Flipped classes and peer marking: incentives, student participation and performance in a quasi-experimental approach

Rabeya Khatoon
Department of Economics, University of Bristol
&
Elinor Jones
Department of Statistical Science, UCL

Developments in Economics Education Conference 2019

Background

Motivated by.....

- Student evaluation comments

  "Not enough problems to practice"

  "I hate having to wait a week to see the answer for this week's tutorial. If possible, make it available right after submission while it is fresh in the mind"

- Large quantitative units, limited resource
- Benefits from peer marking (e.g. Alcalde and Nagel, 2019,)

Existing literature says....

Flipped lectures/classes

- 'required students to watch introductory material prior to a lecture/small group class'(e.g., Becker and Proud, 2018)
- benefits by enhancing personalized and active learning (Keene, 2013, Jensen et al, 2015, Becker and Birdi, 2018)

Online peer assessment

- enables students to promptly receive a greater diversity of feedback (ASKe, 2009) in an anonymous environment (Yang and Tsai, 2010)
- enhances assessment and feedback experience for international students (Chew 2016)
- helps reducing marking load, by up to 70% (Lynch and Schmid, 2017)

Contribution of the study

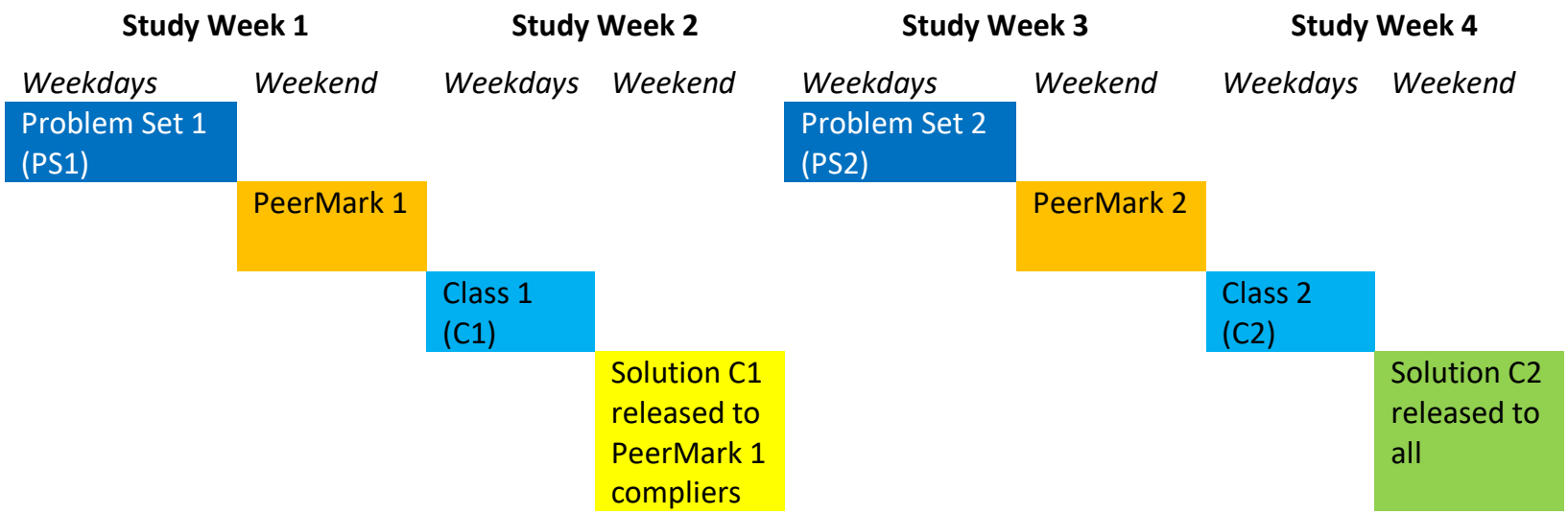
- We propose a new way of flipping quantitative tutorials by active reading and peer marking using virtual learning environment
- We have carefully designed the quasi-experiment with the view to evaluate the effect of peer marking by introducing an incentive
- Our quantitative findings and qualitative analysis based on focus group discussions reveal a number of clear benefits of doing peer mark and suggestions for student buy-in to the system

Study Design

Study Design

Two UK universities involved:

- University A: PG QM unit with 340 students - Autumn term
- University B: UG (yr 2) Statistics unit with 100 students - Spring term



All these are executed using the Virtual Learning Environments: Turnitin in Blackboard and Moodle respectively

Data

Data

- The final exam score of the units are considered as the outcome variable
- To control for their previous scores, the UG result (average % of marks obtained) are considered for the PG students and the UG-yr 1 result (average % of marks obtained) for the UG students
- For about 25% of the PG students, their previous score was reported as CGPA or the UK format of 2:1 or First, which has been converted to % of marks obtained using the GPA pilot project conversion chart of the HEA report (2015)
- Our variable of interest, *'peermarked'* is the total number of peer mark done by each student:

$$peermarked = \begin{cases} 0 & no\ PS\ peermarked \\ 1 & any\ one\ of\ the\ two\ PS\ peermarked \\ 2 & both\ PS\ peermarked \end{cases}$$

Data (Contd...)

- Since *'PeerMark 1'* has the incentive, we measure the *'incentive'*-effect by the dummy variable:
  - peermark1=1 if peermarked PS1, 0 otherwise
- Since only the students who submit a problem set are allowed to peer mark, we define two further dummy variables
  - submit1=1 if submitted PS1, 0 otherwise
  - submit2=1 if submitted PS2, 0 otherwise

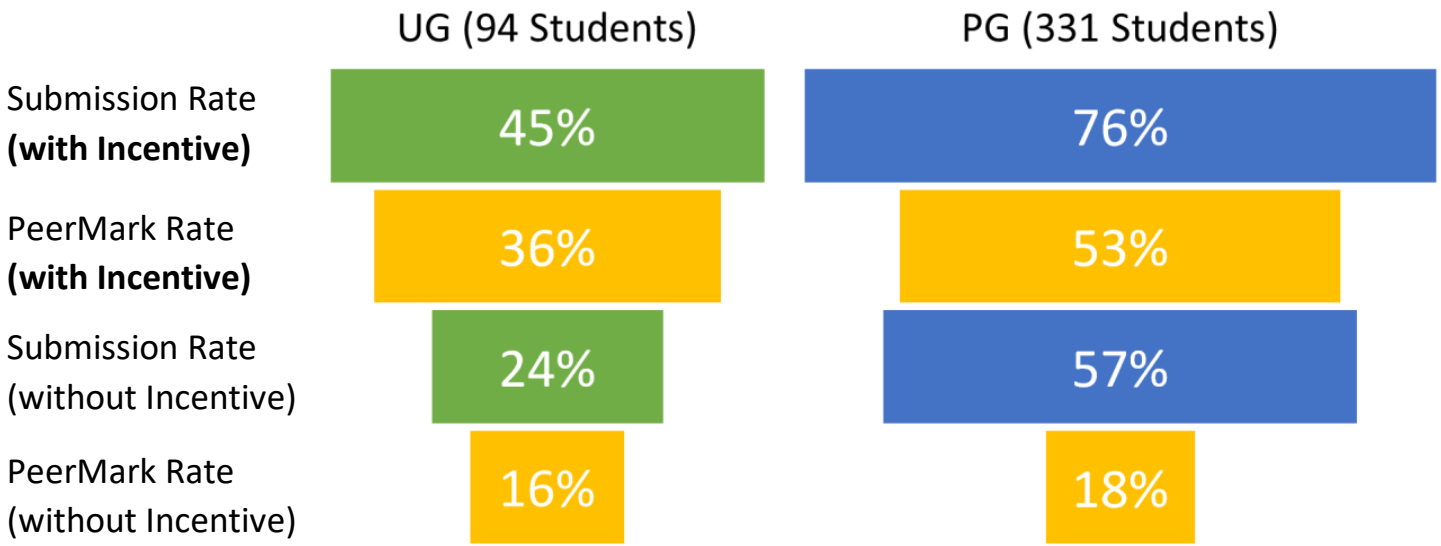
  and their sum, $submit \in [0, 1, 2]$

  Qualitative Information
- We conducted two focus group discussions (FGDs) in the two universities, with 5 students in each
- End of unit student evaluations are considered too, to observe any significant changes/comments compared to the previous year

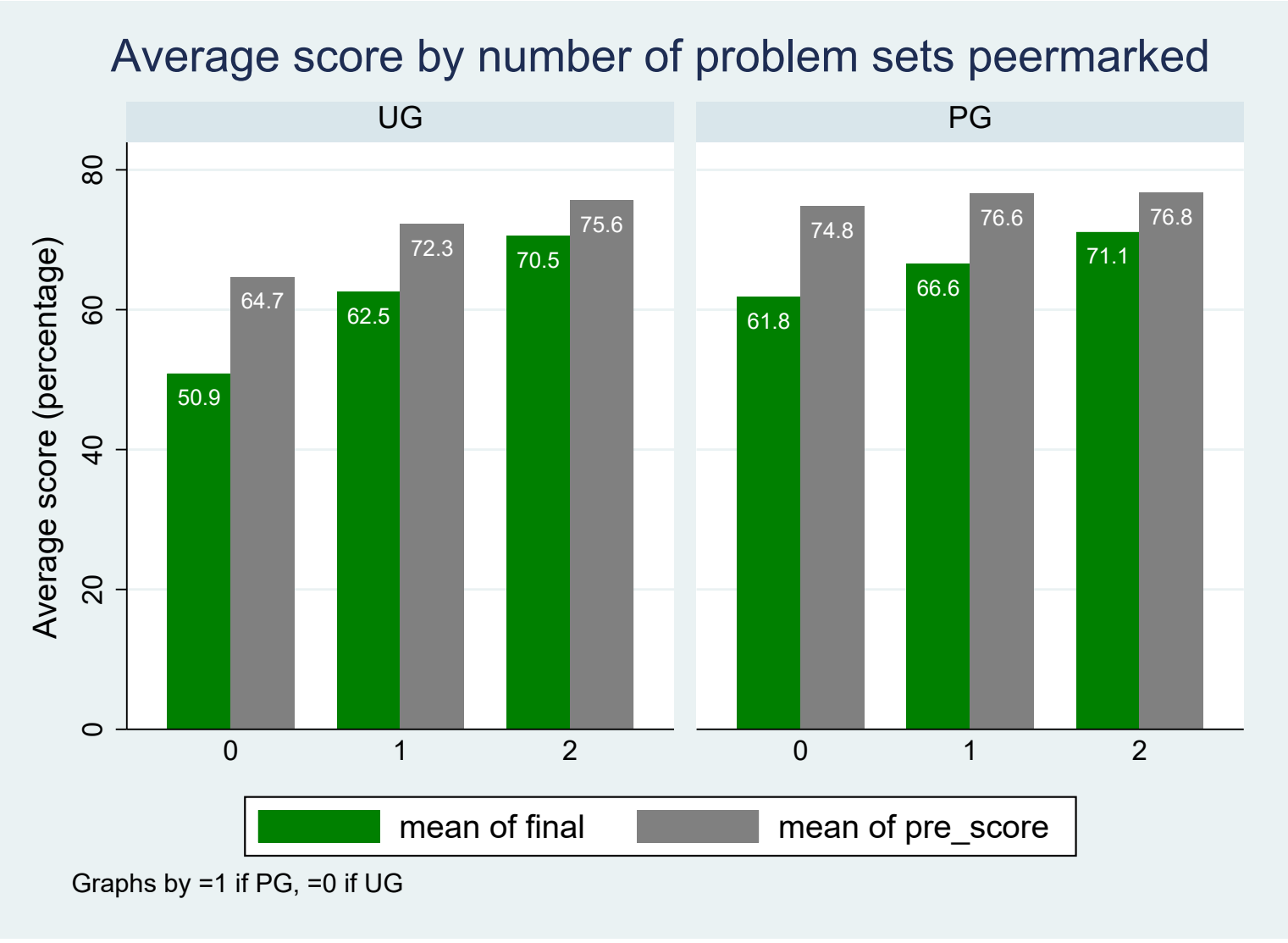
Table 1: Summary Statistics of the Key Variables

Variable	Obs	Mean	Std. Dev.	Min	Max
final	425	63.251	16.610	0	93
level	425	0.779	0.416	0	1
pre_score	385	73.926	11.565	3.5	92.75
submit1	425	0.694	0.461	0	1
submit2	425	0.499	0.501	0	1
submit	425	1.193	0.807	0	2
peermark1	425	0.492	0.501	0	1
peermark2	425	0.174	0.380	0	1
peermarked	425	0.666	0.724	0	2

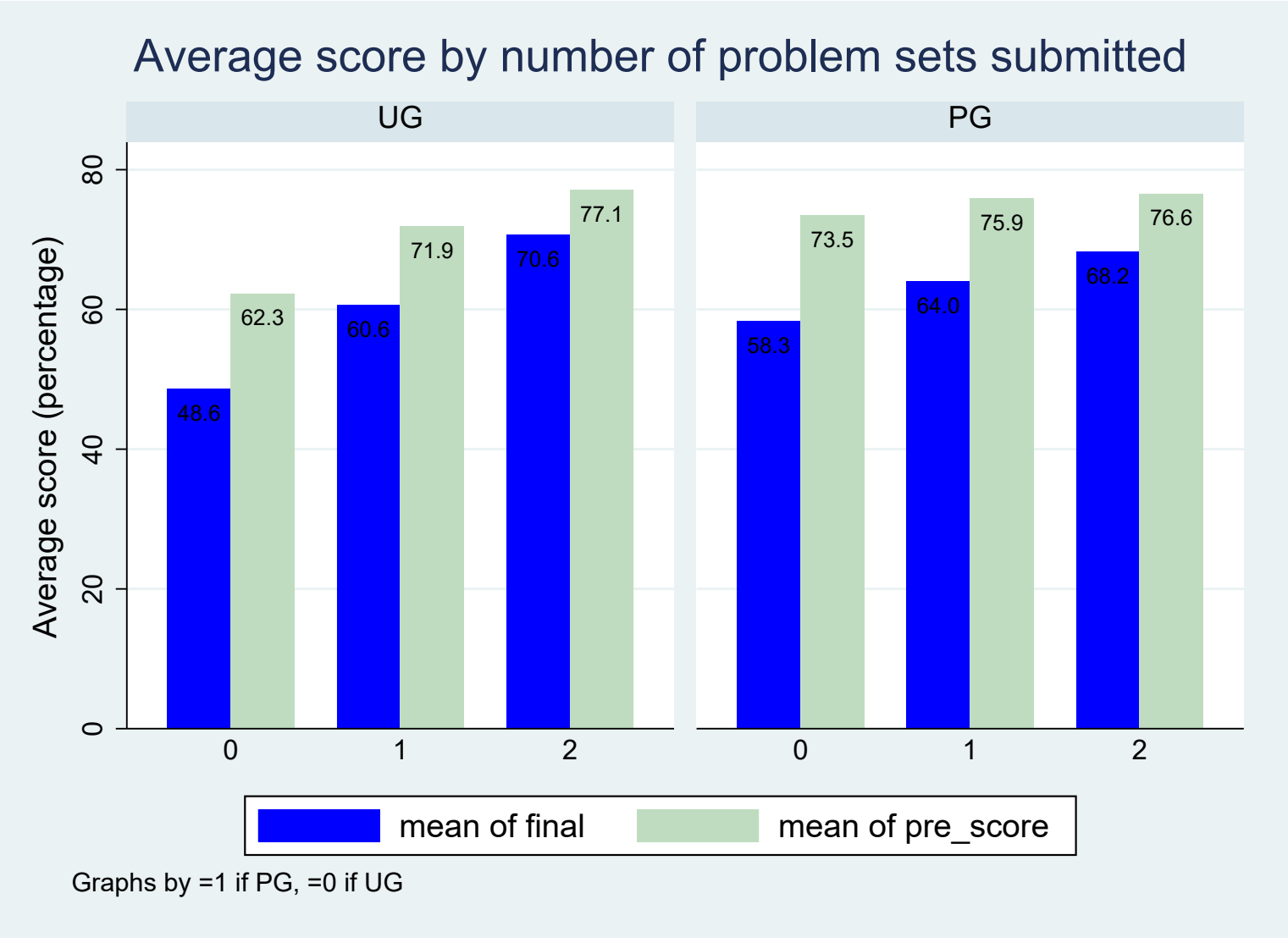
Summary Statistics (Contd...)



Summary Statistics (Contd...)



Summary Statistics (Contd...)



Method and Estimation

Model

Our primary regression model with incentive (peermark1) as IV is:

$$final_i = \beta_0 + \beta_1 peermarked_i + \beta_2 submit_i + \beta_3 prescore_i + \beta_4 level_i + u_i \quad (1)$$

$$peermarked_i = \gamma_0 + \gamma_1 peermark1_i + \gamma_2 submit_i + \gamma_3 prescore_i + \gamma_4 level_i + v_i \quad (2)$$

where,

$final_i$=% of final marks in the unit obtained by student i

$peermarked_i \in [0, 1, 2]$, the number of problem sets peermarked by student i

$submit_i \in [0, 1, 2]$, the number of problem sets submitted by student i

$prescore_i$=% of marks obtained on average in the previous year by student i

$level_i$=1 if PG, 0 if UG

$peermark1_i$=1 if student i peer marked PS1, 0 otherwise

- We also consider the two PS submission variables: $submit1_i$ and $submit2_i$ instead of $submit_i$ for comparison

*peermark*1 as IV

Relevance:

- By construction, number of PS peermarked depends on peermark1 (correlation coefficient 0.87)
- The dependence is higher than *peermark*2, as PS1 has the incentive and PS2 not

From the focus group discussions:

"...the reason I wouldn't is because for problems like 2 the lecturer says the answer will be given to everyone anyway": PG student

"...had to do in order to get something ...the first time because it was the first one, we had this incentive of having the answers if we did it": UG student

Exogeneity:

- The incentive was the online release of C1 solution to those who did peermark1
- Access to C1 solution was not prohibited from other sources, e.g., solving C1 with the tutor in the class, collecting it from friends

"...I know that even though I did not do it, I can get the answers like from my friends who have done that": PG student

Table 2: Estimation Results with *peermark*1 as IV

VARIABLES	(1) OLS	(2) IV	(3) OLS	(4) IV
peermarked	3.481***	3.487**	3.387***	3.136*
	(1.203)	(1.610)	(1.219)	(1.602)
submit1			3.690*	3.875*
			(2.183)	(2.269)
submit2			2.227	2.336
			(1.714)	(1.761)
pre_score	0.405***	0.405***	0.402***	0.402***
	(0.0702)	(0.0697)	(0.0708)	(0.0699)
level	3.969**	3.970**	3.952**	3.901**
	(1.723)	(1.730)	(1.728)	(1.736)
submit	2.856**	2.852*		
	(1.402)	(1.486)		
Constant	25.16***	25.16***	25.18***	25.19***
	(4.897)	(4.867)	(4.891)	(4.843)
Observations	385	385	385	385
R-squared	0.239	0.239	0.239	0.239
1st stage F-stat		664.06		753.49

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.
The sample consists of 94 UG students from University 1 and 291 PG students from University 2 in one academic year.

Table 3: IV Estimation Results: Heterogeneous Effects

	prescore > 69%		*prescore* < 70%	
VARIABLES	PG	UG	PG	UG
peermarked	4.562**	-3.623	2.510	14.79**
	(2.198)	(4.834)	(2.716)	(7.141)
submit	2.055	7.625	2.324	-8.049
	(2.099)	(4.844)	(2.340)	(5.623)
pre_score	0.383	0.893***	0.159	0.590**
	(0.300)	(0.294)	(0.134)	(0.245)
Constant	30.46	-7.792	47.42***	10.54
	(25.43)	(21.22)	(8.612)	(14.78)
Observations	173	48	118	51
R-squared	0.085	0.384	0.059	0.244

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Summary

The 'what', 'why', and 'how'

Benefits of flipping classes by peer mark

Educator perspective

- Since flipping is in an assessment format, it is possible to engage students with minimal incentives
  "... the biggest incentive of doing peer review would be solutions after it"
- Increases the amount of assessment and feedback without increasing marking load significantly (Lynch and Schmid, 2017)
- Fulfills the demand for *"more practice questions"* in quantitative units
- Increases student engagement in the small group classes
- Scope to build up student-textbook connection
- **Improves** student evaluation (from 3.98 to 4.50 in PG, 4.56 to 4.68 in UG)

Benefits (Contd...)

Student perspective

- It increases about 3% of final marks per problem set peermarked
- The system provides students with additional practice assessment

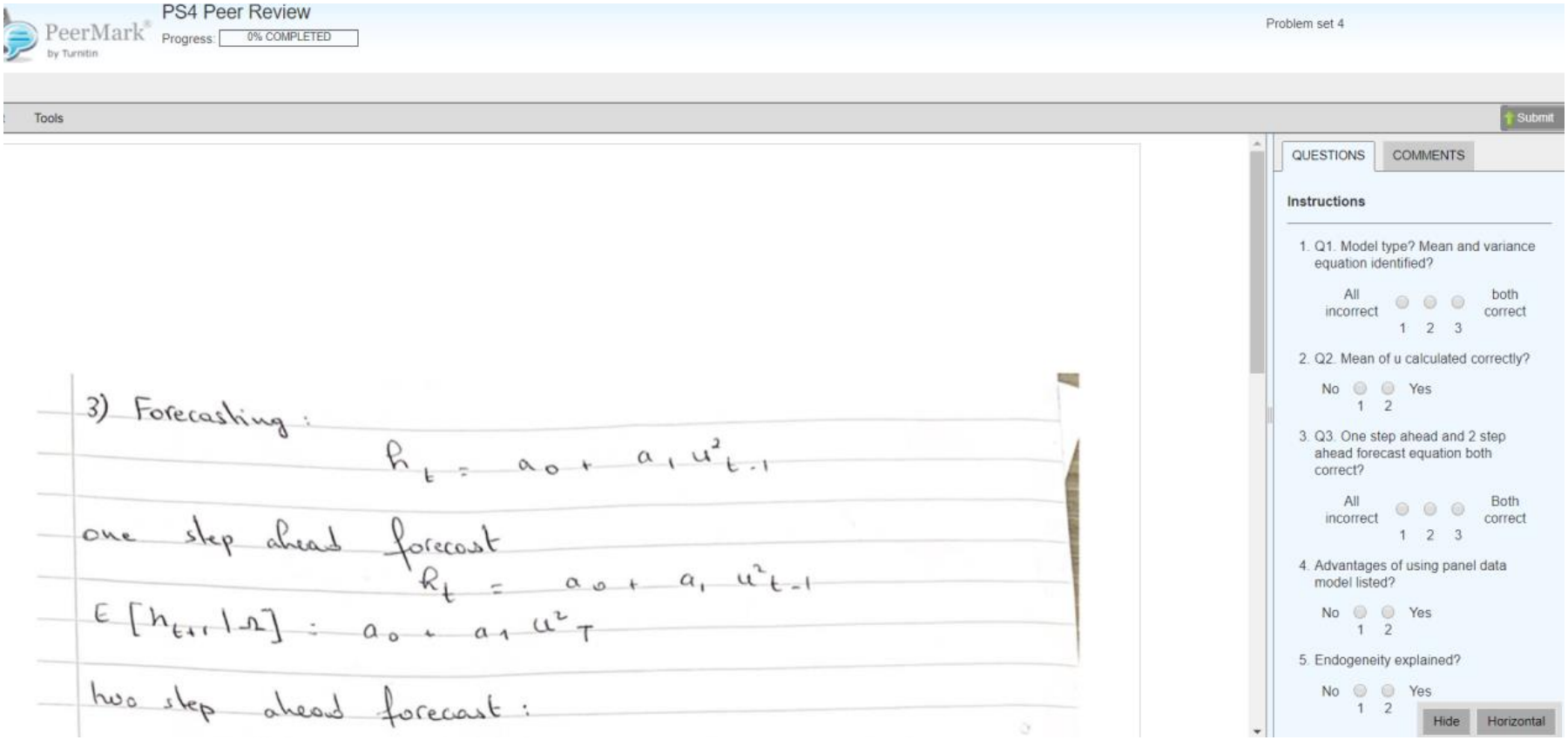
  "I think it was really good, for example, tutorial, when we had the question to do it on our own it was like a mock small exam"
- Enhances timely thinking/reflection, specially for those who generally don't see what feedback they received

  "... doing the marking and I'm looking at the solutions thinking about why an answer is right or wrong, I think it was and additional perk of doing the peer review"
- Helps instant revision

  "Because after I submitted the homework, I can do the peer review so just a good time to help me do revision"
- Improves confidence

  "... during ... doing the peer review and I find the student's answer is the same as me, I have more confidence"

A screenshot of PeerMark using Turnitin



Achieving student buy-in to the system

"encouraging people to do more peer review would be as hard as stopping someone else's smoke"

- A well-managed system and motivation from the lecturer is the key
- They'll all participate if the peer marking is a part of some summative assessment (e.g., Cohen and Williams, 2019)
- Getting additional solution available was an effective incentive, specially, if the problem set felt hard to the students

"if we did have the answers anyway then maybe I wouldn't have done it"
"for some hard questions . . . you just actually want to have a look at the right answers"

- Being able to know the marking criteria by doing peer mark was an incentive too

"I think it was a good opportunity for me to see how another student wrote the answers and also to see the marking scheme"

- Students quickly get demotivated if they don't get marked in the first go or the quality is very poor

"Sometimes like you didn't receive the feedback and . . . it will discourage you"

Possible Challenges

- Time involved in initial setup and providing continuous encouragement
- Increased demand from students to have mark scheme for all exercises

  "Give mark scheme for exercise sheets"
- Some students may struggle with peer marking for relatively harder questions. The lecturer needs to encourage them with care and providing additional resources (e.g., referring to specific textbook examples in the given solutions)

  "... people have no idea what the question's about, how they can be even more helpful on marking someone else's script when they can't even understand the questions themselves"

  "... if the students, ... who I do the peer review with, approach in a different way to solve the problem, I may not figure out why he did it in that way, I would take it wrong "

Thank You