

A NEW ONLINE PROCESS FOR ASSESSING INDIVIDUAL CONTRIBUTIONS TO A TEAM PROJECT: WILL IT WORK?

C.K. Woo

DEPARTMENT OF ASIAN AND POLICY STUDIES



香港教育大學

The Education University
of Hong Kong



Agenda

- Overview
- Research questions
- Material and methods

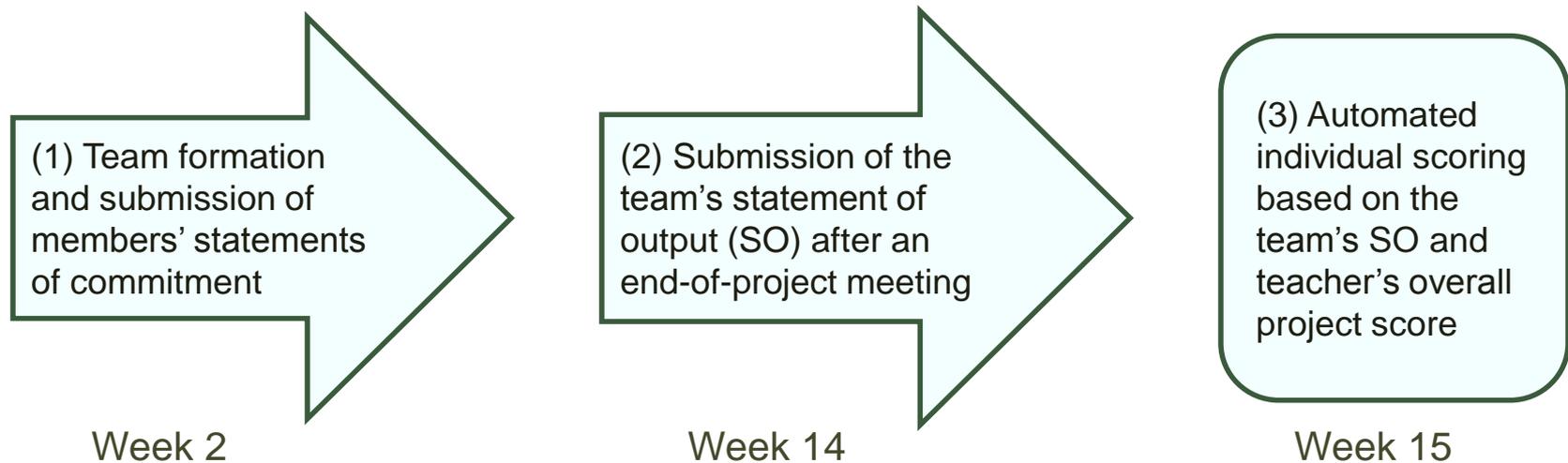


Overview

- Objective: Propose a new online process for assessing individual contributions to a team project, which is often used as a course's learning tool (e.g., a final year project of an undergraduate program)
- This process is fair, easy to use, low cost and perfectly general – a think-outside-the-box solution to a common problem faced by teachers worldwide
- Motivated by a complaint of free riding in my undergrad course: Research Methods, the process integrates my business and litigation experience and applied research in electricity ratemaking, performance-based regulation, and contract negotiations



An online process for a 15-week teaching period



- The current list of interested adopters includes teachers at EdUHK, HK Poly U, UT Austin, and Changzhou University (常州大学)
- Benefits of adoption are: improvement in assessing individual contributions, a new assessment tool, and academic publications with demonstrable research impact
- My questions for you are: (a) what do you think? (b) do you want to try the process?
- Besides an acknowledgement of this presentation, your use of the proposed process for your teaching and research is totally unconditional

Research impact of the proposed process' implementation

Academic publications (even if widely cited) ≠ Research impact

- It promotes the use of individual assessments by addressing student and teacher concerns of free riding and unfair grading, a **worldwide** problem in higher education
- It exemplifies a field-tested integrated approach for consideration by teachers already using or seriously contemplating team projects as an effective learning tool
- It documents a freely available online process for use by teachers **worldwide**, thanks to the process' general applicability that does not depend on student demographics or course attributes
- As such, it answers two key questions:



Research questions

With inter-faculty and inter-university participation, this research project aims to answer:

- What is the current practice of individual assessments in a team project setting?
- Does the current practice comply with the commonly accepted criteria: fairness, transparency, easy implementation, low cost, and achievement of learning goals?
- What is the empirical evidence for the newly proposed process' educational merits?



Examples of current practice: *shortcomings*

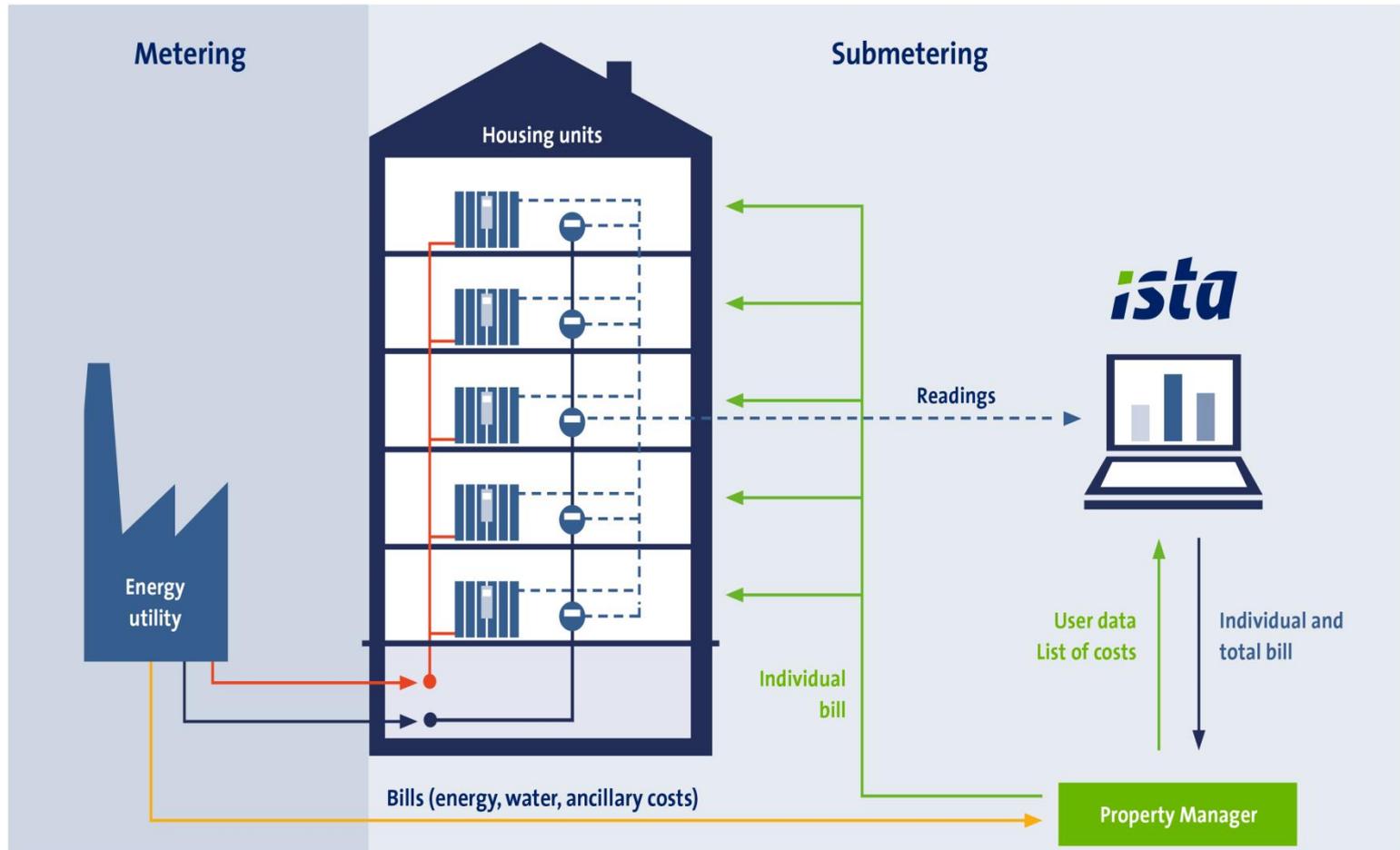
1. Use member-specific intermediate output (e.g., drafts of a power point presentation):
time consuming sans a clear link to the final output that drives a project's overall grade
2. Make an individual member solely responsible for a specific portion of the team project (e.g., the project's literature review): *what if the project's overall quality sucks?*
3. Use a combination of individual assessments based on (1) and (2) and an overall assessment of the team project: *opaque, subjective and hence arbitrary*
4. Adjust the team's overall grade in light of evidence of individual contributions via a declaration of contribution, peer assessment and self-evaluation: *no individual scores*
5. Give the entire team the same grade but prevent free riding by encouraging good team work and close monitoring: *time consuming and no individual scores*

Material and methods

- The Eureka moment
- Electricity ratemaking vs. individual scoring
- Criteria for an acceptable process
- Key components of the proposed process
- Calculation of a team's member-specific scores
- Sensitivity check
- How to determine the process' empirical merits



The Eureka moment



A teacher's scoring of individual contributions to a team project resembles a property manager's allocation of a building's total electricity bill among the individual units.

Electricity ratemaking vs. individual scoring

Electricity ratemaking

- Under master metering, a renter of a 5-unit apartment building has a muted incentive to conserve because the renter pays a fixed *pro rata* share (20%) of the building's total bill
- Solution: individual metering that enables an accurate billing of the renter's own kWh consumption at the applicable tariff
- A sound rate design obeys the Bonbright principles - economic efficiency, fairness: cost-causation and no cross subsidization, low-cost implementation, easy understanding, objectivity, and transparency

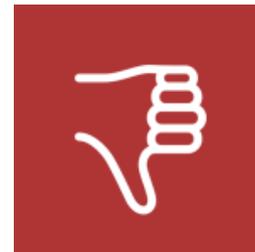
Individual scoring

- Under the single-score-for-all assessment method, a member has the perverse incentive of no penalty of an ill-gotten gain
- Solution: individual scoring, which triggers the questions of how to estimate (record) an individual's contribution (kWh consumption) and how to score (bill) the estimated contribution (at the applicable tariff)
- These questions presage the remarkable but previously unknown connection of electricity ratemaking to academic assessment of individual contributions

Innovative thinking leads to unexpected success (e.g., Viagra based on Pfizer's intended research on hypertension and Post-it Notes based on 3M's failed search for a strong adhesive)

Criteria for an acceptable process

- It should advance a team project's learning goals (e.g., team work), thus precluding an assessment method solely based on individual responsibilities and deliverables
- It should discourage free riding behavior, a direct consequence of the single-score-for-all assessment method's perverse incentive of no penalty for an ill-gotten gain
 - Unintended free ridership: a capable member does most of the work because of the member's concerns of a low grade and other members' slow progress or inability
- It should be easy to implement *sans* costly monitoring and evaluation, thus excluding assessment methods based on each member's log of time spent and intermediate output



Criteria for an acceptable process

- It should be fair: (1) horizontal equity – similar performances imply similar scores; (2) vertical equity – different performances imply different scores; and (3) anonymous equity - individual scores do not depend on student demographics and course attributes, thus imparting fairness among students and across courses
- It should be transparent and objective as required by regulatory governance because an academic assessment method is defined by a set of rules and regulations

By satisfying the above criteria, the proposed process is a superior alternative to an online peer assessment system like CATME, which is difficult to use or WebPA, which is vulnerable to gaming



Key components of the proposed process

1. Statement of commitment (SC): Each member **must**: (a) sign a SC, reflecting the course's expectation of fair contribution and high cooperation; and (b) submit the SCs of all members as part of the team's mandatory declaration of membership (no undue burden)
2. Statement of output (SO): A team's final report **must** include a *pro forma* SO of each member's output contribution to the project's completion and quality as part of the mandatory requirement of team members' reflections (no undue burden)
3. Individual contribution: Based on the SO, the individual contribution of a given member (e.g., A) is a **median** estimate based on **other** members' assessments of A's contribution
4. Individual scoring: Using the results from (3) and a team project's overall score as input, a simple spreadsheet automatically produces a member's individual score

Example of a statement of commitment

As a valuable member of this 5-person team, I commit to actively cooperate and diligently contribute approximately 20% (= 1/5) of the project's deliverables. To ensure the project's timely completion and high quality, my primary areas of responsibility are marked by "✓" below:

- | | |
|--|--|
| <input type="checkbox"/> Topic selection | <input type="checkbox"/> Research plan: what and when to do? |
| <input type="checkbox"/> Literature review | <input type="checkbox"/> Data collection and analysis |
| <input type="checkbox"/> Graphics, tables and charts | <input type="checkbox"/> Discussion of results |
| <input type="checkbox"/> Presentation preparation | <input type="checkbox"/> Final report preparation |
| <input type="checkbox"/> Editing and proofreading | <input type="checkbox"/> Project management and coordination |
| <input type="checkbox"/> Other _____ | |

Name (Student ID) _____; Signature _____; Date _____



Example of a statement of output

Panel A: Initial assessment

Student ID k	Own and others' contributions: C_{km} for $m =$ student ID					
	A	B	C	D	E	Reason
A	0.5	0.2	0.1	0.1	0.1	
B	0.3	0.3	0.2	0.1	0.1	
C	0.2	0.2	0.2	0.2	0.2	
D	0.3	0.2	0.2	0.3	0.0	
E	0.3	0.1	0.2	0.2	0.2	

Panel B: Final assessment determined by consensus _____ or majority vote _____, with changes highlighted in red

Student ID k	Own and others' contributions: C_{km} for $m =$ student ID					
	A	B	C	D	E	Reason
A	0.3	0.2	0.2	0.2	0.1	
B	0.3	0.3	0.2	0.1	0.1	
C	0.2	0.2	0.2	0.2	0.2	
D	0.3	0.2	0.2	0.2	0.1	
E	0.3	0.2	0.2	0.2	0.1	

Process for generating a statement of output

- Step 1: After the project's completion, members hold a 1-hour meeting to reflect on their experiences, including the use of the SCs to frame the bargaining discussion of member-specific contributions.
 - Using documented evidence, each member recaps his/her contributions to the project, thereby discouraging unsubstantiated claims. Each member then announces his/her assessment with reasons of his/her own and other members' contributions.
 - In a 5-person example, A announces $(C_{AA}, C_{AB}, \dots, C_{AE})$, the contributions of members A, B, C, D and E that must sum to 1.0. If A announces a high own contribution (e.g., $C_{AA} = 0.8$), his/her announcement of other members' total contribution is correspondingly low (i.e., $C_{AB} + \dots + C_{AE} = 0.2$).
 - At a team's unlikely request, a teacher attends the meeting as a mediator, akin to the process used in an arbitration. Video recording discourages abusive and collusive behavior. It also generates a complete report of member interactions, useful for resolving a formal complaint and conducting research in teaching and learning.

Process for generating a statement of output

- Step 2: Informed by the results from Step 1 and upon further discussions, each member revises his/her assessments. The members may input their revised assessments into an online spreadsheet to calculate individual scores based on an assumed overall project score (e.g., 80 on a 100-point scale). This informs each member of the revised assessments' impact on individual scores, thus facilitating the bargaining process.
- Step 3: Each member announces his/her final assessments, after seeing other members' revised announcements and the ensuing individual scores. This encourages settlement via peer pressure, desire for consensus, and if necessary, teacher mediation.
- Step 4: If the team fails to reach a settlement despite teacher mediation and video recording, the members' final assessments are to be based on majority voting, which may invite collusion by some members. Hence, a member allegedly injured by the voting outcome can request a formal investigation that requires all members to submit affidavits with evidentiary details, a daunting task that preemptively deters collusive behavior.

What are the benefits of the process?

- The process is perfectly general, applicable to any course of any discipline, thus imparting fairness among students and across courses
- The process is time-efficient because it encourages settlement, thus preempting subsequent complaints of unfair grading that are hard to handle absent its adoption
- Students appreciate the consequences of commitment, diligence, cooperation, interpersonal skills, leadership, resource planning and management, ..., etc.
- Students learn the art of effective negotiation that is supported by sound reasoning and convincing evidence, an important soft skill that prepares students to enter the workplace
- A teacher can use the process' outcome to fairly grade individual contributions in compliance with the overarching principle of horizontal, vertical and anonymous equities
- A teacher can use the process' empirics to publish an academic paper "How to determine individual contributions to a team project: Evidence from **XXXX**"

So many courses, so many papers
So many papers, so little time



Individual contribution calculation for a 5-person team

- A's contribution is $S_A = \text{median of } (C_{BA}, C_{CA}, C_{DA}, C_{EA}) = \text{other members' assessments of A}$. It is less vulnerable to gaming. Suppose A shamelessly gives himself/herself a high assessment. A's dishonest self-claim (e.g., 30% instead of 10%) cannot benefit A because it does not enter into A's calculated contribution.
- Discouragement of gaming may also come from the tit-for-tat behavior of other members: if A gives B a low assessment, B may retaliate with a low assessment of A.
- The median-based calculation is less affected by a member's severely biased assessments of own and others' performances than WebPA's mean-based calculation, thanks to a median being less sensitive to outliers than a mean. A useful analogy is the scoring system for diving in the Olympics Games where the two highest and two lowest scores of seven judges are not used to determine a diver's performance.

Individual score calculation

- A's score is $G_A = \min[(D_A / F) G, \alpha G, 100]$, where $D_A = S_A / (S_A + \dots + S_E)$ = A's adjusted contribution share because S_k may not sum to 1.0; F = equal share = 0.2 for the 5-person team; α = preset scalar > 1 ; and G = overall project score. Thus, (D_A / F) is an estimated extent of A's contribution to G relative to the equal share F .
- G_A is capped at αG or 100 to remedy the odd outcomes in rare but possible scenarios:
 - Scenario 1: The team project's overall score is $G = 20$ and A contributes 100% (i.e., $D_A = 1$ and $D_B = \dots = D_E = 0$). Absent the αG cap, A's final score is 100 [= $(1.0/0.2) 20$], a silly outcome that ignores the project's poor overall quality.
 - Scenario 2: A team's overall score is $G = 90$ and A contributes 100%. Absent the 100 cap, A's project score would be 450 [= $(1.0/0.2) 90$].

How to determine α

- The determination of α may be based on (a) a teacher's view on the maximum individual score as a multiple of G ; or (b) an algorithm that makes the distribution of the teams' averages of individual scores to closely match that of the teams' overall project scores given by the teacher.
- Example of (b): Define $G(j)$ = team j 's overall score and $A(j, \alpha)$ = team j 's equally-weighted average of individual scores for a given α . The variance of $A(j, \alpha)$ around $G(j)$ is $V(\alpha) = \sum_j [A(j, \alpha) - G(j)]^2$. After the process' initial implementation, one can use the course's recorded data to compute $V(\alpha)$ for $\alpha \in \{1.1, \dots, M = \text{maximum multiple of } G\}$. The optimal α is α^* so that $V(\alpha^*) = \min[V(\alpha = 1.1), \dots, V(\alpha = M)]$, an Excel calculation that resembles a bill impact analysis for identifying a rate design *sans* extreme distributional effects.

**A completely solved numerical example:
individual contributions**

Member k	Own and others' contribution assessments by member k				
	A	B	C	D	E
A	0.3	0.2	0.2	0.2	0.1
B	0.3	0.3	0.2	0.1	0.1
C	0.2	0.2	0.2	0.2	0.2
D	0.3	0.2	0.2	0.2	0.1
E	0.3	0.2	0.2	0.2	0.1
S_k	0.3	0.2	0.2	0.2	0.1
D_k / F	1.5	1.0	1.0	1.0	0.5

**A completely solved numerical example:
individual scores**

Overall score G	Individual scores at $\alpha = 1.5$				
	A	B	C	D	E
10	15	10	10	10	5
20	30	20	20	20	10
30	45	30	30	30	15
40	60	40	40	40	20
50	75	50	50	50	25
60	90	60	60	60	30
70	100	70	70	70	35
80	100	80	80	80	40
90	100	90	90	90	45
100	100	100	100	100	50

Does the individual score calculation make sense? Yes

Typical cases: members making contributions of varying amounts

- Case 1: $D_A = 0$ and $G_A = 0$, thus punishing A, a free rider with zero contribution
- Case 2: $D_A = F$ and $G_A = G =$ overall score, reflecting that when A makes a contribution that matches F , A receives G as expected
- Case 3: $D_A > F$, $G_A > G$, thus reflecting that when A makes an above- F contribution, A receives a higher score

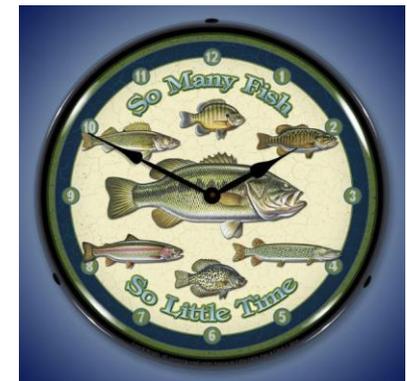
Extreme cases: one member doing the entire project

- Case 4: $D_A = 1$ and $G = 20$, yielding $G_A = 30$ at $\alpha = 1.5$ and $G_B = \dots G_E = 0$.
- Case 5: $D_A = 1$ and $G = 50$, yielding $G_A = 75$ at $\alpha = 1.5$ and $G_B = \dots G_E = 0$.
- Case 6: $D_A = 1$ and $G = 70$, yielding $G_A = 100$ at $\alpha = 1.5$ and $G_B = \dots G_E = 0$.

What are the benefits of the scoring rule?

- A teacher can use the rule to fairly assess individual contributions, thereby obeying the overarching principle of horizontal, vertical and anonymous equities
- It addresses concerns of free riding and unfair grading
- It is perfectly general, applicable to any course of any discipline
- Students appreciate the essence of a well designed incentive scheme that clearly awards a high (low) score for high (low) performance
- Students can use the rule to learn effective negotiation in the end-of-project meeting
- A teacher can use the rule's empirics to publish an academic paper "How to score individual contributions to a team project: Evidence from XXXX"

So many courses, so many papers
So many papers, so little time



Sensitivity check based on a two-person team

Case 1: Two good friends

- Thanks to their friendship, A gives B a high contribution (e.g., 0.6) in the final assessment, and B reciprocates by giving an even higher contribution (e.g., 0.7) to A.
- As $S_A = 0.7$ and $S_B = 0.6$, $D_A = 0.7 / 1.3 = 0.54$; $D_B = 0.6 / 1.3 = 0.46$. Since $F = 0.5$, A's score is $G_A = \min[1.08 G, \alpha G, 100] \geq G_B = \min[0.92 G, \alpha G, 100]$. If the caps are not binding, G_A may exceed G_B by a small amount [= $(1.08 - 0.92) G$]; otherwise, $G_A = G_B$.
- Hence, the process is mildly sensitive to the presence of good friendship. The resulting individual scores, however, mirror mutual consent and are therefore fair.



Sensitivity check based on a two-person team

Case 2: Two lying foes

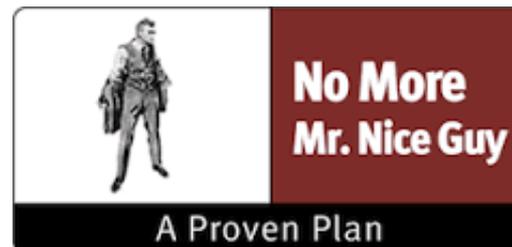


- This case is unlikely when team formation is voluntary and each member has signed a SC.
- Reflecting their mutual dislike, A assigns a low contribution to B (e.g., 0.2) in the initial assessment. B retaliates by assigning an even lower contribution (e.g., 0.1) to A in the final assessment.
- As $S_A = 0.1$ and $S_B = 0.2$, $D_A = 0.1 / 0.3 = 0.33$; $D_B = 0.2 / 0.3 = 0.67$. Since $F = 0.5$, A's score is $G_A = \min[0.67 G, \alpha G, 100] \leq G_B = \min[1.34 G, \alpha G, 100]$. If the caps are not binding, $G_A < G_B$; otherwise, $G_A = G_B$.
- A's mean action may harm A's own score when B retaliates. Reversing the roles of A and B, we infer that B's mean action may harm B's own score.
- With a teacher's mediation and video recording, however, there is an incentive for A and B to moderate their final assessments of each other's contribution, leading to A and B getting reasonable individual scores.

Sensitivity check based on a two-person team

Case 3: Nice A vs. Mean B

- This case is unlikely when team formation is voluntary and each member has signed a SC.
- Nice A honestly gives Mean B a reasonable assessment (e.g., 0.5). However, Mean B maliciously understates A's contribution (e.g., 0.3).
- As $S_A = 0.3$ and $S_B = 0.5$, $D_A = 0.3 / 0.8 = 0.375$; $D_B = 0.5 / 0.8 = 0.625$. Since $F = 0.5$, A's score is $G_A = \min[0.75 G, \alpha G, 100] \leq G_B = \min[1.25 G, \alpha G, 100]$. If the caps are not binding, $G_A < G_B$; otherwise, $G_A = G_B$.
- Mean B's dishonest behavior harms A's score. Recognizing this outcome, A retaliates because being nice does not mean being clueless.
- With a teacher's mediation and video recording, there is an incentive for B to moderate his/her assessment of A's contribution, resulting in A and B getting reasonable individual scores.



Evidence: Collect baseline data at course commencement

Besides student demographics, use the following format to collect the baseline data:

On a scale of 1 (strongly agree) to 5 (strongly disagree), with 6 (unsure) and 7 (not applicable), please opine on the following statements:

- Team projects are useful for achieving my education goals.
- Team projects are a waste of time, without discernible benefits.
- I am aware of free riders in my previously taken courses.
- I am aware of free riders in courses previously taken by other students.
- Free riders lead to unfair grading.
- I have been harmed by free riders.
- Based on my experience with team projects, members tend to contribute equally.
- Based on my experience with team projects, I support a single grade for all members in a given team.
- Based on my experience with team projects, I support individual grades based on member-specific contributions to a team project.

Evidence: Collect comparison data at course completion

- Part 1: Repeat the statements in the last slide to detect any change in student views
- Part 2: Use the following statements in connection to this course:
 - I am aware of free riders in my team project.
 - I have been harmed by free riders in my team project.
 - I am aware of free riders in this course's other team projects.
 - Members have been harmed by free riders in this course's other team projects.
 - Based on the experience with my team project, members tend to contribute equally.
 - Based on my experience with my team project, I support a single grade for all members in a given team.
 - Based on my experience with my team project, I support individual grades based on member-specific contributions to a team project.

A student's responses to the second survey should tie to the same student's responses to the first survey. A possible solution is to give a student the two survey questionnaires with the same anonymous ID number in the first lecture. An alternative is to ask a student to fill in the anonymous ID number in the first survey questionnaire when completing the second survey.

Evidence: Survey teachers at course completion

- On a scale of 1 (strongly agree) to 5 (strongly disagree), with 6 (unsure) and 7 (not applicable), please opine on the following statements in connection to this course:
- The process has advanced a team project's learning goals.
- The process has discouraged free riding behavior.
- The process is easy to implement *sans* costly monitoring and evaluation.
- The process is fair, yielding individual scores that reflect member-specific contributions.
- The process is transparent and objective.
- I will continue to use the process in my future teaching.
- The process needs improvements; please specify _____

The data collected from the three surveys generate the evidence of the proposed process' empirical merits. The data analysis entails graphs and descriptive statistics.

Acknowledgement

- This presentation is based on an on-going paper that I have been writing with Darren Bryant of Education University and Alice Shiu and Christine Armatas of Hong Kong Polytechnic University.
- It benefits from comments of Stephen Cheung, Bob Adamson, Cher Ping Lim, Magdalena Mok and Weng Chung Wang, Education University of Hong Kong; Kim-Sau Chung of Chinese University of Hong Kong; Ira Horowitz, University of Florida; James Smith of Southern Methodist University; and Adonis Yatchew, University of Toronto. It also benefits from discussions with Cat Ng, Ronnie Woo, Sean Lew, and Alisha Woo who share their personal experiences with team projects in their undergrad studies in Hong Kong and North America.
- Without implication, all errors are mine.

C.K. Woo, PhD (Economics, UC Davis)

- He is Professor of Asian and Policy Studies and Associate Dean (Undergraduate Studies and Student Learning, starting Aug-2017) of Faculty of Liberal Arts and Social Sciences at Education University of Hong Kong (<https://www.eduhk.hk/aps/people/prof-woo-chi-keung/>) and Senior Partner (now on leave) of Energy and Environmental Economics, Inc. (www.ethree.com), a 40-person consulting firm headquartered in San Francisco.
- With 30 years of industry experience, he has successfully participated in regulatory proceedings, contract negotiations, and dispute arbitrations in California, Missouri, Texas, British Columbia, Alberta, Ontario, Quebec, and Hong Kong.
- He has published over 140 papers in such scholarly journals as *Energy Policy*, *The Energy Journal*, *Energy Economics*, *IEEE Transactions*, *Journal of Regulatory Economics*, *Energy Law Journal*, *Journal of Public Economics*, *Quarterly Journal of Economics*, and *OMEGA*.
- Recognized by *Who's Who in America*, he is a senior fellow of the United States Association for Energy Economics and an editorial board member of *Energy*, *The Energy Journal* and *Energy Policy*.