THE RISK ELICITATION PUZZLE REVISITED: ACROSS-METHODS (IN)CONSISTENCY?

Paper by Holzmeister and Stefan (2020)

BMMLAB Journal Club 15 February 2021, Karlo Doroc

OUR ROADMAP

THERE IS NO CONSENSUS ON THE BEST WAY TO ELICIT RISK PREFERENCES

- Risk elicitation mechanisms (EMs) are experimental tasks and procedures designed to elicit an individual's level of risk aversion

- Many EMs exist, but there is no consensus on the best one
  - Examples include Holt-Laury test, second-price auctions/BDMs, and certainty equivalent methods

- Choice of EM is often arbitrary, falling back on the procedural invariance axiom: normatively equivalent EMs should have identical risk preference ordering

RISK ELICITATION PUZZLE: HUMAN RISK PREFERENCES ARE UNSTABLE BETWEEN EM

- Dozens of experiments show that the rank-ordering of revealed preferences across individuals vary unpredictably between EMs i.e. they are unstable
  - Within subjects, individuals' risk preferences vary with EMs
  - Between subjects, individuals' degree (or rank) of risk aversion varies with EM

- Pairwise correlations between risk preferences in two EMs tend to be positive, but are typically moderate and can vary dramatically (-0.33 to 0.86)

EXAMPLE EM: SINGLE CHOICE LIST (SPL)

Screen 5: Single Choice List (SCL) – Risk Preference Elicitation

Task 'orange': Your Decision

	Event A		Event B		
No.	**Prob.**	**Payoff**	**Prob.**	**Payoff**	**Your Choice**
1	50.0%	€9.00	50.0%	€9.00	○
2	50.0%	€7.50	50.0%	€12.00	○
3	50.0%	€6.00	50.0%	€15.00	○
4	50.0%	€4.50	50.0%	€18.00	○
5	50.0%	€3.00	50.0%	€21.00	○
6	50.0%	€0.00	50.0%	€24.00	○

EXAMPLE EM: MULTIPLE CHOICE LIST (MPL)

Task 'blue': Your Decision

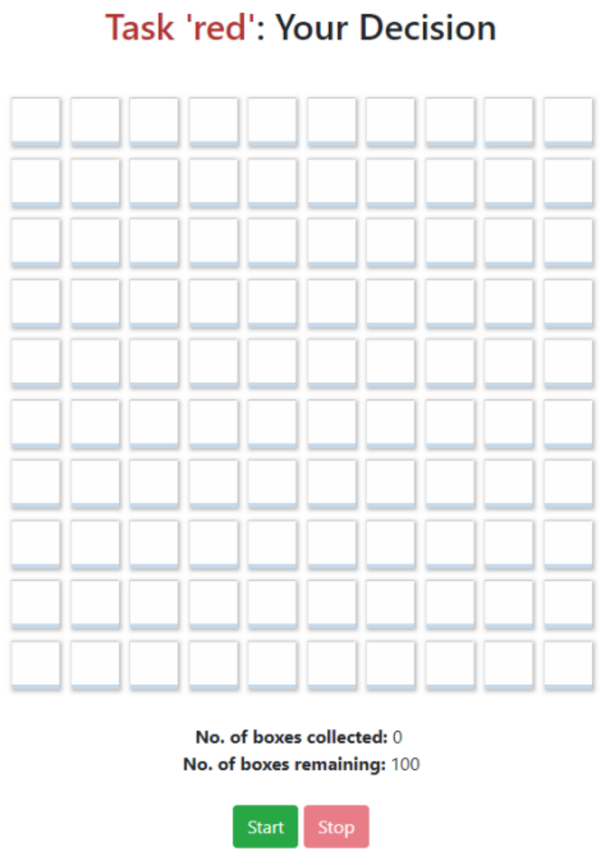
Option A	Option B
€10.00 with a probability of **10.00%**, €8.00 otherwise	€19.25 with a probability of **10.00%**, €0.50 otherwise
€10.00 with a probability of **20.00%**, €8.00 otherwise	€19.25 with a probability of **20.00%**, €0.50 otherwise
€10.00 with a probability of **30.00%**, €8.00 otherwise	€19.25 with a probability of **30.00%**, €0.50 otherwise
€10.00 with a probability of **40.00%**, €8.00 otherwise	€19.25 with a probability of **40.00%**, €0.50 otherwise
€10.00 with a probability of **50.00%**, €8.00 otherwise	€19.25 with a probability of **50.00%**, €0.50 otherwise
€10.00 with a probability of **60.00%**, €8.00 otherwise	€19.25 with a probability of **60.00%**, €0.50 otherwise
€10.00 with a probability of **70.00%**, €8.00 otherwise	€19.25 with a probability of **70.00%**, €0.50 otherwise
€10.00 with a probability of **80.00%**, €8.00 otherwise	€19.25 with a probability of **80.00%**, €0.50 otherwise
€10.00 with a probability of **90.00%**, €8.00 otherwise	€19.25 with a probability of **90.00%**, €0.50 otherwise
€10.00 with a probability of **100.00%**, €8.00 otherwise	€19.25 with a probability of **100.00%**, €0.50 otherwise

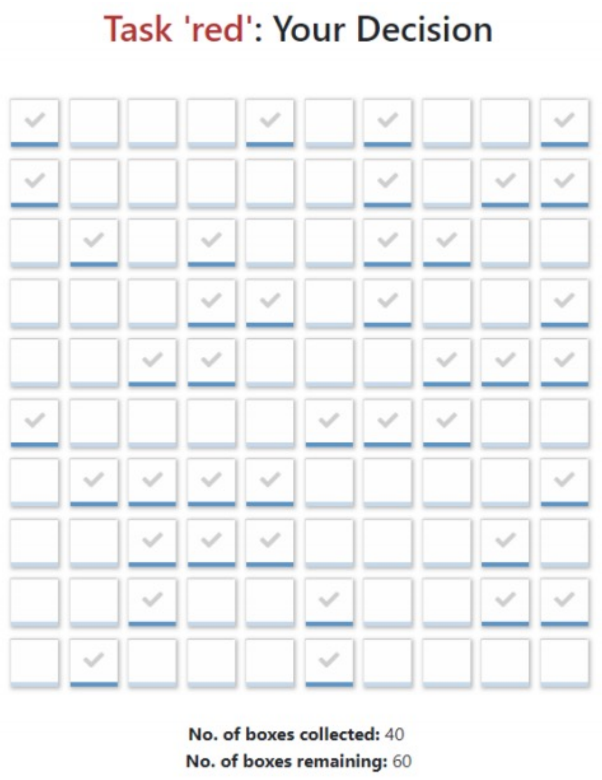
EXAMPLE EM: CERTAINTY EQUIVALENT METHOD (CEM)

Task 'green': Your Decision

Option A			Option B
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€5.00** with a probability of 100% (sure payoff)
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€6.25** with a probability of 100% (sure payoff)
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€7.50** with a probability of 100% (sure payoff)
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€8.75** with a probability of 100% (sure payoff)
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€10.00** with a probability of 100% (sure payoff)
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€11.25** with a probability of 100% (sure payoff)
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€12.50** with a probability of 100% (sure payoff)
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€13.75** with a probability of 100% (sure payoff)
€15.00 with a probability of 50%, €5.00 otherwise	○	○	**€15.00** with a probability of 100% (sure payoff)

EXAMPLE EM: BOMB RISK ELICITATION TASK (BRET)





CAUSES: PROCEDURAL VARIANCE, UNSTABLE PREFERENCES OR INCONSISTENT CHOICES?

- Procedural variance: different EMs fundamentally measure different preference relations
- Unstable risk preferences: an individual's underlying risk preferences are not stable, but instead change over time or across EM
- Inconsistent choices: individuals' observed choices are inconsistent with their own underlying preferences i.e. they make mistakes
  - Note: consistency can only be judged against an underlying theoretical framework. Here, they used CRRA utility

ARE HUMANS AWARE THAT THEIR RISK PREFERENCES VARY ACROSS EM?

- Procedural (in)variance and (un)stable risk preferences cannot be empirically disentangled
  - To test that risk preferences are stable across EMs one must use multiple EMs assuming that each EM is measuring the same preference relations
  - To test procedural invariance one must use multiple EMs assuming that the underlying risk preferences are stable across all tasks and time periods
- As such, the paper focuses on inconsistent choices and an exploratory look at whether humans are aware their risk preferences vary across EM

RESEARCH QUESTIONS

- Which of the following factors are potentially related to the risk elicitation puzzle?
  - Perceived riskiness of the choice?
  - Confidence in the choice?
  - Perceived simplicity of the EM?
  - Perceived complexity of the EM's calculations?
  - Boredom?

METHODS: A CONTROLLED EXPERIMENT

- Controlled experiment with a within-subjects design used to elicit risk preferences across four EMs (BRET, CEM, SPL, MPL)

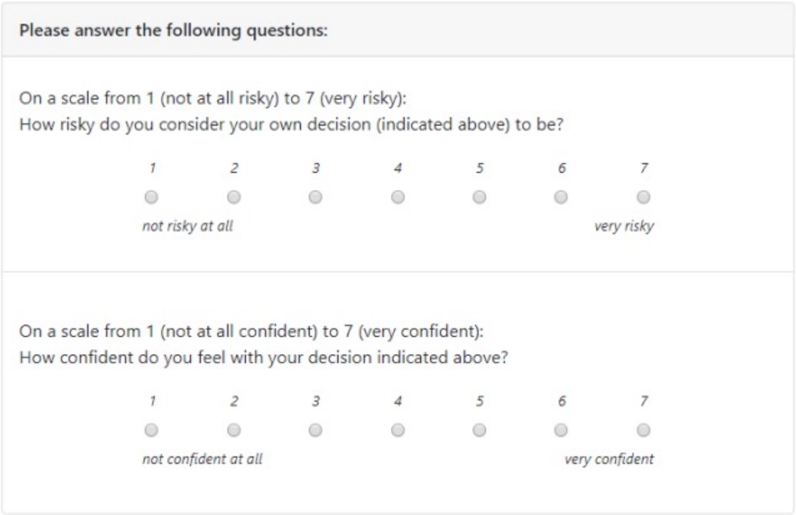
- 198 human volunteers (55% female, mean age = 22.9 years, sd = 2.5) were recruited to participate in ten experimental sessions

- Each EM was presented as a computer game. Participants were paid for their performance, which was randomly selected from only one of the four EMs

- EM sequencing was random and within an EM instance sequencing was random

METHODS: SURVEYS AND QUESTIONNAIRES

- After each instance participants were asked to rate perceived riskiness and confidence
- After all four EMs they were asked whether:
  - The instructions were easy to understand
  - The task required complex calculations
  - The task was boring
- Comprehension and numeracy testing


Please answer the following questions:

On a scale from 1 (not at all risky) to 7 (very risky):
How risky do you consider your own decision (indicated above) to be?

1 2 3 4 5 6 7

not risky at all ... *very risky*

On a scale from 1 (not at all confident) to 7 (very confident):
How confident do you feel with your decision indicated above?

1 2 3 4 5 6 7

not confident at all ... *very confident*

PARTICIPANTS ARE ASSUMED TO FOLLOW A CRRA UTILITY FUNCTION

- **CRRA utility function**:

$$u(x) = \begin{cases} (1-\varphi)^{-1}\, x^{1-\varphi} & \text{if } \varphi \neq 1 \\ ln(x) & \text{if } \varphi = 1 \end{cases}$$

- u is utility, x is payoff, φ is risk aversion (higher = more risk averse, 0 is neutral)

- **Utility of a lottery** j was:
  - $E[u_j] = \sum_k p_k u_k \;\forall\; k \in \{h, l\}$
- h and l are the high (low) lottery payoffs, respectively. Participants are assumed to choose between lottery pairs based on expected utilities.

LATENT CHOICE FUNCTION, PROBIT LINK FUNCTION AND MLE

- **Latent choice function** allows for 'trembles' and mistakes:
  - $\nabla EU_i = E[u_B] - E[u_A] + \sigma\epsilon, \epsilon \sim N(0,1)$
- *B* and *A* are the more (less) risky lottery, respectively. σ is a noise parameter and ϵ a Fechner error term

- **Probit function** links observed choices to latent. $P(B_i \succ A_i) = \Phi(\nabla EU_i)$, where Φ is the standard normal CDF. B is chosen whenever $\Phi(\nabla EU_i) > 0.5$
- **MLE:** $\ln L(\varphi, \sigma|\vec{y}) = \sum_{i=1}^{n} \left(\left[\ln \Phi\left(\nabla E[u_i]\right) \right]^{y_i} + \left[\ln \Phi\left(- \nabla E[u_i]\right) \right]^{1-y_i} \right)$
- Where $y^{\rightarrow}$ is a vector of choices and $y_i = 1$ if the participant chose B, else 0

NEW METRIC: PREFERENCE STABILITY INDEX

- Use MLE to calculate the confidence interval for each individual's risk coefficient, in each EM
- Within subjects, perform 6 unique pairwise comparisons between the intervals obtained for each EM
- An individuals gets a point for each pairwise comparison where the intervals overlap
- Minimum and maximum scores are 0 and 6, respectively

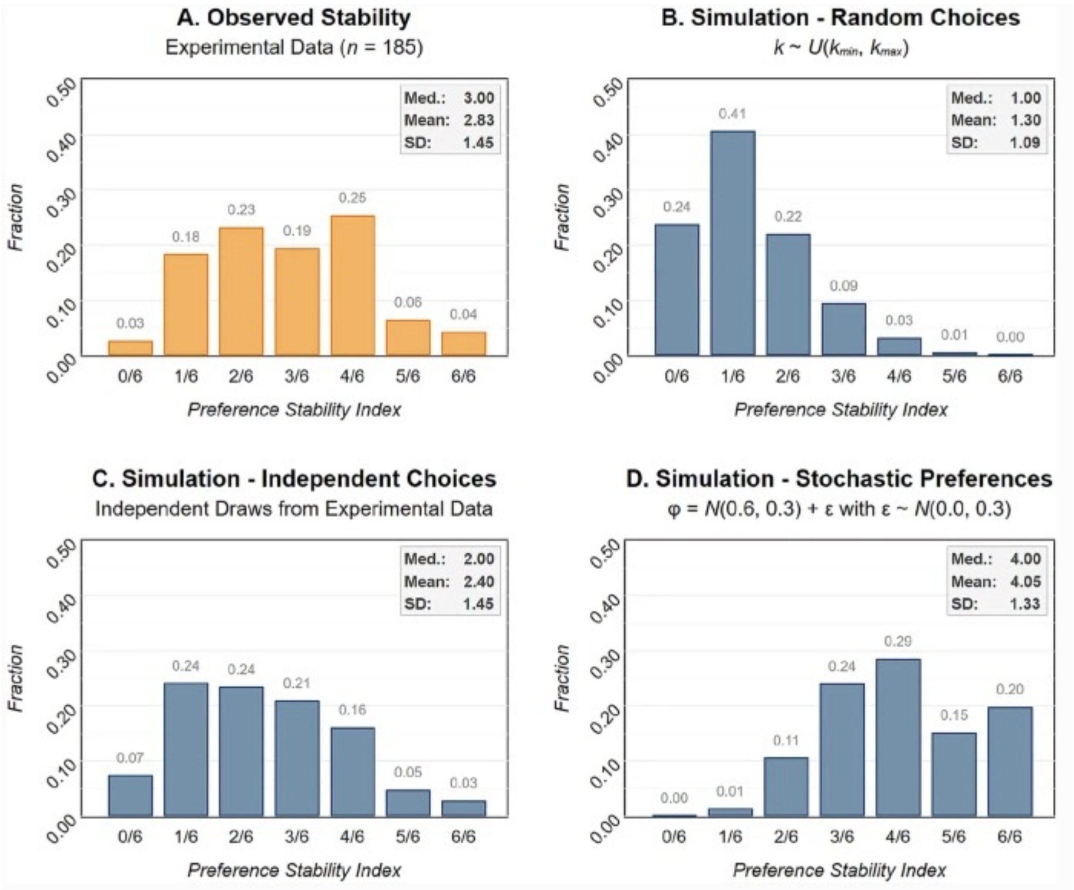
PAIRWISE CORRELATIONS BETWEEN EMS ARE MODERATE AND STATISTICALLY SIGNIFICANT

Table 1 Correlation matrix. The lower triangular matrix reports Spearman rank correlations between the observed number of risky choices in the four tasks; the upper triangular matrix depicts polychoric correlations

	BRET	CEM	MPL	SCL
BRET		0.245 (0.001)	0.350 (0.000)	0.336 (0.000)
CEM	0.222 (0.002)		0.283 (0.000)	0.400 (0.000)
MPL	0.367 (0.000)	0.244 (0.001)		0.387 (0.000)
SCL	0.341 (0.000)	0.338 (0.000)	0.354 (0.000)	

p values are reported in parentheses ($n = 198$). BRET, CEM, MPL, and SCL denote the "bomb" risk elicitation task, the certainty equivalent method, the multiple price list, and the single choice list, respectively

PREFERENCE STABILITY BETWEEN EMS IS LOW AND CONSISTENT WITH INDEPENDENT RANDOM CHOICES



RISK AVERSION AND NOISE PARAMETERS VARIED ACROSS EM

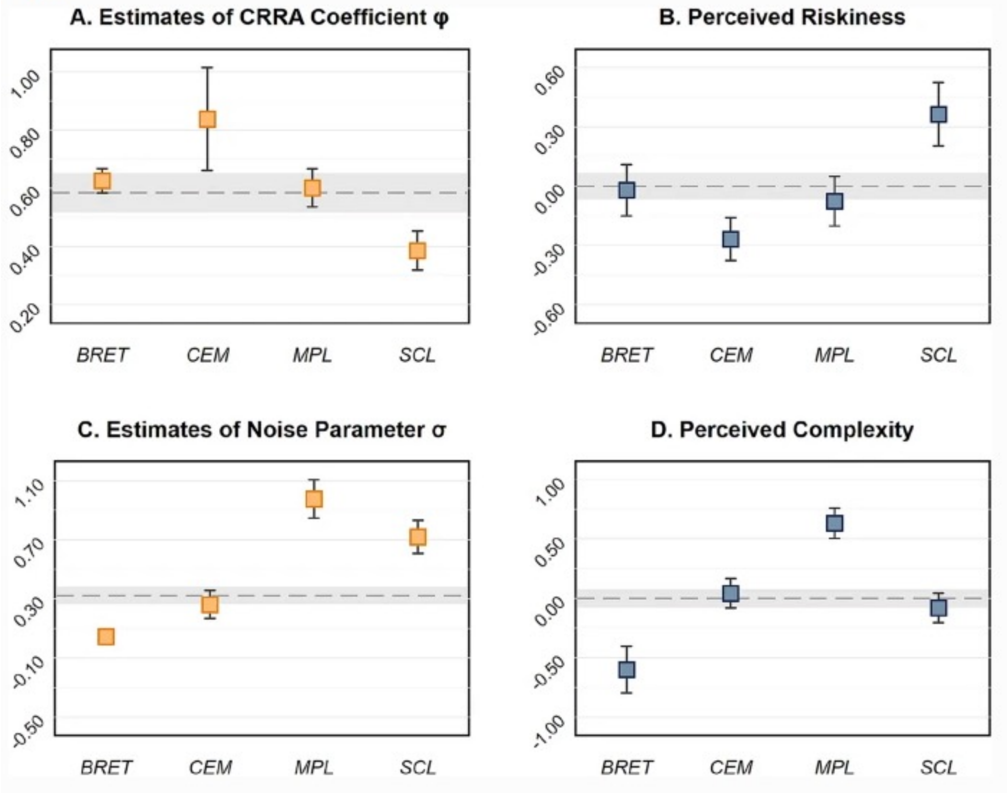
Table 2 **(A)** Maximum likelihood estimates of structural models with Fechner error terms for each of the four risk preference elicitation methods. Standard errors, clustered on the subject level, are reported in parentheses. **(B)** Pairwise differences in point estimates of risk preference parameters φ (lower-triangular matrix) and the standard deviation of noise parameters τ (upper-triangular matrix) between the four risk preference elicitation methods

	BRET	CEM	MPL	SCL
Panel A				
φ	0.626***	0.838***	0.602***	0.387***
	(0.021)	(0.090)	(0.033)	(0.034)
τ	0.046***	0.263***	0.977***	0.720***
	(0.002)	(0.048)	(0.066)	(0.057)
ln *L*	− 5,298	− 458	− 600	− 572
No. of Obs.	19,800	1782	1980	990
Clusters	198	198	198	198
Panel B				
BRET		− 0.217***	− 0.932***	− 0.674***
CEM	0.212*		− 0.715***	− 0.457***
MPL	− 0.025	− 0.237**		0.257**
SCL	− 0.240***	− 0.452***	− 0.215***	

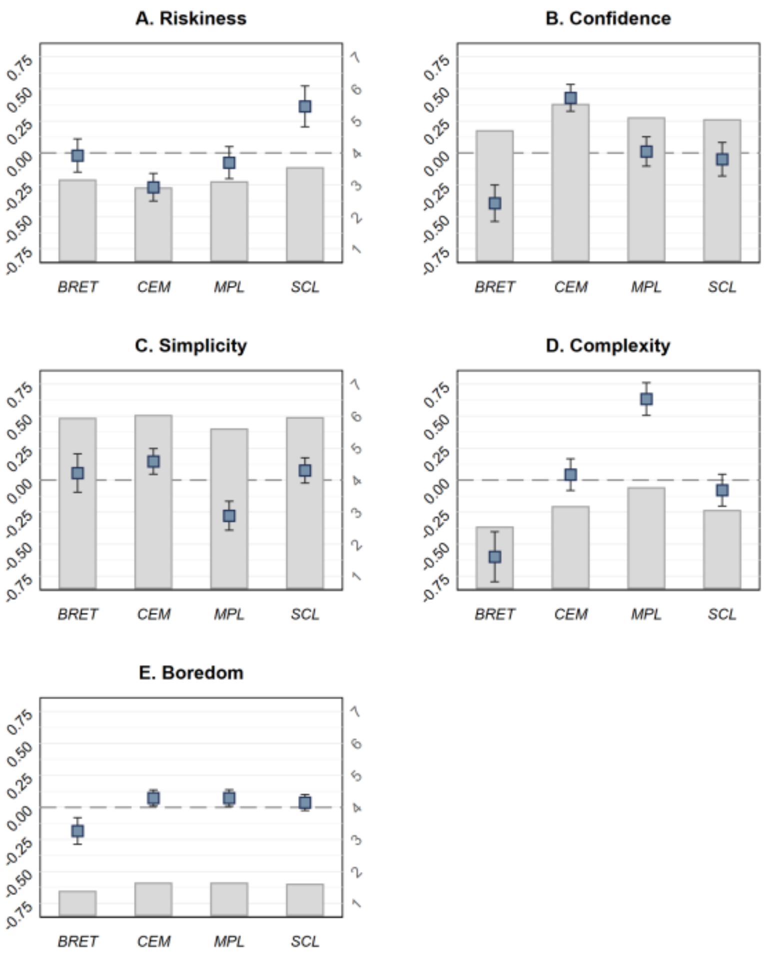
p values are based on pairwise Wald tests. BRET, CEM, MPL, and SCL denote the “bomb” risk elicitation task, the certainty equivalent method, the multiple price list, and the single choice list, respectively. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

PARTICIPANTS KNOW WHAT THEY'RE DOING, BUT OUR MODELS DON'T

Fig. 2



NO CLEAR PATTERN FROM CONFIDENCE, SIMPLICITY OR BOREDOM



AUXILIARY RESULTS

- No correlation between numeracy and preference stability

- No correlation between task comprehension and preference stability

INCONSISTENT CHOICES SEEMS AN IMPLAUSIBLE EXPLANATION FOR THE PUZZLE

- Humans are aware that their choices differ between EMs, so they are not unaware of their own 'inconsistency'
- Given this awareness, it seems implausible that humans are deliberately making choices that contradict their own underlying risk preferences
- Suggests that, rather than choice inconsistency, it is either procedural variance or unstable risk preferences which drive the risk elicitation puzzle

FURTHER QUESTIONS

- How to disentangle procedural invariance and stable risk preferences?
- Can (formally measured) complexity be used to predict risk preferences for a given EM?
- What proportion of variation between EMs can be explained by complexity?

ANY QUESTIONS?