

Revised Close Report: 2011 Household Survey Sample Redesign

Purpose of the paper

- To provide the 2011 Household Survey Sample Redesign project closure report to the MPS Program Board, highlighting the significance of the project, issues, challenges, innovations, opportunities arising and recommendations for the 2016 Redesign project based on lessons learnt.
- In response to an action item from the 16 August 2013 MPS Program Board meeting: [REDACTED], this revised report covers additional aspects and provides additional detail on aspects included in the previous report.

Desired outcome from the Redesign Board

- Board members agree that this report closes this project.
- Board members support the recommendations.
- Board members identify any further items for the 2016 Redesign project team to consider.

Recommendations

Executive Summary and Key Points of Interest

The 2011 Household Surveys Sample Redesign (2011 Redesign) commencement of delivery was delayed by six months. To minimise cost implications, the MPS sample was fully delivered only two months behind schedule through a compressed four month roll out. This compressed roll out schedule introduced volatility in the sample during the phase-in.

The 2011 Redesign had a number of unique challenges:

- a move from the Australian Standard Geographical Classification (ASGC) to the new Australian Statistical Geography Standard (ASGS).
- for the first time the Dual Master Samples (DMS), also known as 'decoupling', was introduced.
- sample preparation moved from the traditional field listing approach to a list-assisted approach using the [REDACTED]
- PSO had an unusually high interviewer workload in 2012 due to a peak in the Household Survey Program [REDACTED]
- the Labour Force Survey Content Review was being progressed in the LFSS area.
- increased budget pressures across the ABS, and staff turn-over.
- introduction of MPS e-Form coinciding with sample roll-out.

Innovations of note in the 2011 Redesign included:

- use of ESRI ArcGIS software for automated Interviewer checklist routing. The intelligence gathered may lead to Enterprise solutions.
- [REDACTED]
- introduction of [REDACTED] (both a challenge and an opportunity).

Introduction

1 The Household Surveys Sample Redesign ('Redesign') occurs every five years following the Census of Housing and Population. Three areas are most heavily involved in the Redesign:

- Labour Force & Supplementary Surveys (LFSS) has the responsibility "to lead, coordinate and drive the redesign project" [REDACTED] [REDACTED] including key aspects such as working with Household Survey Methodology (HSM) to specify sample design objectives, communication of design outputs and providing secretariat for the various meetings and associated groups;
- HSM provides the methodological design and sample selection; and
- Data Acquisition and Collection Branch (in the 2011 redesign, the Population Survey Program (PSP) section, since disbanded) has carriage of sample preparation and implementation, including specification and project management of associated input/output systems.

All PLaSS Household Survey Centres are also key stakeholders as the redesign develops the frame for all household surveys.

2 As there was considerable overlap between membership of the Redesign and MPS Program Boards, a combined meeting was held on July 1, 2013. At that meeting it was agreed that it would be the last for the Redesign Board, with the MPS Program Board taking carriage of reception and acceptance of the project Closure Report. The Project Manager position ceased in August 2013, so the Close Report was presented to the MPS Program Board on 16 August 2013 [REDACTED] At that meeting the

"Board generally agreed with the content of the report but noted that it does not sufficiently highlight the significance of the exercise, the considerable issues which arose or the impact on other aspects of the program. It was further noted that there have been additional issues with sample phase-in during the current month requiring extra work and testing, and the project is not closed or fully implemented at this stage. It was recognised that the Project Manager position has ceased and only important changes should be made to the report..."

This Revised Close Report is in response to that Board directive.

Key Deliverables

3 The key deliverables for the 2011 Redesign were the complete phase in of the MPS sample by June 2013 [REDACTED]
[REDACTED] The August 2011 Redesign Board meeting approved the MPS [REDACTED] design objectives [REDACTED]

4 The MPS Sample was initially scheduled to be introduced over eight months from November 2012 to June 2013. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Significance of the Project: Issues and Challenges

7 Through the life of the project, there were circumstances and challenges that placed extraordinary pressure on the project delivery, timetable and resources and made the 2011 Sample Redesign unique from previous Redesign projects. Against this backdrop some innovative solutions emerged and, with compromises in the margin of error of the LFS estimates, the Redesign was delivered close to schedule.

Geography: introduction of ASGS

8 Unique complexities for the 2011 Redesign commenced with the requirement for sample area design and selection to move from the [Australian Standard Geographical Classification \(ASGC\)](#) (cat. no. 1216.0) to the new [Australian Statistical Geography Standard \(ASGS\)](#) (cat. no. 1270.0.55.001 to 1270.0.55.006), which resulted in a different geographic structure for the base frame. In August 2010 [REDACTED] the Working Group noted significant concerns regarding Redesign finances and resources, particularly around the introduction of the new geography standard. It was noted that this Redesign was not a simple re-run of the previous Redesign, given the ASGS change and the Sample Management System (SMS) redevelopment (discussed later in this paper).

9 The change in geography added significant complexity, with overlap issues and extra work required. This included:

- an impact on the ability and procedure of selecting a sample with minimal sample overlap between the 2006 and 2011 samples, and
- the geographic size of mesh blocks and/or difficulty identifying their boundaries meant they were not always suitable geographic areas to bound clusters of dwellings for enumeration [REDACTED]

[REDACTED] Methodology had noted concern on this issue early in the ASGS introduction planning [REDACTED]

10 An example of the overlap issues was minimising impact of the NATSIHS (& INPAS) which had been selected under the old ASGC but in the field during the Redesign phase in. This contributed to considerable overlap and difficulty in finding unused sample in NT. The solution included an agreement, due to the NT's high sample fraction, to relax the five year rule to three years for [REDACTED] in NT and to only partially avoid overlap (rather than 100% avoiding overlap) with old sample areas [REDACTED]. The strategy adopted to minimise overlap (discussed further in 'Sample Selection' below) requires some manual effort to manage when specific dwellings rotate into the MPS and [REDACTED] samples over the life of the sample.

Dual Master Samples (DMS)

11 For the first time the Dual Master Samples (DMS) [REDACTED] also known as 'decoupling', was introduced into the Redesign frame. Previously, MPS and [REDACTED] were selected from the same frame, but the decoupling of these two samples will provide more opportunities for household surveys into the future (see 'Innovations and Opportunities' below) but also introduced unique complexities for the project. A population of 12,000 or more was agreed as the cut off point for whether the sample would be decoupled [REDACTED], with a coupled sample continuing in Rest of State Self-Representing Area (SRA) Small Towns, Rest of State SRA Rural, Rest of State non-SRA Least Remote and Rest of State non-SRA Remote areas.

12 The new methodology and design, while delivering potential opportunities for household surveys, also added complexity to the Redesign. The BFUs (Base Frame Units) were larger than anticipated in some rural/remote [non-SRA] areas. To some extent, this was a direct consequence of some mesh blocks being too large to bound clusters. To mitigate this, significant resources were expended by PSP (with Geography) to increase efficiency of within block travel by splitting the BFUs which required additional SIMS and Mapping developments. Mesh blocks were used as the base unit of the sample design, but are not designed for enumeration, and as such PSP has had significant issues during the sample preparation phase identifying boundaries, especially in rural areas. There has also been issues associated with size of some mesh blocks or the extent of mesh block pooling in metro areas. This is discussed further in 'Sample Selection' below.

[REDACTED]

[REDACTED]

[REDACTED]

Sample Selection issues: splitting of BFU's, Growth Strata, Indigenous Community Frame, Census Data, Sample Overlap.

16 As noted earlier, unplanned work relating to splitting of BFUs and mesh blocks took considerable resources by PSP. The 2011 Household Survey Sample Redesign pooled meshblocks to form sampling units. [REDACTED]

[REDACTED] FSUs (First Stage Units) are area selection units for the MPS master sample. In regions with low dwelling density, FSU's were further combined to form FSU2's as an extra stage of selection. A MPS sample in one month comprises one cluster from each selected MPS FSU. FSUs are an aggregate of BFUs so that these units contain sufficient dwellings to last the five year life of the sample design. The formation of these sampling units was performed automatically by HSM using AZTool, [REDACTED]

17 During the office listing process by PSP, it became apparent that some of the BFUs were significantly larger than practical, especially in rural areas. [REDACTED]

[REDACTED] PSP's effort to split the BFUs was challenged by the necessity of splitting along meshblock boundaries rather than the previous Redesigns where PSP would determine boundaries more logical to enumeration. Approximately 550 BFUs required splitting. The meshblock boundaries are not always easy to determine, especially in rural areas. As mentioned earlier in this paper, Methodology had noted concern on mesh block boundaries early in the ASGS introduction planning [REDACTED]. Some meshblocks themselves were also larger than ideal for enumeration. Approximately 300 meshblocks required splitting for enumeration efficiency. Further details are in the paper to the Redesign Board on September 2012: [REDACTED]

18 Growth Strata identification, as in previous Redesigns, is based on information sourced from local and state governments and was extremely variable in quality and timeliness of delivery across states, and required more resources and time to capture and prepare than planned.

19 Late in the Redesign the shortcomings of using recent data from Indigenous communities to model the relationship between estimate quality and sample design parameters were identified. As a result the Indigenous Community Frame used in 2006, and being used as the basis for the 2011 Redesign, was delayed in roll out. [REDACTED]

[REDACTED] The Board also acknowledged the importance of work on further investigations for improving Aboriginal and Torres Strait Islander enumeration, both for Aboriginal and Torres Strait Islander estimates and for NT labour force estimates.

[REDACTED]

[REDACTED]

22 For the first time a Redesign was faced with the prospect of considerable overlap of sample selection in the Northern Territory. This was a result of the continued use of a high sampling fraction in the Northern Territory, first introduced in the 2006 Redesign [REDACTED] and continued in the 2011 Redesign [REDACTED]. In 2001 the sampling fraction for the Northern Territory was [REDACTED]. The NT sample for the 2011 Redesign was approximately 15% larger than that of the previous Redesign, with the resulting issues of respondent burden and sample overlap.

However, balancing constraints of sample size and quality aims, the Board agreed to a [REDACTED] sampling fraction for NT (Agenda Item 5: [Notes Link](#)).

23 The change in geography added another layer of complexity to this issue, with partial overlap between new and previous areas. First identified as an issue at the June 2011 Board meeting [REDACTED] [REDACTED] a preliminary paper was presented in March 2012 [REDACTED]

[REDACTED]

howe followed at this point, with cluster timing management on an adhoc basis and the relaxing of the 5 year Survey Charter being incorrectly attributed to those previously in the MPS in some instances. The PSO systems were expected to have been able to handle this aspect but have not yet had that functionality available, requiring manual work arounds that further strain limited PSO resources, with results that are not always timely nor accurate. See also the 'System Development' section of this paper.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

24 Mid-redesign it was agreed to move sample preparation from the traditional field listing approach to a list-assisted approach using [REDACTED] to align with the ABS commitment to a future adoption of a register based approach to household surveys.

[REDACTED]

[REDACTED]

26 In August 2011 [REDACTED] it was noted to the Redesign Board the possibility of major delays to redesign implementation as many elements of the project including sample frame and maintenance procedures, overlap control methods, systems and documentation were redesigned and rebuilt to accommodate the list-assisted approach. The Board acknowledged the certainty of significant impacts this change would have on the implementation of the Redesign [REDACTED] Population Survey Operations (PSO) system improvements (such as the upgrade from SMS to SIMS) and mapping software were most directly impacted by this change, though an Innovation in mapping and routing software, with possible enterprise benefits, was a result (see ESRI ArcGIS in 'Innovations and Challenges' below). Cost savings associated with the use of a list-assisted frame did not materialise, however knowledge and experience was gained about when and how list-assisted methods can be applied that will be beneficial for the future.

27 A [REDACTED] list-assisted frame was introduced with anticipated savings [REDACTED]

[REDACTED] It was anticipated that in urban areas the interviewer need only receive a list of selected dwellings in the area, and not need to undertake any block listing or check listing related work. Other marginal savings were also expected including the reduction of block listing and check listing training required for interviewers. However, [REDACTED] quality did not meet the pre-Census expectations generated from the initial Geography exploration [REDACTED] This meant that in areas where block-listing was not performed, significant office work was required to produce an initial dwelling list. Following a small QA study of the lists produced from the office-listing procedure and desktop geographical analysis of listing and listing routing quality, it was decided to checklist all areas for the sample roll out, and accordingly there was no immediate cost savings compared to previous frames. Post roll out there is an opportunity to investigate the introduction of an earlier QA process to determine if there are area types that are of sufficient 'good' quality to not checklist and thus harness some of the anticipated savings. Investigation into this opportunity has not yet commenced (see Innovations and Opportunities below). There was also uncertainty on the direction of the Redesign (list-based or list-assisted or area-based) until early 2012. The late decision making contributed to the delayed roll out of the sample. A QA process to identify areas of sufficient quality to not check-list was proposed in September 2012 [REDACTED] [REDACTED], but as predicted by [REDACTED] [REDACTED] the resources were insufficient to act on the plan. Further details on the opportunity in this space is noted in the 'Innovations and Opportunities' section of this paper.

[REDACTED]

Accelerated Roll-out

37 There was significant work associated with the decision to accelerate the roll-out of the new sample from eight months to four months. This required additional work on system changes to allow an additional rotation group to be rotated in at the same time, changes to the RSEs during the roll-out including a new SE model and the potential impact on the estimates and their volatility during the new sample roll out. To keep some costs in the 2012-13 budget year, efforts were made to flip-in the SD areas of the sample in June 2013, with the balance in July.

38 Under previous Redesigns the two flip-in months spread the sample roughly equally. However due to time constraints in June, the flip-in areas were less than optimal in the 2011 Redesign, with only the smaller areas (TAS, ACT, NT) being flipped in June, and the rest of the sample in July. This added further complexity in the July data. In addition, the increased rotation resulted in a larger margin of error on the estimates during the period, and the flip in of all of the large states in one month brought the matched sample down to around 60% (usually around 80%) thus increasing the margin of error while also having a potential impact on regional estimates. This is not the preferred position in a usual Redesign roll out and should be avoided in the future.

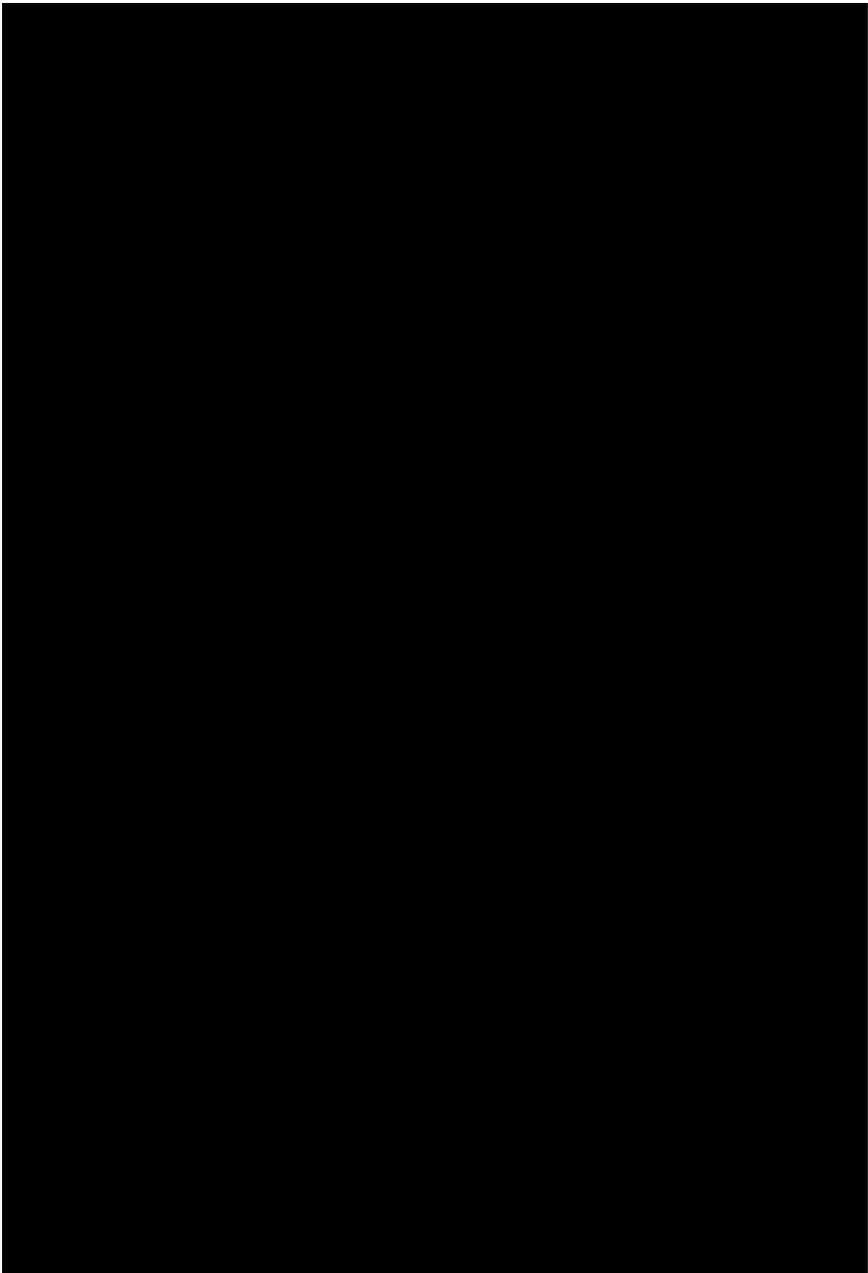
[REDACTED]

Environmental Scan and Workload Impacts

[REDACTED]

40 In this period electronic self-enumeration (eForm) was also being introduced for the MPS, diverting further resources within the LFSS and PSP teams. In addition, the LFS Content Review was being progressed in the LFSS. Coupled with all of these pressures was increased budget pressures, directly resulting in a reduction in the LFSS Redesign project team from the usual 2-3 FTE to a single team member for the last 14 months of the project, as well as significant turn over in the HSM and PSP teams working on the Redesign.

41 Given the length of the project and the Environmental Scan, it is of little surprise that the Redesign had resource and staffing issues. [REDACTED]

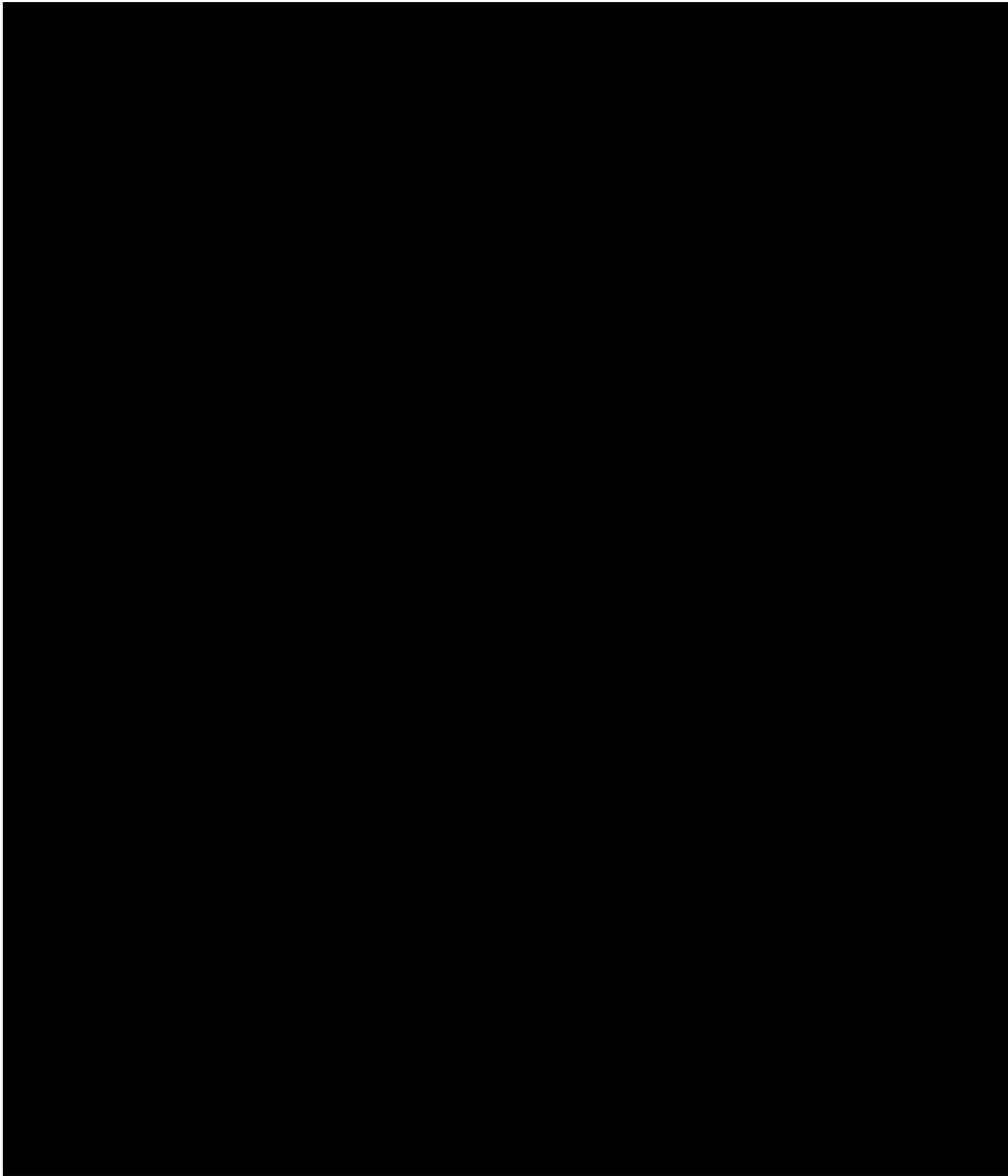


44 The automated results were still individually reviewed as part of quality assurance but much less manual work was needed with a smaller number of changes in listing required. PSO Redesign staff have estimated that the original process of blocking the 2011 Redesign took 15-30 minutes set-up for each block. The new process is estimated at 5-10 minutes, or less. Accordingly there is a significant improvement and saving from this innovation.

45 Beyond the Redesign, the use of ESRI software has potential ABS enterprise benefits. There is now a wider interest for using the Network Analyst software (an extension of ArcGIS used in conjunction with a Navteq road network dataset) to assist with providing efficient travel routing for both households survey interviewers and Census field officers.

[REDACTED]

[REDACTED]



[REDACTED]

[REDACTED]

[REDACTED] generated cost saving by removal of check listing:

50 As noted earlier, based on preliminary reports on the quality of [REDACTED] data the 2011 Redesign had expected that there would be a cost saving through removal of some PSO block listing and check listing in the majority of areas (predominantly urban). Once [REDACTED] was integrated into the Redesign it became apparent that [REDACTED] quality did not meet earlier expectations and block listing and check listing was still required. However, in an effort to retain some of these savings and ensure annual sample maintenance costs remain stable, DACB will raise with HSM the need for a project to identify areas where check listing will not be required due to sufficient [REDACTED] quality, based on the first six months after the commencement of the roll-out of the new sample.

Sample decoupling - [REDACTED] design flexibility:

51 The 2011 Household Sample Redesign introduced a dual master sample framework, whereby a fixed set of areas provide sample for MPS [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Future efficiency for PSO:

55 The new SIMS system is being integrated with the PSO automated workload allocation system (AWAT) input/output workload data, providing increased administrative efficiency of Interviewer allocations in the future including reduced effort in preparing inputs to AWAT.

Targeted follow-up strategies:

56 The move to a sample built from finer level sampling units (mesh blocks) provides better support for analyses and methods which attempt to use information on the Census characteristics of the sample areas (e.g. targeted follow-up strategies). This opportunity could be considered in Responsive Design work. [REDACTED]

Areas explored but discarded or did not materialise:

57 A number of options were considered for the Redesign but after consideration either discarded, shelved or did not materialise:

- Less frequent enumeration of various MPS subpopulations. This was discussed and considered by the Redesign Board [REDACTED] [REDACTED]. The Board was hesitant to support reducing frequency of Indigenous Community Frame (ICF) selections as the ICF population changed LF status more frequently than other parts of the population. The Board also discussed that operationally, the impacts of reducing enumeration on collection activity may be considerable with relationships with Communities weakened with less frequent contact. SD options were to be further investigated. In the end, this discussion was taken off the Redesign and given to the Household Surveys Review Steering Committee (HSRSC) on [REDACTED] [REDACTED]. [REDACTED] of these discussions is now being considered again by both the Population and Social Statistics Futures review [REDACTED] and the NCATSIS holistic program review [REDACTED]. The SD component is being considered by the Population and Social Statistics Futures review.
- Sub-state estimates. Raised in September 2010 [REDACTED] [REDACTED] the Board affirmed in November 2010 that sub-state estimates should not drive the Redesign but should be considered as an option "to potentially meet user-funded requirements for sub-state data in future (such as top-ups)" [REDACTED].
- Relaxing Equal Probability of Selection (EPS) constraints within a State. Traditionally the Redesign sample has had an EPS constraint. Discussed in March 2012 (Agenda Item 5: [Notes Link](#)), the Board agreed that "due to the minimal financial savings, and the potential quality implications of relaxing the constraint" the EPS constraint would be retained.



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]