RGC Ref. No.: UGC/FDS14/P06/17 (please insert ref. above)

## RESEARCH GRANTS COUNCIL COMPETITIVE RESEARCH FUNDING SCHEMES FOR THE LOCAL SELF-FINANCING DEGREE SECTOR

### FACULTY DEVELOPMENT SCHEME (FDS)

#### **Completion Report**

(for completed projects only)

Submission Deadlines:	1.	Auditor's report with unspent balance, if any: within six months of
	2.	the approved project completion date. Completion report: within <u>12</u> months of the approved project completion date.

### Part A: The Project and Investigator(s)

#### 1. Project Title

Alternative Item Count Techniques for Sensitive Survey Questions: Theories and Methods

### 2. Investigator(s) and Academic Department(s) / Unit(s) Involved

Research Team	Name / Post	Unit / Department / Institution
Principal Investigator	TANG, Man-Lai/Professor	Mathematics, Statistics and Insurance /The Hang Seng University of Hong Kong

#### 3. Project Duration

			Date of RGC /
	Original	Revised	Institution Approval
			(must be quoted)
Project Start Date	01/01/2018		
Project Completion Date	31/12/2020	30/06/2021	24/11/2020
Duration (in month)	36 months	42 months	24/11/2020
Deadline for Submission of Completion Report	31/12/2021	30/06/2022	24/11/2020

#### Part B: The Final Report

#### 5. Project Objectives

- 5.1 Objectives as per original application
  - 1. (New Poisson ICT with Non-Compliance) Develop a new Poisson ICT that takes the non-compliance from the respondents into consideration;
  - 2. (*New Inflated-Zero Poisson ICT*) Develop a new Poisson-type ICT that allows the outcome of the innocuous statement follows the more realistic inflated-zero Poisson distribution;
  - 3. (*Regression Modeling with Sensitive Outcome*) Develop a regression model for binary sensitive outcomes; and
  - 4. (Availability of Free Computer Programs for Practitioners) Implement the proposed methods in non-commercial computer programs and make them available to practitioners.
- 5.2 Revised objectives

Date of approval from the RGC:	9/11/2017
Reasons for the change:	It is difficult to accomplish all the objectives due to budget cut.

- 1. (New Poisson ICT with Non-Compliance) Develop a new Poisson ICT that takes the non-compliance from the respondents into consideration;
- 2. (New Inflated-Zero Poisson ICT) Develop a new Poisson-type ICT that allows the outcome of the innocuous statement follows the more realistic inflated-zero Poisson distribution; and

- 3. (Regression Modeling with Sensitive Outcome) Develop a regression model for binary sensitive outcomes.
- 5.3 Realisation of the objectives

(Maximum 1 page; please state how and to what extent the project objectives have been achieved; give reasons for under-achievements and outline attempts to overcome problems, if any)

Together with my collaborators, we have successfully developed a new Poisson ICT that takes the non-compliance from the respondents into consideration. In addition, we have developed a regression model for binary sensitive outcomes based on this model. The result has been published in Statistics in Medicine (2020, 39, 4480 - 4498) (i.e., fulfilled Objectives 1 & 3). We have considered various randomized response models (i.e., the Warner's randomized response model, unrelated question model, item count technique model and cheater detection model) which are not influenced by the zero-inflation issue. Most importantly, sample size formulas that control, with a given assurance probability, the width of a confidence interval within the planned range were developed. The result has publication Psychometrika been accepted for in (2022,https://doi.org/10.1007/s11336-022 -09854-w) (i.e., fulfilled Objective 2). Simulation studies were conducted to confirm that our proposed estimates perform satisfactorily (i.e., fulfilled Objectives 1, 2 and 3).

Objectives	Addressed	Percentage Achieved
(as per 5.1/5.2 above)	(please tick)	(please estimate)
1. (New Poisson ICT with Non-Compliance)ICT with new Poisson ICT that takes the 	~	100%
2. ( <i>New Inflated-Zero Poisson ICT</i> ) Develop a new Poisson-type ICT that allows the outcome of the innocuous statement follows the more realistic inflated-zero Poisson distribution	~	100%
3. ( <i>Regression Modeling with Sensitive</i> <i>Outcome</i> ) Develop a regression model for binary sensitive outcomes	~	100%

5.4 Summary of objectives addressed to date

#### 6. Research Outcome

6.1 Major findings and research outcome (Maximum 1 page; please make reference to Part C where necessary)

The Poisson item count technique (PICT) is a survey method that was recently developed to elicit respondents' truthful answers to sensitive questions. It simplifies the well-known item count technique (ICT) by replacing a list of independent innocuous questions in known proportions with a single innocuous counting question. However, ICT and PICT both rely on the strong "no design effect assumption" (i.e., respondents give the same answers to the innocuous items regardless of the absence or presence of the sensitive item in the list) and "no liar" (ie, all respondents give truthful answers) assumptions. To address the problem of self-protective behavior and provide more reliable analyses, we introduced a noncompliance parameter into the existing PICT. Based on the survey design of PICT, we considered more practical model assumptions and developed the corresponding statistical inferences. Simulation studies were conducted to evaluate the performance of our method. Finally, a real example of automobile insurance fraud was used to demonstrate our method. These results have been published in *Statistics in Medicine* (2020, **39**, 4480 - 4498).

We reviewed four randomized response models, namely, the Warner's randomized response model, unrelated question model, item count technique model and cheater detection model. These models were considered since they are not influenced by the zero-inflation issue. Noticing that it is an important step in survey sampling to determine the required number of participants, we derive sample size formulas based on confidence interval estimation of prevalence for these models. Specifically, our sample size formulas control, with a given assurance probability, the width of a confidence interval within the planned range.

Simulation results demonstrate that all formulas are accurate in terms of empirical coverage probabilities and empirical assurance probabilities. All formulas are illustrated using a real-life application about the use of unethical tactics in negotiation. These results was accepted to be published in *Psychometrika* (2022, <u>https://doi.org/10.1007/s11336-022-09854-w</u>).

6.2 Potential for further development of the research and the proposed course of action (*Maximum half a page*)

N.A.

#### 7. Layman's Summary

(Describe <u>in layman's language</u> the nature, significance and value of the research project, in no more than 200 words)

One of the most important challenges in modern survey measurement is the elicitation of truthful answers to sensitive questions about behavior and attitudes (e.g., abortion, illegal drug use and racial prejudice). It has long been well known that accessing information regarding a sensitive characteristic in a population usually induces two notorious issues, namely non-response bias (i.e., respondents refuse to collaborate in the fear of the protection of their confidentiality) and response bias (e.g., respondents answer the sensitive questions but give false answers), which usually induce estimate's efficiency loss, inflated sampling variance, and biased estimates. Therefore, techniques that guarantee anonymity, minimize the respondents' feelings of jeopardy, and encourage honest answers are of great demand. In this project, we several models (i.e., Poisson item count technique model, the Warner's randomized response model, unrelated question model, item count technique model and cheater detection model) for sensitive survey questions.

#### Part C: Research Output

8. Peer-Reviewed Journal Publication(s) Arising <u>Directly</u> From This Research Project

(Please attach a copy of the publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The	Latest Stat	us of Public	ations		Title and Journal / Rook	Submitted			
Year of Publicatio n	Year of Acceptanc e (For paper accepted but not yet published)	Under Review	Under Preparation (optional)	Author(s) (denote the correspond -ing author with an asterisk <sup>*</sup> )	Book (with the volume, pages and other necessary publishing details specified) Poisson item count	Submitted to RGC (indicate the year ending of the relevant progress report)	Attached to this Report (Yes or No)	Acknowledged the Support of RGC (Yes or No)	Accessible from the Institutional Repository (Yes or No) Yes (https://resea rch.hsu.edu h
2020				Wu, Q, Tang, M. L.*, Fung, D. W. H. & Tian, G. L.	technique with noncomplia nce, Statistics in Medicine, <b>39(29)</b> , 4480 - 4498. https://doi. org/10.100 2/sim.8736		Yes	Yes (Annex I)	k/project/?pr oject title=A lternative%2 OItem%20Co unt%20Tech niques%20fo r%20Sensitiv e%20Survey %20Questio ns:%20Theor ies%20and% 20Methods )
	2022			Qiu, S.F., <b>Tang, M.</b> L.*, Tao, J. R. & Wong, R.S.	Sample size determinati on for interval estimation of the prevalence of a sensitive attribute under randomized response models, <i>Psychometr</i> <i>ika</i> (https://doi. org/10.100 7/s11336-0 22-09854- w)		Yes	Yes (Annex II)	Yes (( <u>https://rese</u> arch hsu.edu. <u>hk/project/?p</u> roject title= <u>Alternative%</u> 20Item%20C ount%20Tec <u>hniques%20f</u> or%20Sensiti ve%20Surve y%20Questi ons:%20The ories%20and %20Methods )

# 9. Recognized International Conference(s) In Which Paper(s) Related To This Research Project Was / Were Delivered

(Please attach a copy of each conference abstract)

Month / Vear /			Submitted to RGC (indicate the year ending of the relevant progress	Attached to this Report	Acknowledged the Support of RGC	Accessible from the Institutional Repository
Place	Title	Conference Name	report)	(Yes or No)	(Yes or No)	(Yes or No)
N.A.						

# **10.** Whether Research Experience And New Knowledge Has Been Transferred / Has Contributed To Teaching And Learning

(Please elaborate)

No. The results that have been developed are too technical to students in my University.

#### 11. Student(s) Trained

(Please attach a copy of the title page of the thesis)

Name	Degree Registered for	Date of Registration	Date of Thesis Submission / Graduation
N.A.			

#### 12. Other Impact

(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, teaching enhancement, etc.)

N.A.

#### **13.** Statistics on Research Outputs

	Peer-reviewed Journal Publications	Conference Papers	Scholarly Books, Monographs and Chapters	Patents Awarded	Other Rese Output (please spe	earch s cify)
No. of outputs arising directly from this research project	2	0	0	0	Туре	No.

# 14. Public Access Of Completion Report

(*Please specify the information, if any, that cannot be provided for public access and give the reasons.*)

Information that Cannot Be Provided for Public Access	Reasons
N.A.	