RGC Ref. No.:
UGC/FDS14/P01/14
(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 the approved project completion date.*
2. *Completion report: within **12** months of the approved project completion date.*

Part A: The Project and Investigator(s)

1. Project Title

New items count techniques for surveys with sensitive 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 and Statistics/ The Hang Seng University of Hong Kong
Co-Investigator(s)	Nil	Nil
Others	Nil	Nil

3. Project Duration

	Original	Revised	Date of RGC / Institution Approval *(must be quoted)*
Project Start Date	01/12/2014		
Project Completion Date	30/11/2016	30/11/2017	18/05/2017
Duration *(in month)*	24 months	36 months	18/05/2017
Deadline for Submission of Completion Report	30/11/2017	30/11/2018	18/05/2017

Part B: The Final Report

5. Project Objectives

5.1 Objectives as per original application

1. Develop a new ICT for surveys with sensitive questions that produce binary measures;
2. Develop a new ICT for surveys with sensitive questions that produce count measures. Bayesian approach will also be investigated;
3. Develop a new ICT that relaxes the independence and compliance assumptions; and
4. Develop non-commercial computer programs for practitioners. We believe that the proposed models and procedures will complement the existing ICTs, and several high-quality articles will accompany the completion of this project.

5.2 Revised objectives

Date of approval from the RGC:	10/2014
Reasons for the change:	It is difficult to accomplish all the objectives due to budget cut.

1. Develop a new ICT for surveys with sensitive questions that produce binary measures;
2. Develop a new ICT for surveys with sensitive questions that produce count measures. Bayesian approach will also be investigated;
3. Develop non-commercial computer programs for practitioners. We believe that the proposed models and procedures will complement the existing ICTs, and several high-quality articles will accompany the completion of this project.

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)

We will state how and to what extent the above objectives have been achieved as follows:

1. (Objective 1). A Poisson item count technique (PICT) for surveys with sensitive questions that produce binary or quantitative measures has been developed in Liu et al. (2017). Unlike the original item count techniques (ICTs), only one single non-sensitive question (whose outcomes follow Poisson distribution) is required. Besides, the notorious ceiling effect of ICTs was successfully solved.
2. (Objective 2). In Liu et al. (2017), we also considered the non-sensitive question to follow the negative binomial distribution. By doing so, the new negative binomial item count technique (NBICT) can be readily adopted for sensitive question with count measures.
3. (Objectives 1 and 2). It is well known that surveys with sensitive questions usually produce incomplete data. In Poon et al. (2015) and Qiu et al. (2016), we have developed confidence intervals for partially validated series (which is one popular type of incomplete data). We believe that the confidence intervals proposed in this paper can be readily adopted for the sensitive parameters in surveys with sensitive questions.

4. (Objectives 1 and 2). We are planning to build up a non-parametric model for sensitive questions later. For this purpose, we attempted to develop a randomized quantile regression estimation for heteroskedastic non-parametric model in Xiong et al. (2017). We believe that the estimation method developed in this paper will be very useful for modeling sensitive questions.
5. (Objective 3). We have implemented all the proposed methodologies described above in C or/and R. All programs will be available for users upon request.

5.4 Summary of objectives addressed to date

Objectives *(as per 5.1/5.2 above)*	**Addressed** *(please tick)*	**Percentage Achieved** *(please estimate)*
1. Develop a new ICT for surveys with sensitive questions that produce binary measures	✔	100%
2. Develop a new ICT for surveys with sensitive questions that produce count measures. Bayesian approach will also be investigated	✔	100%
3.Develop non-commercial computer programs for practitioners. We believe that the proposed models and procedures will complement the existing ICTs, and several high-quality articles will accompany the completion of this project	✔	100%
4.		

6. Research Outcome

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

This project accomplished all proposed objectives and four articles were accepted or/and published in some international statistics journals. Briefly, a Poisson item count technique (PICT) for surveys with sensitive questions that produce binary and quantitative measures has been developed in Liu et al. (2017). Unlike the original item count techniques (ICTs), only one single non-sensitive question (whose outcomes follow Poisson distribution) is required. Besides, the notorious ceiling effect of ICTs was successfully solved. Besides, we also considered the non-sensitive question to follow the negative binomial distribution. By doing so, the new negative binomial item count technique (NBICT) can be readily adopted for sensitive question with count measures. All proposed methodologies were implemented in C or/and R. All programs will be available for users upon request.

In general, surveys with sensitive questions can be formulated as incomplete data problems. In Poon et al. (2015) and Qiu et al. (2016), we have developed confidence intervals for partially validated series (which is one popular type of incomplete data). We believe that the confidence intervals proposed in this paper can be readily adopted for the sensitive parameters in surveys with sensitive questions.

We are also planning to build up a non-parametric model for sensitive questions later. For this purpose, we attempted to develop a randomized quantile regression estimation for heteroskedastic non-parametric model in Xiong et al. (2017). We believe that the estimation method developed in this paper will be very useful for modeling sensitive questions.

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

All the proposed techniques assume that all respondents will comply with the survey design and are willing to give the truthful answers. This is the famous no liar assumption. It is of future research interest to develop PICT and/or NBICT for sensitive surveys with non-compliance.

It is noteworthy that in both PICT and NBICT it is assumed that the non-sensitive question is assumed to follow either Poisson or negative binomial distributions. To broaden the application of the methods, it is of practical interest to extend the distributions which allow inflated zeros.

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)

Obtaining valid answers to questions that are of a sensitive or embarrassing nature is an age-old problem in survey research (e.g. tax evasion in business investigations; or domestic violence in public opinion research). Non-response (i.e. the respondent's refusal to respond) and response bias (the respondent providing untruthful responses) often occur when sensitive questions are asked directly. These two main sources of non-sampling errors significantly affect sampling estimates and, by extension, statistical inference. In particular, non-response bias results in too small a sample size, which reduces an estimate's efficiency, and response bias results in a biased parameter estimate.

In this project, we develop Poisson and negative binomial item count techniques (for both binary and count sensitive questions) that aim to maximize cooperation, minimize the respondent's sense that he/she is in jeopardy, and guarantee anonymity when sensitive issues are being studied. Programs that implement the proposed methodologies will be made available to practitioners upon request.

Part C: Research Output

8. Peer-Reviewed Journal Publication(s) Arising Directly 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 Status of Publications				**Author(s)** *(denote the correspond-ing author with an asterisk*)*	**Title and Journal / Book** *(with the volume, pages and other necessary publishing details specified)*	**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)*
Year of Publication	**Year of Acceptance** *(For paper accepted but not yet published)*	**Under Review**	**Under Preparation** *(optional)*						
2015				Poon, Wai-Yin, Qiu, Shi-Fang*, and Tang Man-Lai	Confidence interval construction for the Youden index based on partially validated series. *Computational Statistics and Data Analysis*, **84**, 116 – 134.	2016	Yes	Yes	Yes
2016				Qiu, Shi-Fang*, poon, Wai-Yin, and Tang, Man-Lai	Confidence intervals for an ordinal effect size measure based on partially validated series. *Computational Statistics and Data Analysis*, **103**, 170 – 192.	2016	Yes	Yes	Yes
2017				Xiong, Wei, Tian, Maozai, and Tang, Man-Lai	Randomized quantile regression estimation for heteroskedastic non parameteric model. *Communications in Statistics – Therory and Methods*, **46**, 5147 – 5179.	2016	Yes	Yes	Yes

2017				Liu, Yin, Tian, Guo-Liang, Wu, Qin, and Tang, Man-Lai	Poisson-Poisson Item Count Techniques for Surveys with Sensitive Discrete Quantitative Data. *Statistical Papers, DOI 10.1007/s00362-017-0895-7*		Yes	Yes	Yes

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 / Year / Place	**Title**	**Conference Name**	**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)*
29/06/2015 – 03/07/2015 Smolenice, Slovakia	Non-randomized response models for sensitive surveys with noncompliance	PROBSTAT 2015 – the Seventh International Conference on Probability and Statistics	2015	Yes	Yes	Yes

10. Whether Research Experience And New Knowledge Has Been Transferred / Has Contributed To Teaching And Learning
(Please elaborate)

NA

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**
NA			

12. Other Impact

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

NA

13. 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**
NA	

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

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report - Attachment
(for completed projects only)

RGC Ref. No.:	UGC/FDS14/P01/14
Principal Investigator:	TANG MAN LAI
Project Title:	New items count techniques for surveys with sensitive questions: theories and methods

Statistics on Research Outputs

	Peer-reviewed Journal Publications	Conference Papers	Scholarly Books, Monographs and Chapters	Patents Awarded	Other Research Outputs (Please specify)
No. of outputs arising directly from this research project [or conference]	4	1	0	0	0