



Office for
National Statistics

**1st Meeting of the Methodological Assurance Review
Panel**

Agenda & Minutes

**15 & 16 October 2018, Office for National Statistics
London**

Office for National Statistics Methodological Assurance Review Panel Meeting

Agenda

15 & 16 October 2018

Drummond Gate London

Chair: Sir Bernard Silverman

Day 1

Time	Item	Presenter
10:30-10:45 15 mins	Welcome & Introductions, role of panel and housekeeping	Bernard Silverman & Rachel Skentelbery
1 10:45-12:15 90mins	Statistical Disclosure Control: Statistical disclosure control (SDC) for 2021 UK Census (EAP101)	Keith Spicer
12:15-13:00 45mins	Lunch	
2 13:00-14:30 90mins	Hard to Count Index: Hard to Count Index for the 2021 Census (EAP102)	Ercilia Dini
14:30-14:45 15mins	Break	
3 14:45-16:15 90mins	Census Coverage Survey: 2021 Census Coverage Survey Design Strategy (EAP103)	Adriana Castaldo
16:15-16:45 30mins	Summaries & Actions	Bernard Silverman & Rachel Skentelbery

Day 2

Time	Item	Presenter
09:30-09:35 5mins	Introduction & Opening	Bernard Silverman
4 09:35-09:55 20mins	Transformation of Population Statistics (EAP104)	Becky Tinsley, Owen Abbott & Peter Jones

5 09:55-11:20 85mins	Population Coverage Survey: Admin Data Census Population Coverage Survey (EAP104)	Peter Jones
11:20-11:30 10mins	Break	
6 11:30-12:55 85mins	Coverage Estimation approach: Coverage Estimation Strategy for the 2021 Census of England and Wales (EAP105)	Viktor Račinskij
12:55-13:40 45mins	Lunch	
7 13:40-15:05 85mins	Census Adjustment: The 2021 Census Coverage Adjustment Strategy (EAP106)	Alison Whitworth
15:05-15:30 25mins	Summaries & Actions	Bernard Silverman & Rachel Skentelbery

Attendee List

External Panel Members

Sir Bernard Silverman (Chair)
Prof Natalie Shlomo (External Panel Member)
Dr Nik Lomax (External Panel Member)
Dr Oliver Duke-Williams (External Panel Member)

Office for National Statistics

Rachel Skentelbery (Vice-Chair, Chief Methodologist)
Cal Ghee (ONS Panel Member)
Owen Abbott (ONS Panel Member)
Sarah Henry (ONS Panel Member)
Keith Spicer (Presenter)
Ercilia Dini (Presenter)
Adriana Castaldo (Presenter)
Becky Tinsley (Presenter)
Peter Jones (Presenter)
Viktor Račinskij (Presenter)
Alison Whitworth (Presenter)
James Redmore (Secretariat)

Apologies

Prof David Martin (External Panel Member)

Actions

Agenda item¹	Action
[1,1]	A1 – ONS to reconsider the policy of not releasing detailed information about SDC parameters.
[1,1]	A2 – A paper on methods transparency should be brought to a future meeting
[1,1]	A3 – SDC team to revise their paper to provide more information on the swapping of non-risky records, the benefits of having two perturbation (pre-tabular and post-tabular) methods and whether differential privacy can be considered for the flexible table builder.
[1,1]	A4 – SDC to consider perturbation at all geographies.
[1,2]	A5 – The HtC team to add international practices and patterns to the paper and consider whether those could be applied to their hard to count index.
[1,2]	A6 – ONS to continue working to obtain data from other suitable sources such as electoral register data and/or commercial sources.
[1,3]	A7 – The Census team to provide information on how homeless populations will be accounted for.
[1,3]	A8 – The CCS team to provide more information on the size of postcodes in the CCS and consider whether subsampling within large postcodes would be feasible.
[1,3]	A9 – The CCS team to complete the research to determine whether the AZTool is providing a good estimate of cost based on validation against actual 2011 CCS costings.
[1,3]	A10 – The CCS team to evaluate theoretically the different sample designs to ensure that the simulations produce the same results, which will provide more evidence on the likely intra-cluster correlations for the cluster designs as well as the sample allocation.
[1,3]	A11 – The CCS team to revise the paper to include more information on international practices of post-enumeration surveys for provision at a future meeting.
[1,5]	A12 – ADC team to present more detailed explanations of how the population spine is being constructed to support administrative data population estimates.

¹ [M,N] – M denotes the panel number, N denotes its position on the agenda.

[1,5]	A13 – ADC team to explore the ability to validate records through the use of wider sources of administrative data, rather than using survey data and dependent interviewing.
[1,6]	A14 – The panel requested that the CCE team provide some academic papers on this approach compared with the DSE approach for 2011 vs 2021.
[1,6]	A15 – The CCE team are to present work on over-coverage estimation at one of the 2019 panels
[1,7]	A16 – The panel requested the CCA team provide more detailed evidence on the combinatorial optimisation approach described in the paper and comparison with the 2011 approach.
[1,7]	A17 – The panel agreed that further testing and development of the CCA strategy is required, including conducting research into the accuracy of imputation and generation of uncertainty measures.
[1,8]	A18 – ONS are to arrange a meeting for a discussion on transparency.

Minutes

1. Statistical Disclosure Control for the 2021 Census – Keith Spicer

Background

This paper outlines the methodology that is proposed to protect the confidentiality of individual respondents to the 2021 Census. The main focus of this paper is the protection of confidentiality within tabular outputs, using a combination of pre-tabular record swapping and controlled random perturbation to output tables.

Discussion & Suggestions

1.1 – Primary concern for this paper was around ensuring transparency. Panel members generally agreed that due to the sensitive nature of the SDC methods this was of the upmost importance, but that ONS should reconsider the policy of not releasing detailed information about SDC parameters. Panel members also agreed that a separate panel discussion on methods transparency was required.

1.2– The panel recommended that there should be swapping of non-risky records as well as risky records.

1.3– The panel recommended that consideration given to perturb *all* tables at every geography, even if that meant inconsistent national totals – it would emphasise to users that they had been protected.

1.4 – Harmonisation of SDC methods across the UK was discussed.

1.5 – The panel discussed whether both types of Statistical disclosure control were necessary. Could either record swapping be sufficient in and of itself, and the same for Cell-Key method?

Actions

A1 – ONS to reconsider the policy of not releasing detailed information about SDC parameters.

A2 – A paper on methods transparency should be brought to a future meeting.

A3 – SDC team to revise their paper to provide more information on the swapping of non-risky records, the benefits of having two perturbation (pre-tabular and post-tabular) methods and whether differential privacy can be considered for the flexible table builder.

A4 – SDC to consider perturbation at all geographies.

2. Hard to Count Index for the 2021 Census – Ercilia Dini

Background

This paper develops a hard to count (HtC) index for the 2021 Census, to be used for both planning the census field activities and as a stratification variable in the Census Coverage Survey. Due to the changing of the 2021 Census to a predominately online data collection a HtC index was developed to reflect that change in identifying areas at risk of non-response. The HtC index methodology allows it to be split into two domains - a Digital Domain and a willingness to respond domain. The methodology used for the latter domain is based on the Index of Multiple Deprivation. This paper identifies areas that are hard to count due to their unwillingness to respond or lack of internet access.

Discussion & Suggestions

2.1 – The panel discussed concerns about how accurate the HtC index was for the digital domain. The panel suggested it needed more data to be sufficiently accurate, although recognised that finding appropriate sources was not easy.

2.2 – The panel were initially confused with how willingness and unwillingness were presented graphically and asked for future presentations to be clearer.

2.3 – The panel were concerned that the model with its current data sources doesn't account for those who choose not to use the internet and recommended exploring further data sources. In addition, panel members asked for further exploration of inputs that were not geographic such as ethnic group and age.

2.4 – The panel discussed the possibility of reducing the bottom 2% of households in the index to something stricter, such as 1%, as this would allow for identifying even harder to count households.

2.5 – There were further discussions by the panel around what interventions could increase response rate in low response areas.

2.6 – The panel suggested HtC consider how to integrate other data and research into both the domains to increase likely accuracy.

Actions

A5 – The HtC team to add international practices and patterns to the paper and consider whether those could be applied to their hard to count index.

A6 – ONS to continue working to obtain data from other suitable sources such as electoral register data and/or commercial sources.

3. 2021 Census Coverage Survey Design Strategy – Adriana Castaldo

Background

This paper is centred around the development of the sample design for the Census Coverage Survey (CCS). The CCS is a survey of approximately 1% of households in England and Wales, carried out independently from the Census and its primary function is to capture and estimate those who have been missed by the Census. The paper proposes a sample design strategy for the 2021 CCS, based on the experience and lessons from the 2011 sample design. The key recommendation is to continue to sample postcodes as the sampling unit.

Discussion & Suggestions

3.1 – Discussion centred around sampling postcodes over addresses, and the problematic elements of sampling postcodes, specifically the need to then survey every address within a chosen postcode. The panel expressed concern for very large postcodes this would be inefficient, and a subsampling approach could be considered.

3.2 – The panel raised concerns about the use of the AZTool and discussed whether or not it is accurate for estimating cost, requesting more research into whether or not it's the best tool to use.

Actions

A7 – The Census team to provide information on how homeless populations will be accounted for.

A8 – The CCS team to provide more information on the size of postcodes in the CCS and consider whether subsampling within large postcodes would be feasible.

A9 – The CCS team to complete the research to determine whether the AZTool is providing a good estimate of cost based on validation against actual 2011 CCS costings.

A10 – The CCS team to evaluate theoretically the different sample designs to ensure that the simulations produce the same results, which will provide more evidence on the likely intra-cluster correlations for the cluster designs as well as the sample allocation.

A11 – The CCS team to revise the paper to include more information on international practices of post-enumeration surveys for provision at a future meeting.

4&5. Transformation of Population Statistics & Administrative Data Census (ADC) Population Coverage Survey – Pete Jones

Background

The ONS is working towards transforming Population Statistics System, including future census population estimates from administrative data. This paper sets out the design considerations for underpinning a Population Coverage Survey (PCS) to support population estimates from administrative data.

Discussion & Suggestions

5.1 – The panel discussed the need for a coverage survey and based on the evidence presented agreed that a coverage survey would be necessary to deliver population estimates of required quality.

5.2 – There was discussion around explaining coverage patterns using alternative sources of data and demographic analysis, in particular the need to deal with over-coverage.

5.3 – ONS should consider commissioning some of the research around over-coverage to academics, for example those engaged in research with other NSIs.

5.4 – The panel requested the ADC team to consider other approaches to adjusting over-coverage emerging in other NSIs.

5.5 – The panel agreed with the principle that a future population coverage survey could be integrated with the redesigned ONS Labour Market Survey.

5.6 – Detailed discussion about dependent interviewing. Panel not averse to the use of admin records in the field to measure over-coverage but need more evidence of the benefits and that other alternatives aren't viable.

5.7 – Discussion of what is the long-term vision for the ONS in terms of surveys and the use of administrative data – where does ONS want to be in the next 10 years with its programme of surveys.

Actions

A12 – ADC team to present more detailed explanations of how the population spine is being constructed to support administrative data population estimates.

A13 – ADC team to explore the ability to validate records through the use of wider sources of administrative data, rather than using survey data and dependent interviewing.

6. Coverage Estimation Strategy for the 2021 Census of England and Wales – Viktor Račinskij

Background

The purpose of this paper is to develop the census coverage estimation (CCE) strategy for the 2021 census. The CCE has four key components: estimation of under-coverage errors; estimation of over-coverage errors; residual bias adjustment; and the coverage estimation of communal establishments. This model is supported by other work such as the census coverage survey (CCS). A successful CCS is imperative for the success of the CCE.

Discussion & Suggestions

6.1 – Panel members were generally impressed with the proposed modelling approach but wanted to see some refinements to the model including deciding on which x-variables would be included in the model, the impact of adding the random effect and clearly defining the model elements.

6.2 – The panel recommended that the CCE team try different modelling techniques under different simulation settings, for example, comparing the use of a Probit model as an alternative to the Logit model that was used.

6.3 – The panel raised concerns about how to present the approach to the public, in order to be transparent. This is an issue for the ONS to consider.

6.4 – The panel also requested ONS consider the best way to publish the model, deciding whether or not parameters from the model should be released to the public.

Actions

A14 – The panel requested that the CCE team provide some academic papers on this approach compared with the DSE approach for 2011 vs 2021.

A15 – The CCE team are to present work on over-coverage estimation at one of the 2019 panels.

7. The 2021 Census Coverage Adjustment Strategy – Alison Whitworth

Background

The purpose of this paper is to describe the census coverage adjustment (CCA) strategy which aims to amend the unit level census database so that it is consistent with the population estimates from the coverage estimation process, and so robust estimates for lower level geographies can be obtained. It builds on the 2011 methodology and the lessons learnt from its implementation, which has practical difficulties.

Discussion points

7.1 – Much of the discussion about this paper centred around weighting vs imputation. Some panel members needed to be convinced as to why this process was necessary, rather than calculating and releasing coverage weights. Due to the necessary number of variables needed and the complex relationships between them this technique is not viable. The panel members agreed that imputing rather than weighting in this process was necessary.

7.2 – Some concerns were also raised by the panel regarding tabulating the results, particularly about how uncertainty would be communicated to users.

7.3 – Questions were raised about the continued role of CANCEIS in the adjustment process. Assurance was given on its role for item imputation.

7.4 – The approach was robustly challenged by the panel members, but after significant discussion panel members felt it was the only feasible approach to this problem.

Actions

A16 – The panel requested the CCA team provide more detailed evidence on the combinatorial optimisation approach described in the paper and comparison with the 2011 approach.

A17 – The panel agreed that further testing and development of the CCA strategy is required, including conducting research into the accuracy of imputation and generation of uncertainty measures.

8. Any Other Business

The panel were pleased and impressed by the scientific sophistication of the work and its approach. The Panel expressed their appreciation for all the hard work and effort presenters have put into their work and papers.

Panel members would appreciate more international comparisons.

The secretariat was asked to ensure that each team is to prepare and send them to the panel members further in advance.

Actions

A18 – ONS are to arrange a meeting for a discussion on transparency.