

Implementation of Online Questionnaire in General Household Survey in Hong Kong

Olivia Or¹

¹ Census and Statistics Department, Government of the Hong Kong Special Administrative Region of the People's Republic of China

Abstract:

The General Household Survey (GHS) of Hong Kong is a monthly household survey on the characteristics of the labour force and household income. All along, data collection of GHS was conducted via face-to-face interviews through field visits and computer-assisted telephone interviews (CATI). Since July 2017, the Census and Statistics Department (C&SD) has started to introduce online questionnaire (OQ) in GHS on a trial basis to enhance its services to survey respondents. Considering the importance of GHS statistics such as unemployment and underemployment rates, the tight monthly cycle of the survey, and the operational and technical complexities involved in adding a new data collection mode on top of the existing ones, prudent change management strategies taking care of the concerns and benefits of all major stakeholders were crucial to ensure successful implementation. Starting with a very small subsample of cases, OQ was introduced by phase along with close monitoring of possible effects on the survey estimates. The trial results were satisfactory, with favourable take-up rate, low error rate in the submitted data and little impact on GHS key statistics. OQ was fully implemented in January 2019 and continued to receive favourable participation from respondents. Being a convenient and privacy-assured channel for provision of data, OQ helps entice sampled households to respond to GHS and increases the contact rate of those households who are more difficult to reach by the conventional data collection methods. The overall response rate of GHS was improved and the average cost of data collection per enumerated case reduced. This paper shares the experience of the successful implementation of OQ as a new data collection mode in a labour force household survey.

Keywords: Change management; Data collection mode; Labour force survey; Household survey; Data quality

1. Introduction:

The General Household Survey (GHS) of Hong Kong is a monthly household survey on the characteristics of the labour force and household income. In each month, a random sample of about 9,000 households is enumerated, of which about half are new cases sampled for the first time and half are repeated cases that had been sampled 3 months ago. All along, data collection of GHS new cases and repeated cases without telephone numbers (including cases which could not be enumerated or the respondents concerned were unwilling to provide telephone numbers in the first round) was conducted via face-to-face interviews through field visits, while repeated cases with telephone numbers were enumerated through computer-assisted telephone interviews (CATI).

With wider penetration and advancement of information technology, adopting online questionnaire (OQ) in statistical surveys has become more feasible in recent years. The successful experience of the 2016 Population By-census in adopting electronic questionnaire demonstrated that the society was ready to respond to statistical survey via online mode and this self-reporting channel could in fact better suit the changing life-style of some households who are more tech savvy, usually return home late and prefer not to be disturbed at home.

In view of this, the Census and Statistics Department (C&SD) started to introduce OQ in GHS on a trial basis in July 2017, as a measure to enhance its services to survey respondents. This paper summarises the experience of the successful implementation of OQ as a new data collection mode in a labour force household survey.

2020 Asia–Pacific Statistics Week

A decade of action for the 2030 Agenda: Statistics that leaves no one and nowhere behind

15-19 JUNE 2020 | Bangkok, Thailand

2. Methodology:

GHS is a monthly survey. Moving-quarter statistics compiled from GHS, including size of labour force, unemployment, underemployment and household income, are vital statistical indicators of the economic condition and are published monthly. It is thus imperative to ensure that data collection remains uninterrupted such that these data series will continue to be available as scheduled. On the other hand, adding a new data collection mode to existing ones is operationally and technically complex. If a full-blown implementation approach was adopted, it would take a long time to get all operational procedures well planned and all staff trained up at the same time, and it would be prone to service disruption when unexpected problems with the new system or operations occurred. In particular, in view of the short data collection period of the survey, any hiccups in the new operations might pose high risks to the timely dissemination of survey results.

In view of the above considerations, it was decided to roll out the OQ on a trial basis first, starting with a small number of cases and then gradually increasing the scale of the trial month by month. The trial was divided into 2 stages. The first stage of trial implementation focused on testing the adequacy of the OQ computer systems and operational arrangements, effectiveness of the OQ questionnaire design and receptiveness of GHS respondents. The second stage involved gradual rollout of OQ to representative sub-samples of the GHS sampled cases of increasing size month by month, focusing on studying the take-up rate and impacts on key survey estimates and manpower resources requirement.

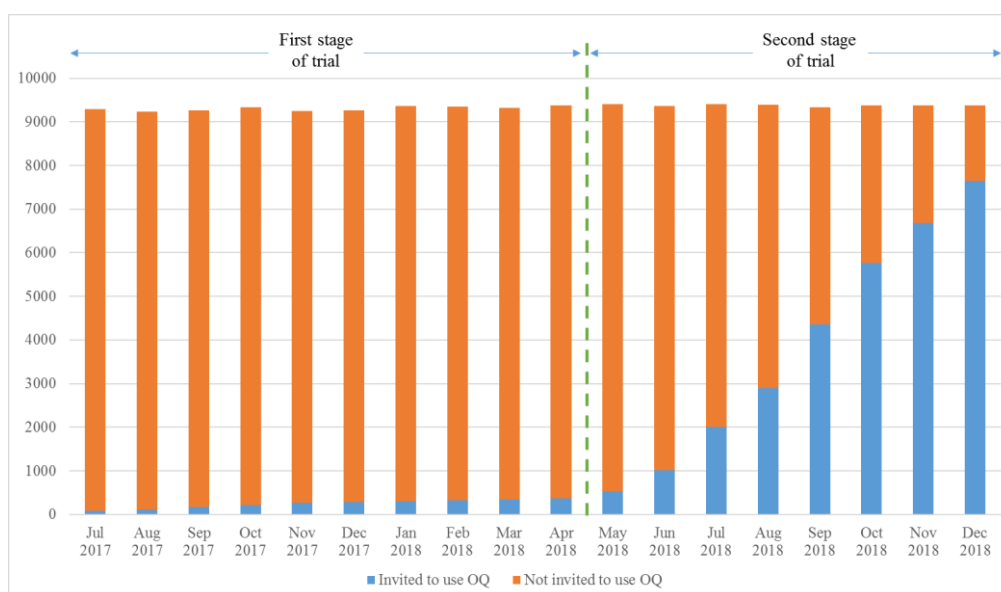


Figure 1: Number of GHS sampled cases invited to use OQ during trial implementation (July 2017 – December 2018)

The first stage of trial implementation lasted from July 2017 to April 2018 for 10 months. Started off with only 81 cases (less than 1% of all sampled cases in a month) invited to use the OQ in the first month of trial, more cases were invited in each subsequent month, reaching 364 cases in the 10th month (about 4% of all sampled cases in a month). The selection of cases was mainly based on field officers assigned to the cases. Only those well-trained ones participated at first, while training was extended to the rest of the field officers progressively. A small scale in the first stage also allowed prompt support to be provided to fix any technical or operational problems encountered. Based on the experience and feedback received from both field officers and respondents, operational procedures and system functionalities were refined, such as simplifying the activation page, modifying the question flow and deriving some data fields automatically based on answers to other questions rather than asking directly for the data.

2020 Asia–Pacific Statistics Week

A decade of action for the 2030 Agenda: Statistics that leaves no one and nowhere behind

15-19 JUNE 2020 | Bangkok, Thailand

The second stage of trial implementation commenced in May 2018 and lasted for 8 months up to December 2018. The monthly sample of the GHS consists of 18 replicates of quarters, each being a representative sample of quarters in Hong Kong. In the first month of the second stage of trial implementation, one replicate was invited to participate in GHS via OQ. Introducing OQ to one out of 18 replicates allowed expanding OQ adoption in a controlled manner. In case something went wrong, the remaining 17 replicates could still be used to compile representative survey results. This by-replicate approach also allowed comparison of OQ replicates and non-OQ replicates to study the impacts of introducing OQ on key survey estimates.

The number of replicates invited to OQ then gradually increased every month, until all 18 replicates were invited in the last month of trial implementation in December 2018 (Figure 1). As the take-up rate was satisfactory and no impact to major labour force statistics was observed, the OQ was eventually fully implemented in January 2019.

It is important to ensure that the data collected via this new data collection mode has high quality as those collected via the usual modes (i.e. face-to-face interview and CATI). During the first stage of trial implementation, all OQ cases were selected for quality checking (QC) in order to fully assess their data quality by verifying the key information (such as employment status) provided on the OQ, and to identify potential areas for improvement in the computer system. Towards the second stage of trial implementation, a QC plan was formulated to select a certain proportion of the OQ cases for QC. The QC plan took into consideration which data items in the submitted OQs are subject to higher risk of reporting errors and thus which categories of the submitted OQs should be subject to higher QC rate, response burden of OQ respondents as well as time and manpower required for the checking.

3. Results:

Take-up of OQ

The take-up of OQ was satisfactory from the beginning of trial implementation. Figure 2 shows the overall take-up rate of OQ (in terms of number of OQ activated cases over all cases invited to use OQ). In the first stage of trial implementation, OQ cases were selected based on field officers, and thus the coverage of cases in terms of geographical areas varied month by month. Therefore, the OQ take-up rate showed fluctuation as some geographical areas with higher elderly population might not use the OQ as much as other areas with more younger households. The average take-up rate in the first stage of trial was 21.4%, notably higher than the planned level.

OQ take-up rate remained above 20% towards the second stage of trial implementation, with an average of 21.3%. This was very encouraging in light that GHS is a voluntary survey with a short OQ reporting period of about 10 days only. The take-up rate was also comparable to those of 20-25% in the Labour Force Surveys of Australia and Canada, which are both mandatory surveys.

After full implementation, the OQ take-up rate was further boosted up with better communications with sampled households who seemed reluctant to take a face-to-face survey. The average OQ take-up rate in 2019 reached 26.0%.

Error rate and enquiries

The error rate in the OQ data was in general low (< 5%) and not significantly different from non-OQ cases. Errors in OQ data were largely attributed to misinterpretation of the concepts involved and were later remedied by fine-tuning the wording of questions and explanatory notes in the OQ. Enquiries received during the trial implementation period also provided valuable information on users' experience. Enhancements were made promptly on the questions, system functionalities and performance. Throughout the trial implementation, the number of OQ-related enquiries (as a percentage of cases invited to use OQ) had been dropping to a very low level as seen in Figure 3.

2020 Asia-Pacific Statistics Week

A decade of action for the 2030 Agenda: Statistics that leaves no one and nowhere behind

15-19 JUNE 2020 | Bangkok, Thailand

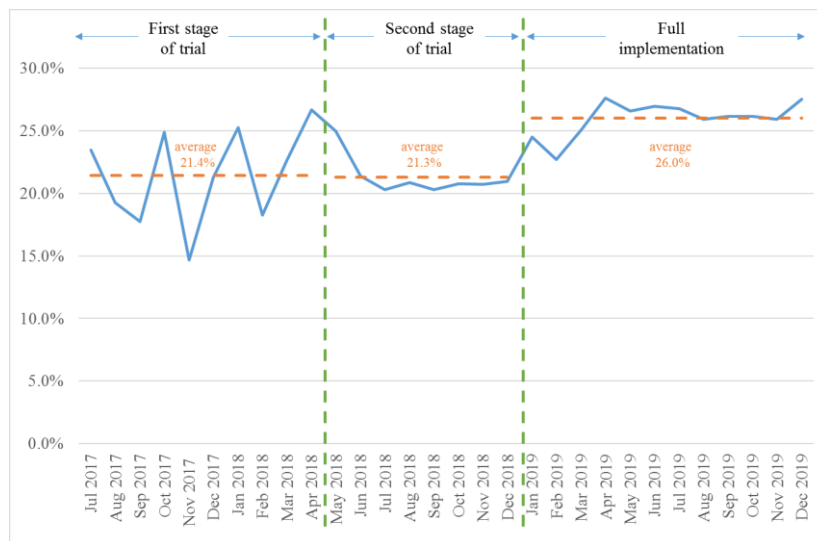
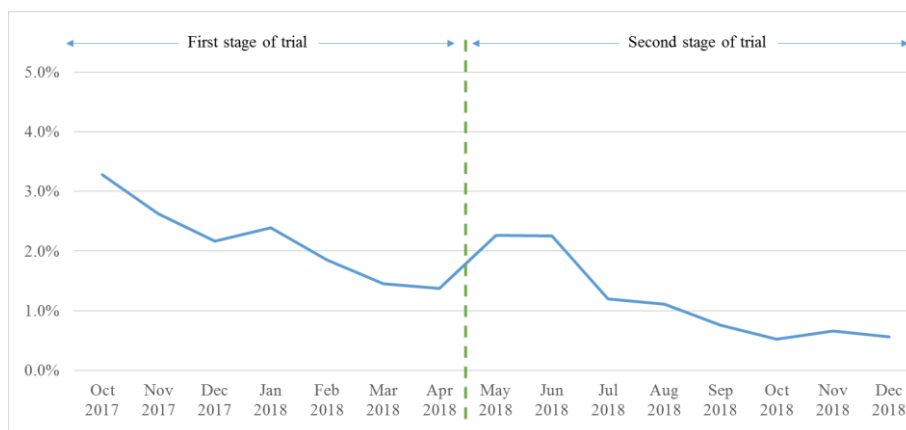


Figure 2: OQ take-up rate during trial implementation and after full implementation (July 2017 – December 2019)



Note: The mild increase in May and June 2018 as the trial entered the second stage was due to system problems, which had been quickly rectified.

Figure 3: Number of OQ-related enquiries as a percentage of cases invited to use OQ during trial implementation (October 2017 – December 2018)

Increase in response rate

Since the introduction of OQ, it was noticed during the trial period that the average response rate of OQ invited cases was consistently higher than that of non-OQ cases by 2-3 percentage points. Apparently, OQ has the added benefit of facilitating data reporting of some sampled households who have difficulties to respond to GHS under the conventional interviewing modes, such as persons being at home only at a late time of the night or weekends (and cannot be reached by our field officers), persons not willing to disclose their personal information to strangers face-to-face and persons feeling intruded when being called or visited by strangers. This helped to improve the overall response rate and data quality of GHS.

Impact to GHS statistics

In the second stage of trial implementation, replicates, each being an independent and representative sub-sample, that were invited to OQ (OQ replicates) were compared with those not invited to OQ (non-OQ replicates) to assess if the introduction of this new mode would pose any impact to the key statistics compiled from the GHS. In general, there were no significant differences observed in the demographic distributions of persons enumerated between the OQ replicates and non-OQ replicates (Figure 4). The

2020 Asia–Pacific Statistics Week

A decade of action for the 2030 Agenda: Statistics that leaves no one and nowhere behind

15-19 JUNE 2020 | Bangkok, Thailand

unemployment rates and underemployment rates compiled with OQ and non-OQ replicates in the moving quarters in the second stage of trial implementation also showed no notable differences upon statistical testing.

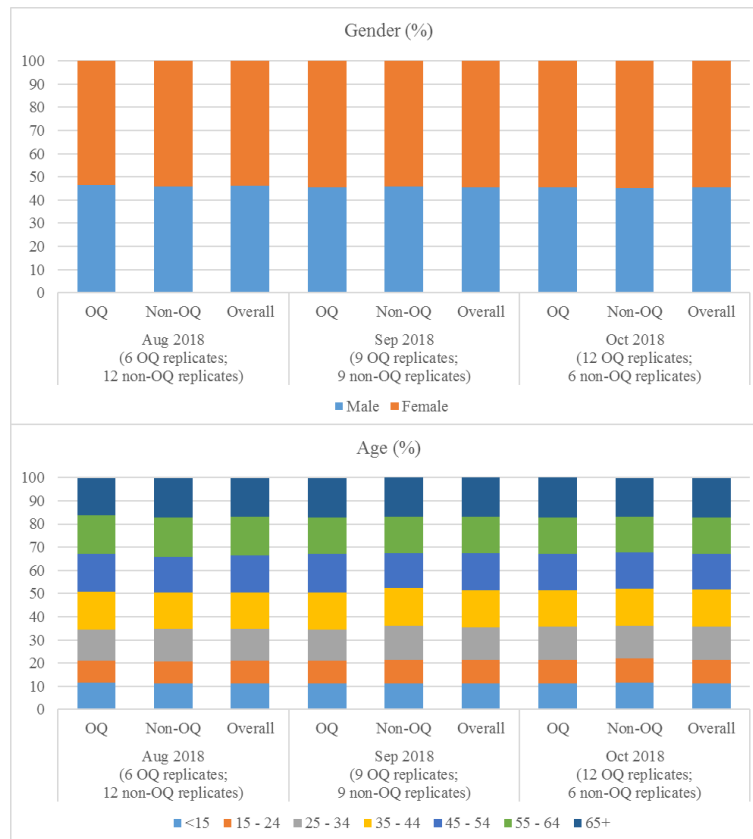


Figure 4: Demographic distributions in enumerated persons in OQ and non-OQ replicates, August – October 2018

4. Discussion, Conclusion and Recommendations:

The implementation of OQ in GHS was successful, with favourable responses from respondents and minimal disruption to monthly survey operations. A number of key success factors in the implementation are identified.

Phased implementation approach

Instead of a full-blown implementation, a long trial implementation period was adopted with gradual increase in the scale month by month. Although it took a long time (1.5 years), impacts to monthly survey operations and statistics compiled were minimised. This approach also allowed a longer time frame for testing and refining the system functionalities, question wordings and flow, as well as the operational procedures involved in incorporating this new mode of data collection.

Training of staff

As the trial period was long, the core team responsible for OQ-related enquiries and support received sufficient training from working with trial cases, and became familiarised with the new system and operational procedures. Those field officers who are not directly responsible for providing support related to OQ were also trained up, so that they understood the implementation plan well, and could answer any questions that respondents might have and promote the use of OQ to respondents who did not prefer to answer the survey face-to-face.

2020 Asia–Pacific Statistics Week

A decade of action for the 2030 Agenda: Statistics that leaves no one and nowhere behind

15-19 JUNE 2020 | Bangkok, Thailand

Well-designed system and procedures with user perspective

All systems and procedures were developed and designed with the mind to facilitate data reporting by GHS respondents using OQ conveniently and confidently. Desktop research and experiments in this aspect had been conducted and reference had been drawn to the experiences of other national statistical offices in the stage of system design to ensure that the interface of the computer system is user-friendly and the question wordings and instructions are suitably tuned from the perspective of the respondents. In addition, all communication materials, such as notification letters, promotion leaflets, and email and SMS messages were carefully crafted to effectively prompt the respondents to use OQ and to provide clear instructions for them.

Data privacy and security

Data privacy and security was one of the main concerns of the public in using the OQ. With this in mind, the GHS OQ system was designed with full regard to data privacy and security, on top of conforming to all relevant government standards and guidelines. Messages to emphasise the security of using OQ were included in various communication materials with respondents where appropriate. This helped strengthen the confidence of the respondents in using OQ. Indeed, respondents had seldom raised concerns on data privacy and security of OQ during the trial period.

Continuous improvement and fine-tuning

The OQ system and working procedures were continuously improved and fine-tuned throughout the trial period, as the phased implementation approach provided scope and time for this. Respondents were asked on their user experience when we conducted data verification, and enquiries on OQ were properly recorded and analysed to identify problems and enhancement opportunities. The GHS OQ system was also designed with maximum flexibility from the onset, such that question wordings and instructions can be enhanced with little effort and time.

Conclusion remarks

OQ effectively provides an additional convenient and round-the-clock channel for the sampled households to respond to GHS. It minimises disturbance to them that may be caused by our field visits and telephone calls. With a phased and gradual implementation approach, meticulously designed computer system and operational procedures, and clear communications with both internal staff and the public, the implementation of OQ was smooth and successful, with no disruption to monthly survey operations and statistics compilation.

This new data collection mode also proved to be particularly useful during the recent COVID-19 outbreak when face-to-face enumeration had to be suspended. Since early February 2020, additional efforts have been put in place to boost up the take-up rate of OQ, including issuing press releases to the public and appeal letters/ reminder letters/ SMSs to sampled cases with clear and targeted message to solicit their participation through OQ. With these new measures, the OQ take-up rate further increased to a record high of 37.7% in the survey month of April 2020. As a result, the overall response rate of GHS, which relied mostly on OQ and CATI during the recent months, could still be maintained at an acceptable level in face of fieldwork suspension. This has enabled the compilation and dissemination of labour force statistics to continue as scheduled without any interruption.

References:

1. Christensen, Ann I., Ekholm, Ola, Glumer, Charlotte, & Juel, Knud. (2013). Effect of survey mode on response patterns: comparison of face-to-face and self-administered modes in health surveys. *European Journal of Public Health*, Vol. 24, No. 2, 327-332.
2. de Leeuw, Edith D., Hox, Joop J., & Dillman, Don A. (2008). *International Handbook of Survey Methodology*. Psychology Press.
3. Hsu, Joanne W., & Mcfall, Brooke H. (2015). Mode effects in mixed-mode economic surveys: Insights from a randomized experiment. *Finance and Economics Discussion Series 2015-008*.