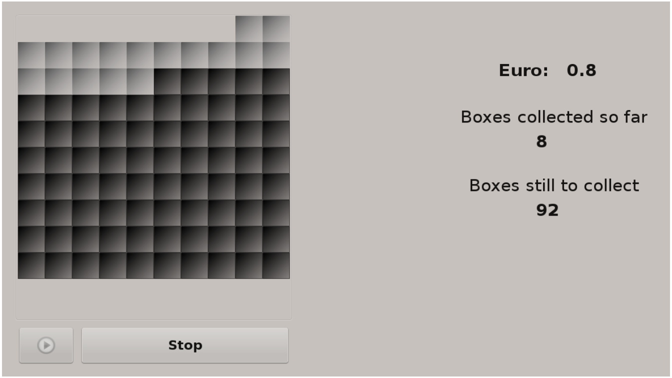


Figure : The Safe-BRET interface after 8 seconds

- Manipulation: subjects told the time bomb cannot be in the first 25 boxes;
- Only those that would choose $n < 25$ (a negligible fraction) are affected. Those choosing $n \geq 25$ should make the same choice in the two treatments according to expected utility theory.

Safe-BRET: results by gender

	BRET			Safe-BRET		
	N	Stopping point	Mann-Whitney	N	Stopping point	Mann-Whitney
Males	105	46.38	p=0.657	73	49.79	p= 0.079
Females	164	46.65		118	46.72	
Cohen's d		-0.02			0.25	

Table : Results by gender of the baseline BRET and the Safe-BRET treatments

- The BRET without safe option display the same behavior across gender.
- Including a safe option among the possible alternatives determines gender differences to emerge.
- It is an indirect effect: females are not more likely to choose the safe option (n=25): 7.6% vs 6.8%, p=.54

EG without a safe option

EG without a safe option: experiment

Panel (a): Original EG

	Event	Probability	Outcome
1	A	50%	4 €
	B	50%	4 €
2	A	50%	6 €
	B	50%	3 €
3	A	50%	8 €
	B	50%	2 €
4	A	50%	10 €
	B	50%	1 €
5	A	50%	12 €
	B	50%	0 €

Panel (b):EG nosafe

	Event	Probability	Outcome
2	A	50%	6 €
	B	50%	3 €
3	A	50%	8 €
	B	50%	2 €
4	A	50%	10 €
	B	50%	1 €
5	A	50%	12 €
	B	50%	0 €

EG with a safe option: certainty effect

Females are significantly more likely to choose the safe alternative: 20.5% vs. 6%, p=.01

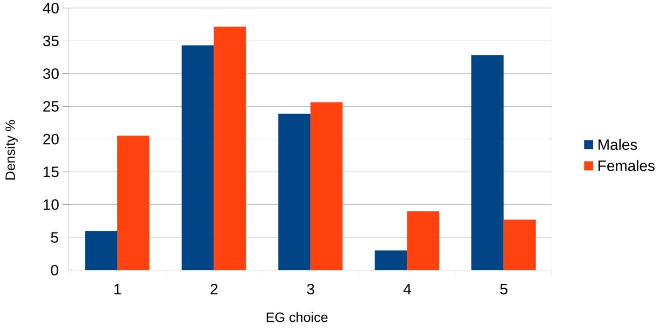


Figure : Choices by gender in the EG task with a safe option

EG: indirect effect

However, there is a strong indirect effect that prevails

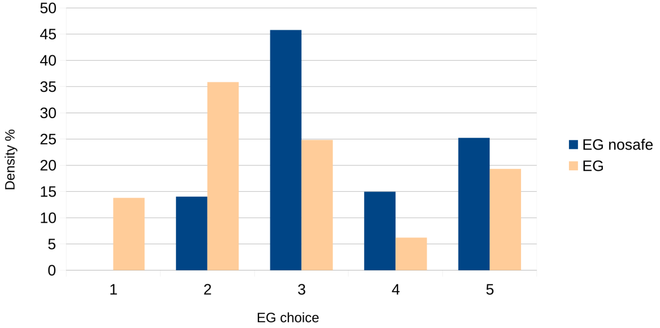


Figure : Choices in the EG nosafe and EG safe tasks

EG: overall results

EG						
		N	Average Choice	Std. Dev.	Cohen's d	Mann Whitney
Safe	**Males**	67	3.22	1.38	.58	.001
	Females	78	2.56	1.14		
	Diff. (M-F)		.66			
Non-Safe	**Males**	47	3.85	1.04	.59	.004
	Females	60	3.25	.93		
	Diff. (M-F)		.60			

In the EG task results lead weak support to our conjecture (by means of a certainty effect only)

Moving probabilities

EG without fixed probabilities

Probability manipulations on EG

EG nosafe - 50/50

	Event	Probability	Outcome
2	A	50%	6 €
	B	50%	3 €
3	A	50%	8 €
	B	50%	2 €
4	A	50%	10 €
	B	50%	1 €
5	A	50%	12 €
	B	50%	0 €

EG nosafe - 40/60

	Event	Probability	Outcome
2	A	40%	6 €
	B	60%	3 €
3	A	40%	8 €
	B	60%	2 €
4	A	40%	10 €
	B	60%	1 €
5	A	40%	12 €
	B	60%	0 €

Table : Variations of the EG task along the probability dimension

Probability manipulations on EG: results

		N	choice	MW	Cohens
EG NoSafe 50-50	Males	47	3.85	.00	.59
	Females	60	3.25		
EG NoSafe 40-60	Males	38	3.81	.01	.56
	Females	48	3.18		

Table : Overall results by gender of the probabilities manipulations

Conclusion

Summary

Gender differences in risk aversion are not ubiquitous

they are task-dependent and are usually small in magnitude

There is a clear correlation between the elicitation method used and the likelihood of observing gender differences

Explanations

1. The availability of a safe option explains part of such differences. The way in which the effect of a safe option operates is also task specific.
2. Fixed (50%-50%) probabilities do not play a significant role
3. A great deal of variance is left unexplained

The role played by a safe alternative: Consequences

The fact that gender differences in risk taking are affected by the presence a certain outcome requires a careful theoretical interpretation

Preferences for a safe outcome, whatever the ultimate cause, cannot be given an Expected Utility representation. They constitute a violation of the independence axiom that implies linearity in probabilities, something that excludes that subjects disproportionately prefer a change of probability around zero.

Most of the literature providing evidence about gender differences in risk attitudes assumes implicitly or explicitly an Expected Utility framework. Within this context it is inappropriate to talk about different risk aversion as long as the gender gap stems from a safe alternative.

Discussion

Critics might say that in the real world situations rarely exclude a safe option and women simply are more prudent when there is actually something at stake. So **who cares about such technicalities?**

If the goal is to rank subjects according to their willingness to tolerate risk
$\rightarrow$ **OK!**
But we have to be careful not to give an interpretation of risk aversion in an Expected Utility framework

If the goal is to derive a cardinal measure of risk aversion, to estimate parameters of a utility function, or in any case to work with preferences characterized by an expected utility representation
$\rightarrow$ **Don't!**
Or use a task like the BRET in which safe options are certainly excluded and not this 'preference for certainty' effect.

Thanks!

Appendix

Safe-BRET: under the hood

- In the case of the Safe-BRET, subjects are told that the bomb cannot be in the first 25 boxes. In other words, choosing up to 25 boxes is safe. Each lottery is characterized by:

$$L^k = \begin{cases} \left[k \text{ with prob. } 1 \right. & \text{if } k \leq 25 \\ \left[\begin{array}{l} 0 \text{ with prob. } \frac{k-25}{75} \\ k \text{ with prob. } \frac{100-k}{75} \end{array} \right. & \text{if } 25 \leq k \leq 100 \end{cases}$$

- Example: at $k = 20$, $L = \{100\% : 20\}$
- Example: at $k = 40$, $L = \{80\% : 40 ; 20\% : 0\}$

Results by gender: no gender difference in Risk Aversion

We find NO gender difference in the Baseline (both dynamic and static).

		N	BRET Mean	BRET Mann-Whitney z	SOEP Mean	SOEP Mann-Whitney z
Static	Males	30	44.23	$\text{Prob} > \|z\| = 0.66$	4.63	$\text{Prob} > \|z\| = 0.97$
	Females	54	43.44		4.65	
Dynamic	Males	105	46.38	$\text{Prob} > \|z\| = 0.66$	5.33	$\text{Prob} > \|z\| = 0.04$
	Females	164	46.65		4.83	

Table : Stopping time in the baseline treatments, breakdown by gender

- Gender difference often found (Eckel and Grossman, Charness and Gneezy)
- Always treated as 'a fact', but many studies *do not* find it
- Turns out to be task-specific
- Why is it not there in the BRET?

BRET with loss frame

Framing the BRET for losses

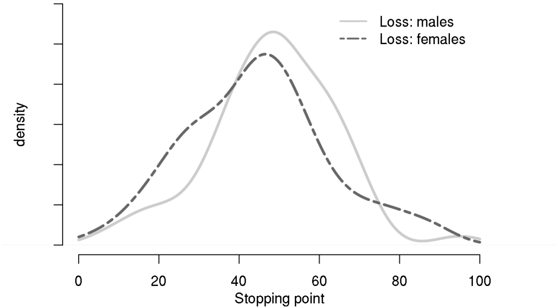
We ran a specific *Loss Aversion* treatment - changing just the **frame**

- Subjects on arrival find 2.5e on their desk
- These are on top of the 2.5e show-up fee
- The 2.5 are at stake in a framed BRET
- in which information is conveyed as losses/gains around a reference point

e.g.: at $k = 16 \rightarrow$ 'you are losing 0.9e w.r.t. the starting endowment'
e.g.: at $k = 37 \rightarrow$ 'you are gaining 1.2e w.r.t. the starting endowment'

Results should not change, unless if subjects are *loss averse*.
This is a rather weak manipulation: It is a pure framing effect as there is no way of securing a sure amount.

BRET with losses: results by gender



		BRET		SOEP	
	N	Mean	Mann-Whitney z	Mean	Mann-Whitney z
Males	60	48.16	Prob $> \|z\| = 0.057$	5.61	Prob $> \|z\| = 0.004$
Females	69	43.43		4.56	

Agnew, J. R., Anderson, L. R., Gerlach, J. R., Szykman, L. R., 2008. Who chooses annuities? an experimental investigation of the role of gender, framing, and defaults. The American Economic Review 98 (2), pp. 418–422.

Andersen, S., Harrison, G., Lau, M., Rutström, E., 2006. Elicitation using multiple price list formats. Experimental Economics 9, 383–405.

Anderson, L., Freeborn, B., 2010. Varying the intensity of competition in a multiple prize rent seeking experiment. Public Choice 143, 237–254.

Baker, R. J., Laury, S. K., Williams, A. W., 2008. Comparing small-group and individual behavior in lottery-choice experiments. Southern Economic Journal 75 (2), 367–382.

Brañas-Garza, P., Rustichini, A., 2011. Organizing Effects of Testosterone and Economic Behavior: Not Just Risk Taking. PLoS ONE 6 (12), e29842.

Carlsson, F., He, H., Martinsson, P., Qin, P., Sutter, M., 2012. Household decision making in rural china: Using experiments to estimate the influences of spouses. Journal of Economic Behavior & Organization 84 (2), 525–536.

Chakravarty, S., Harrison, G. W., Haruvy, E. E., Rutström, E. E., 2011. Are you risk averse over other people's money? Southern Economic Journal 77 (4), 901 – 913.

Chen, Y., Katuščák, P., Ozdenoren, E., 2013. Why Can't a Woman Bid More Like a Man? Games and Economic Behaviour 77 (1), 181–213.

Dave, C., Eckel, C., Johnson, C., Rojas, C., 2010. Eliciting risk preferences: When is simple better? Journal of Risk and Uncertainty 41 (3), 219–243.

Drichoutis, A. C., Koundouri, P., 2012. Estimating risk attitudes in conventional and artefactual lab experiments: The importance of the underlying assumptions. Economics - The Open-Access, Open-Assessment E-Journal 6 (38), 1–15.

Eckel, C. C., Wilson, R. K., 2004. Is trust a risky decision? Journal of Economic Behavior & Organization 55 (4), 447 – 465.

Ehmke, M., Lusk, J., Tyner, W., 2010. Multidimensional tests for economic behavior differences across cultures. The Journal of Socio-Economics 39 (1), 37 – 45.

Harrison, G. W., Johnson, E., McInnes, M. M., Rutström, E. E., June 2005. Risk Aversion and Incentive Effects: Comment. American Economic Review 95 (3), 897–901.

Harrison, G. W., Lau, M. I., Rutstrm, E. E., Tarazona-Gmez, M., 2013. Preferences over social risk. Oxford Economic Papers 65 (1), 25–46.

Holt, C., Laury, S., 2002. Risk aversion and incentive effects. American Economic Review 92 (5), 1644–1655.

Houser, D., Schunk, D., Winter, J., 2010. Distinguishing trust from risk: An anatomy of the investment game. Journal of Economic Behavior & Organization 74 (1-2), 72 – 81.

Masclet, D., Colombier, N., Denant-Boemont, L., Lohéac, Y., 2009. Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. Journal of Economic Behavior & Organization 70 (3), 470 – 484.

Menon, M., Perali, F., 2009. Eliciting risk and time preferences in field experiments: Are they related to cognitive and non-cognitive outcomes? are circumstances important? Rivista Internazionale di Scienze Sociali 117 (3), 593–630.

Mueller, J., Schwieren, C., 2012. Can personality explain what is underlying women's unwillingness to compete? Journal of Economic Psychology 33 (3), 448 – 460.

Ponti, G., Carbone, E., 2009. Positional learning with noise. Research in Economics 63 (4), 225 – 241.

Viscusi, W., Phillips, O., Kroll, S., 2011. Risky investment decisions: How are individuals influenced by their groups? Journal of Risk and Uncertainty 43 (2), 81–106.